

Thema:

Image Retrieval aus Bilddatenbanken

Ausarbeitung

im Rahmen des Seminars Unterstützung von Landminendetektion durch
Bildauswertungsverfahren und Robotereinsatz

im Fachgebiet praktische Informatik

Themensteller: Prof. Dr. Xiaoyi Jiang
Prof. Dr. Wolfram-M. Lippe
Betreuer: Dr. Dietmar Lammers
Dipl.-Inform. Steffen Wachenfeld

vorgelegt von: Jens Feldkamp
Graelstraße 27
48153 Münster
0251 / 52 57 11
feldkamp@uni-muenster.de

Abgabetermin: 2004-01-05

Inhaltsverzeichnis

| | |
|--|-----|
| Inhaltsverzeichnis | II |
| Abbildungsverzeichnis | III |
| Abkürzungsverzeichnis | IV |
| 1 Einleitung..... | 1 |
| 2 Image Retrieval..... | 2 |
| 2.1 Grundlagen des Content-based Image Retrieval | 2 |
| 2.2 Color-based Image Retrieval | 3 |
| 2.2.1 Das RGB-Farbmodell..... | 3 |
| 2.2.2 Farb-Histogramme | 4 |
| 2.2.3 Räumliche Verteilung von Farbe | 8 |
| 2.3 Shape-based Image Retrieval | 9 |
| 2.3.1 Kantenextraktion | 9 |
| 2.3.2 Eigenschaftsvektoren | 10 |
| 2.3.3 Eigenschaftsgrafiken | 13 |
| 2.3.4 Transformationen | 16 |
| 2.4 Texture-based Image Retrieval..... | 16 |
| 2.4.1 Texturmerkmale | 17 |
| 2.4.2 Grauwert-Kookkurrenzmatrix (GLCM)..... | 17 |
| 2.4.3 Weitere Verfahren des Texture-based Image Retrieval | 19 |
| 2.5 Performance-Messung von CBIR-Algorithmen | 20 |
| 2.6 Optimierungsmöglichkeiten | 21 |
| 2.6.1 Ablage von Bildmerkmalen als Meta-Information | 21 |
| 2.6.2 Buckets | 21 |
| 2.6.3 Bild-Vorverarbeitung | 22 |
| 2.7 Beispiele für konkrete CBIR-Systeme..... | 22 |
| 3 Techniken zur Minendetektion und -identifikation | 23 |
| 3.1 Fotografie..... | 23 |
| 3.2 Infrarot-Kameras..... | 23 |
| 3.3 Röntgengeräte | 24 |
| 3.4 Weitere Techniken..... | 25 |
| 4 Anwendungsmöglichkeiten von CBIR bei der Minenräumung | 26 |
| 5 Fazit | 28 |
| Literaturverzeichnis | 29 |
| Anhang | 32 |
| A Internet-Adressen einiger frei zugänglicher CBIR-Systeme | 32 |

Abbildungsverzeichnis

| | | |
|------------|--|----|
| Abb. 2.1: | RGB-Farbwürfel | 4 |
| Abb. 2.2: | Beispiel für ein Farb-Histogramm | 5 |
| Abb. 2.3: | Vergleich dreier Farb-Histogramme | 6 |
| Abb. 2.4: | Getrennte Histogramme für die drei RGB-Farbkanäle | 8 |
| Abb. 2.5: | Verschiedene Farb-Verteilungen bei identischem Farb-Histogramm | 8 |
| Abb. 2.6: | Aufteilungsmöglichkeiten | 9 |
| Abb. 2.7: | Beispiel-Bild vor und nach Sobel-Operation | 10 |
| Abb. 2.8: | Polarkoordinaten-Netz beim Formmatrix-Verfahren | 11 |
| Abb. 2.9: | Kontur C | 12 |
| Abb. 2.10: | Skeleton eines Rechtecks | 13 |
| Abb. 2.11: | Vereinfachung der Form beim Chain Encoding | 14 |
| Abb. 2.12: | Mögliche Kodierung beim Chain Encoding | 14 |
| Abb. 2.13: | Beispiel-Form und resultierende Turning-Funktion | 14 |
| Abb. 2.14: | Beispiele für Texturen | 16 |
| Abb. 2.15: | Beispiel-Bild für GLCM | 17 |
| Abb. 2.16: | Matrix H der absoluten Häufigkeiten der Grauwert-Paare | 18 |
| Abb. 2.17: | GLCM-Matrix für das Beispiel-Bild | 18 |
| Abb. 3.1: | XBT-Bilder zweier Minen-Typen | 24 |

Abkürzungsverzeichnis

| | |
|------|---------------------------------|
| CBIR | Content-based Image Retrieval |
| CMY | Cyan, Magenta, Yellow |
| GLCM | Grey Level Co-Occurrence Matrix |
| GMRF | Gauss-Markov Random Field |
| HSV | Hue, Saturation, Value |
| IR | Image Retrieval |
| QBIC | Query By Image Content |
| RGB | Rot, Grün, Blau |
| UXO | Unexploded Ordnance |
| XBT | X-Ray Backscatter Technology |

1 Einleitung

Ziel dieser Ausarbeitung ist es, Verfahren und Techniken des Content-based Image Retrieval (CBIR) vorzustellen und anschließend zu untersuchen, ob und wie sie sich für die (teil-)automatisierte Landminenräumung eignen. Speziell wird hierbei darauf eingegangen, welchen Beitrag sie zu den zwei Teilaufgaben der Minensuche, nämlich zum einen dem Auffinden von verdächtigen Objekten und zum anderen der anschließenden Typisierung des gefundenen Gegenstandes, leisten können.

Besonders wünschenswert wäre eine Senkung der Fehlalarm-Rate, die bei heutigen Minenräum-Verfahren immer noch sehr hoch ist und den Aufwand für die Räumung eines Minenfeldes stark erhöht. Daher wird im Besonderen untersucht, wie CBIR-Verfahren helfen könnten, ungefährliche Gegenstände (wie z.B. Blechdosen) von Minen zu unterscheiden und verschiedene Minen-Modelle zu erkennen.

Zu erwähnen wäre hierbei noch, dass sich die meisten hier gemachten Aussagen auch auf die Räumung von Blindgängern und unverschossener Munition (sog. unexploded ordnances, UXOs) übertragen lassen.

In Kapitel 2 werden zunächst einige grundlegende Begriffe, Unterbereiche und Problemfelder des CBIR erklärt und anschließend verschiedene beim CBIR genutzte Verfahren erläutert.

Kapitel 3 beschreibt kurz einige Verfahren zur Minensuche, während in Kapitel 4 versucht wird, zu beurteilen, ob und wie die mit Hilfe dieser Methoden erzeugten Bilder mittels CBIR weiterverarbeitet werden können, um zur Erreichung der o.g. Ziele beizutragen.

Kapitel 5 schließlich beinhaltet ein kurzes Fazit der zuvor angestellten Überlegungen.

An dieser Stelle soll noch kurz darauf hingewiesen werden, dass es aufgrund der Kürze dieser Ausarbeitung hier lediglich möglich ist, einen kurzen Überblick über die existierenden Verfahren zu liefern. Eine wirklich detaillierte Darstellung der bereits entwickelten Verfahren sowie ihrer Kombinationsmöglichkeiten würde ein Vielfaches an Raum beanspruchen.

2 Image Retrieval

Unter Image Retrieval (IR)-Algorithmen versteht man solche, die zu einer vorgegebenen Suchanfrage ‚passende‘ Bilder in einer Bilddatenbank finden und sie als Ergebnis der Suche zurückliefern. Im Gegensatz zum klassischen Information- und Document-Retrieval besteht hierbei die besondere Schwierigkeit darin, dass die betrachteten Daten in keiner Weise vorstrukturiert sind. Textdokumente sind im Allgemeinen vom Autor in irgendeiner Weise geordnet (enthalten nur sinnvolle Worte, besitzen Überschriften etc.), Bild-Dateien hingegen sind schlicht Arrays verschiedenfarbiger Pixel (oder vergleichbare Repräsentationen) ohne offensichtliche inhärente Bedeutung.

Man unterscheidet beim Image Retrieval grundsätzlich solche Verfahren, die auf textuellen Meta-Informationen beruhen, welche zu den einzelnen Bildern in der Datenbank abgelegt sind, und solchen, die sich direkt auf den Inhalt der Bilder an sich beziehen.

Ein Beispiel für die erste Gruppe wäre etwa ein Programm, das vom Anwender den Namen eines Künstlers erfragt und dann aus der dazugehörigen Datenbank alle Bilder heraussucht, in deren Meta-Informationen als Urheber eben dieser Künstler eingetragen ist. Diese Verfahren umgehen natürlich das oben beschriebene Problem der unstrukturierten Bild-Informationen, gehören aber eher in den Bereich des Text Retrievals, welches nicht Gegenstand dieser Arbeit ist. Außerdem sind diese Algorithmen für die automatische Minensuche und -identifikation kaum geeignet, da sie immer eine Eingabe durch den Menschen erfordern.

Interessanter in diesem Zusammenhang sind solche Verfahren, die sich auf den Bild-Inhalt selbst beziehen und daher auch Content-based Image Retrieval (CBIR)-Systeme genannt werden. Diese sollen im Folgenden genauer betrachtet werden.

2.1 Grundlagen des Content-based Image Retrieval

Auf dem Gebiet des Content-based Image Retrieval (CBIR) unterscheidet man grundsätzlich drei Abstraktionsebenen, auf denen entsprechende Verfahren arbeiten.

Auf der ersten Ebene werden nur ‚primitive‘ Eigenschaften der Bilder betrachtet, also solche, die ohne menschliche Mithilfe und ohne die Verwendung externer Informationen ermittelt werden können. Zu dieser Gruppe gehören Farben, Texturen, Formen oder die räumliche Anordnung von Bildmerkmalen. Auf dieser Ebene sind ausschließlich Suchanfragen der Form „Finde ein Bild mit einer großen blauen Fläche am oberen Bildrand“ oder – häufig – „Finde ein Bild, das so ähnlich aussieht wie dieses“ möglich.

Die zweite Ebene hingegen umfasst auch abgeleitete Attribute, wie zum Beispiel die Erkennung eines abgebildeten Objektes. Zur Bestimmung dieser Attribute sind immer Zusatzinformationen nötig, die nicht im Bild selbst enthalten sind. Dies kann zum einen menschliches Wissen sein (z. B. wenn der Mensch das zu untersuchende Bild betrachtet und dann dem Computer mitteilt, dass darauf ein bestimmtes Objekt zu sehen ist), oder auch sonstige Informationen, die etwa einem geeigneten Algorithmus mitteilen, woran er das Objekt auf einem Bild erkennt. Unterscheiden lässt sich hierbei noch, ob das System nur nach Bildern mit Objekten aus einer abgegrenzten Gruppe (z. B. „Türme“) su-

chen können soll, oder sogar nur nach solchen, die einen individuellen Vertreter einer solchen Gruppe („Eiffelturm“) zeigen.

Auf dem höchsten Abstraktionsniveau schließlich werden solche Attribute verarbeitet, deren Behandlung ein hohes Maß an Wissen über die dargestellten Objekte oder Situation erfordern. Hier wären dann auch Suchanfragen wie „Finde Bilder, die Volkstanzgruppen zeigen“ oder sogar „Finde Bilder, die Freude ausdrücken“ möglich.

Obwohl für die meisten Anwendungen Algorithmen wünschenswert wären, die auf der dritten oder zumindest der zweiten Ebene arbeiten, gibt es bisher fast ausschließlich solche auf der ersten Ebene. Die weitergehende Verarbeitung von Bildinformationen ist heute – bis auf wenige Spezialfälle auf stark begrenzten Anwendungsgebieten – noch nicht (effizient) möglich. Der Umstand, dass Computer bisher kaum in der Lage sind, Bildinhalte (Objekte, Situationen) wirklich als solche zu erkennen und ihre Bedeutung (speziell in dem jeweiligen Kontext) zu verstehen, wird in der Literatur auch oft als die semantische Lücke (semantic gap) bezeichnet.

CBIR, wie es heutige Systeme bieten, lässt sich also etwa so definieren: CBIR bezeichnet den Prozess der Extraktion gewünschter Bilder aus einer großen Sammlung auf der Basis von Eigenschaften (wie Farbe, Texturen oder Formen), die automatisch aus den Bildern selbst abgeleitet werden können. Diese Eigenschaften können primitiv oder semantisch sein, der Gewinnungsprozess muss jedoch vorwiegend automatisch erfolgen.¹

Ein Großteil der aktuell eingesetzten CBIR-Systeme stützen sich auf die Bildattribute Farbe, Textur und Form. Daher sollen diese drei Aspekte in den folgenden Abschnitten vorgestellt werden.

2.2 Color-based Image Retrieval

Die in einem Bild enthaltenen Farben sind ein wichtiges Merkmal, das für das Finden ähnlicher Bilder große Bedeutung hat. Farb-basierte CBIR-Verfahren versuchen daher, eine geeignete Beschreibung der Farbverteilung zu erzeugen und dann anhand dieser in vertretbarer Zeit Bilder mit ähnlichen Charakteristiken zu finden.

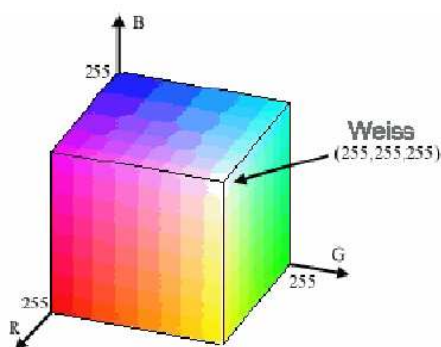
2.2.1 Das RGB-Farbmodell

Damit Farbinformationen der realen Welt überhaupt von Computern verarbeitet werden können, muss zunächst eine geeignete Repräsentation für die verschiedenen Farbwerte geschaffen werden. Es gibt mehrere sog. Farbmodelle, die dies leisten (z. B. RGB, CMY, HSV, usw.). Die Vorstellung dieser Farbmodelle steht jedoch nicht im Mittelpunkt dieser Arbeit, weshalb hier auf eine ausführliche Beschreibung verzichtet werden soll. Im Folgenden wird hier das RGB-Farbmodell verwendet, da es den meisten Lesern am vertrautesten sein dürfte. Die vorgestellten Verfahren des CBIR funktionieren prin-

¹ Vgl. Eakins, Graham (1999).

ziell jedoch auch mit anderen Farbmodellen – zumal sich das RGB-Modell teilweise recht einfach in andere Modelle überführen lässt.²

Das RGB-Modell ist ein sogenanntes additives Farbmodell. Das bedeutet, dass hier durch die Addition der drei Grundfarben Rot, Grün und Blau Weiß erzeugt wird (denn weißes Licht ist ja nichts anderes als Licht, in dem alle sichtbaren Wellenlängen vorkommen). Die drei Grundfarben addieren sich also zu einer Gesamtfarbe.³ Um nun einen bestimmten Farbton zu erhalten, gibt eine Variable für jede Grundfarbe an, wie stark sie in dem betreffenden Farbton enthalten ist. Alle so darstellbaren Farbtöne lassen sich als sogenannter RGB-Würfel darstellen. Jede Achse gibt dabei die Intensität einer der drei Grundfarben an.



Quelle: Tomberge (2002), S.4

Abb. 2.1: RGB-Farbwürfel

Wie in Abb. 2.1 zu erkennen, wird die Intensität jeder Grundfarbe mit einem Wert zwischen 0 und 255 angegeben. Daraus resultiert, dass mit dem RGB-Modell $256^3 \approx 16,7$ Millionen verschiedene Farben darstellen lassen.

2.2.2 Farb-Histogramme

Eine Möglichkeit, zwei Bilder bezüglich der Farbe zu vergleichen, sind Farb-Histogramme. Der Begriff Histogramm stammt aus dem Bereich der Statistik und bezeichnet die grafische Darstellung einer Häufigkeitsverteilung. Ein Farb-Histogramm stellt also dar, wie oft ein bestimmter Farbton in einem Bild vorkommt.⁴

Es wäre nun theoretisch möglich, zu einem Bild ein Farb-Histogramm zu erstellen, welches die Häufigkeit jedes im RGB-Modell darstellbaren Farbtons angibt. Dieses Vorgehen hätte jedoch zwei Nachteile: Zum einen würde die Speicherung eines solchem Histogramms relativ viel Speicher beanspruchen (da 16,7 Millionen Einzelwerte gespeichert werden müssten). Zum anderen wären solch genaue Histogramme beim späteren Vergleich untereinander extrem empfindlich auch gegenüber minimalen Farbabweichungen zwischen zwei ansonsten identischen Bildern, wie sie zum Beispiel durch unterschiedliche Belichtung entstehen können.

² Siehe hierzu auch Manjunath u.a. (2001), S. 706 und Block, Lucks, Tippmann (2001).

³ Vgl. Block, Lucks, Tippmann (2001).

⁴ Vgl. Tomberge (2002), S. 6.

Aus diesen Gründen reduziert man die betrachteten Farbtöne, indem man ähnliche Farben zu Farbklassen zusammenfasst. Dieser Vorgang funktioniert ähnlich wie die Quantisierung beim Digitalisieren analoger Daten und lässt sich mathematisch so ausdrücken:

Sei C ein Farbraum (z. B. derjenige, der alle mit dem RGB-Modell darstellbaren Farben enthält) und $P = \{c_1, c_2, \dots, c_i, \dots, c_n \mid c_i \in C\}$ eine Untermenge von C (die Menge der sog. Referenzfarben), wobei $n \ll |C|$ sein sollte. Dann ordnet die Funktion $Q : C \rightarrow P$ jedem Farbton aus C einen (möglichst ähnlichen) Farbton aus P zu.⁵

Nun kann ein Farb-Histogramm zu beliebigen Bildern erstellt werden. Besitzt ein Pixel des Bildes eine Nicht-Referenzfarbe, wird es derjenigen Referenzfarbe zugerechnet, die die Funktion Q seiner Farbe zuweist.

Ein Farb-Histogramm h ist also ein Vektor $[h_1, \dots, h_n]$, bei dem jedes h_j die Anzahl der Bild-Pixel mit der Farbe c_j enthält.

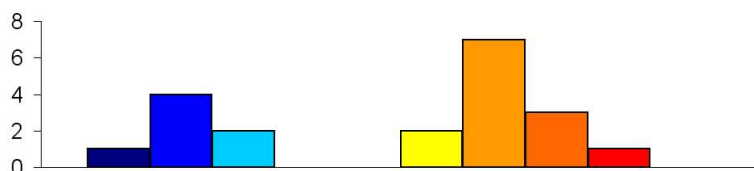


Abb. 2.2: Beispiel für ein Farb-Histogramm

Vergleich von Farb-Histogrammen

Um die Ähnlichkeit zweier Farb-Histogramme (und damit auch die eventuelle Ähnlichkeit der zugrundeliegenden Bilder) zu bestimmen, benötigt man eine Funktion, die zwei Histogrammen eine reelle Zahl (den Abstand der Histogramme) zuweist. Eine solche Funktion nennt man Metrik. Mathematisch ausgedrückt:

Sei H eine Menge von Farb-Histogrammen. Eine Funktion $d : H^2 \rightarrow \mathfrak{R}$ mit den Eigenschaften

- 1) $d(x, y) \geq 0$; $d(x, y) = 0 \Leftrightarrow x = y$ Positive Definitheit
- 2) $d(x, y) = d(y, x)$ Symmetrie
- 3) $d(x, z) \leq d(x, y) + d(y, z) \forall x, y, z \in H$ Dreiecksungleichung

heißt Metrik; H und d gemeinsam metrischer Raum.

Offensichtlich gibt es sehr viele Funktionen, die diese Bedingungen erfüllen, darunter sehr einfache, die aber für Zwecke des CBIR viel zu ungenau sind:

$$d_1(x, x) = 0, d_1(x, y) = 1 \Leftrightarrow x \neq y$$

⁵ Ciocca, Schettini, Zuffi (2001), S. 3.

Wenn die zu vergleichenden Bilder eine einheitliche Größe haben (oder ihre Farb-Histogramme auf eine einheitliche Größe normiert sind, wovon hier ausgegangen wird) und die Histogramme auf derselben Menge P (mit $|P| = n$) der Referenzfarben beruhen, kommen folgende Metriken in Frage:

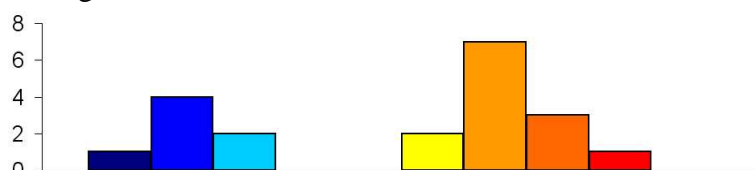
$$d_2(x, y) = \sum_{i=1}^n |x_i - y_i| \quad \text{Minkowski-Form Metrik } L_1$$

$$d_3(x, y) = \sum_{i=1}^n (x_i - y_i)^2 \quad \text{Minkowski-Form Metrik } L_2 \text{ (auch Euklidische Distanz)}$$

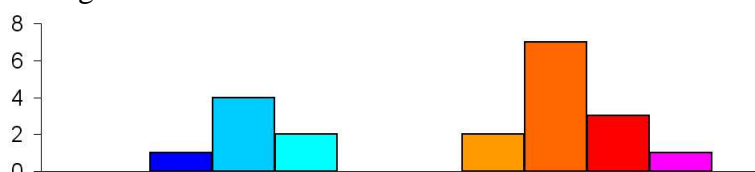
Beide Metriken zeigen allerdings im praktischen Einsatz keine besonders guten Resultate⁶, da sie sehr empfindlich auf Verschiebungen der Farbpalette, wie sie z. B. durch unterschiedliche Beleuchtung entstehen können, reagieren und dann auch für eigentlich ähnliche Bilder viel zu hohe Abstandsmaße liefern.

Als Beispiel hierfür sollen die drei folgenden Histogramme dienen. Das schon bekannte Farb-Histogramm soll mit zwei anderen verglichen werden. Wie der menschliche Betrachter leicht sieht, ist das zweite Histogramm mit dem ersten identisch, von einer Verschiebung der Farben abgesehen. Das dritte jedoch weicht ziemlich stark vom ersten ab, die Blau-Anteile fehlen völlig.

Histogramm 1



Histogramm 2



Histogramm 3

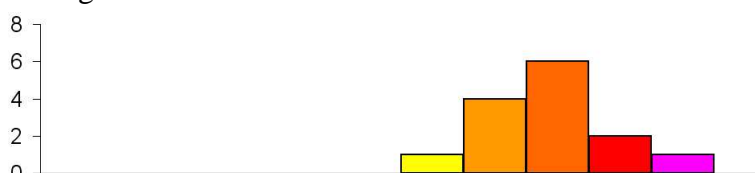


Abb. 2.3: Vergleich dreier Farb-Histogramme

Trotzdem liefert die L_1 -Metrik für die Histogramme 1 und 2 einen Wert von 22, für 1 und 3 hingegen nur 16. Bei der L_2 -Metrik ist der Unterschied noch gravierender: 464 für den Vergleich von Histogramm 1 und 2, nur 42 für die Histogramme 1 und 3. Beide Metriken kommen hier also zu dem Schluss, dass Histogramm 3 ähnlicher sei als Histogramm 2, und das, obwohl Histogramm 2 nur leicht farbverschoben ist. Die Ursache hierfür liegt darin, dass diese einfachen Metriken nur die Häufigkeiten identischer

⁶ Vgl. Ciocca, Schettini, Zuffi (2001), S. 27.

Referenzfarben vergleichen, eventuelle Parallelen bei ähnlichen Farben jedoch ignorieren.

Um diese Schwäche zu reduzieren oder gar zu vermeiden, wurde eine große Anzahl an anderen Metriken entwickelt, die hier nicht alle im Detail vorgestellt werden können.⁷ Als ein Beispiel soll jedoch der Ansatz von Hafner et al.⁸ dienen:

$$d_4(x, y) = \sqrt{\sum_i \sum_j a_{ij} (x_i - y_j)^2},$$

$$\text{mit } a_{ij} = 1 - \frac{d_{ij}}{\max(d_{ij})}$$

und d_{ij} als euklidischer Abstand der beiden Farben in einer Variante des Munsell-Farbraumes⁹.

Dadurch, dass hier die Häufigkeit jeder Referenzfarbe mit derjenigen jeder anderen Farbe verglichen wird, wird erreicht, dass auch Übereinstimmungen oder Unterschiede in der Häufigkeit ähnlicher Farben berücksichtigt werden. Farbverschiebungen fallen so beim Histogramm-Vergleich nicht mehr so stark ins Gewicht.

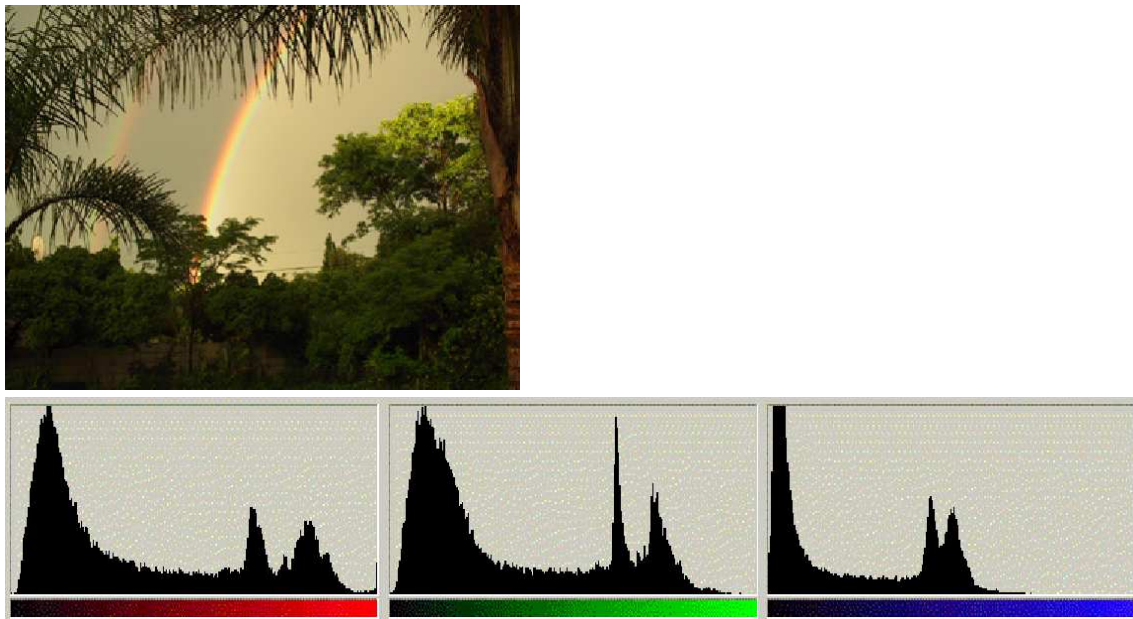
Weitere Histogramm-Arten

Obwohl Farb-Histogramme ein wichtiges Instrument des CBIR sind, hat sich gezeigt, dass es unter Umständen sinnvoll sein kann, weitere Histogramme in die Betrachtung mit einzubeziehen. Dies können z. B. Helligkeits-, Farbton- oder Sättigungshistogramme sein. Ebenso wäre es möglich, getrennte Histogramme für die drei Farb-Kanäle zu erstellen.

⁷ Für weitere Details vgl. Ciocca, Schettini, Zuffi (2001), S. 7 ff.

⁸ Vgl. Hafner u.a. (1995).

⁹ Farbraum, der auf einem perzeptuellen Farb-Modell beruht. Es berücksichtigt vor allem die Wahrnehmung der Farben durch den menschlichen Beobachter. Zu Einzelheiten siehe z. B. Domik (1997).



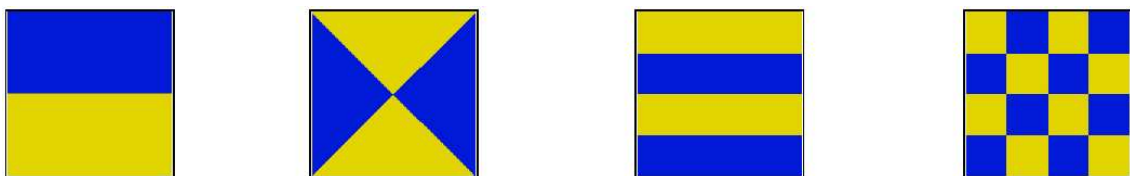
Quelle: Tomberge (2002), S. 9.

Abb. 2.4: Getrennte Histogramme für die drei RGB-Farbkanäle

Welche Histogramme genau sinnvoll sind, hängt von den zu untersuchenden Bildern ab und muss in jedem Anwendungsfall neu – eventuell auch experimentell – ermittelt werden.

2.2.3 Räumliche Verteilung von Farbe

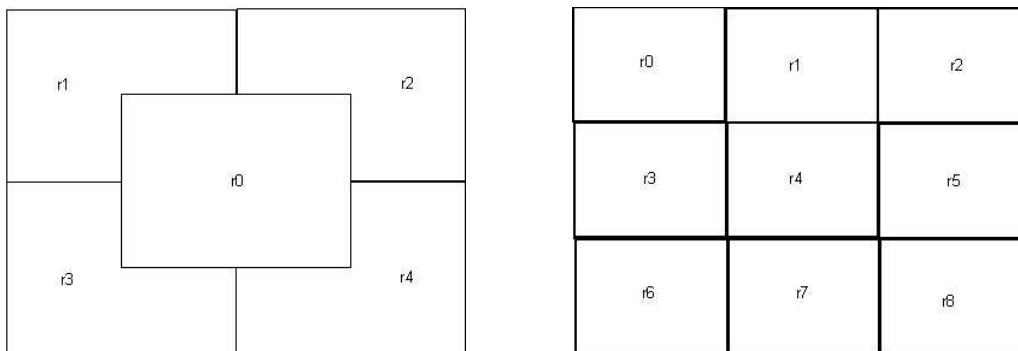
Ein großes Problem beim Image-Retrieval allein mit Hilfe von Farb-Histogrammen ist, dass diese keinerlei Aussage über die räumliche Verteilung der Farbwerte im Bild machen. Die Bilder in Abb. 2.5 haben sicher identische Farb-Histogramme, sind aber trotzdem recht unterschiedlich.



Quelle: Ciocca, Schettini, Zuffi (2001), S. 10.

Abb. 2.5: Verschiedene Farb-Verteilungen bei identischem Farb-Histogramm

Um dieses Problem zu beheben, werden die untersuchten Bilder oft zunächst in unterschiedliche Regionen aufgeteilt. Zwei Möglichkeiten, dies zu tun, sind in Abb. 2.6 dargestellt.



Quelle: Trölitersch (2002), S. 22, 23.

Abb. 2.6: Aufteilungsmöglichkeiten

Für die entstandenen Bildregionen können anschließend individuelle Eigenschaftsmaße (z. B. Histogramme oder dominante Farbe¹⁰) ermittelt werden. Bei der späteren Ähnlichkeitsbestimmung zweier Bilder kann dann jede Bild-Region mit der entsprechenden im anderen Bild verglichen werden.¹¹ So wird auch die Farb-Verteilung im Bild berücksichtigt, was oft zu besseren Ergebnissen beim Vergleich und damit beim Image Retrieval führt.

2.3 Shape-based Image Retrieval¹²

Die Form von Objekten ist neben der Farbe für Menschen ein wichtiges Merkmal bei der Erkennung der Art eines Gegenstandes. Daher liegt es nahe, diese Eigenschaft auch beim Image Retrieval einzusetzen.

Dazu sind mehrere Aufgaben zu lösen: Zunächst die Extraktion der Form aus dem Bild selbst, also etwa die Erkennung der Umrisse und Kanten. Ist dies geschehen, ist es notwendig, eine geeignete Repräsentationsform für die extrahierten Merkmale zu finden und schließlich auch ein Ähnlichkeitsmaß, um die Ähnlichkeit zweier Bilder zu bestimmen.

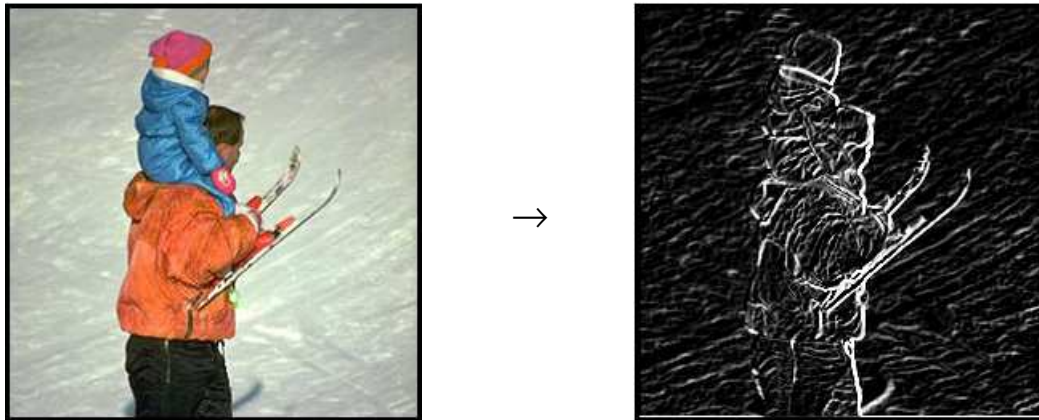
2.3.1 Kantenextraktion

Die Extraktion von Formen aus Bildern ist eine ausgesprochen komplexe Aufgabe. Während es Menschen im Allgemeinen sehr leicht fällt, auf Bildern die Grenzen eines abgebildeten Gegenstandes und andere Formen zu erkennen, ist dies Computern nicht so ohne Weiteres möglich. Es wurde allerdings eine Vielzahl von Algorithmen entwickelt, denen dies (in gewissen Grenzen) gelingt. Als Beispiel sei hier der Sobel-Operator genannt, der große Farbunterschiede zwischen benachbarten Pixeln als Kanten interpretiert (siehe Abb. 2.7).

¹⁰ Zu Einzelheiten zur Bestimmung der dominanten Farbe in einem Bildbereich vgl. Manjunath u.a. (2001), S. 709f.

¹¹ Eine Beispiel-Metrik für den regionen-basierten Vergleich zweier Bilder findet sich in Trölitersch (2002), S. 22.

¹² Vgl. Dlugosz (2002), S. 2-12.



Quelle: Wu (1998), S. 11.

Abb. 2.7: Beispiel-Bild vor und nach Sobel-Operation

Aufgrund ihrer Zahl und Komplexität können diese Algorithmen hier jedoch nicht im Detail vorgestellt werden (sie sind auch kein Problem des CBIR im eigentlichen Sinne). Im Folgenden wird daher davon ausgegangen, dass die in den untersuchten Bildern enthaltenen Formen bereits in Form einer Binärgrafik – ähnlich der in Abb. 2.7 – vorliegen, die angibt, welcher Pixel zur Form gehört, und welcher nicht.

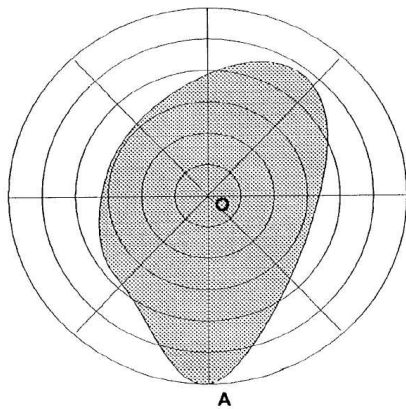
Diese Art der Repräsentation von Kanten ist jedoch zur Speicherung und zum späteren Vergleich nur mäßig geeignet. Daher wurden andere Verfahren entwickelt, die auf diesen Gebieten bessere Eigenschaften haben, von denen einige im Folgenden exemplarisch vorgestellt werden sollen. Zentrale Anforderung an alle diese Repräsentationsformen und Vergleichsmethoden ist die Invarianz gegenüber Translation (Verschiebung), Skalierung (Größenänderung) und Rotation der abgespeicherten bzw. verglichenen Formen.

2.3.2 Eigenschaftsvektoren

Bei der Benutzung von Eigenschaftsvektoren wird versucht, die Eigenschaften der untersuchten Form in ein Format zu bringen, das sich als einfacher Vektor oder auch als Matrix abspeichern lässt. Man unterscheidet dabei interne und externe Verfahren. Interne Verfahren berücksichtigen die von der Form eingeschlossene Fläche, externe nur ihren Umriss oder sogar nur bestimmte Punkte.

Formmatrix-Verfahren

Das Formmatrix-Verfahren ist ein internes. Hierbei wird über die Form (wie in Abb. 2.8 zu sehen) ein Polarkoordinatensystem gelegt. Danach wird für jeden Schnittpunkt zweier Linien untersucht, ob er innerhalb oder außerhalb der Form liegt. Gehört er noch dazu, wird ihm eine 1 zugeordnet, sonst eine 0. Die so gesammelten Werte werden anschließend in einer Matrix abgespeichert, der sog. Formmatrix (shape matrix) F . Dabei repräsentieren die Zeilen der Matrix die radialen Geraden, die Spalten hingegen die konzentrischen Kreise.



Quelle: Dlugosz (2002), S. 4.

Abb. 2.8: Polarkoordinaten-Netz beim Formmatrix-Verfahren

Eine mögliche Repräsentation der oben dargestellten Form in einer Matrix wäre also z. B. die folgende:

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Die Gerade \overline{AO} ist dabei als die erste Zeile ausgewählt worden, die übrigen folgen im Uhrzeigersinn. Die Spalten repräsentieren die Kreise von innen nach außen.

Diese Form der Darstellung ist sowohl translations-, als auch rotationsinvariant (die Spalten der Matrix können einfach umsortiert werden) sowie auch skalierungsinvariant (wenn unabhängig von der Größe der Form dasselbe Koordinatensystem darüber gelegt wird).¹³

Digitale Momente

Ein weiteres Beispiel eines internen Verfahrens ist das der digitalen Momente. Hierbei wird die Form als zweidimensionale Dichte-Funktion dargestellt:

$$\rho(x, y) = \frac{1}{a(A)} \mathbf{1}_A(x, y)$$

Dabei ist $A \subset \mathfrak{R}^2$ die Menge der zur Fläche gehörenden Punkte und $a(A)$ ihr geometrischer Flächeninhalt. Die Division durch den Flächeninhalt ist notwendig, um eine Funk-

¹³ Vgl. Dlugosz (2002), S. 4.

tion zu erhalten, die der Definition einer Dichtefunktion entspricht.¹⁴ Die bivarianten Momente der Dichtefunktion sind folglich:

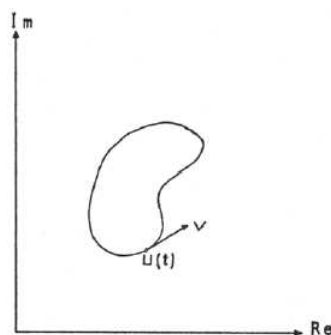
$$m_{pq} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} x^p y^q \rho(x, y) dx dy, \text{ mit } p, q = 0, 1, 2, \dots$$

Es kann gezeigt werden, dass eine Form durch die Menge dieser Momente eindeutig identifiziert werden kann. Durch entsprechende Modifizierung des Ansatzes kann sogar erreicht werden, dass die resultierenden Momente translations-, skalierungs- und rotationsinvariant werden.¹⁵

Beim praktischen Einsatz eines solchen Verfahrens muss jedoch berücksichtigt werden, dass die Berechnung der Momente sehr aufwändig ist, was bei Echtzeit-Anwendungen zu Problemen führen kann.

Fourierkoeffizienten

Nur als Beispiel für eine externe Methode genannt, aber aus Platzgründen nicht detailliert beschrieben, sei das Fourierkoeffizienten-Verfahren.¹⁶ Hier wird eine geschlossene Kontur C durch eine komplexwertige Funktion u ausgedrückt. Diese Funktion erhält man, indem man einen Punkt betrachtet, der sich auf C mit einer bestimmten Geschwindigkeit v entlangbewegt. Die Position dieses Punktes auf der Kontur zu einem bestimmten Zeitpunkt t ist dann genau u(t).



Quelle: Dlugosz (2002), S. 13.

Abb. 2.9: Kontur C

Da u offensichtlich periodisch ist, lässt es sich auch als komplexwertige Fourierreihe ausdrücken. Man kann nun einen Satz von Koeffizienten für diese Reihe berechnen, die dann die Kontur C beschreiben. Auch diese Darstellungsweise ist translations-, skalierungs- und rotationsinvariant.

¹⁴ $\rho(x,y) \geq 0 \forall x, y \in D$, vollständiges Integral ergibt 1.

¹⁵ Vgl. Dlugosz (2002), S. 16.

¹⁶ Zu weiteren Einzelheiten zu diesem Verfahren siehe Dlugosz (2002), S. 13ff.

Vergleich zweier Formen auf Basis von Eigenschaftsvektoren

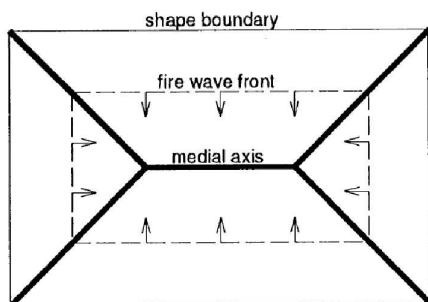
Um zwei Formen anhand der aus ihnen erzeugten Eigenschaftsvektoren zu vergleichen, benutzt man, ähnlich wie schon in Abschnitt 2.2 beschrieben, Metriken auf den jeweiligen Vektorräumen. Auch hier wird oft die euklidische Norm oder auf ihr basierende Metriken eingesetzt (siehe Abschnitt 2.2).

2.3.3 Eigenschaftsgrafiken

Bei dieser Gruppe von Verfahren dient nicht ein Vektor zur Repräsentation der Eigenschaften der Form, sondern andere, einfachere Formen. Auch hier gibt es, wie schon bei den Eigenschaftsvektoren, interne und externe Verfahren:

*Skeleton*¹⁷

Bei diesem Verfahren stellt man sich die betrachtete Form als Fläche aus brennbarem Material vor. Wird nun an allen Rändern gleichzeitig Feuer gelegt, wird sich das Feuer immer weiter zur Mitte der Form hin ausbreiten. Treffen sich dabei zwei Feuerfronten, löschen sie sich gegenseitig aus. Die Menge aller Punkte, an denen sich zwei Fronten treffen (quench points), bildet das Skelett der Form.



Quelle: Dlugosz (2002), S. 8.

Abb. 2.10: Skeleton eines Rechtecks

Das Skelett ist nun charakteristisch für die Form, kann gespeichert und mit den Skeletten anderer Formen verglichen werden.

Weitere, einfachere interne Verfahren

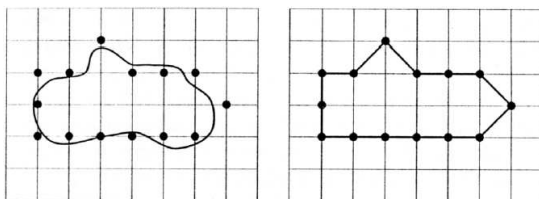
Abgesehen vom Skeleton-Ansatz gibt es noch eine ganze Reihe einfacherer geometrischer Ansätze. So kann man z.B. den kleinsten Kreis bestimmen, der die Form ganz umschließt oder den größten Kreis, der gerade noch ganz in die Form passt. Oft wird auch versucht, sich komplexeren Formen durch die Kombination mehrerer einfacher geometrischer Figuren (Kreise, Ovale, Rechtecke, Dreiecke etc.) anzunähern. Position, Größe und evtl. Orientierung dieser Figuren können dann gespeichert und beim Ver-

¹⁷ Auch als Prairie Fire, Symmetric Axis oder Medial Axis Transform bezeichnet.

gleich mit anderen Formen als Repräsentationen für die komplexe Form eingesetzt werden.

Chain Encoding

Wie schon bei den Eigenschaftsvektoren betrachten externe Verfahren ausschließlich den Umriss der untersuchten Formen. Das chain encoding ist eine eher einfache Methode, bei der ein relativ grobes Raster über die Form gelegt wird. Anschließend werden jeweils die Schnittpunkte der Geraden markiert, die dem Umriß am nächsten liegen.



Quelle: Dlugosz (2002), S. 9.

Abb. 2.11: Vereinfachung der Form beim Chain Encoding

| | | |
|---|---|---|
| 3 | 2 | 1 |
| 4 | x | 0 |
| 5 | 6 | 7 |

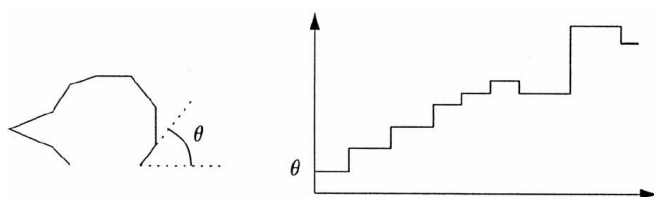
Quelle: Vgl. Dlugosz (2002), S. 9.

Abb. 2.12: Mögliche Kodierung beim Chain Encoding

Um sie abzuspeichern, kann z. B. die in Abb. 2.12 dargestellte Codierung benutzt werden. Für die Form aus Abb. 2.11 ergäbe sich dann folgende Zahlenfolge (beginnend in der oberen linken Ecke): 0, 1, 7, 0, 0, 7, 5, 4, 4, 4, 4, 4, 2, 2.

Turning-Funktion

Ein weiteres Beispiel für ein externes Verfahren ist die Turning-Funktion. Bei dieser Methode wandert man quasi auf dem Umriss der Form einmal herum und erstellt dabei einen Graphen, indem man für gerade Strecken auf der Kontur entsprechend lange zur x-Achse parallele Strecken einfügt. Wird eine Ecke im Umriss erreicht, richtet sich das ‚Höhenniveau‘ der folgenden Strecke nach dem Winkel zwischen dem letzten und dem nun folgenden Abschnitt (siehe Abb. 2.13).



Quelle: Dlugosz (2002), S. 9.

Abb. 2.13: Beispiel-Form und resultierende Turning-Funktion

Vergleich zweier Formen auf Basis von Eigenschaftsgrafiken

Anders als bei Eigenschaftsvektoren sind die Möglichkeiten zum Vergleich bei Eigenschaftsgrafiken sehr vom benutzten Verfahren abhängig. Sie arbeiten nicht auf Vektoren, sondern direkt auf Punktmengen oder Polygonen. Man unterscheidet hierbei drei Gruppen von Distanzmaßen:

Einige Distanzmaße vergleichen direkt Punktmengen miteinander. Beispiele hierfür sind Bottleneck Matching, Minimum Weight Matching, Uniform Matching, Minimum Deviation Matching oder auch die Hausdorff-Distanz, die für zwei Mengen A und B wie folgt definiert ist:¹⁸

$$\delta_H(A, B) = \max(h(A, B), h(B, A))$$

$$h(A, B) = \max(\min\|a - b\|)$$

mit $a \in A, b \in B$.

Für Kurvenrepräsentationen (wie z. B. die Turning-Funktion) gibt es meist spezielle Distanzmaße, die speziell für die spezifische Repräsentationsform zugeschnitten sind. Die Fréchet-Distanz für zwei Kurven mit den Repräsentationen P und Q und der Länge M bzw. N ist hingegen allgemeingültig definiert.¹⁹ Zur Berechnung dieses Distanzmaßes betrachtet man zwei Punkte, die auf den beiden zu vergleichenden Kurven von einem Ende zum anderen entlang wandern, und zwar mit einer beliebigen (auch nicht-konstanten) Geschwindigkeit. Für jeden Zeitpunkt t ($t \in [0;1]$) geben beide Funktionen $\alpha(t)$ für die Kurve P und $\beta(t)$ für die Kurve Q an, welche Strecke die Punkte bereits zurückgelegt haben, es muss also gelten:

$$\alpha: [0;1] \rightarrow [0;M]; \beta: [0;1] \rightarrow [0;N]; \alpha(t), \beta(t) \text{ monoton steigend}$$

Die (absolute) Position der beiden Punkte zum Zeitpunkt t wird durch die Funktionen $P(\alpha(t))$ bzw. $Q(\beta(t))$ angegeben. Die Fréchet-Distanz ist nun definiert als:

$$\delta_F(P, Q) = \min(\max d(P(\alpha(t)), Q(\beta(t))))$$

D. h., man misst den maximalen euklidischen Abstand der beiden Punkte – während sie auf ihren Kurven entlang wandern – für ein Funktionenpaar $((\alpha_1(t); \beta_1(t)))$. Dann ersetzt man die Funktionen $\alpha_1(t)$ und $\beta_1(t)$ durch andere Funktionen, welche o. g. Voraussetzungen erfüllen, und misst erneut. Dies wiederholt man mit allen möglichen Funktionen für $\alpha(t)$ und $\beta(t)$ (in der Praxis natürlich nur mit einer endlichen Anzahl von Funktionen). Das Minimum aller ermittelten Maxima ist schließlich die Fréchet-Distanz.

Die dritte Gruppe der Vergleichsmethoden basiert auf dem Vergleich bestimmter Regionen der Form, also z. B. Überlappungsfläche, geometrische Differenz usw..

¹⁸ Vgl. Luckhaus-Ferrero, Rösner (2003), S. 2.

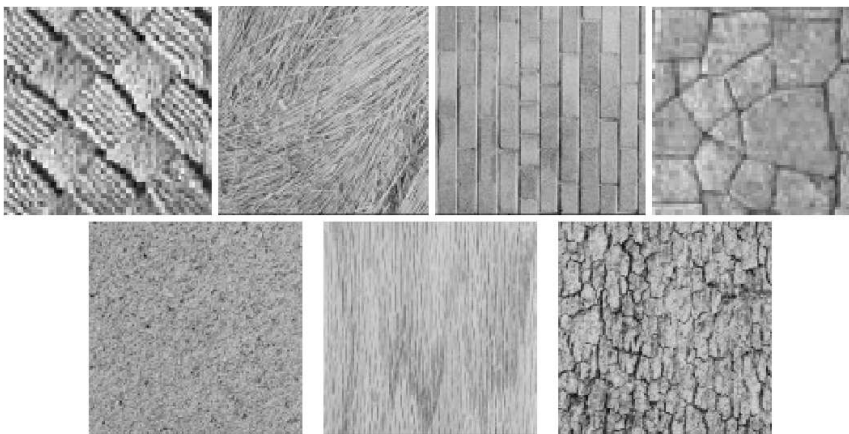
¹⁹ Vgl. Pelletier (2002).

2.3.4 Transformationen

Der Vollständigkeit halber sollen hier kurz die Transformationsverfahren erwähnt werden. Sie basieren hauptsächlich auf der Idee, als Ähnlichkeitsmaß den Aufwand anzusehen, den man treiben müsste, um eine der beiden zu vergleichenden Formen in die andere zu überführen (also quasi zu verbiegen). Zu erwähnen wären hier z. B. Modales Matching oder der Verformungsenergie-Ansatz. Diese Ansätze sind allerdings mathematisch relativ komplex, weshalb hier auf eine detailliertere Beschreibung verzichtet werden soll.

2.4 Texture-based Image Retrieval

Obwohl (oder gerade weil) Texturen ein intuitiv verständliches Konzept sind, gibt es keine allgemein anerkannte Definition des Begriffs. Eine mögliche Definition ist jedoch, dass unter dem Begriff die räumliche Variation der Rückstreu-Eigenschaften einer Oberfläche verstanden wird.²⁰ Im Bezug auf Bilder ist vielleicht am ehesten noch die Aussage zutreffend, dass Texturen diejenigen Informationen im Bild sind, welche noch übrigbleiben, nachdem man Farb- und Form-Informationen entfernt hat.



Quelle: Ma (2002), S. 2.

Abb. 2.14: Beispiele für Texturen

Trotz dieses etwas spärlichen theoretischen Unterbaus kann die Untersuchung von in Bildern enthaltenen Texturen oft hilfreich sein, etwa um Flächen ähnlicher Farbe zu differenzieren (z. B. Meer/Himmel). So kann die Textur-Analyse nicht nur dazu benutzt werden, in einer Datenbank ähnliche Bilder zu finden, sondern auch, um verschiedene Regionen eines Bildes zu identifizieren (z. B. zur Objekterkennung).

Allgemein unbestritten ist, dass Texturen aus Basiselementen – den sogenannten Texeln – bestehen, die je nach Textur regelmäßig oder zufällig angeordnet sein können. Natürliche Texturen sind dabei meist zufällig, während künstliche oft gewissen Gesetzmäßigkeiten folgen. Einige Texturen sind hierarchisch aufgebaut. Bei einer Ziegelmauer bilden die Ziegel in der Gruppe ein Muster, während jeder einzelne Ziegel auch eine eigene Textur besitzt.

²⁰ Vgl. Kaleschke (2003), S. 77.

2.4.1 Texturmerkmale

Um Texturen zu klassifizieren, ist es notwendig, bestimmte Eigenschaften einer Textur zu definieren, die bei einer konkreten Textur gemessen oder zumindest anderweitig beurteilt werden können. Dazu existieren mehrere Ansätze:

Rao und Lohse unterscheiden hier Periodizität (periodische Verteilung von Grundmustern), Gerichtetheit (eine Holzmaserung ist gerichtet, Sand hingegen nicht) und Komplexität (die Schwierigkeit, die Textur zu beschreiben). Tamura, Mori und Yamawaki hingegen charakterisieren Texturen nach Grobheit (Größe der Grundmuster der Textur), Kontrast, Gerichtetheit, Linienhaftigkeit, Regelmäßigkeit und Rauheit.²¹

2.4.2 Grauwert-Kookkurrenzmatrix (GLCM)

Eine Möglichkeit zur Beschreibung von Texturen ist die Grauwert-Kookkurrenzmatrix (Grey Level Co-Occurrence Matrix, GLCM) von Haralick et al.²² Sie drückt aus, mit welcher relativen Häufigkeit zwei Grautöne in einer bestimmten Anordnung in einem Bild vorkommen:

$$C_{ij} = \frac{H_{ij}}{\sum_{i=1}^G \sum_{j=1}^G H_{ij}},$$

mit G Anzahl der verschiedenen Grautöne im Bild, H_{ij} Absolute Häufigkeit des gemeinsamen Auftretens der Grautöne i und j.²³

Das Vorgehen zur Erstellung der Matrix lässt sich am besten am Beispiel erklären:

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 3 |
| 1 | 2 | 3 | 4 | 4 |
| 2 | 3 | 4 | 4 | 3 |
| 3 | 4 | 3 | 3 | 2 |
| 3 | 3 | 2 | 2 | 1 |

Abb. 2.15: Beispiel-Bild für GLCM

Abb. 2.15 zeigt ein Beispiel-Bild, das vier verschiedene Grau-Werte enthält. Um nun eine GLCM zu diesem Bild zu berechnen, muss zunächst einmal festgelegt werden, in welchem Abstand bzw. welcher Position die zu vergleichenden Grauwerte zueinander liegen sollen. Für das Beispiel wird hier angenommen, dass nur direkt nebeneinander liegende Grauwerte interessant sein sollen, der Offset-Vektor $d = (dx;dy)$ wird hier also auf $(1;0)$ gesetzt. Alternativ sind natürlich auch andere Offset-Vektoren möglich, für die sich dann auch jeweils eine andere GLCM ergibt. Außerdem wird die relative Posi-

²¹ Vgl. Novychny (2002), S. 4.

²² Vgl. Haralick, Shanmugam, Dinstein (1973).

²³ Vgl. Kaleschke (2003), S. 78.

tion der zu vergleichenden Grauwerte auch oft über einen Abstand d und einen Winkel θ (typischerweise gleich 0° , 45° , 90° oder 135°) angegeben.

Zur Berechnung der einzelnen Werte der Matrix H wird nun für jedes Grauwert-Paar (i, j) überprüft, wie oft die beiden Werte direkt nebeneinander vorkommen. Exemplarisch für die Werte 1 und 2: Der Grauwert 2 kommt genau zweimal direkt rechts neben dem Grauwert 1 vor. Die übrigen Werte errechnen sich entsprechend und sind in folgender Abbildung aufgeführt:

| | | | |
|---|---|---|---|
| 1 | 2 | 0 | 0 |
| 1 | 1 | 3 | 0 |
| 0 | 2 | 3 | 3 |
| 0 | 0 | 2 | 2 |

Abb. 2.16: Matrix H der absoluten Häufigkeiten der Grauwert-Paare

Um nun die GLCM-Matrix zu erhalten, müssen alle Werte aus der Matrix in Abb. 2.16 durch die Anzahl der Grauwert-Paare (hier durch 20) dividiert werden:

| | | | |
|------|------|------|------|
| 0,05 | 0,1 | 0 | 0 |
| 0,05 | 0,05 | 0,15 | 0 |
| 0 | 0,1 | 0,15 | 0,15 |
| 0 | 0 | 0,1 | 0,1 |

Abb. 2.17: GLCM-Matrix für das Beispiel-Bild

Um die im Bild enthaltenen Texturen zu beschreiben, wird meist nicht die GLCM selbst, sondern auf ihr basierende statistische Informationen benutzt. Einige Möglichkeiten hierfür sind im Folgenden aufgeführt:

$$\text{Entropie: } \sum_{i=1}^G \sum_{j=1}^G C_{ij} \log C_{ij}$$

$$\text{Unähnlichkeit: } \sum_{i=1}^G \sum_{j=1}^G C_{ij} |i - j|$$

$$\text{Kontrast: } \sum_{i=1}^G \sum_{j=1}^G C_{ij} (i - j)^2$$

$$\text{Korrelation: } \frac{\sum_{i=1}^G \sum_{j=1}^G (i - \mu_x)(j - \mu_y) C_{ij}}{\sigma_x \sigma_y},$$

mit μ_x , μ_y Mittelwerte und σ_x , σ_y Standardabweichungen von

$$C(x) = \sum_{j=1}^G (x, j) \text{ bzw. } C(y) = \sum_{i=1}^G (i, y).$$

Wünschenswert wäre hierbei natürlich, dass diese Maße invariant gegenüber Verschiebungen und Skalierungen des mittleren Grauwertes wären. Beide dieser Anforderungen erfüllt jedoch nur die Entropie, alle anderen GLCM-Merkmale sind nur invariant gegenüber der Verschiebung.

Die GLCM bilden ausschließlich die räumlichen Textureigenschaften, nicht jedoch die Form der Textel ab. Außerdem ist es auch sehr schwer, einen geeigneten Offset-Vektor bzw. ‚gute‘ Werte für d und θ zu finden. Auch mehrere GLCMs für verschiedene Werte zu berechnen ist wegen des notwendigen Rechenaufwandes unpraktikabel. Daher ist die GLCM nur für die Texturklassifikation, nicht aber für die Textursegmentierung geeignet.

2.4.3 Weitere Verfahren des Texture-based Image Retrieval

Zwei weitere Möglichkeiten zum textur-basierten Vergleich von Bildern sollen nur noch kurz angesprochen werden, da für eine detaillierte Beschreibung der mathematisch relativ komplexen Verfahren hier der Platz fehlt.

Gauss-Markov Random Field (GMRF)

Bei der GMRF-Methode wird angenommen, dass die Intensitäten von benachbarten Pixeln über Gauss-Verteilungen miteinander korreliert sind. Das Modell besitzt bestimmte Parameter, die beschreiben, wie stark und auf welche Weise die Intensität eines bestimmten Pixels von den Pixeln in einer symmetrischen Nachbarschaft beeinflusst wird.

Beim GMRF-Verfahren wird nun versucht, diese Parameter für eine gegebene Textur so zu bestimmen, dass sie die Intensitäts-Verteilungen in der Textur möglichst gut nachbilden. Dies geschieht oft im Zuge eines iterativen Prozesses, bei dem zunächst Start-Parameter geschätzt werden. Anschließend wird das Ergebnis – z. B. mit der Methode der kleinsten Quadrate – mit dem Original verglichen, die Parameter auf Basis dieses Vergleiches angepasst und so fort, bis eine befriedigende Übereinstimmung erreicht ist.²⁴

Ist dies für beide Bilder bzw. Texturen geschehen, können die ermittelten Parameter verglichen und so die Ähnlichkeit der beiden Texturen abgeschätzt werden.

*Gabor Filter (Wavelet)*²⁵

Der Gabor Filter basiert auf der Fourier-Transformation für Bilddaten.²⁶ Mit seiner Hilfe ist es möglich, Kanten und ähnliche Strukturen in Bildern zu erkennen. Besonders vorteilhaft für die Textur-Analyse ist hierbei, dass der Gabor-Filter bei der einmaligen An-

²⁴ Zu Einzelheiten der Parameter-Schätzung vgl. Bader, Jájá, Chellappa (1994).

²⁵ Vgl. Zhang, D. u. a.

²⁶ Für Einzelheiten zur Fourier-Transformation vgl. Gliet, Kritzler, Pieke, Schreiber (2003).

wendung nur solche regelmäßigen Strukturen erkennt, die in eine bestimmte Richtung laufen und eine bestimmte Größe besitzen.

Allerdings lässt sich die Richtung und die Größe der erkannten Strukturen über bestimmte Parameter des Gabor-Filters steuern. Wendet man den Filter also mehrfach auf ein Bild an und verwendet dabei jedes Mal andere Parameter-Werte, erkennt der Filter jeweils nur diejenigen Texturen, die eben die jeweilige Richtung und Grobheit besitzen. So kann eine Eigenschafts-Repräsentation für das Bild erstellt werden, die Auskunft über die darin enthaltenen Texturen gibt (aufgeschlüsselt nach ihrer Richtung und Grobheit). Möglich ist auch, zusätzlich noch zu speichern, wo im Bild sich welche Textur befindet.

Über den Vergleich dieser Merkmale lässt sich dann die Ähnlichkeit zweier Bilder bestimmen.

2.5 Performance-Messung von CBIR-Algorithmen

Zur Messung der Qualität der im CBIR verwendeten Verfahren werden oft zwei aus dem Information Retrieval stammende Messgrößen verwendet, nämlich Precision und Recall.

Sei D eine Menge von Daten, also z. B. eine Menge von Bildern in einer Datenbank. Diese Menge kann ein Mensch nun für eine gegebene Suchanfrage q in zwei Teilmengen einteilen, die Menge der relevanten Bilder $R(q)$ und die Menge der nicht relevanten Bilder $\bar{R}(q)$. Nimmt man nun an, die Anfrage q werde an ein CBIR-System gestellt und dieses liefere eine bestimmte Menge von Bildern $A(q)$ als Antwort zurück. Dann sind Precision und Recall wie folgt definiert:

$$p = \frac{|A(q) \cap R(q)|}{|A(q)|} \quad \text{Precision}$$

$$r = \frac{|A(q) \cap R(q)|}{|R(q)|} \quad \text{Recall}$$

Precision gibt also an, welcher Anteil der als relevant zurückgelieferten Bilder auch wirklich relevant waren, während Recall misst, welcher Anteil der relevanten Bilder gefunden wurde. Ideal wäre ein Algorithmus, der bei beiden Merkmalen eine 1 erreicht. Er würde alle relevanten Bilder zurückliefern, aber kein einziges nicht relevantes.

Leider besitzt dieser Ansatz zur Performance-Messung einige Schwächen: Zum einen ist es bei Bildern auch für Menschen viel schwieriger, zu bestimmen, welche Bilder für eine gegebene Suchanfrage relevant sind, als z. B. bei Texten, weil die Bedeutung eines Bildes fast nie eindeutig ist. Zum anderen liefern CBIR-Verfahren meist keine undifferenzierte Menge angeblich relevanter Bilder zurück, sondern eine geordnete Liste, in der die Bilder nach Ähnlichkeit zum Suchbild sortiert sind. Die Länge dieser Liste ist oft auf k Bilder begrenzt (sonst würden schließlich alle Bilder der Datenbank – sortiert nach ihrer Ähnlichkeit zum Suchbild – zurückgegeben). Ist k nun kleiner als die Menge

der relevanten Bilder in der Datenbank, kann das Recall-Maß nicht mehr sinnvoll angewendet werden.

Trotz dieser Einschränkungen werden Recall und Precision oft eingesetzt. Darüber hinaus gibt es jedoch viele weitere Ansätze zur Messung der Qualität von CBIR-Systemen.²⁷

2.6 Optimierungsmöglichkeiten

Um die Suche nach Bildern in Bilddatenbanken effizienter zu machen, gibt es verschiedene Ansätze. Drei von ihnen sollen hier kurz vorgestellt werden:

2.6.1 Ablage von Bildmerkmalen als Meta-Information

Es ist unsinnig, bei jeder neuen Suchanfrage die benutzten Eigenschafts-Repräsentationen für die Bilder in der Datenbank neu zu berechnen. Günstiger ist es, Histogramme, Eigenschaftsvektoren usw. einmalig (beim Einfügen des Bildes in die Datenbank) zu bestimmen und dann als Meta-Informationen zusammen mit dem Bild abzulegen.

So steigt zwar der Speicherbedarf für die Datenbank, die Antwortzeit für eine Suchanfrage dürfte sich jedoch drastisch verkürzen, weil die zur Charakterisierung der Bilder eingesetzten Algorithmen nur einmal auf das Bild aus der Suchanfrage und nicht auf jedes zu vergleichende Bild in der Datenbank angewendet werden müssen.

Einzig bei den Transformations-Verfahren des Shape-based Image Retrieval²⁸ ist dieses Vorgehen nicht möglich, da sie auf die beiden zu vergleichenden Bilder gemeinsam angewendet werden.

2.6.2 Buckets

Ein weiterer naheliegender Ansatz ist, bei einer Suchanfrage das Suchbild nicht nacheinander mit allen Bildern in der Datenbank zu vergleichen, sondern eine Art Suchbaum aufzubauen. Hierzu werden die Bilder bereits beim Einfügen in die Datenbank untereinander verglichen und in Gruppen (sog. Buckets) eingeteilt. Für jeden Bucket wird dann ein Bild ausgewählt, das ‚seine‘ Gruppe ähnlicher Bilder repräsentieren soll. Dieses Verfahren kann bei einer sehr großen Zahl von Bildern in der Datenbank auch mehrstufig angewendet werden, so dass die Repräsentationsbilder wiederum untereinander verglichen und dann ihrerseits in Buckets aufgeteilt werden. So kann ein Baum von Buckets entstehen.

Bei einer Suchanfrage wird das Suchbild dann zunächst mit den Repräsentationsbildern auf der obersten Ebene verglichen. Wird dabei ein ähnliches Bild gefunden, wird der

²⁷ Vgl. Smeulders u. a. (2000), S. 23f.

²⁸ Siehe Abschnitt 2.3.4.

Vergleich mit allen Bildern desjenigen Buckets, den das Bild repräsentiert, wiederholt, bis schließlich ein oder mehrere hinreichend ähnliche Bilder gefunden wurden.

Denkbar ist hierbei, auf den höheren Stufen der Bucket-Hierarchie die Schwellenwerte, ab denen ein Bild als ähnlich eingestuft wird, zunächst relativ niedrig anzusetzen und sie dann, während man beim Vergleichen die Hierarchie herabwandert, stufenweise zu erhöhen. So könnten zunächst auch relativ unähnliche Bilder in der Gruppe der ‚verdächtigen‘ Bilder verbleiben, um dann ggf. erst bei späteren Vergleichen auf tieferer Ebene ausgeschlossen zu werden. So kann vermieden werden, dass Bilder, die eigentlich dem Suchbild ähnlich sind, im Suchergebnis nicht auftauchen, weil ihr Bucket auf einer höheren Stufe von der weiteren Suche ausgeschlossen wurde, weil das jeweilige Repräsentationsbild nicht hinreichend ähnlich war.

Allerdings bedeutet die Absenkung der Schwellenwerte auf höheren Ebenen der Hierarchie natürlich auch, dass mehr Vergleiche durchgeführt werden, weil weniger Buckets von vornherein ausgeschlossen werden. Hier muss also ein Mittelweg zwischen geringem Aufwand für die Suche einerseits und Vollständigkeit des Suchergebnisses andererseits gefunden werden.

2.6.3 Bild-Vorverarbeitung

Sehr hilfreich für den Vergleich von Bildern auf Basis der in den Abschnitten 2.2 bis 2.4 vorgestellten Verfahren ist eine geeignete Vor-Verarbeitung sowohl der Bilder in der Datenbank als auch der Suchbilder. Ähnlichkeiten oder Unterschiede zweier Bilder können von entsprechenden Algorithmen sehr viel leichter erkannt werden, wenn beide Bilder identisch (oder wenigstens sehr ähnlich) im Bezug auf Größe, Belichtung, Betrachtungswinkel der abgebildeten Objekte usw. sind. Da Techniken, die eine weitgehende Angleichung dieser Parameter ermöglichen, jedoch bereits in einer anderen Ausarbeitung zu diesem Seminar ausführlich behandelt werden, soll hier nicht weiter auf diesen Aspekt eingegangen werden.²⁹

2.7 Beispiele für konkrete CBIR-Systeme

Heute ist bereits eine ganze Reihe von CBIR-Systemen im Einsatz, die u. a. die hier vorgestellten Verfahren nutzen. Für eine Vorstellung oder sogar einen Vergleich bezüglich der eingesetzten Methoden und Qualität der erzielten Ergebnisse fehlt hier jedoch der Platz.³⁰

Anhang A enthält jedoch die Internet-Adressen einiger frei zugänglicher CBIR-Systeme. Der geneigte Leser mag sich so einen eigenen Eindruck vom Stand der Technik verschaffen.

²⁹ Zu Details hierzu vgl. Gliet, Kritzler, Pieke, Schreiber (2003).

³⁰ Einen Vergleich aktueller CBIR-Systeme findet sich in Veltkamp, Tanase (2002).

3 Techniken zur Minendetektion und -identifikation³¹

In diesem Kapitel soll kurz die Funktionsweise einiger Techniken zur Entdeckung und Identifikation von Minen und UXOs vorgestellt werden. Die detailliertere Darstellung beschränkt sich dabei auf solche Techniken, welche Bilder liefern, die sich mit Methoden der Bildverarbeitung – und hier speziell des Image Retrievals – weiterverarbeiten lassen.

3.1 Fotografie

Befindet sich die Mine an der Erd-Oberfläche, kann der Einsatz ‚normaler‘ Fotografie sinnvoll sein. Allerdings reicht schon eine dünne Erd-Schicht aus, um die Mine für diese Art von Sensoren unsichtbar zu machen. Auch Pflanzen können die Mine verdecken und so eine Entdeckung per Fotografie verhindern.

Daher wird Fotografie im sichtbaren Spektrum wohl eher bei der Identifikation einer bereits gefundenen und auch schon freigelegten Mine zum Einsatz kommen.³²

3.2 Infrarot-Kameras

Mithilfe von Infrarot-Kameras lassen sich sehr schnell und einfach Temperaturbilder von größeren Gebieten erzeugen. In Bezug auf die Minenräumung sind hierbei drei Effekte relevant:

Zum ersten unterscheidet sich eine Mine stark von normalem Erdboden, was ihre Wärmefähigkeit, -aufnahme, -abgabe und -speicherfähigkeit betrifft. Dadurch entstehen vor allem bei Temperaturschwankungen (z. B. am Vormittag, wenn die Sonne ein Gelände erwärmt) Temperaturunterschiede zwischen Mine und dem umgebenden Material (Erde). Auch vergrabene Minen können so teilweise erkannt werden, weil die Mine den Erdboden in ihrer direkten Umgebung kühlt oder wärmt.

Zweitens ist es beim Vergraben von Minen unerlässlich, zunächst ein Loch zu graben, die Mine hineinzulegen und sie danach wieder mit Erde zu bedecken. Da diese Deckschicht aus Erde anschließend aus naheliegenden Gründen nicht (z. B. durch einfaches Festtreten oder mit technischen Hilfsmitteln) wieder verdichtet werden kann, besitzt die oberste Erdschicht an dieser Stelle anschließend eine andere Dichte als im umgebenden Gelände. Dieser Dichteunterschied wirkt sich ebenfalls auf die Wärmeleit- und -speicherfähigkeit des Materials aus, weshalb auch solche Spuren von vergrabenen Minen noch für eine gewisse Zeit (einige Wochen) durch Infrarot-Kameras entdeckt werden können.

Der dritte Ansatz bezieht auch noch die Polarisation des emittierten und reflektierten Infrarotlichtes mit ein, da diese bei natürlichem Erdboden und künstlich hergestellten Objekten unterschiedlich ist.

³¹ Vgl. Bruschini u. a. (1999), S. 7-17.

³² Siehe Kapitel 4.

Wie schon erwähnt, wird die Infrarot-Fotografie bisher fast ausschließlich dazu verwendet, auf großflächigen Aufnahmen (z.B. vom Flugzeug aus) Minen zu erkennen. Für diese Aufgabe sind jedoch Methoden der Objekterkennung (die hier nicht Gegenstand der Betrachtung sind) wesentlich besser geeignet als solche des CBIR. Denkbar wäre natürlich, ein detailliertes Infrarot-Bild einer bereits gefundenen Mine aufzunehmen und dieses dann mittels Image Retrieval weiterzuverarbeiten. Ob dieses Vorgehen jedoch Vorteile gegenüber der Verwendung von Fotografie im sichtbaren Spektrum brächte, erscheint zweifelhaft.

3.3 Röntgengeräte

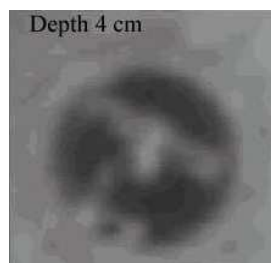
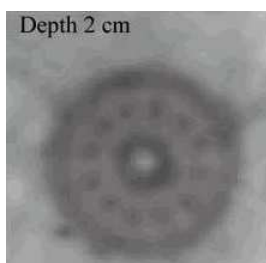
Auch Röntgenstrahlen werden zur Minendetektion eingesetzt. Da es bei der Minensuche normalerweise aber nicht möglich ist, wie bei medizinischen Röntgen-Aufnahmen, einen Röntgen-Film hinter/unter dem zu untersuchenden Objekt anzubringen, arbeiten die hier benutzten Geräte anders. Sie beschießen den Untergrund mit Röntgen-Strahlen und messen dann die vom Boden zurückgestreute Strahlung. Daher nennt man diese Technik auch XBT (X-Ray Backscatter Technology).



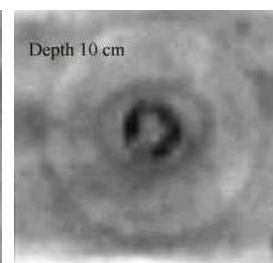
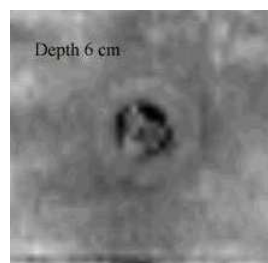
PPM-2 Anti-Personen-Mine,
ehemalige DDR, Durchmesser ca. 12 cm



TM-62M Panzer-Mine,
ehemalige UDSSR, Durchmesser ca. 32 cm



XBT-Bilder einer vergrabenen PPM-2
in Tiefen von 2 und 4 cm



XBT-Bilder einer vergrabenen TM-62M
in Tiefen von 6 und 10 cm

Quelle: Fotos (v.l.n.r.) NPA (PPM-2), Niemann u. a. (2002), NPA (TM-62);
XBT-Aufnahmen: Niemann u. a. (2002)

Abb. 3.1: XBT-Bilder zweier Minen-Typen

Da für dieses Verfahren jedoch schädliche Röntgen-Strahlung freigesetzt werden muss und die benötigten technischen Geräte (noch) sehr groß und schwer sind, eignet es sich eher für die nähere Untersuchung eines schon gefundenen Objektes als für das Absuchen großer Gebiete. Allerdings sind auch schon tragbare Geräte entwickelt worden, die aber keine Bilder erzeugen.³³

³³ Vgl. Carts-Powell (2000).

Röntgen-Verfahren haben zwei große Vorteile: Zum einen können sie Bilder auch von vergrabenen Objekten erzeugen. Dies ist sehr hilfreich, weil das Ausgraben einer Mine oder eines UXOs (mal abgesehen vom Arbeitsaufwand) selbstverständlich immer eine Gefahr für den menschlichen Minenräumer bedeutet. Außerdem können, wie in Abb. 3.1 zu erkennen, zweidimensionale Abbildungen für verschiedene Tiefen gewonnen werden. So kann man nicht nur Informationen über die äußere Form des Objektes, sondern auch über seinen inneren Aufbau gewinnen.

3.4 Weitere Techniken

Neben den schon genannten gibt es noch weitere bilderzeugende Techniken zur Minensuche und -identifikation (z. B. Metalldetektoren). Die meisten von ihnen erzeugen zunächst nur punktuelle Messwerte. Führt man jedoch viele einzelne Messungen an jeweils leicht versetzten Orten durch, lassen sich diese Einzelwerte zu zweidimensionalen ‚Karten‘ des untersuchten Geländes zusammensetzen. Verschiedene Farb- oder Grautöne könnten auf diesen Karten die unterschiedlichen Messwerte repräsentieren und so durchaus hilfreiche Visualisierungen liefern.

Die Auflösung der resultierenden Bilder ist aufgrund der Funktionsweise der zugrundeliegenden Messtechniken jedoch viel zu gering, um mittels Image Retrieval weiterverarbeitet zu werden. Trotzdem könnten auch diese Techniken, kombiniert mit Objekterkennungs-Algorithmen, zur Minensuche eingesetzt werden. Da Objekterkennung hier jedoch – wie bereits mehrfach erwähnt – nicht behandelt wird, sollen die verschiedenen Messverfahren an dieser Stelle nicht näher betrachtet werden.³⁴

³⁴ Für detaillierte Beschreibungen vgl. Bruschini u. a. (1999), S. 7-17, Quantum Magnetics (2002) und Nebbia u. a. (1999).

4 Anwendungsmöglichkeiten von CBIR bei der Minenräumung

Wie in Abschnitt 3.2 bereits kurz angesprochen, eignen sich die in Kapitel 2 vorgestellten Verfahren kaum zur Entdeckung von Minen. Hier sind Verfahren der Objekterkennung wesentlich geeigneter.³⁵

Ist allerdings erst einmal ein verdächtiges Objekt gefunden, kann das CBIR sehr nützlich sein, um das gefundene Objekt zu identifizieren. Voraussetzung hierfür ist allerdings, dass mittels der in den Abschnitten 3.1 bis 3.3 beschriebenen Verfahren detailliertere Aufnahmen des Objektes erstellt werden. Ist dies geschehen, können diese Bilder – eventuell nach vorheriger Übertragung und/oder Vorverarbeitung – als Suchbild für eine Anfrage an ein CBIR-System genutzt werden. Enthält die zugehörige Bild-Datenbank ein ähnliches Bild, kann dieses dem menschlichen Minenräumer zur Verfügung gestellt werden, eventuell zusammen mit weiteren in der Datenbank abgelegten Informationen wie etwa Kategorisierungs-Informationen (z. B. Panzermine, Antipersonenmine, UXO, Blechdose, ...), Typenbezeichnungen oder sogar Hinweisen zur Entschärfung.

Selbst wenn anhand des Suchbildes keine eindeutige Identifizierung des Objektes möglich ist, ist es für den menschlichen Minenräumer sicher eine große Hilfe, wenn das CBIR-System ihm zu dem Bild der vorgefundenen Mine eine Auswahl von vielleicht zehn ähnlich aussehenden Minen-Modellen zurückliefert. In einem solchen Fall wäre er sicher leicht in der Lage, die gefundene Mine durch Vergleich des gerade aufgenommenen Bildes mit denen aus der Datenbank zu identifizieren.

Nun soll noch kurz auf konkrete Kombinationsmöglichkeiten der in Kapitel 2 genannten CBIR-Verfahren mit den in Kapitel 3 vorgestellten Bildgewinnungs-Verfahren eingegangen werden:

Fotografie

Bilder aus dem Spektrum des sichtbaren Lichts lassen sich sehr gut mit den vorgestellten CBIR-Verfahren weiterverarbeiten (schließlich wurden diese genau zu diesem Zweck entwickelt). Es lassen sich Algorithmen aller drei CBIR-Bereiche (Color-, Shape- und Texture-based Image Retrieval) benutzen. Am Erfolg versprechendsten ist hier sicher eine Kombination mehrerer Verfahren. Welche Algorithmen allerdings speziell für die Identifikation verschiedener Minen geeignet sind, müsste zunächst in entsprechenden Testreihen festgestellt werden.

Infrarot

Da Infrarot-Aufnahmen natürlich nicht im sichtbaren Spektrum aufgenommen werden, enthalten sie im allgemeinen keine Farb-Informationen im eigentlichen Sinne. Zwar werden auch Infrarot-Aufnahmen oft eingefärbt, um sie für den menschlichen Beobachter leichter erfassbar zu machen. Dabei werden verschiedene Temperaturen der abgebildeten Objekte durch Farben symbolisiert (z. B. Blau für kalt oder Rot für warm). Dabei

³⁵ Zu Details hierzu vgl. Grosse-Lordemann, Lambers (2003).

handelt es sich jedoch um relativ willkürliche Zuordnungen von Farben durch den Menschen (oder durch von ihm geschaffene Geräte) zu Temperaturwerten und nicht um wirkliche Eigenschaften des abgelichteten Gegenstandes an sich.

Daher erscheinen auf diesem Gebiet die Verfahren des Color-based Image Retrieval relativ ungeeignet. Ein und dasselbe Objekt kann schließlich auf einer Infrarot-Aufnahme rot aussehen, auf einer anderen aber blau, nur weil seine Temperatur sich auf den Aufnahmen unterscheidet oder die Kamera unterschiedlich eingestellt war. Die Bildersuche auf Basis von Form oder Textur sollte jedoch auch mit Infrarot-Aufnahmen ohne Weiteres möglich sein. Viele Verfahren dieser Kategorien arbeiten sowieso auf Basis von Graustufen-Bildern, so dass die fehlenden Farben kaum stören sollten.

Die einzige Einschränkung müsste vielleicht bei sehr heißen Objekten getroffen werden. Diese wirken auf Infrarot-Aufnahmen oft wie ein blendender Scheinwerfer im sichtbaren Bereich. Dadurch kann es zur ‚Überstrahlung‘ von Texturen und Kanten kommen, was zu Problemen bei der Weiterverarbeitung im Rahmen des CBIR führen kann. Allerdings sind die wenigsten Minen beheizt, so dass dieser Effekt in diesem speziellen Anwendungsbereich kaum praxisrelevant sein dürfte.

Röntgen-Geräte

Wie in Abb. 3.1 zu erkennen, liefern auch diese Bilder keine Farbinformationen, so dass auch hier das Color-based Image Retrieval ausscheidet. Auch die texturbasierte Analyse könnte durch das doch relativ starke Rauschen auf XBT-Aufnahmen stark behindert werden. Hier wäre in jedem Fall eine sehr intensive Vorverarbeitung der Suchbilder nötig.

Dadurch, dass Röntgengeräte aber auch Informationen über den inneren Aufbau der abgebildeten Gegenstände liefern können, bieten sie wahrscheinlich ein großes Potential für das Shape-based Image Retrieval. Mittels entsprechender Bilder in der Datenbank könnten verschiedene Objekte und Minentypen nicht nur aufgrund ihrer äußeren Form, sondern auch durch Vergleich der inneren Struktur unterschieden werden.

5 Fazit

Es bleibt festzuhalten, dass das CBIR sehr wohl einen hilfreichen Beitrag zur Minenräumung leisten kann. Für das Erkennen von Minen auf entsprechenden Bildern (z.B. aus dem Flugzeug aufgenommenen Infrarot-Aufnahmen) ist es zwar nur mäßig geeignet, hier sollten eher Objekterkennungs-Algorithmen eingesetzt werden.

Bei der Klassifizierung und Identifizierung eines bereits gefundenen Gegenstandes kann es jedoch gute Dienste leisten. Dabei ist es oft gar nicht nötig, dass ein CBIR-System ein gefundenes Objekt vollkommen selbständig erkennt. Sehr hilfreich für menschliche Minensucher wäre es auch schon, wenn ein CBIR-System aus der Datenbank eine kleine Gruppe ähnlich aussehender Bilder herausucht, um ihnen somit die Klassifikation zu erleichtern.

Besonders in Verbindung mit XBT-Aufnahmen könnte so der mit Fehlalarmen verbundene Aufwand beträchtlich gesenkt und somit die Effizienz des gesamten Minenräumvorgangs wesentlich gesteigert werden.

Literaturverzeichnis

- Bader, D.; Jájá, J.; Chellappa, R.: Scalable data parallel algorithms for texture synthesis using Gibbs random fields. 1994.
<http://www.umiacs.umd.edu/research/EXPAR/papers/ITIP-543/ITIP-543-html.html>. Abrufdatum 2003-12-15.
- Block, K.; Lucks, A.; Tippmann, M.: Das RGB-Farbmodell. 2001.
<http://olli.informatik.uni-oldenburg.de/Grafiti3/grafiti/flow12/page3.html>. Abrufdatum 2003-12-15.
- Bruschini, C. u. a.: Study on the State of the Art in the EU related to humanitarian demining technology, products and practice. EUDEM: The EU in humanitarian DEMining Final Report. Brüssel 1999.
- Carts-Powell, Y.: Portable mine detector borrows x-ray astronomy methods. 2000.
<http://www.spie.org/web/oer/march/mar00/cover2.html>. Abrufdatum 2003-12-15.
- Ciocca, G.; Schettini, R.; Zuffi, S.: A survey for methods for colour image indexing and retrieval in image databases. Mailand 2001.
- Dlugosz, S.: Shape-based Image Retrieval. Ausarbeitung zum Seminar Content-Based Retrieval of Image and Video-Databases. Universität Münster, Münster 2002.
- Domik, G.: Perzeptuelle Farbräume. 1997. http://www.uni-paderborn.de/fachbereich/AG/agdomik/computergrafik/cg_skript/html/node142.htm. Abrufdatum 2003-12-15.
- Eakins, J.; Graham, M.: Content-based Image Retrieval. A report to the JISC Technology Applications Programme. 1999.
<http://www.unn.ac.uk/iidr/report.html>. Abrufdatum 2003-12-15.
- Gliet, J.; Kritzler, M.; Pieke, B.; Schreiber, T.: Bildaufbereitung für die Objekterekennung. Ausarbeitung zum Seminar Unterstützung von Landminendetektion durch Bildauswertungsverfahren und Robotereinsatz. Universität Münster, Münster 2003.
- Grosse-Lordemann, C.; Lambers, M.: Objekterkennung in Bilddaten. Ausarbeitung zum Seminar Unterstützung von Landminendetektion durch Bildauswertungsverfahren und Robotereinsatz. Universität Münster, Münster 2003.
- Hafner, J. u. a.: Efficient color histogram indexing for quadratic form distance functions. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. Hrsg.: Institute of Electrical and Electronics Engineers Inc. San Jose 1995 (Band 17, Ausgabe 7), S. 729-736.
- Haralick, R. M.; Shanmugam, K.; Dinstein, I.: Textural features for image classification. In: IEEE Transactions on Systems Man and Cybernetics. Hrsg.: Institute of Electrical and Electronics Engineers Inc. 1973 (Band 3, Ausgabe 6), S. 610-621.
- Kaleschke, L.: Fernerkundung des Meereises mit aktiven und passiven Mikrowellensensoren. Dissertation, Universität Bremen, Bremen 2003.

- Luckhaus-Ferrero, R.; Rösner, S.: Gesichtserkennung. Foliensatz zum Seminar Multimedia Retrieval. Universität Hildesheim, Hildesheim 2003.
- Ma, J.: Texture-based Image Retrieval. Ausarbeitung zum Seminar Content-Based Retrieval of Image and Video-Databases. Universität Münster, Münster 2002.
- Manjunath, B. S. u. a.: Color and Texture Descriptors. In: IEEE Transactions on Circuits and Systems for Video Technology. Hrsg.: Institute of Electrical and Electronics Engineers Inc. 2001 (Band 11, Ausg. 6), S. 703-717.
- Nebbia, G. u. a.: EXPLODET: Advanced nuclear techniques for humanitarian demining. In: Mine'99 Proceedings. Euroconference on Sensor systems and signal processing techniques applied to the detection of mines and unexploded ordnance. Hrsg.: Universität Firenze, Firenze 1999, S. 78-81.
- Niemann, W. u. a.: Detection of buried landmines with X-ray backscatter technology. 2002. <http://www.ndt.net/article/ecndt02/96/96.htm>. Abrufdatum 2003-12-15.
- Novychny, O.: Texture-based Image Retrieval. Ausarbeitung zum Seminar Content-Based Retrieval of Image and Video-Databases. Technische Universität Berlin, Berlin 2002.
- NPA (Norwegian Peoples Aid) (PPM-2): Landmines in Angola: PPM-2. http://www.angola.npaid.org/mine_germany_ppm_2.htm. Abrufdatum 2003-12-15.
- NPA (Norwegian Peoples Aid) (TM-62): Landmines in Angola: TM-62M and the TM-62 series. http://www.angola.npaid.org/mine_russia_tm_62m.htm. Abrufdatum 2003-12-15.
- Pelletier, S.: Computing the Fréchet distance between two polygonal curves. 2002. <http://www.cim.mcgill.ca/~stephane/cs507/Project.html>. Abrufdatum 2003-12-15.
- Quantum Magnetics: Quadrupole Resonance. 2002. http://www.qm.com/core_technology/quadrupole_resonance_body.htm. Abrufdatum 2003-12-15.
- Smeulders, A. u. a.: Content-based Image Retrieval at the end of the early years. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. Hrsg.: Institute of Electrical and Electronics Engineers Inc. 2000 (Band 22, Ausgabe 12), S. 1-32
- Tomberge, P.: Color-based Image Retrieval. Ausarbeitung zum Seminar Content-Based Retrieval of Image and Video-Databases. Universität Münster, Münster 2002.
- Trörlitzsch, J.: Color-based Image Retrieval. Foliensatz zum Seminar Content-Based Retrieval of Image and Video-Databases. Technische Universität Berlin, Berlin 2002.
- Veltkamp, R.; Tanase, M.: Content-based Image Retrieval systems: A survey. 2002. <http://give-lab.cs.uu.nl/cbirsurvey/cbir-survey.pdf>. Abrufdatum 2003-12-16.
- Wu, J. C.: Presentation on Template-based Image Retrieval. Foliensatz. Tsing Hua Universität Hsinshu, Hsinshu 1998.

Zhang, D. u. a.: Content-based Image Retrieval using Gabor texture filters.
<http://www.gscit.monash.edu.au/~dengs/resource/papers/pcm00.pdf>.
Abrufdatum 2003-12-15.

Anhang

A Internet-Adressen einiger frei zugänglicher CBIR-Systeme

- FIDS (Flexible Image Database System):
<http://www.cs.washington.edu/research/imagedatabase/demo/fids/>
- ImageRETRO (Image RETrieval by Reduction and Overview):
<http://python.wins.uva.nl:6022/>
- LCPD (Leiden 19th Century Portrait Database):
<http://ind156b.wi.leidenuniv.nl:2000/>
- MARS (Multimedia Analysis and Retrieval System):
<http://www-db.ics.uci.edu/pages/demos/index.shtml>
- PicToSeek:
http://zomax.wins.uva.nl:5345/ret_user/
- SQUID (Shape Queries Using Image Databases):
<http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/squid.html>