

2 Flächenorientierte Segmentierung

Unter **Bildsegmentierung** versteht man die Zerlegung eines Bildes in sogenannte Regionen. Diese sollen Objekten der abgebildeten realen Welt, oder zumindest Teilen dieser Objekte entsprechen. Das Ergebnis der Segmentierung ist eine **symbolische Beschreibung** des Bildinhaltes anhand der Regionen und ihrer Merkmale. Diese Beschreibung kann dann von den nachfolgenden Verfahren weiterverarbeitet werden.

Die Segmentierung eines Bildes ist einer der wesentlichen Verarbeitungsschritte bei der automatischen Auswertung eines Bildes. Dies gilt von einfachen industriellen Anwendungen (z.B. Erkennung von Objekten auf einem Fliesband) bis hin zu anspruchsvollen Bildinterpretationsaufgaben in natürlicher Umgebung.

Im weiteren werden nur Methoden vorgestellt, die ausschließlich auf der Basis der Bildinformation, d.h. der Bildmatrix arbeiten. Andere Segmentierungs-Ansätze, die Zusatzwissen beispielsweise über den Bildinhalt (z.B. "Klötzchen-Welt") oder über die dem Szenenaufbau zugrunde liegende Semantik machen (z.B. "Himmel ist oben"), werden hier nicht betrachtet.

Bevor die Verfahren näher erläutert werden, soll zunächst mit Hilfe eines **Homogenitäts-Prädikates P** die Definition für die Segmentation eines Bildes angegeben werden.

(2-1)

Definition 2.1: Die **Regionen** R_i bilden bezüglich des Homogenitäts-Prädikates P eine Segmentation des Bildes B, wenn die nachfolgenden Bedingungen erfüllt sind:

- (α) $R_i \neq \emptyset$ und R_i zusammenhängend für alle i und $R_i \cap R_j = \emptyset$ für alle i, j mit $i \neq j$;
- (β) $\cup R_i = B$ (über alle i);
- (γ) $P(R_i) = \text{true}$ für alle i ;
- (δ) $P(R_i \cup R_j) = \text{false}$ für alle i, j ($i \neq j$), R_i und R_j benachbart.

Das Homogenitäts-Prädikat P stellt dabei das eigentliche Kriterium zur Abgrenzung der Regionen dar. Es kann sich z.B. auf Grau- bzw. Farbwerte oder auf Textureigenschaften (siehe Kapitel 4) beziehen und wird üblicherweise mit Hilfe eines Toleranz-Schwellwertes realisiert.

Für die Bildsegmentierung gibt es (noch) keine geschlossene Theorie, aber eine Vielzahl unterschiedlicher Vorgehensweisen. Dennoch lassen sich die Segmentierungsverfahren (und entsprechend die Repräsentationen der Regionen) in zwei grundlegend unterschiedliche, duale Ansätze aufteilen:

- 1.) Bei der **konturorientierten Segmentierung** werden die Übergänge zwischen den Regionen, also die Inhomogenitäten zwischen unterschiedlich eingefärbten oder texturierten Bildbereichen ermittelt. Bei dieser Vorgehensweise werden in der Regel zuerst Kantendetektoren eingesetzt, und anschließend wird versucht, aus den gefundenen Kantenelementen die geschlossenen Konturen der Regionen zu rekonstruieren. Die Regionen können unmittelbar mit Hilfe ihrer Konturen repräsentiert werden.
- 2.) Bei der **flächenorientierten Segmentierung** werden homogene Bildbereiche zu Regionen zusammengefasst. Die Regionen werden also über ihr "Inneres", d.h. die Menge aller zugehörigen Pixel ermittelt. (Beispielsweise können alle gleich eingefärbten, zusammenhängenden Pixel zu einer Region zusammengefasst werden.)

Vom Prinzip her sind diese beiden Ansätze gleichwertig und sollten dasselbe Ergebnis liefern. In der Praxis hat es sich aber gezeigt, dass in Abhängigkeit von der Aufgabenstellung und den Gegebenheiten im Bild oft einer der beiden Ansätze geringeren Aufwand und/oder bessere Ergebnisse verspricht.

Bei den flächenorientierten Segmentierungsverfahren, auf die in diesem Kapitel näher eingegangen wird, unterscheidet man die merkmalsbasierten Bildpunktklassifizierungs-Verfahren (siehe Kapitel 2.1) und die topologisch basierten Flächenwachstums- bzw. Verschmelzungs- sowie die Zerteilungs-Verfahren (siehe Kapitel 2.2 bzw. 2.3).

2.1 Bildpunktklassifizierungs-Verfahren

Bei der Gruppe der **Bildpunktklassifizierungs-Verfahren** werden je Bildpunkt bestimmte Merkmale gemessen, wie z.B. Grau- oder Farbwert. Bildpunkte mit ähnlichen Merkmalswerten werden in Klassen zusammengefaßt; z.B. die Klasse der weißen Bildpunkte. Die zusammenhängenden Mengen von Bildpunkten gleicher Klassenzugehörigkeit bilden die Regionen der Segmentation.

2.1.1 Bildpunktklassifizierungs-Verfahren mit einer festen Schwelle

Einfache Bildpunktklassifizierungs-Verfahren basieren auf **Binarisierungs-** bzw. **Schwellwert-Verfahren** für Grauwertbilder. Bei diesen Verfahren wird ein Schwellwert S ausgewählt, und die Bildpunkte (i,j) werden nach folgender (Binarisierungs-)Vorschrift in zwei Klassen eingeteilt, die im folgenden durch zwei Grauwerte g_1 und g_2 repräsentiert werden:

$$g'(i,j) = \begin{cases} g_1 & \text{falls } g(i,j) \leq S \\ g_2 & \text{sonst.} \end{cases} \quad (2-2)$$

Das Ergebnis der Binarisierung ist ein Binärbild; siehe Abbildung 2.1. Abschließend müssen bezüglich der beiden Klassen, d.h. der Grauwerte g_1 und g_2 die Zusammenhangskomponenten (ZHKs) gebildet werden; vgl. "Integrierte Lehrveranstaltung". Die Zusammenhangskomponenten bilden die Regionen der Segmentation. (Entsprechendes gilt für die Kapitel 2.1.2 und 2.1.3).

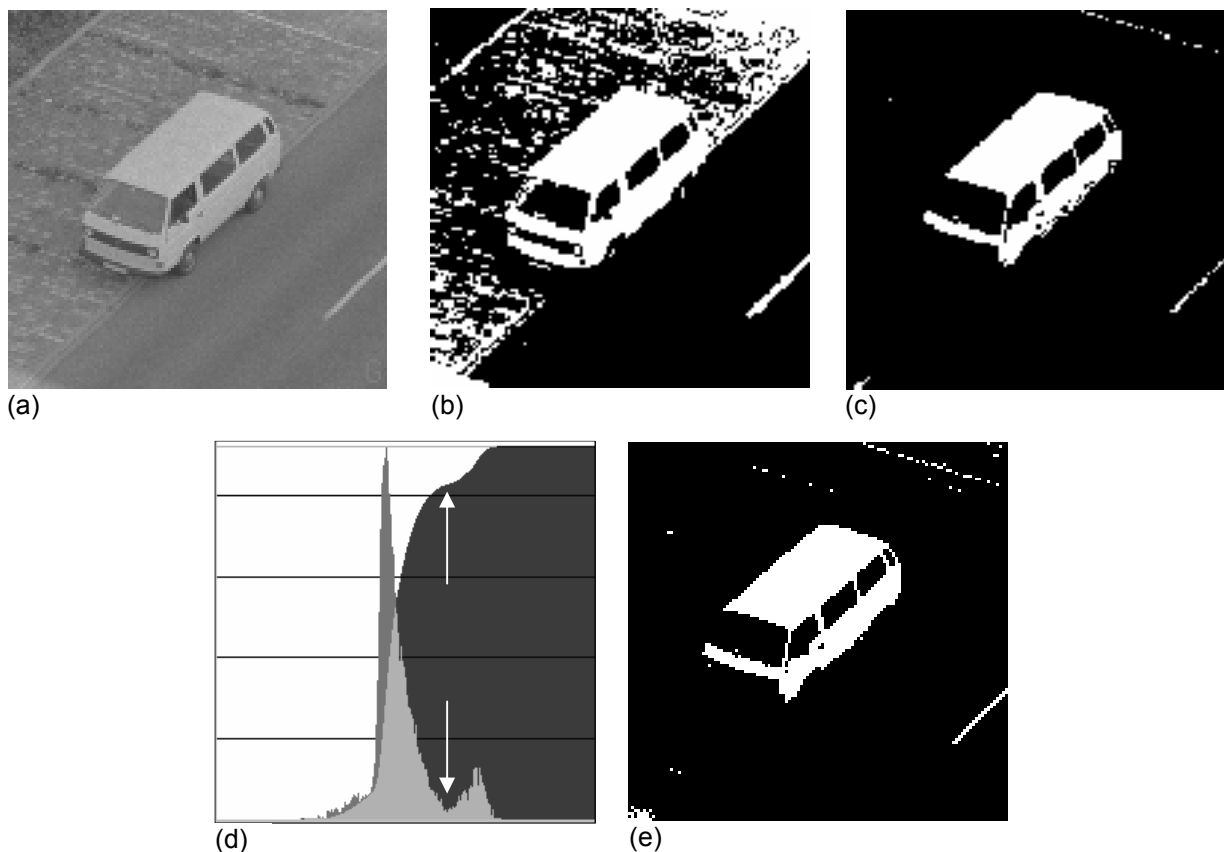


Abb. 2.1: Beispiel für eine Binarisierung: (a) Originalbild; (b) Binarisierung mit zu niedrigem Schwellwert; (c) Binarisierung mit zu hohem Schwellwert; (d) Überlagerung von Histogramm und kumulativen Häufigkeiten; (e) optimal binarisiertes Bild.

Bei dem beschriebenen Vorgehen ist die Wahl des Schwellwertes S entscheidend für das Ergebnis. Wenn der Schwellwert zu niedrig gewählt wird, so ist es zwar wahrscheinlich, dass z.B. interessierende helle Objekte vollständig vorhanden sind, aber es können sich im Umfeld

sehr viele Störungen abzeichnen, die zum Teil sogar mit den Objekten verschmelzen; siehe Abbildung 2.1 (b). Wird der Schwellwert andererseits zu hoch gewählt, so verschwinden zwar die meisten Störungen in der Umgebung der Objekte, aber in den Objekten können Löcher auftreten und wesentliche Teile können fehlen; siehe Abbildung 2.1 (c).

Zur Ermittlung eines günstigen Schwellwertes zieht man häufig das Histogramm $h(g)$ des Bildes heran; siehe Abbildung 2.1 (d) mittel- und hellgrau. (Zur Erinnerung: Die Histogramm-Funktion $h(g)$ gibt zu jedem Grauwert g an, wie oft er im Bild auftritt).

Bei zahlreichen Anwendungen ist das Grauwert-Histogramm **bimodal**, d.h. es besteht im wesentlichen aus zwei Peaks, die durch ein Tal getrennt sind. Dieses Tal ist in der Regel eine gute Wahl für den Binarisierungs-Schwellwert S ; siehe unterer Pfeil in Abbildung 2.1 (d).

Die automatische Suche nach dem Histogramm-Tal erfolgt aber zweckmäßigerweise nicht im Histogramm $h(g)$, sondern in den daraus abgeleiteten kumulativen Häufigkeiten:

Die **kumulative Häufigkeit** (auch **Summenhäufigkeit**)

$$k(g) = \sum_{\tilde{g} \leq g} h(\tilde{g}) \quad (2-3)$$

gibt zu jedem Grauwert g an, wie oft alle Grauwerte, die kleiner oder gleich diesem Grauwert sind, im Bild auftreten. Mit $h(\tilde{g})$ wird dabei die Häufigkeit des Grauwertes \tilde{g} , d.h. der Histogramm-Eintrag an der Stelle \tilde{g} bezeichnet.

Die "flachste" Stelle in den kumulativen Häufigkeiten (die weder am Anfang noch am Ende des Spektrums liegt) gibt die gesuchte Stelle an. Zu diesem Zweck bildet man die Differenzen von nicht unmittelbar benachbarten Positionen der kumulativen Häufigkeiten und sucht das Minimum über alle diese Differenzen $d(i)$:

$$\min_{i \in [i_0, g_{MAX} - i_0]} d(i) := |k(i - i_0) - k(i + i_0)| \quad (2-4)$$

Die Konstante i_0 , die festlegt, wie weit die zu vergleichenden Positionen der kumulativen Häufigkeiten auseinander liegen, wird üblicherweise auf einen Wert zwischen 2 und 5 gesetzt. Unter Umständen kann es sinnvoll sein, den Funktionsverlauf der kumulativen Häufigkeiten vorher einer eindimensionalen Glättung zu unterziehen.

Abbildung 2.1 (d) zeigt neben dem Histogramm auch die kumulativen Häufigkeiten des Eingangsbildes in hell- bzw. dunkelgrau. Die flachste Stelle in den kumulativen Häufigkeiten, die genau über der tiefsten Stelle des Histogramm-Tales liegt, ist durch einen Pfeil markiert. Die Binarisierung mit diesem Schwellwert liefert das Binärbild aus Abbildung 2.1 (e).

Ein anderer Weg zur Bestimmung eines günstigen Binarisierungs-Schwellwertes S wird z.B. in Sonka et al. "Image Processing, Analysis and Machine Vision" Kapitel 5.1.1 geschildert.

Verfahren dieser Art heißen **Binarisierungsverfahren mit einer festen Schwelle**. Die Festlegung der Schwelle ist ein globales Kriterium, das räumliche Zusammenhänge nicht berücksichtigt. Das Homogenitäts-Prädikat aus der Definition 2.1 bedeutet hier: "Zugehörigkeit zum selben Histogramm-Peak".

2.1.2 Bildpunktklassifizierungs-Verfahren mit einer variablen Schwelle

In vielen Fällen ändert sich z.B. die Beleuchtung innerhalb einer abgebildeten Szene (etwa Beleuchtungsabfall zum Rande), so daß mit einer festen Schwelle keine befriedigenden Ergebnisse erzielt werden können. Hier kann das Bild in regelmäßige rechteckige Teilbilder zerlegt werden. Je Teilbild wird ein Schwellwert bestimmt. Um beim Übergang von einem Teilbild zu einem anderen einen abrupten Wechsel des Schwellwertes zu vermeiden, wird ein "weicher" Übergang durch Interpolation zwischen den verschiedenen Werten von S erreicht. Verfahren dieser Art werden als Binarisierungsverfahren **mit einer variablen oder dynamischen Schwelle** bezeichnet.

2.1.3 Bildpunktklassifizierungs-Verfahren mit mehreren festen Schwellen

Wenn das Histogramm eines Grauwertbildes mehrere deutlich ausgeprägte Peaks enthält, spricht man von einem **multimodalen** Histogramm. In solchen Fällen kann eine monoton steigende Folge von Schwellenwerten S_1 bis S_n in den Tälern zwischen den Peaks definiert werden. Eine entsprechende Erweiterung der Binarisierungsvorschrift (2-2) liefert eine Einteilung der Bildpunkte in $(n+1)$ Klassen. Segmentationsverfahren dieser Art heißen **Mehr-Schwellen-Verfahren**.

Abbildung 2.2 zeigt ein Beispiel für ein Mehr-Schwellen-Verfahren mit zwei Schwellenwerten. Mit der ersten Schwelle (linker Pfeil im Histogrammbild 2.2 (b)) erhält man im Bild den dunklen Kanal sowie z.B. Schattenflächen neben den Häusern; siehe Abbildung 2.2 (c) in schwarz. Eine Schwellwertbildung mit dem zweiten Schwellwert (rechter Pfeil) liefert die Flächen der hellen Gebäude, die im Bild enthalten sind; siehe Abbildung 2.2.(c) in weiß.

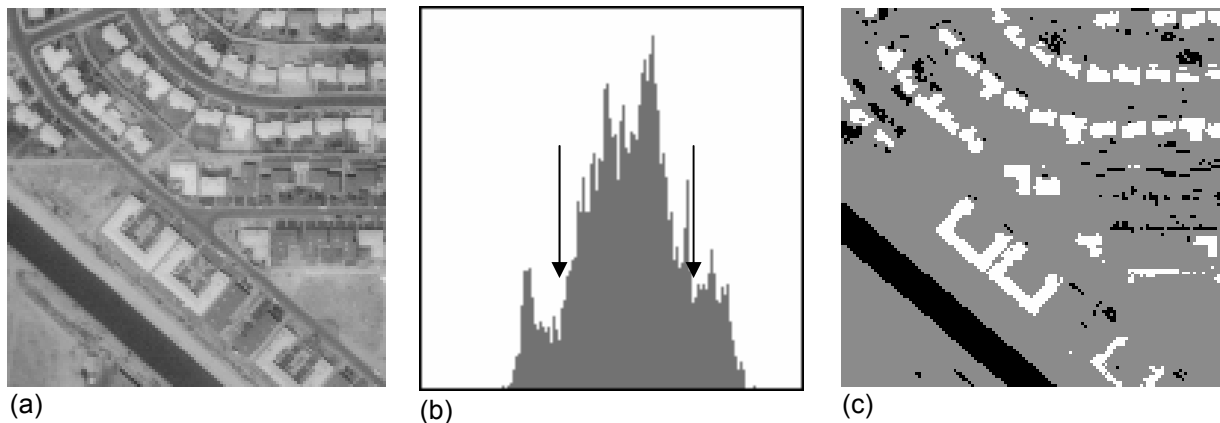


Abb. 2.2: Beispiel für ein Mehr-Schwellen-Verfahren mit zwei Schwellenwerten.

Schlussbemerkung zu Kapitel 2.1: Alle bislang geschilderten Verfahren waren **eindimensionale Klassifizierungs-Verfahren**, da der Merkmalsraum nur die Grauwerte erfasste, also eindimensional war. Die Verfahren lassen sich prinzipiell auf mehrdimensionale Merkmalsräume erweitern, wenn statt der Grauwerte Farbwerte (z.B. R, G, B) oder neben den Grauwerten der einzelnen Bildpunkte ein weiteres Merkmal, wie z.B. die Grauwert-Varianz innerhalb einer kleinen Umgebung jeden Bildpunktes miterfasst wird.

Da der Rechenaufwand bei der Auswertung mehrdimensionaler Histogramme im Vergleich zur erreichbaren Verbesserung aber unverhältnismäßig stark ansteigt, werden für höhere Dimensionen (≥ 2) in der Regel andere Klassifizierungs-Verfahren eingesetzt, die in einem späteren Kapitel beschrieben werden.

2.2 Flächenwachstums-Verfahren

Im Gegensatz zu den nach einem globalen Kriterium vorgehenden Bildpunktklassifizierungs-Verfahren arbeiten die **Flächenwachstums-Verfahren (region growing)** nach lokalen oder regionalen Kriterien. Dabei spielt die gegenseitige räumliche Anordnung der Bildpunkte bei der Bildung der Regionen eine entscheidende Rolle.

Der Grundgedanke bei allen Flächenwachstums-Verfahren ist der, benachbarte einzelne Bildpunkte oder auch einen einzelnen Bildpunkt und eine benachbarte Fläche zusammenzufassen, wenn sie ähnliche Merkmalswerte aufweisen. Auf diese Weise werden sukzessive die Regionen der Bildsegmentation aufgebaut.

Die Verkettung zweier benachbarter Bildpunkte bzw. die Anlagerung eines Bildpunktes an eine bereits ermittelte Fläche erfolgt dabei in Abhängigkeit von dem in Definition 2.1 aufgeführten Homogenitäts-Prädikat. Beispiele für solche Prädikate sind:

- punktbezogenes Prädikat: zwei benachbarte Bildpunkte X und Y haben einen ähnlichen Grauwert, d.h. wenn für einen Schwellwert S gilt:

(2-5)

$P(X \cup Y) = \text{true}$ genau dann, wenn $|\text{GW}(X) - \text{GW}(Y)| < S$;

- umgebungsbezogenes Prädikat: zwei benachbarte Bildpunkte haben bzgl. einer jeweiligen kleinen Umgebung ähnliche Merkmalswerte (z.B. mittlerer Grauwert, Grauwert-Varianz);
- flächenbezogenes Prädikat: ein Bildpunkt und eine (bereits ermittelte) benachbarte Fläche haben ähnliche Merkmalswerte (z.B. Ähnlichkeit des Bildpunkt-Grauwertes und des mittleren Flächen-Grauwertes). Entsprechendes gilt für zwei benachbarte Flächen.

Viele Flächenwachstums-Verfahren verarbeiten ein Bild systematisch in einem Durchlauf z.B. von links oben nach rechts unten. Ein aktuell zu bearbeitender Bildpunkt wird dabei entweder einer bereits ermittelten (links und/oder oben liegenden) Fläche zugeschlagen, oder er initialisiert eine neue Fläche.

Andere Flächenwachstums-Verfahren verteilen sogenannte Keimzellen im Bild. Durch sukzessive Anlagerung benachbarter Bildpunkte wachsen diese Keimzellen, bis keine "ähnlichen" Bildpunkte mehr angelagert werden können und damit die Regionen der Segmentation ermittelt sind. Bei diesen Methoden muss sichergestellt werden, dass alle Bildteile von den Keimzellen aus erreichbar sind.

Im folgenden wird zunächst eine Methode vorgestellt, die eine Mischform aus sukzessiven Wachstumsverfahren und Keimzellen-Wachstumsverfahren darstellt.

2.2.1 Hysterese-Schwelwertverfahren

Hysterese-Schwelwertverfahren stellen eine Verknüpfung von "normalen" Schwelwertverfahren, die global arbeiten, und Flächenwachstumsverfahren, die bei ihrem Wachstum den räumlichen Zusammenhang von Bildpunkten berücksichtigen, dar.

Die Hysterese-Schwelwertverfahren werden im weiteren für die Segmentierung heller Objekte vorgestellt. Eine Segmentierung dunkler Objekte ergibt sich entsprechend.

Die Grundidee des Verfahrens zur Segmentierung heller Objekte basiert auf zwei Annahmen:

- Nur wirklich helle Pixel gehören mit Sicherheit zu den interessierenden Objekten.
- Etwas dunklere Pixel gehören nur dann zu einem interessierenden Objekt, wenn sie an bereits ermittelte Objektflächen angrenzen.

Für die konkrete Umsetzung des Verfahrens benötigt man zwei Schwellwerte S_1 und S_2 mit $S_1 > S_2$. Mit diesen beiden Schwellwerten verfährt man wie folgt:

- 1.) Zunächst wird das Bild systematisch zeilenweise abgearbeitet, bis ein erster Bildpunkt auftritt, für dessen Grauwert g gilt: $g > S_1$. Dieser Bildpunkt wird als **Keimzelle** der ersten Zusammenhangskomponente definiert und erhält die Marke 1.
- 2.) Die N4- bzw. N8-Nachbarn der Keimzelle werden untersucht, ob für ihren Grauwert g gilt: $g > S_2$. Falls dies der Fall ist, so erhält der betreffende Punkt die Marke der zugehörigen Keimzelle und die Untersuchung wird (rekursiv) auf die N4- bzw. N8-Nachbarn dieses Punktes ausgedehnt. Auf diese Weise wächst die Zusammenhangskomponente, bis keine Bildpunkte mit einem Grauwert $g > S_2$ mehr angelagert werden können.
- 3.) Nun wird die systematische zeilenweise Suche nach einer neuen Keimzelle für die nächste Zusammenhangskomponente fortgesetzt. Es wird also der nächste Bildpunkt gesucht, der noch nicht markiert ist und dessen Grauwert $g > S_1$ ist. Dieser Bildpunkt erhält die nächste freie Marke; er wird als Keimzelle definiert und die Auswertung wird mit 2.) fortgesetzt.
- 4.) Das Verfahren endet, wenn der systematische Durchlauf die rechte untere Bildecke erreicht. Abschließend werden noch alle nicht markierten Bildpunkte auf die Hintergrund-Marke 0 gesetzt.

Das Verfahren liefert eine Zerlegung des Bild-Vordergrundes mit den interessierenden hellen Objekten in durchnummerierte Zusammenhangskomponenten. Der Bildhintergrund (mit der Marke 0) ist nicht notwendig eine Zusammenhangskomponente, so dass das Ergebnis noch keine Segmentation des Bildes im Sinne von Definition 2.1 darstellt. Zu diesem Zwecke müsste erst noch eine Zusammenhangskomponenten-Markierung für den Hintergrund durchgeführt werden.

Abbildung 2.3 zeigt ein Beispiel für die Anwendung eines Hysterese-Schwelwertverfahrens. Im Gegensatz zum Ergebnis der Standard-Binarisierung aus Abbildung 2.1, die hier in Abbildung 2.3 (b) noch einmal gezeigt ist, liefert das Hysterese-Schwelwertverfahren nur die Bus-Karosserie, wobei hier sogar noch weitere Bildpunkte korrekt ergänzt wurden.

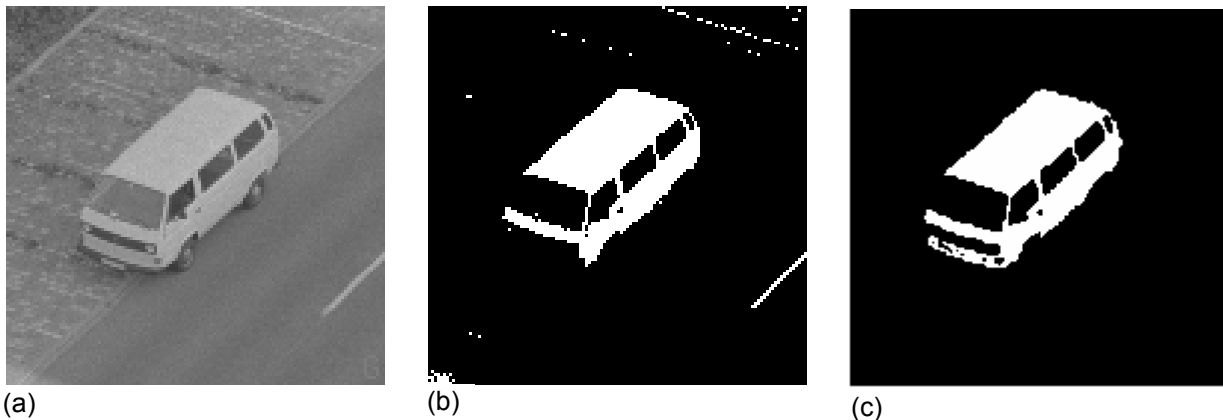


Abb. 2.3: Beispiel für die Anwendung eines Hysterese-Schwelwertverfahrens im Vergleich zu einer Standard-Binarisierung: (a) Originalbild; (b) Ergebnis einer Standard-Binarisierung; (c) Ergebnis eines Hysterese-Schwelwertverfahrens.

2.2.2 Systematisches Flächenwachstumsverfahren

Das systematische Flächenwachstumsverfahren stellt eine Verallgemeinerung des Verfahrens zur Markierung von Zusammenhangskomponenten (ZHKs) dar; vgl. "Integrierte Lehrveranstaltung".

In einem ersten Durchgang wird das Bild zeilenweise abgearbeitet:

- 1.) Das erste Pixel initiiert die erste ZHK und erhält die vorläufige Marke 1. Danach wird zum rechten Nachbarpixel gewechselt und dieses wird als aktuelles Pixel gesetzt.
- 2.) Die oberhalb bzw. links (innerhalb des Bildes) liegenden N4- bzw. N8-Nachbarn des aktuellen Pixels werden untersucht, ob ihre vorläufige Marke auf das aktuell untersuchte Pixel übertragen werden kann. Dazu wird anhand des Homogenitäts-Prädikates überprüft, ob das aktuelle Pixel zu einem seiner Nachbarn bzw. der durch diesen Nachbarn vorläufig markierten Fläche "passt".
 - Wenn dies nicht der Fall ist, so erhält das aktuelle Pixel eine neue vorläufige Marke.
 - Wenn dies für genau einen Nachbarn der Fall ist, so übernimmt das aktuelle Pixel die vorläufige Marke dieses Nachbarn.
 - Wenn dies für mehrere Nachbarn mit verschiedenen Markierungen der Fall ist, so wird eine dieser Marken für das aktuelle Pixel übernommen und alle betroffenen Marken werden als äquivalent vermerkt.
- 3.) Als nächstes wird das rechte Nachbarpixel untersucht und dieses wird als aktuelles Pixel gesetzt. Falls das Zeilenende erreicht ist, wird an den Beginn der Folgezeile gesprungen. Weiter mit 2.).
- 4.) Der erste Durchgang endet, wenn die rechte untere Bildecke erreicht ist.

Es folgt ein vollständiger zweiter Durchgang durch das Bild, in dem die Äquivalenzen ausgewertet werden und alle Pixel, die äquivalente vorläufige Marken tragen, dieselbe end-

gültige Marke erhalten. Weiterhin wird darauf geachtet, dass diese endgültigen Marken eine lückenlose (bei 1 beginnende) Folge von Marken bilden.

Als Ergebnis erhält man ein in seine Zusammenhangskomponenten zerlegtes Bild, das eine Segmentation des Bildes im Sinne von Definition 2.1 darstellt.

Abbildung 2.4 zeigt ein Ergebnis des systematischen Flächenwachstumsverfahrens unter Verwendung der N8-Nachbarschaft. Es wurde ein flächenbezogenes Homogenitäts-Prädikat verwendet, bei dem ein Pixel die vorläufige Marke eines Nachbarn übernimmt, wenn der Grauwert des betrachteten Pixels nur geringfügig vom mittleren Grauwert aller bereits mit dieser Marke vorläufig markierten Pixel abweicht. (Nach jeder Zuweisung einer Marke muss der zugehörige Mittelwert angepasst werden).

Zunächst wurde das ursprüngliche Bild einer Glättung unterzogen, um im Rahmen dieses Beispiels nicht zu viele, nachher graphisch kaum unterscheidbare Marken zu erhalten; siehe Abbildung 2.4 (a). In Abbildung 2.4 (b) ist das Ergebnis des ersten Durchgangs gezeigt. Die Farb- d.h. Markenwechsel innerhalb des Busdaches finden überall dort statt, wo die Kontur einen "Sprung" um mindestens zwei Pixel nach links macht, so dass die untersuchte Nachbarschaft des "vordersten" Dach-Pixels vollständig außerhalb des Busses liegt. Abbildung 2.4 (c) zeigt das Ergebnis des zweiten Durchgangs mit dem vollständig segmentierten Bild.

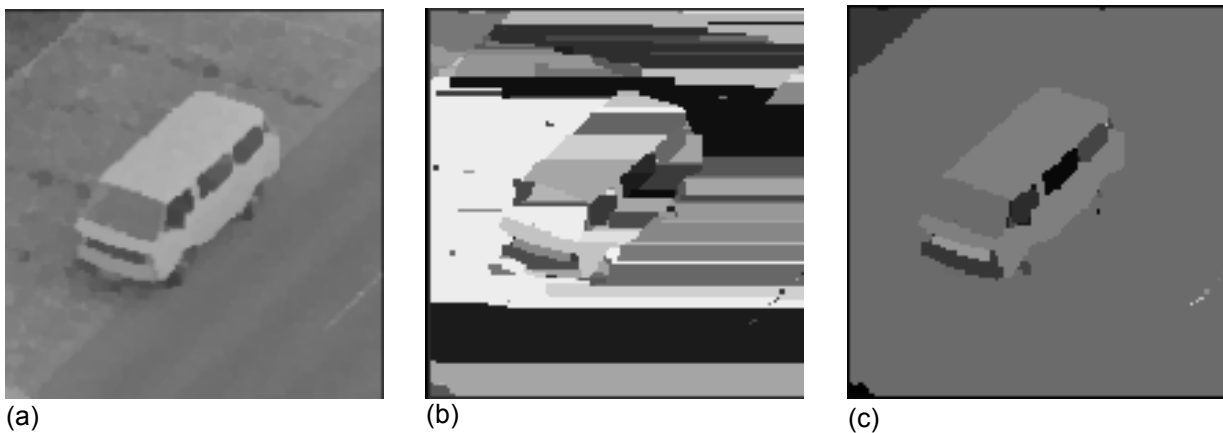


Abb. 2.4: Beispiel für die Anwendung des systematischen Flächenwachstumsverfahrens: (a) Originalbild; (b) Ergebnis des ersten Durchgangs; (c) Ergebnis des zweiten Durchgangs mit dem vollständig segmentierten Bild. (Siehe auch Farbbild-Anhang am Ende der Datei zu diesem Kapitel).

Schlussbemerkung: Statt eines systematischen zeilenweisen Vorgehens hätte man auch hier wieder einen rekursiven Ansatz wählen können, bei dem jedes "neue" Pixel in alle Richtungen wächst, bis unter Verwendung des Homogenitäts-Prädikates kein weiteres Pixel mehr angelagert werden kann. Dann wird wieder zeilenweise nach einem "neuen", d.h. noch nicht markierten Pixel gesucht; usw.. Wenn man die "neuen" Pixel der Reihe nach durchnummeriert, so erhält man wie oben eine Segmentation des Bildes mit einer lückenlosen Folge von Marken.

2.2.3 Keimzellen Wachstumsverfahren

Beim Keimzellenwachstum werden die Keimzellen z.B. systematisch im Bild verteilt. Jede Keimzelle erhält eine eindeutige vorläufige Marke.

In einem ersten Schritt werden alle Nachbarpixel von Keimzellen untersucht, ob sie anhand des Homogenitäts-Prädikates an die Keimzellen angelagert werden können. In diesem Falle erhalten die angelagerten Pixel die vorläufige Marke der zugehörigen Keimzelle.

Dieses Vorgehen wird für die Nachbarn der neu markierten Pixel fortgesetzt, usw., bis kein Pixel mehr vorläufig markiert werden kann.

Abschließend wird mit Hilfe des Homogenitäts-Prädikates untersucht, ob benachbarte, vorläufig markierte Bildbereiche zusammengefasst werden können. Es wird wieder darauf geachtet, dass sich eine lückenlose Folge von endgültigen Marken ergibt.

Falls bei dem geschilderten Vorgehen alle Pixel des Bildes einer Keimzelle zugeordnet wurden, so stellt das Endergebnis eine Segmentation des Bildes dar. (Es kann vorkommen, dass in kleinen, homogenen Bildbereichen keine Keimzelle gesetzt wurde und diese Bereiche daher unmarkiert bleiben, da sie im Rahmen des Wachstumsprozesses keiner Keimzelle zugeordnet werden können).

Abbildung 2.5 zeigt verschiedene Schritte des Keimzellenwachstumsverfahrens. In Abbildung 2.5 (a) ist das Bild mit der systematischen Überlagerung der Keimzellen zu sehen; an die Keimzellen wurde bereits eine Pixelschicht angelagert. In Abbildung 2.5 (b) ist der Wachstumsprozess beendet, da keine Bildpunkte mehr angelagert werden können. Abbildung 2.5 (c) zeigt das Endergebnis nach der Verschmelzung benachbarter "passender" Bildteile. Die weißen Bildstellen wurden keiner Zusammenhangskomponente zugeordnet; hier saß weder eine Keimzelle noch konnte eine Keimzelle in diese Bereiche "hineinwachsen".

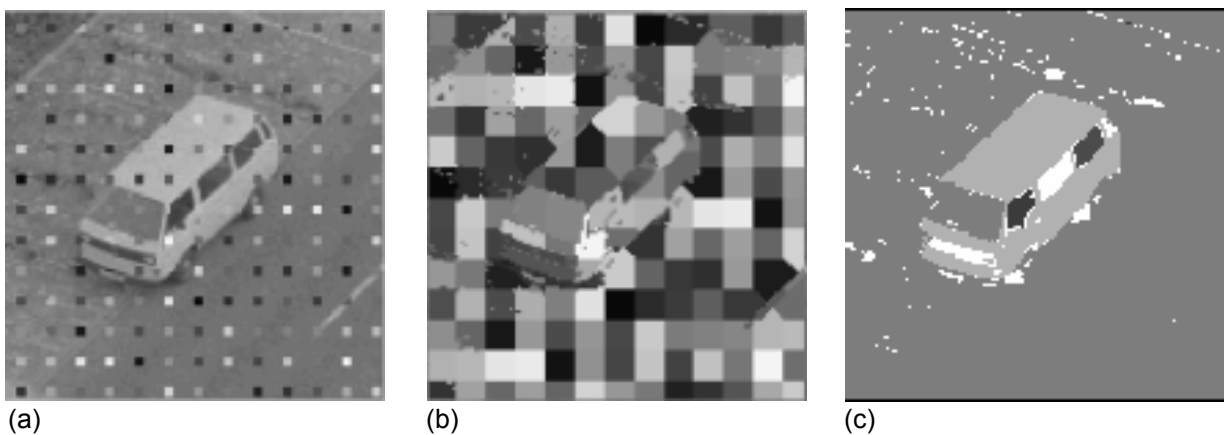


Abb. 2.5: Beispiel für die Anwendung des Keimzellen-Flächenwachstumsverfahrens: (a) Originalbild mit überlagerten Keimzellen und erster angelagerter Pixelschicht; (b) Zwischenergebnis nach der Beendigung des Wachstumsprozesses; (c) Endergebnis nach der Verschmelzung passender Bildbereiche. (Siehe auch Farbbild-Anhang am Ende der Datei zu diesem Kapitel).

Schlussbemerkung: Eine Verbesserung des Vorgehens kann man z.B. dadurch erzielen, dass man die Keimzellen nicht systematisch, sondern gezielt innerhalb von homogenen Bildbereichen platziert. Dies kann z.B. dadurch geschehen, dass man für das Bild erst eine Kantendetektion durchführt und dann das Kantenbild binarisiert. Auf das Binärbild wird anschließend eine Abstandstransformation angewendet, und die Keimzellen werden dann in den lokalen Maxima des resultierenden Abstandsgebirges platziert. (Dies sind Bildpositionen, die in homogenen Bildbereiche liegen und dabei "möglichst weit" von den Kanten entfernt sind).

2.2.4 Flutungs-Verfahren und Wasserscheiden-Transformation

Ein generelles Problem aller bislang geschilderten Verfahren ist, dass die interessierenden Objekte an allen Seiten relativ deutlich zu ihrer Umgebung abgegrenzt sein müssen. Wenn dies nicht der Fall ist, so ist es bei den Bildpunktklassifizierungs-Verfahren kaum möglich, einen geeigneten Schwellwert für die Trennung von Objekt und Hintergrund zu finden; siehe z.B. Abbildung 2.1 (b). Und die Flächenwachstumsverfahren neigen dazu, an kritischen Stellen mit einem fließenden Grauwertübergang zwischen Objekt und Hintergrund "auszulaufern" und benachbarte Bildbereiche mitzuerfassen.

Setzt man andererseits die Schwellwerte für die Binarisierung bzw. das Homogenitäts-Prädikat zu niedrig an, so erhält man bei der Binarisierung i.d.R. nur einen Teil der interessierenden Objekte (siehe Abbildung 2.1 (c)) und die Flächenwachstumsverfahren liefern eine **Übersegmentation**, bei der das Bild in viel zu viele Teilflächen zerlegt wird.

Das im folgenden geschilderte Vorgehen versucht, auf einer speziell erzeugten Übersegmentation aufbauend, zu entscheiden, welche zuvor getrennten Bildteile wieder zusammengefasst werden dürfen. Das Gesamt-Verfahren besteht aus zwei Teilen, die die anschaulichen Namen "Flutungs-Verfahren" und "Wasserscheiden-Transformation" tragen.

Das **Flutungsverfahren**, das die Übersegmentation liefert, lässt sich am besten erklären, wenn man das Eingangs-Grauwertbild als Grauwertgebirge interpretiert wird. Dunkle Objekte, die im weiteren zunächst die interessierenden Objekte sein sollen, bilden Talkessel im Grauwertgebirge. Zur Verdeutlichung soll das Verfahren anhand des eindimensionalen Grauwertprofils (Grauwertverlauf) erläutert werden, das sich entlang einer Testzeile ergibt, die die interessierenden Objekte schneidet; siehe Abbildung 2.6 (a). Die dunklen Objekte zeichnen sich im Grauwertprofil als Täler ab und man kann z.B. versuchen, sie mit Hilfe einer Binarisierung zu ermitteln.

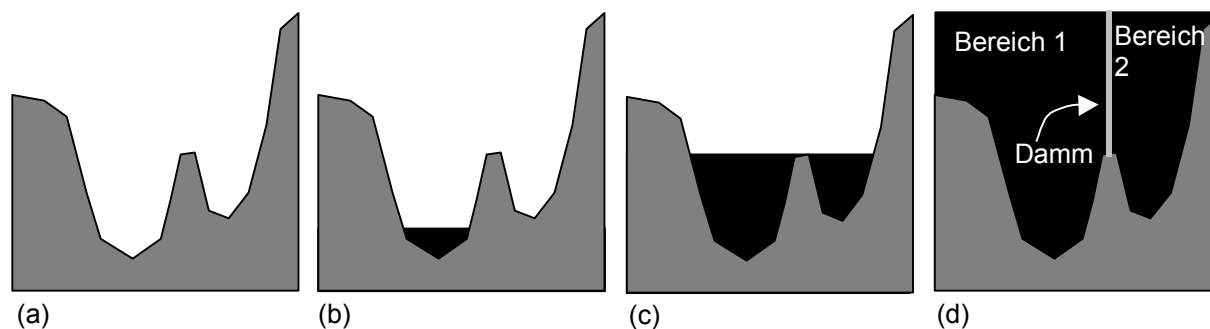


Abb. 2.6: Prinzip des Flutungs-Verfahrens.

Da man keine Vorkenntnisse über einen geeigneten Binarisierungsschwellwert hat, beginnt man mit einem niedrigen Schwellwert und erzeugt durch stetes Anheben des Schwellwertes eine Folge von Binarisierungen. Bei diesem Vorgang werden zuerst die tiefsten Täler "geflutet", d.h. als unerschwellig ermittelt; siehe Abbildung 2.6 (b). Bei einer weiteren Anhebung der Schwelle kommen neue geflutete Bereiche hinzu und bereits geflutete Bereiche vereinigen sich; siehe Abbildung 2.6 (c). Im vorliegenden Fall verhindert man jedoch die Vereinigung, indem man an der Trennlinie, d.h. der Wasserscheide des Grauwertgebirges einen "Damm" zwischen den beiden Bereichen errichtet, der dann beim weiteren Anheben der Schwelle ebenfalls erhöht wird. Wenn man den Binarisierungsschwellwert bis zum maximal zulässigen Grauwert anhebt, also das Grauwertgebirge vollständig überflutet, so erhält man eine mosaikartige Zerlegung des Bildes in durch Dämme getrennte Bildbereiche; siehe Abbildung 2.6 (d). Dieses Ergebnis stellt in der Regel eine Übersegmentation dar, da viele eigentlich zusammengehörige Bildbereiche durch Dämme getrennt wurden.

Im folgenden Schritt wird nun mit Hilfe der **Wasserscheiden-Transformation** versucht, gezielt Dämme im Grauwertgebirge wieder zu entfernen, um dadurch Bildbereiche, die zum selben Objekt gehören, wieder zu vereinigen. Auch hier soll das Grundprinzip des Verfahrens wieder an einem eindimensionalen Beispiel erläutert werden. Die einzelnen beim Fluten durch Dämme getrennten Bildbereiche seien jeweils durch ein charakterisierendes Merkmal (z.B. ihren mittleren Grauwert) repräsentiert; siehe Abbildung 2.7 (a).

Zu dem übersegmentierten Mosaikbild wird zunächst ein **Regionen-Adjazenz-Graph (RAG)** erzeugt; siehe oberer Graph in Abbildung 2.7 (b). Dazu wird jeder Bereich des Mosaikbildes durch einen Knoten repräsentiert, und in die Knoten wird das charakterisierende Merkmal der Region eingetragen. Anschließend werden Knoten, die benachbarten Bildbereichen entsprechen, durch eine Kante verbunden.

Aus dem RAG wird dann ein **Kontur-Graph** erzeugt, dessen Knoten genau den Kanten des RAG entsprechen; siehe unterer Graph in Abbildung 2.7 (b). In die Knoten des Kontur-Graphen wird die absolute Werte-Differenz zwischen den beiden zugehörigen Knoten des RAG, eingetragen. Zwei Knoten des Kontur-Graphen werden genau dann durch eine Kante verbunden, wenn die ihnen entsprechenden Kanten im RAG einen gemeinsamen Knoten haben, d.h. wenn die beiden Konturstücke zu einem gemeinsamen Bildbereich gehören.

Dann wird der Kontur-Graph "geflutet", indem das Flutungs-Verfahren auf die Bearbeitung von Graphen erweitert wird: die Knoten des Kontur-Graphen werden hinsichtlich ihres Wertes einer Binarisierung mit einem ständig zunehmenden Schwellwert unterzogen. Dabei werden wieder zuerst die tiefsten Stellen (siehe "Minima" in Abbildung 2.7 (b)) als unterschwellig erkannt und geflutet. Diese Stellen können dann bei der Anhebung des Schwellwertes "wachsen", indem auch benachbarte Knoten überflutet werden; siehe links in Abbildung 2.7 (b)). Und auch hier kann es wieder auftreten, dass zwei bereits geflutete Bereiche des Graphen sich vereinigen, wenn der Schwellwert den Wert "dazwischenliegender" Knoten überschreitet. Entsprechend zum oben geschilderten Vorgehen wird die Verschmelzung auch hier verhindert, indem in den trennenden Knoten ein Damm errichtet wird und derartige Knoten als "Wasserscheide" markiert werden; siehe Abbildung 2.7 (b).

Wenn die Binarisierung mit zunehmendem Schwellwert abgeschlossen ist, werden abschließend alle Konturstücke, die im gefluteten Kontur-Graphen keine Wasserscheiden darstellen, im RAG und im Mosaikbild gelöscht. Im RAG entspricht dies einer Zusammenfassung von Knoten (siehe Abbildung 2.7 (c)), und im Mosaikbild entstehen neue, größere "zusammengesetzte" Bereiche; siehe Abbildung 2.7 (d). Ein weiteres prinzipielles Beispiel für eine zweidimensionale Anwendung des Verfahrens ist in Abbildung 2.8 gezeigt.

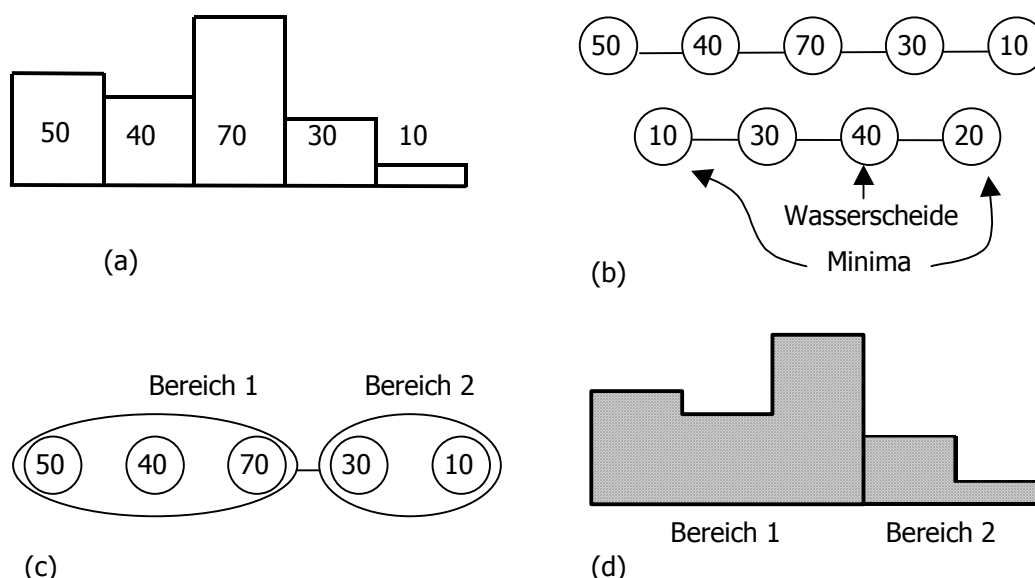
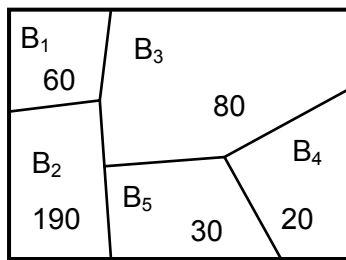


Abb. 2.7: Prinzip des Flutungs-Verfahrens und der Wasserscheiden-Transformation.

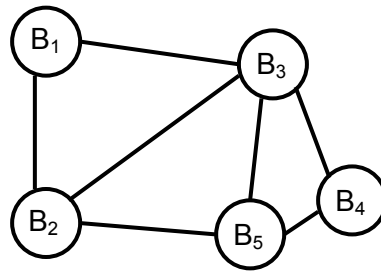
Durch die Anwendung von Flutung und Wasserscheiden-Transformation ist es möglich, im übersegmentierten Bild gezielt Flächen zusammenzufassen, die mit hoher Wahrscheinlichkeit zu einem gemeinsamen Objekt gehören. Wenn durch dieses Vorgehen die gewünschten Objekte noch nicht in zusammenhängender Form erkannt wurden, so kann der Ablauf (mehrfach) iterativ wiederholt werden, wobei jedes Mal weitere Zusammenfassungen erfolgen.

Das Verfahren, das bisher nur für dunkle Objekte eingeführt wurde, lässt sich einfach auf beliebig eingefärbte Objekte erweitern. Um diese Fälle auf den oben geschilderten Fall zurückzuführen, betrachtet man nicht das ursprüngliche Grauwertbild, sondern ein daraus (z.B. mit Hilfe eines Kantendetektors) abgeleitetes Gradientenbild. Homogene Bereiche des Originalbildes werden dabei auf relativ einheitlich dunkel eingefärbte Flächen abgebildet. Diese bilden "Talkessel" im Gradientenbild, die wie oben beschrieben geflutet werden können. Die Kanten zwischen den Flächen des ursprünglichen Grauwertbildes bilden sich im Grauwertgebirge des Gradientenbildes als "Höhenzüge" ab, die die Talkessel voneinander trennen. Auf diesen Höhenzügen entstehen dann die Dämme, die die Bereiche des übersegmentierten Mosaikbildes erzeugen.

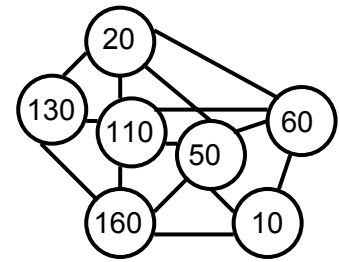
Im Farbbild-Anhang ist die iterative Anwendung von Flutungs-Verfahren und Wasserscheiden-Transformation auf ein Farbbild gezeigt.



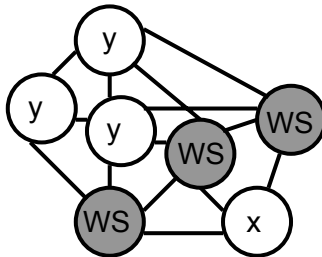
(a) Ausgangsbild mit Bereichen und zugehörigen Grauwerten



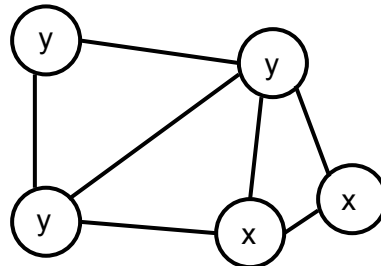
(b) Regionen-Adjazenz-Graph



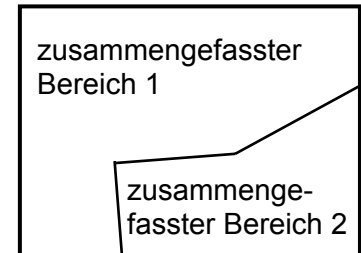
(c) Kontur-Graph



(d) gefluteter Kontur-Graph



(e) Regionen-Adjazenz-Graph mit zusammengefassten Bereichen



(f) Ergebnisbild

Abb. 2.8: Prinzip des Flutungs-Verfahrens und der Wasserscheiden-Transformation;
(Quelle: <http://iristown.engr.utk.edu/~koschan/paper/FarbW2000.pdf>)

Das Flutungs-Verfahren und die Wasserscheiden-Transformation haben sich in der Praxis als leistungsfähiges Werkzeug zur Bildsegmentierung erwiesen, wobei aber im voraus nicht angegeben werden kann, wie viele Iterationen durchgeführt werden müssen.

2.3 Zerteilungs- und Zerteilungs-/Verschmelzungs-Verfahren

Die in Kapitel 2.2 vorgestellten Flächenwachstums-Verfahren werden auch als Verschmelzungs-Verfahren (**merge**) bezeichnet, da Bildpunkte und/oder Bildteilbereiche beim Wachstumsprozess verschmolzen werden. Diese Verfahren arbeiten "bottom up", indem sie von einer Initialzerlegung ausgehend (z.B. jeder Bildpunkt entspricht einer Fläche) immer weniger, aber größere Flächen erzeugen. Dabei ist für jedes Zwischenergebnis die dritte Bedingung in Definition 2.1 erfüllt, die vierte jedoch noch nicht. Wenn keine Anlagerung eines Bildpunktes bzw. Verschmelzung zweier Flächen mehr möglich ist, sind die Flächen maximal und auch die vierte Bedingung ist erfüllt. Die so erhaltenen Flächen sind dann die Regionen der Bildsegmentation.

Im Gegensatz zu den Verschmelzungs-Verfahren stehen die **Zerteilungs-Verfahren** (**split**), die "top-down" arbeiten. Zu Beginn wird das ganze Bild als eine Fläche angesehen. Wenn für diese Fläche das Homogenitäts-Prädikat nicht erfüllt ist, wird sie in Teilflächen aufgespalten; z.B. regelmäßige Vierteilung wie beim Aufbau eines Quadrees. Die Homogenitäts-Überprüfung und die Aufspaltung werden für die jeweils neu erzeugten Flächen so lange wiederholt, bis alle Flächen das Homogenitäts-Prädikat erfüllen. Zerteilungs-Verfahren gehen also von der vierten Bedingung in Definition 2.1 aus und streben die Erfüllung der dritten Bedingung an.

Häufig werden die beiden Methoden des Zerteilens und Verschmelzens miteinander zu den sogenannten **Zerteilungs-/Verschmelzungs-Verfahren** (**split and merge**) kombiniert. Ausgangspunkt der Segmentierung ist eine beliebige rasterförmige Aufteilung des Bildes in gleich große Teilbilder. Diese Zerlegung befriedigt in der Regel weder die dritte noch die vierte Bedingung in Definition 2.1. Sukzessive werden nun "inhomogene" Flächen in vier Flächen zerteilt und benachbarte "ähnliche" Flächen verschmolzen, bis alle Bedingungen erfüllt sind.

3 Repräsentationen von Regionen

In diesem Kapitel wird gezeigt, wie Segmentierungsergebnisse, die mit Hilfe der flächenorientierten Segmentierungsverfahren aus Kapitel 2 erzeugt wurden, geeignet repräsentiert werden können.

Die vorgestellten Repräsentationsformen beziehen sich im Weiteren meist auf eine einzige Region und deren Komplement. Die Methoden können aber entsprechend auch auf den Normalfall, in dem ein Bild aus zahlreichen Regionen besteht, angewendet werden.

Bei den Regionen-Repräsentationen wird zwischen der Repräsentation geometrischer (Kapitel 3.1) und der Repräsentation topologischer Eigenschaften (Kapitel 3.2) unterschieden.

In Kapitel 3.1.1 werden die Repräsentationsformen vorgestellt, die sich auf das Ergebnis einer flächenorientierten Segmentierung beziehen. Konturorientierte Repräsentationsformen, die sich auf die Regionen-Kontur stützen, werden in "Computer Vision" Kapitel 4 im Zusammenhang mit der konturorientierten Segmentierung behandelt.

Da es bei bestimmten Aufgabenstellungen zweckmäßig ist, zwischen flächen- und konturorientierter Repräsentation zu wechseln, wird in Kapitel 3.1.2 die Ermittlung der Kontur für eine flächenhaft vorliegende Region beschrieben. Der umgekehrte Weg muss nicht behandelt werden, da die "Flutung" einer vektoriiell beschriebenen Kontur unter dem Stichwort Vektor-Rasterkonvertierung bereits in "Graphische Datenverarbeitung 1" behandelt wurde.

3.1 Geometrische Repräsentationen von Regionen

Im Gegensatz zur topologischen Repräsentation von Regionen, bei der z.B. Nachbarschafts-Beziehungen erfasst werden, sollen hier geometrische Eigenschaften, wie beispielsweise die Größe einer Region, aus der Repräsentation ableitbar sein.

Das Ziel der unterschiedlichen Repräsentationsformen ist dabei eine im Hinblick auf die nachfolgenden Verarbeitungsschritte optimierte Erfassung der Regionen.

3.1.1 Flächenorientierte Repräsentationen von Regionen

Wie oben beschrieben sollen sich die vorgestellten Repräsentationsformen auf eine Region R und ihr Komplement $\neg R$ ($\text{not } R$) beziehen. Die Region ist als Ergebnis einer flächenorientierten Segmentierung als Menge markierter Pixel gegeben. Im weiteren wird zwischen vollständigen und unvollständigen Repräsentationen unterschieden.

3.1.1.1 Vollständige flächenorientierte Repräsentationen

Unter der **vollständigen Repräsentation** einer Region versteht man eine Repräsentationsform, die eine pixelgenaue Rekonstruktion der Region und ihres Komplementes erlaubt.

3.1.1.1.1 Bildmatrix

Die flächenhafte Repräsentation einer Region mit Hilfe einer **Bildmatrix** kann als die natürliche Repräsentation der Region angesehen werden. Da laut Voraussetzung nur eine Region und ihr Komplement auftritt, genügt zur Repräsentation eine binäre Bildmatrix mit $M \times N$ Pixeln P ; (M steht für die Spalten- und N für die Zeilenzahl). Die Pixel der Region sollen im Weiteren den Wert 1 und die Pixel des Komplements den Wert 0 haben.

Zur Abspeicherung der Bildmatrix sind also $M \times N$ Bits erforderlich. (Falls K Regionen, d.h. K Marken auftreten würden, so würde man $M \times N \times \text{Id}(K)$ Bits benötigen, wobei "Id" den Logarithmus dualis, also den Logarithmus zur Basis 2 bezeichnet.

3.1.1.1.2 Aufzählung

Bei der **Aufzählung** werden die einzelnen Pixel der Region R (z.B. zeilenweise) mit ihren Koordinaten erfasst und in einer Liste abgelegt. Bei dieser Repräsentationsform, die sich nur für "kleine" Regionen empfiehlt, ist ein besonders einfacher Zugriff auf aufeinanderfolgende Pixel der Region möglich.

3.1.1.1.3 Lauflängen

Die Erfassung einer Region mit Hilfe von **Lauflängen** (*run length*, siehe auch "Graphische Datenverarbeitung 1", Kapitel 7) stützt sich auf die räumliche Kohärenz der Regionen in Zeilenrichtung.

Da es sich im vorliegenden Fall um Binärbilder handelt, genügt es, für jede Zeile die Lauflängen aufeinander folgender Läufe aus Komplement-Pixeln, Regionen-Pixeln, Komplement-Pixeln, Regionen-Pixeln usw. anzugeben. (Falls eine Bildzeile bereits mit einem Pixel aus der Region beginnt, so hat der erste Lauf, der sich ja auf das Komplement der Region bezieht, die Länge 0).

Eine Lauflängen-Repräsentation lohnt sich nur bei großen und wenig "zerfransten" Regionen.

3.1.1.1.4 Quadrees

Im Gegensatz zu den letzten Repräsentationsformen, die das Bild zeilenweise erfassen, wird bei **Quadrees** die Kohärenz in Zeilen- und Spaltenrichtung genutzt. Der Einfachheit halber wird vorausgesetzt, dass das Bild quadratisch ist ($M=N$) und dass die Kantenlänge eine Zweierpotenz ist ($M=2^n$, mit einer natürlichen Zahl n).

Das Bild wird durch einen Baum vom Grad vier beschrieben; die Blätter des Baumes sind Quadrate, die entweder vollständig aus Pixeln der Region oder vollständig aus Pixeln des Regionen-Komplementes bestehen.

Die Konstruktion des Quadrees erfolgt rekursiv:

- der Wurzelknoten des Baumes repräsentiert das ganze Bild. Wenn das Bild einheitlich eingefärbt ist, so besteht der Baum nur aus dem Wurzelknoten;
- andernfalls werden an den Wurzelknoten vier Nachfolgerknoten angehängt, von denen jeder einem (quadratischen) Viertel des Bildes entspricht;
- nun werden alle Knoten untersucht:
 - Knoten, bei denen alle Pixel des zugehörigen quadratischen Bildausschnitts vollständig in der Region bzw. vollständig in ihrem Komplement liegen, bilden ein Blatt des Baumes;
 - zu allen übrigen Knoten werden wieder vier Nachfolgeknoten definiert, die vier quadratischen Teilen des betreffenden Bildausschnittes entsprechen. Diese vier Folgeknoten werden wieder untersucht, usw..

Die Abarbeitungsreihenfolge bei der Unterteilung eines Bildes in vier quadratische Teilbilder ist üblicherweise: links/oben, rechts/oben, links/unten und zuletzt rechts/unten; siehe Schema rechts.

0	1
2	3

Abbildung 3.1 zeigt ein Beispiel für die Quadtree-Repräsentation eines Binärbildes.

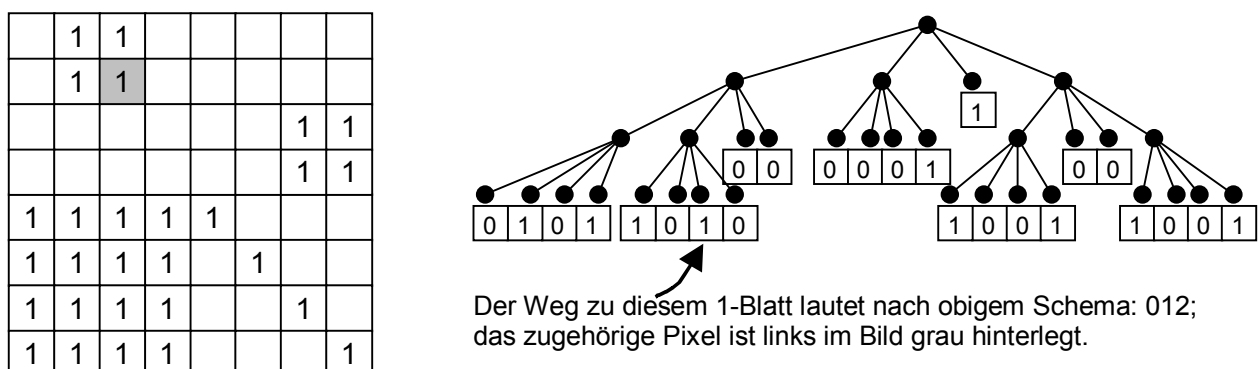


Abb. 3.1: Beispiel für die Quadtree-Repräsentation eines Binärbildes.

Quadtree-Repräsentationen spielen in der Bildverarbeitung und der Computer Graphik eine große Rolle, da sie besonders geeignet sind, großflächige und kompakte, einheitlich eingefärbte Bildbereiche günstig zu repräsentieren.

Zur schnellen Navigation in Quadtrees wurden spezielle Algorithmen entworfen, die es z.B. erlauben, besonders effizient auf benachbarte Pixel zuzugreifen, auch wenn diese in einem anderen Blatt des Baumes liegen.

3.1.1.2 Unvollständige flächenorientierte Repräsentationen

Der Vorteil der oben vorgestellten Repräsentation ist, dass sie die pixelgenaue Rekonstruktion des Bildes erlauben. Auf der anderen Seite erfassen sie aber auch jede Störung und jedes Detail, was je nach Anwendung überflüssig bis hinderlich sein kann. Im Folgenden werden zwei **unvollständige flächenorientierte Repräsentationen** angegeben, die das komplette Bild in schlechterer Auflösung zeigen, bzw. nur das "Skelett" der Region erfassen.

3.1.1.2.1 Bildmatrix mit geringerer Auflösung

Die Verkleinerung eines Bildes (Binärbild, Regionenbild oder Grauwertbild) zu einem Bild mit **geringerer Auflösung** stellt den einfachsten Fall einer unvollständigen Repräsentation dar. Da eine Verkleinerung gleichzeitig eine Glättung der Bildfunktion bewirkt (siehe "Graphische Datenverarbeitung 1"; Kapitel 7) können bei jeder Verkleinerung bestimmte Details unwiderruflich verloren gehen, so dass eine pixelgenaue Rekonstruktion des Originalbildes nicht mehr möglich ist.

Ein häufig verwendeter Spezialfall bei der Verkleinerung eines Bildes ist die Erzeugung einer Bildpyramide, bei der das ganze Bild iterativ in i- und j-Richtung auf die halbe Kantenlänge verkleinert wird, indem Blöcke aus je vier Pixeln zu einem Pixel zusammengefasst werden. Bei quadratischen Bildern, deren Kantenlänge eine Zweierpotenz ist, endet die iterative Verkleinerung bei der Bildpyramiden-Berechnung mit einem ein Pixel großen Bild (das im Falle von Grauwertbildern den mittleren Grauwert des ursprünglichen Bildes angibt).

Der Speicherbedarf einer kompletten Bildpyramide für ein Bild der Kantenlänge $M=N=2^n$ ist nur um etwa ein Drittel größer, als der für die Abspeicherung des Originalbildes.

3.1.1.2.2 Verdünnung

Bei Regionen, die aus langgestreckten Teilen bestehen, ist es oft für die weitere Verarbeitung sinnvoll, statt der kompletten Region nur ihre Skelettlinie zu erfassen. Ein bekanntes Beispiel hierfür ist die optische Zeichenerkennung (OCR für **O**ptical **C**haracter **R**ecognition), bei der die Strichstärke, mit der ein Zeichen geschrieben ist, nicht relevant für die Bedeutung des Zeichens ist.

Bei einer **Verdünnung** werden sukzessive immer wieder Randschichten des Objektes abgetragen, bis nur noch das "Skelett" des Objektes übrig ist. Dabei muss aber (im Gegensatz zur Rangfolge-Operation Erosion; siehe "Graphische Datenverarbeitung 1") beachtet werden, dass eine Region nicht in mehrere Regionen zerfällt oder verschwindet, und dass ein Pixel breite Linien nicht verkürzt werden.

Eine ausführliche Darstellung der Verdünnung ist in "Computer Vision" Kapitel 4 enthalten. Dort wird diese Operation eingesetzt, um die Ergebnisse einer Kanten- bzw. Liniendetektion, die in der Regel einen breiten Streifen von Kanten- bzw. Linienelementen liefert, zu verdünnen, um letztlich zu einer ein Pixel breiten Kontur zu kommen.

Ein Beispiel zur Verdünnung aus dem Bereich der optischen Zeichenerkennung ist in Abbildung 3.2 dargestellt. Abbildung 3.2 (a) zeigt einen segmentierten Einzelbuchstaben und in Abbildung 3.2 (b) ist das Ergebnis der Verdünnung dieser Region zu sehen.



Abb. 3.2: Beispiel für eine Verdünnung:
 (a) Originalbild; (b) Ergebnis einer N8-Verdünnung.

3.1.2 Konturverfolgung

Im folgenden wird beschrieben, wie man aus dem Ergebnis einer flächenorientierten Segmentierung, bei der eine Region ja durch die Menge aller zugehörigen Pixel gegeben ist, eine konturorientierte Repräsentation erhält, die aus der Menge aller aufeinanderfolgenden Konturpixel besteht.

Ein **Konturpixel** wird dabei als ein Pixel der Region definiert, das in seiner N4-Nachbarschaft (!) ein Pixel aus dem Hintergrund enthält.

Die Ermittlung der Regionen-Kontur erfolgt durch **Konturverfolgung**; dabei wird eine geordnete und geschlossene Folge von Konturpixeln ermittelt. Es muss auch hier wieder zwischen Regionen mit N4- und N8-Zusammenhang unterschieden werden.

Ein **Konturpixel** wird als ein Pixel der Region definiert, das in seiner N4-Nachbarschaft ein Pixel aus dem Hintergrund enthält. (Dies gilt unabhängig davon, ob die Region N4- oder N8-zusammenhängend ist!)

Zuerst wird der Algorithmus zur Konturverfolgung einer Region mit N8-Zusammenhang vorgestellt:

- 1.) Als erstes wird ein Startpunkt K_0 gesucht, der zur Region gehört und der ein oberes N8-Randpixel ist. Vom Startpunkt ausgehend wird die Region dann im Gegenuhrzeigersinn umlaufen, um die Kontur-Folgepunkte K_1, K_2 usw. zu suchen.
- 2.) Um das zweite Konturpixel K_1 zu finden, wird die rechts gezeigte Maske so über K_0 gelegt, dass K_0 im Zentrum der Maske liegt. Da K_0 ein oberer Randpunkt ist, ist gewährleistet, dass der Maskenpunkt P_0 außerhalb der Region liegt. Nun sucht man die Maskenpunkte P_i für $k=1, 2, 3$ usw. ab, bis das erste Pixel gefunden wird, das wieder innerhalb der Region liegt. Dieses Pixel ist das Folgepixel K_1 der Regionen-Kontur. Die Position des Maskenpunktes P_{k-1} , der ja noch außerhalb der Region lag, wird ebenfalls als $P_0(\text{next})$ vermerkt.
- 3.) Zur Suche nach einem Konturpixel K_{i+1} wird die oben gezeigte Maske über das zuletzt gefundene Kontur-Pixel K_i gelegt. Dann wird die Maske so um ihren Mittelpunkt gedreht, dass der mit P_0 bezeichnete Punkt auf der Position $P_0(\text{next})$ zu liegen kommt, die bei der Suche nach K_i das letzte außerhalb der Region liegende Pixel bezeichnete (siehe oben). Dann wird von P_0 ausgehend das erste Pixel P_k gesucht das wieder innerhalb der Region liegt und dieses wird als Konturpixel K_{i+1} ausgewählt. Weiter am Anfang von 3.).

P_1	P_0	P_7
P_2	K_i	P_6
P_3	P_4	P_5

Der Algorithmus endet, wenn der Startpunkt K_0 und sein Folgepunkt K_1 wieder in genau dieser Reihenfolge ermittelt werden. (Es reicht nicht, nur zu überprüfen, ob der Startpunkt wieder erreicht wird, denn wenn der Startpunkt z.B. an einer ein Pixel breiten "Einschnürung" der Region liegt, würde andernfalls der rechts von der Einschnürung liegende Teil der Region nicht ermittelt werden).

Der Algorithmus ist auch geeignet, um die Kontur von "Löchern" in Konturen zu ermitteln. Während im "Normalfall" die Konturpixel die Region in mathematisch positivem Sinne umlaufen (siehe Beispiel unten), ist dies bei der Konturierung von Löchern umgekehrt.

Das einfache Beispiel in Abbildung 3.3 zeigt die Ermittlung der N8-Kontur einer quadratischen Region, die aus vier Bildpunkten besteht.

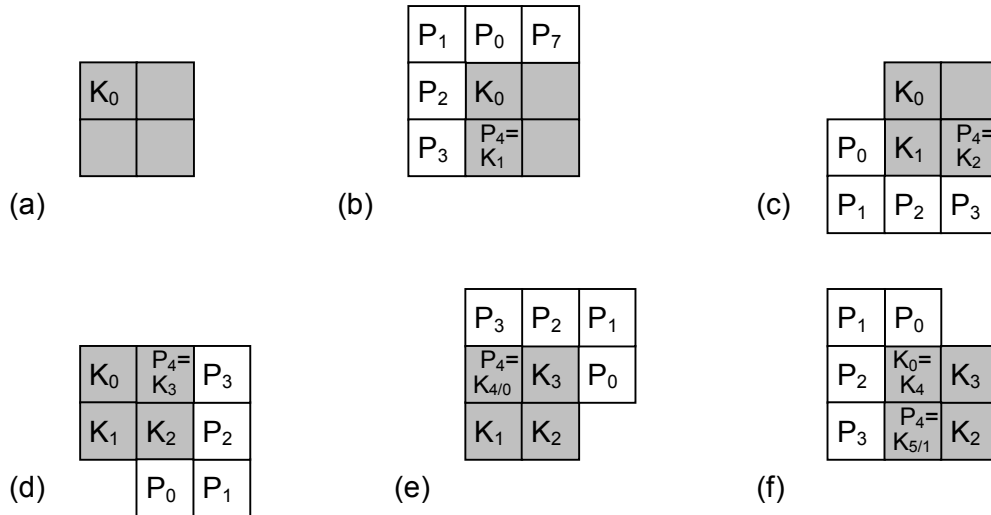


Abb. 3.3: Beispiel für die Ermittlung der N8-Kontur.

Zur Ermittlung einer N4-Kontur muss der oben beschriebene Algorithmus wie folgt abgeändert werden:

- von P_0 ausgehend werden nur die Maskenelemente mit geradem Index untersucht, und es wird das erste Element P_{2^*k} ($k=1, 2, 3$) gesucht, das wieder innerhalb der Region liegt.
- das nächste Konturelement K_{i+1} ergibt sich nun entweder zu
 - $K_{i+1} = P_{2^*k}$, falls das zuvor übersprungene Maskenelement P_{2^*k-1} nicht zur Region gehört; die Position von P_{2^*k-1} wird als $P_0(\text{next})$ vermerkt;
- oder zu
 - $K_{i+1} = P_{2^*k-1}$, falls das zuvor übersprungene Maskenelement P_{2^*k-1} auch zur Region gehört; die Position von P_{2^*k-2} wird als $P_0(\text{next})$ vermerkt.

Abbildung 3.4 zeigt ein Beispiel für eine N8-Konturverfolgung. In Abbildung 3.4 (a) ist das Ergebnis einer Binarisierung dargestellt und Abbildung 3.4 (b) zeigt als Überlagerung die ermittelte N8-Kontur.

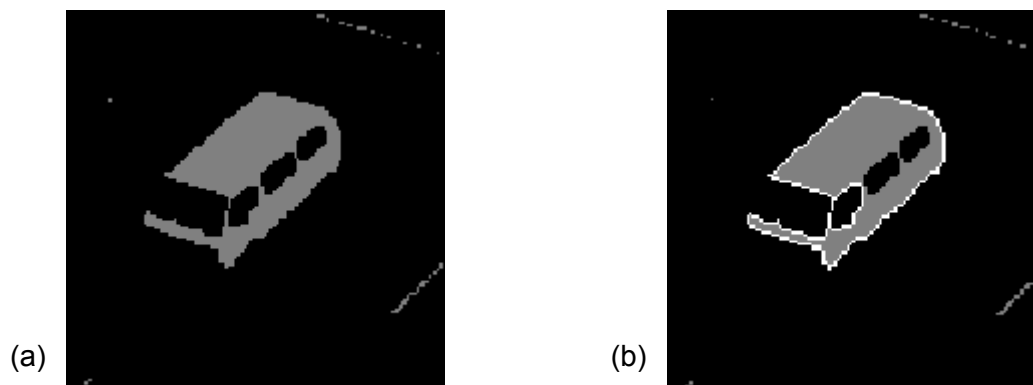


Abb. 3.4: Beispiel für die Konturverfolgung:
 (a) Ergebnis einer Binarisierung;
 (b) überlagertes Ergebnis einer N8-Konturverfolgung.

3.2 Topologische Repräsentationen von Regionen

Wesentlich für die nachfolgende Bildinterpretation ist, dass die Ergebnisse der Segmentierung alle wichtigen geometrischen, topologischen und photometrischen Eigenschaften repräsentieren. Mit möglichst geringem Aufwand sollen zu jeder Region die zugehörigen Merkmalswerte (z.B. Grau- oder Farbwert, Textur, Form usw.) zugreifbar sein. Besonders wichtig ist, dass topologische Nachbarschaftsbeziehungen zwischen Regionen einfach rekonstruierbar sind (z.B. Region R_i benachbart zu Region R_j).

Eine gängige Methode zur Abspeicherung solcher Nachbarschaftsbeziehungen ist der **Regionen-Adjazenz-Graph (RAG)**; vgl. Kapitel 2.2.4 und siehe Abbildung 3.5 (a). Dieser ist ein ebener, d.h. überschneidungsfreier Graph, dessen Knoten für die Regionen der Segmentierung stehen. Eine Kante verbindet zwei Knoten des Graphen genau dann, wenn die beiden zugehörigen Regionen benachbart sind. Der RAG ist homomorph zur sogenannten **Regionen-Adjazenz-Matrix (RAM)**, in der mit 1 bzw. 0 kodiert wird, ob zwei Regionen benachbart sind oder nicht (siehe Abb. 3.5 (b)). Um den vollständigen Zusammenhang im Bild zu repräsentieren, wird eine Region R_0 eingeführt, die das gesamte Bild umgibt, die also **nicht** zum Bild gehört. Auf diese Weise ist einfach erkennbar, welche Regionen den Bildrand berühren.

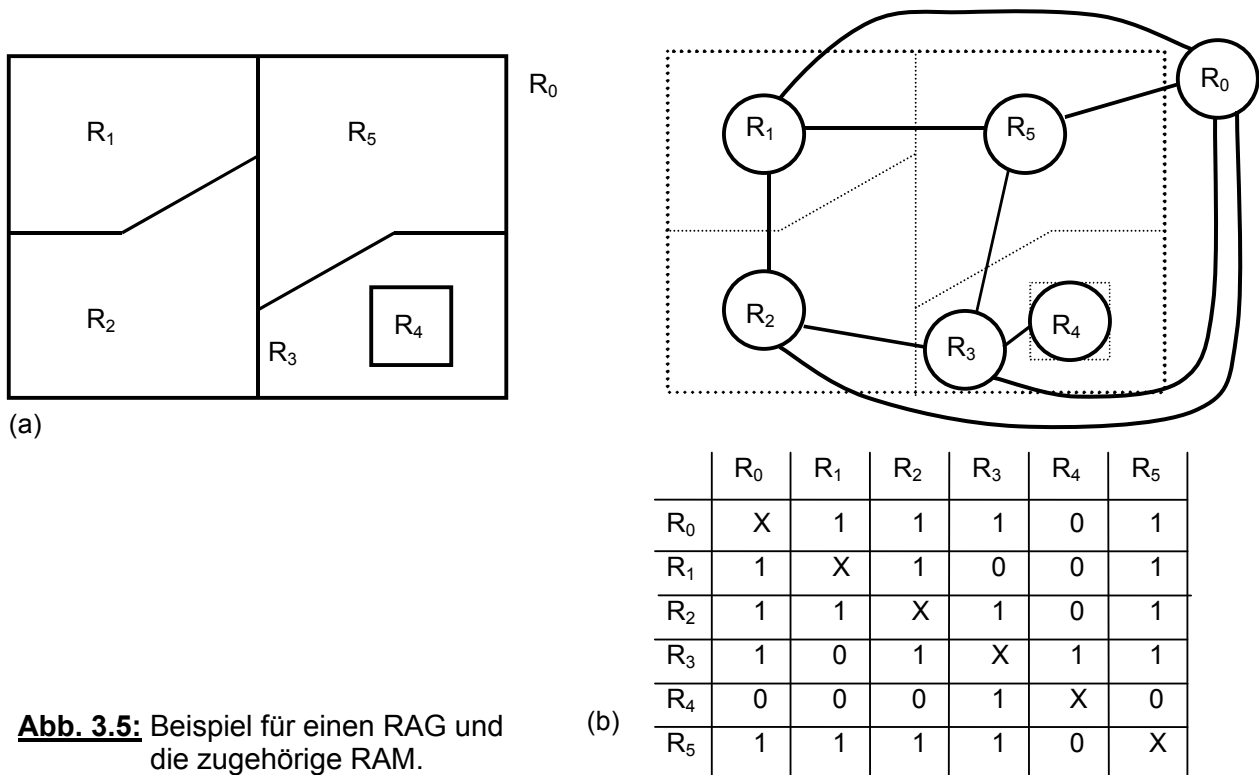


Abb. 3.5: Beispiel für einen RAG und die zugehörige RAM.