

Techniken der Farbsegmentierung ausgehend von Fuzzy-Clusteranalyse im Farbraum

Josef Richardt

Gesellschaft zur Förderung angewandter Informatik e.V. (GFal)
Rudower Chaussee 30, 12489 Berlin-Adlershof
E-Mail: richardt@gfai.de

1. Einleitung

Unter Farbsegmentierung versteht man grob gesagt die Erkennung zusammenhängender Flächen im Bild, die sich durch eine bestimmte Farbe auszeichnen. Das Erkennen solcher Flächen beherrscht ein Mensch intuitiv und mühelos, bei einer Automatisierung solcher Erkennungsvorgänge tritt jedoch das Problem auf, dass im Normalfall schon bei einfachen Anwendungsfällen "die Farbe" einer bestimmten Teilfläche im Bild nicht so eindeutig definiert ist, wie es intuitiv zu sein scheint. Diese Einschätzung wird deutlich, wenn man anhand von Bildbeispielen das Histogramm der Farben des Bildes im RGB-Farbraum betrachtet. Ein Einblick in den dreidimensionalen Farbraum ist auf der Basis von 3D-Grafik (OpenGL) in der Tat möglich.

Ausgehend von einer Zerlegung der Farben eines Bildes im Farbraum (Clusteranalyse) lassen sich Farbsegmente im Bild bestimmen. In diesem Vortrag werden dazu in der GFal entwickelte Techniken der Farbsegmentierung vorgestellt und diskutiert. Die dabei verfolgte Anwendungsrichtung besteht darin, dass im Bereich der Lebensmittelindustrie eine Erkennung der "Farbe der Bildobjekte" eine Grundaufgabe im Rahmen einer automatischen Qualitätsüberwachung ist. Was hierbei als "Farbe der Bildobjekte" zu verstehen ist, erweist sich bei genauerer Untersuchung als ein recht unscharfer Begriff.

2. Fuzzy-Farbhistogramme und deren Visualisierung mit OpenGL

Im RGB-Farbmodell wird die Farbe eines Pixels repräsentiert durch einen Vektor im dreidimensionalen RGB-Raum. Wenn man die RGB-Vektoren aller Pixel eines Bildes im RGB-Raum darstellt, erhält man dort eine Punktwolke. Diese Punktwolke lässt sich schlecht mit einer mit 3D-Grafik sichtbar machen, weil im Normalfall die Anzahl der Punkte zu hoch ist. Eine anschauliche Darstellung lässt sich dennoch gewinnen, indem der dreidimensionale RGB-Raum mit einer Gitterstruktur gerastert wird. Dazu werden die Bereiche der RGB-Werte (dies sind die Bereiche $0 \leq R \leq 255$, $0 \leq G \leq 255$, $0 \leq B \leq 255$ im Fall einer 8-bit-Auflösung) mit einer konkreten, aber änderbaren Schrittweite eingeteilt, und die Punkte zwischen den Gitterpunkten werden mit einer Fuzzy-Zugehörigkeit den benachbarten Gitterpunkten zugeordnet. Auf diese Weise erhält man für alle Gitterpunkte im RGB-Raum eine Belegung, die zum Ausdruck bringt, wie viele RGB-Werte eines Bildes sich in der Nähe des jewei-

ligen Gitterpunktes befinden. Es handelt sich hierbei um ein 3D-Histogramm. Dieses kann visualisiert werden mit Hilfe von 3D-Computergrafik (OpenGL). Man erhält somit einen Einblick in den RGB-Raum. Dabei sieht man in jedem Gitterpunkt eine mehr oder weniger große Häufung der an dieser Stelle des Farbraumes vorkommenden Farben.

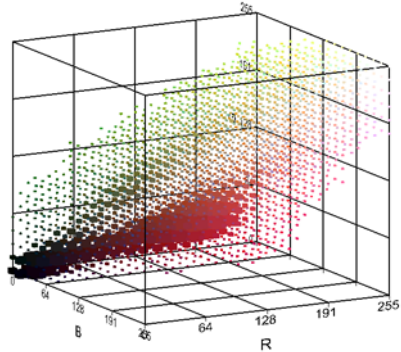


Abbildung 1: Ein Bild mit Erdbeeren wurde in den RGB-Farbraum abgebildet. Beim hier verwendeten Gitter ist jede der Achsen R, G, B in 15 gleiche Teile untergliedert. Die Punkte im Farbhistogramm werden in einer (2D-) Größe dargestellt, die proportional zur Anzahl der zu diesem RGB-Wert gehörigen Pixel ist.

Für die Visualisierung des Histogramms, d. h. der summierten Zugehörigkeitswerte im RGB-Raum, sind Techniken zur Skalierung der Darstellung nötig. Die Ansicht des Farbhistogramms hängt von einer Vielzahl von Parametern ab. Hierzu gehören unter anderem die Schrittweite der Rasterung der R-, G- und B-Komponente und auch die Skalierung der Größe eines im 3D-Raum angezeigten Gitterpunktes. Hinzu kommt die Verwendung von Quantilen, um besonders selten auftretende Farben von der Darstellung ausschließen zu können. So wurden beim zuvor gezeigten Farbhistogramm 10% aller Farben im Bild nicht angezeigt. Und zwar wurden die am schwächsten besetzten Gitterpunkte bei der Darstellung ausgeblendet, damit die Punktwolke nicht noch verschwommener erscheint, als sie ohnehin schon ist.

Die anschauliche, dreidimensionale Darstellung eines Farbhistogramms erlaubt Einblicke in die farbliche Zusammensetzung eines Bildes. Hierbei wird anschaulich deutlich, warum es so schwierig ist, "die verschiedenen Farben" in einem Bild auf algorithmischem Weg voneinander zu unterscheiden. Ein ganz typisches Phänomen ist, was man bei obigem Bild der Erdbeeren feststellen kann. Zwar nimmt der menschliche Betrachter eine gewisse Eindeutigkeit der Farben wahr (so treten im obigen Bild der Erdbeeren offenbar die zwei Farben "rot" und "grün" auf), dagegen kann man im Farbhistogramm keine klare Trennung der verschiedenen Farben erkennen. Typisch ist, dass die Punktwolken der verschiedenen Farben gleitend ineinander übergehen. Demzufolge erscheinen "die verschiedenen Farben" im Farbraum gewissermaßen als eine Fuzzy-Set im Sinne der Fuzzy-Theorie. Das heißt, ein Ort im RGB-Raum lässt sich oft nur unscharf den einzelnen Farben zu-

ordnen. Bei einer Zerlegung des Farbraums, die auf eine Unterscheidung der verschiedenen Farben im Bild zielt, kommen daher Techniken der Fuzzy-Clusteranalyse und auch statistische Methoden der Datenanalyse ins Spiel.

Ein zweites Phänomen, das die Farberkennung erschwert, wird ebenfalls bei der Betrachtung der Farbhistogramme sichtbar. Eine ungleichmäßige Ausleuchtung der Bildobjekte ist im Falle von biologischen Produkten unvermeidbar. So müssen sich zwangsläufig zwischen den Erdbeeren im obigen Bild schattige Stellen ergeben, und auf der Wölbung einer Frucht wird diese heller aussehen als zum Rand hin. Demzufolge wird die Farberkennung in einem Bild stark überlagert durch das Problem der schwankenden Helligkeit der Farbe. Dieses Phänomen wird im Farbhistogramm sichtbar dadurch, dass die Farbcluster typischerweise eine langgezogene Form haben, wobei es Ausläufer in Richtung schwarz und weiß gibt.

2. Clusteranalyse im Fuzzy-Farbhistogramm

Techniken der Clusteranalyse dienen dazu, die verschiedenen Häufungen der Farben im Farbraum zu unterscheiden. So ist es im obigen Beispiel von Abbildung 1 mit Hilfe der Clusteranalyse möglich, im Farbraum die zwei Klassen "rot" und "grün" zu bestimmen. Ganz wesentlich hierbei ist, dass es offensichtlich auch Punkte im Farbraum gibt, die man besser keiner der beiden Klassen "rot" und "grün" zuordnet, weil eine eindeutige Zuordnung reine Willkür wäre.

Im Ergebnis der Clusteranalyse werden die einzelnen Farben bestimmten Farbclustern zu geordnet. Ein Farbcluster entspricht einer im Bild gehäuft auftretenden Farbe. Durch eine Rückprojektion der im Farbraum bestimmten Farbcluster in das Bild erhält man eine Farbsegmentierung. Am Beispiel der Farbklassen "rot" und "grün" von Abbildung 1 wird dies in der folgenden Abbildung gezeigt. Eine solche Farbsegmentierung ist ohne Beachtung von Ortsinformation zustande gekommen, sie liefert aber für viele Fälle bereits eine brauchbare Zerlegung des Bildes in verschiedenfarbige Bereiche. Hierbei müssen die Pixel, die sich nicht eindeutig zuordnen lassen, von der Farbanalyse zunächst ausgeklammert werden.

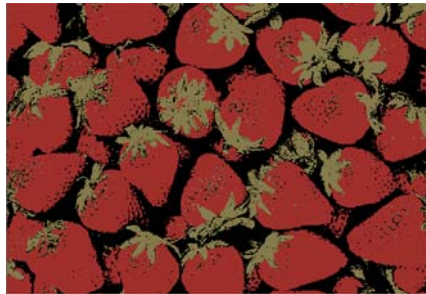


Abbildung 2: Das rechte Bild zeigt die zwei Farbklassen "rot" und "grün", die durch eine Clusteranalyse im Farbraum (dazu Abb. 1) gewonnen wurden. Hierbei sind die Pixel schwarz dargestellt, die keiner der beiden Farbklassen angehören.

Die Methode der Rückprojektion zur Farbsegmentierung sagt noch nichts darüber aus, wie die Zerlegung des Farbraumes mit Hilfe einer Clusteranalyse ausgeführt wurde. Für diese Clusteranalyse kommen traditionell eine Vielzahl von Techniken in Frage. Techniken, die mit scharfen Grenzen zwischen den Klassen arbeiten, sind dabei wenig empfehlenswert, da solche scharfen Grenzen im realen Bildmaterial selten existieren. Die in dem hier zugrundeliegendem Forschungsprojekt¹ verwendeten Clusterverfahren basieren auf dem Fuzzy-C-Means-Verfahren und auf dem EM-Algorithmus. Zu beiden Verfahren existiert eine umfangreiche wissenschaftliche Fachliteratur. Die am Ende dieses Artikels angegebenen Literatur enthält dazu weitere Hinweise.

Der EM-Algorithmus erweist sich im wesentlichen als eine Verallgemeinerung des Fuzzy-C-Means-Verfahrens. Bei ihm wird ein Cluster im Merkmalsraum beschrieben durch den Mittelwert m und die Kovarianzmatrix Σ der Vektoren (hier konkret Farbvektoren) des Clusters. Dabei wird der Abstand $d(v)$ eines Vektors v von diesem Cluster berechnet als Mahalanobisabstand:

$$d(v) = (m - v)^T \Sigma^{-1} (m - v)$$

Das Fuzzy-C-Means-Verfahren verwendet die Kovarianzmatrix nicht und berechnet den Abstand eines Vektors v vom Clustermittelpunkt m einfach als das Quadrat des euklidischen Abstandes. Beiden Verfahren ist gemeinsam, dass sie Iterationsverfahren sind: In einem Berechnungsschritt werden ausgehend von den Merkmalen eines Clusters (m , Σ) unscharfe oder probabilistische Zugehörigkeiten zu den verschiedenen Clustern berechnet, und in einem zweiten Schritt bestimmt man die Merkmale der Cluster (Mittelwert und ggf. Kovarianz) aus den Zugehörigkeiten. Eine mathematisch detaillierte Beschreibung findet man in [Richardt u.a. 2002].

3. Einbeziehung von Ortsinformation zum Bestimmen von Farbsegmenten

Aus der Analyse der Farbhistogramms wird erkennbar, wie wichtig es ist, dass eine Farbsegmentierung adaptiv arbeiten muss, denn eine klar definierte Farbe mit eindeutigen Grenzen für die RGB-Werte gibt es in realen Bildern so gut wie nicht. Andererseits sind mit dem oben beschriebenen Verfahren der Rückprojektion aus dem Farbraum die Möglichkeiten einer Farbsegmentierung, die ohne ein Ortsinformation auskommt, weitgehend ausgereizt. Will man tiefer in die Farbanalyse eines Bildes einsteigen, wird es nötig, auch Ortsinformation zu verwenden. Als ein möglicher Ansatz hierzu wurde ein Region-Growing-Verfahren entwickelt, das die zuvor beschriebene Techniken auch im Rahmen der lokalen Farbanalyse im Bild verwendet. Ziel der Entwicklung war ein fuzzifizierter Region-Growing-Algorithmus, der die im Bild vorliegenden Farbschwankungen und -störungen adaptiv beachten kann.

Mit dieser Intention wurde ein Region-Growing-Verfahren implementiert, das nicht auf den einzelnen Pixel des Bildes arbeitet, sondern auf größeren quadratischen

¹ Das noch laufende Projekt wird im Rahmen des InnoWatt-Förderprogramms durch das BMWA gefördert (Reg-Nr. 25 / 04).

Zellen (Kacheln). In diesen Kacheln wird die lokale "Statistik" der Farbe des Bildes bestimmt. Hierzu gehört die Berechnung des Mittelwertes und die Berechnung der Kovarianzen der Farben innerhalb einer Kachel. Diese Werte werden verwendet als Entscheidungskriterien bei der Ausbreitung der Zellen im Rahmen des "Regionenwachstums" (Region Growing). Für die einzelnen Pixel ergeben sich Zugehörigkeiten der Pixel zu der lokal (pro Kachel) gültigen Farbe daraus, dass man eine Ausreißeranalyse mit den Methoden der Statistik bzw. der Ausgleichsrechnung durchführt. Hierbei werden die Pixel, deren Farbe nicht der lokal dominierenden Statistik (Mittelwert und Kovarianz) entspricht, als nicht dazugehörig behandelt. Auf diese Weise resultiert trotz der Bildstörungen eine relativ robuste Bestimmung der Zugehörigkeiten der einzelnen Pixel, und die zugrundeliegende Kachelung des Verfahrens ist auf der Ebene der Zuordnung der einzelnen Pixel nicht mehr sichtbar. Ein Beispiel hierfür zeigt die folgende Abbildung.

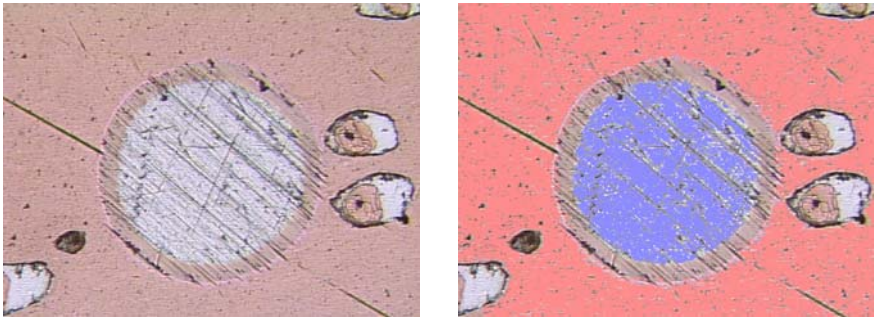


Abbildung 3: Die Abbildung zeigt einen kreisförmigen Kalottenschliff durch Materialschichten. Hierbei sind die Abmessungen des Kreisringes zu bestimmen, um eine Schichtdicke berechnen zu können. Als Vorstufe zur Detektion des Kreisringes wurde hier das Fuzzy-Region-Growing-Verfahren angewendet. Dieses ist funktionsfähig trotz der in den Bildgebieten enthaltenen Störungen.

Literatur

Duda u. a. 2002 Duda, Richardt O. ; Hardt, Peter E. ; Storck, David G.: *Pattern Classification*. Wiley & Sons, 2002 – ISBN 0-471-05669-3

Richardt u. a. 2002 Richardt, J. ; Streckenbach, T.; Gruner, C.: *Modular formalism of Deterministic Annealing, EM Algorithm, Clustering and Regression*. In: www.gfai.de . URL: <http://www.gfai.de/forschungsbereiche/fuzzy/pubs/publist.html> (27.09.2002)