

# Farbbildverarbeitung für großformatige Dokumente

G. Eberhardt; J. Saedler; J. Schlegel  
Graphikon GmbH  
Mandelstraße 16  
D-10409 Berlin  
Tel.: (0 30) 4 21 04-6; Fax: (0 30) 4 21 04-750  
eMail: [post@graphikon.de](mailto:post@graphikon.de)  
URL: <http://www.graphikon.de>

## 1. Einleitung

Durch die enormen Datenmengen in großformatigen Dokumenten, die einige Gigabyte erreichen können, mussten vor allem Verfahren entwickelt werden, die hinsichtlich der Verarbeitungsgeschwindigkeit optimiert sind. Für die Weiterverarbeitung aufzubereiten sind Darstellungen farbiger Szenen, die möglichst fotorealistisch wiederzugeben sind, Abbildungen, bei denen farbige Szenen und graphische Strukturen benachbart oder überlagert sind oder solche wie Karten oder Formulare, denen wenige Objektfarben zugrunde liegen und aus denen die Objekte fehlerfrei separiert werden sollen.

Der Beitrag behandelt Verfahren zur Ableitung von Regionen und graphischen Strukturen aus farbigen Dokumenten und Plänen. Hier sind die Objekte meist Flächen und "dünne Linien", denen jeweils nur ein Farbwert zugeordnet sein soll. Durch die Herstellungsprozesse und Alterungen der Dokumente treten aber auch in den Flächen verschiedenste Störungen auf. Unvermeidbar sind Mischfarben an den Regionsgrenzen und bei "dünnen" Linien, die nur aus Regionsgrenzen bestehen sowie bei Überlagerungen verschiedener Objekte. Gegenüber der Grauwertverarbeitung ist eine einfache Segmentierung nach Hintergrund und Objekten nicht möglich, sondern es erfolgt eine Segmentierung in gleichwertig zu behandelnde, benachbarte oder überlagerte Objekte.

Um die erforderlichen Verarbeitungsgeschwindigkeiten zu erreichen, werden die Farbwerte möglichst schnell reduziert, bevorzugt Look-up-Tabellen angewendet und möglichst wenig Nachbarschaftsoperationen auf Pixelebenen durchgeführt. Dazu wird das Abbild im ersten Schritt möglichst grob in Bereiche zerlegt, die ein Uniformitätskriterium erfüllen oder die sich durch ihre Attribute von benachbarten Regionen relevant unterscheiden. Wobei diese ein einziger Farbwert, eine große Anzahl kleiner Regionen mit unterschiedlichen Farbwerten oder ortsabhängige Änderungen von Farbwerten sein können. Bereiche mit letzteren Eigenschaften werden in weiteren Schritten, auch unter Einbeziehung geometrischer oder topologischer Relationen und Vorkenntnissen, detaillierter analysiert.

## 2. Vorverarbeitung

Aspekte zur modernen Farbverarbeitung, -klassifizierung und -filterung sind in der Literatur besonders in [5] dargestellt. Im Ergebnis des Scannens liegen für jede Grundfarbe (R, G, B) je 8 Bit vor. Diese möglichen 16,7 Millionen Farben werden für eine effektive Weiterverarbeitung in den **HSV-Farbraum** übertragen, in dem die Buntart, die Sättigung und Intensität die Dimensionen darstellen. Der HSV-Farbraum zeichnet sich durch beson-

dere Anschaulichkeit und Berechenbarkeit aus. Die Farbklassifikation kann auf den Größenvergleich mit Grenzwerten in den drei Dimensionen zurückgeführt werden. Darüber hinaus ist für die Farbbildverarbeitung die Möglichkeit der Bezugnahme auf die absolute Sättigung wesentlich, da im Falle eines kleinen Abstandes des Abbildungspunktes einer Farbe in der Buntheitsebene vom Unbuntpunkt kleine Änderungen der RGB-Werte zu großen, sprunghaften Änderungen der Buntart führen. Auch der Nachteil der hohen Korrelation der Werte in den Farbkanälen R, G, B der gescannten Dokumente wird deutlich gemindert.

Die Umrechnung erfolgt je nach Genauigkeitsanforderungen über vorgefertigte Look-up-Tabellen oder mittels Hashtabellen. Die Look-up-Tabellentechnik erfordert einen hohen Speicherplatzbedarf, sodass praktisch nur mit 18 Bit/Pixel gerechnet werden kann, was nur zum Teil zu akzeptierbaren Ungenauigkeiten führt. Für technische Dokumente ist die Transformation über Hashtabellen besonders gut geeignet, da dort zwar große Datenmengen, aber nur relativ wenige Farbwerte vorhanden sind. Vorteilhaft ist es dann, dass nur die wirklich vorhandenen Farbwerte mit einer Tabellentechnik umgerechnet werden und dies, ohne dass Genauigkeiten verloren gehen.

Die **Farbklassifikation** dient zur Segmentierung des Dokumentes in relevante Regionen sowie zur Reduzierung von Farbwerten und dabei gleichzeitig zur Minderung von Störungen. Dadurch werden die Daten auf maximal 256 unterschiedliche Farbenklassen reduziert. Bei der Farbverarbeitung wird nicht nur auf die traditionelle Trennung von Buntheit und Helligkeit, sondern ähnlich dem menschlichen Farbempfinden auch auf die Unterscheidung des Farbtons und der Sättigung unter Verwendung von Polarkoordinaten orientiert.

Der erste Programmschritt umfaßt die Möglichkeit des Festlegens und Definierens von Farbklassen. Das Anlernen von Farbklassen ist sowohl durch interaktives Anklicken in einem digitalen Abbild des Beispieldokumentes als auch durch Eingabe der Grenzwerte möglich, wobei die Ergebnisse unmittelbar visuell bewertet werden können. Im Ergebnis entsteht eine Belegung der Look-up-Tabellen, die sowohl die Klassifizierung als auch die Transformation enthält. In Abhängigkeit von Häufigkeiten und Verteilungen wurde der Raum, in dem alle darin abgebildeten Farbwerte einem ausgewählten Farbwert zugeordnet werden, durch einen beim Anlernen zu ermittelnden Rotationsellipsoiden beschrieben.

Für eine effektive Klassifizierung wurde eine Kombination aus überwachtem und unüberwachtem Lernen aufgebaut. Dabei werden die Farbwerte in der ersten Stufe durch das überwachte Lernen so reduziert, dass sie für einen Bediener noch überschaubar sind. In einem weiteren Teilschritt wird das manuelle Zusammenfassen von so entstandenen Klassen ermöglicht. Dabei ist eine eindeutige Zuordnung nicht zwingend. Im Ergebnis können mehrere Binärlayer gebildet werden, um die Probleme durch Mischfarben an Kanten oder durch Konvergenzprobleme in mehreren Farbebenen zu reduzieren oder auszugleichen.

Ein überwachtes Lernen (teach-in) wurde für die folgenden Dialogfunktionen implementiert:

- neue Farbklassen anlegen oder vorhandene Farbklassen löschen
- Farbklassen zum Anlernen auswählen
- Farbklassen zurücksetzen
- direkte interaktive Manipulation der Parameter der Klassen als Möglichkeiten für erfahrene Bediener mit genauen Vorstellungen über die Farbräume
- Anlernen mit dem Mauszeiger und Anklicken der zugehörigen Bildposition

- Auswahl von Bildbereichen zum Anlernen und zur Ableitung von statistischen Werten aus diesen, wie Mittelwerte, Streuungen, Kovarianz und Eigenvektoren der Rotationsellipsoide zur Definition der Farbräume der Klassen.

Für ein unüberwachtes Lernen (Clustering) wird im Abbild des Dokumentes nach natürlichen Gruppierungen (Cluster) gesucht. Ausgang dafür bilden Farbhistogramme. Als Abstandsmaß bei der Zuordnung wurde die Mahalanobis-Distanz verwendet. Weiterhin wurden Verfahren zum Zusammenfassen von Klassen (Merging) abgeleitet, um geeignete Reduzierungen der Einträge zu ermöglichen.

Im zweiten Klassifizierungsschritt erfolgt dann die automatische Zuordnung jedes im Dokument auftretenden Farbwertes zu einer Farbklasse und damit zu einem einzigen Farbwert (Ort im Farbraum). Dabei wird er dem Zentrum der Farbklasse zugeordnet, in dem er sich befindet. Liegt er in keiner Farbklasse, können diese vergrößert oder der Farbwert einer Rückweisungsklasse zugeordnet werden. Es wurden 3 Varianten realisiert:

- (1) Mit fester Farbreduktion, die eine schnelle Verarbeitung erlaubt, da es möglich ist, die Look-up-Tabelle im Cache unterzubringen. Sie ist relativ ungenau, aber für Dokumente mit wenigen, gut ausgeprägten Bereichen ausreichend.
- (2) Mit zweistufigen Abbildungen mit Genauigkeiten von 8, 10, 12 Bit/Farbkanal, die weniger Maschinenbefehle, aber mehr Speicherzugriffe als (1) erfordern, dies aber im Cache ermöglichen und eine bessere Anpassung der Look-up-Tabellen erlauben und damit eine höhere Genauigkeit als Verfahren (1) erreichbar machen; die dadurch erforderliche größere Look-up-Tabelle kann nicht im Cache gehalten werden.
- (3) Als Hashmap, bei der Genauigkeit, Geschwindigkeit und Speicherbedarf durch Parametrisierung der Hashfunktion gut steuerbar sind, dies aber durch Verluste bei der Verarbeitungsgeschwindigkeit erkauft wird.

Darüber hinaus wurde für Karten mit relativ wenigen Farbwerten ein Verfahren zum Anlernen und Modifizieren von Farbpaletten entwickelt. Nach der automatischen Reduktion der Farbwerte auf eine vorgegebene Anzahl werden die Farbwerte in ausgewählten Ausschnitten von Regionen interaktiv einem neuen Farbwert aus einer Palette zugeordnet. Die Abbildung dieses neuen Farbwertes wird sofort für den gesamten Bildausschnitt angezeigt. Unerwünschte Veränderungen in Regionen (besonders in Grenzbereichen) können dann durch Löschen/Hinzunahme von weiteren Farbwerten sofort korrigiert werden.

Die Klassifizierung der Farbwerte weiterer Karten kann unter Nutzung dieser Paletten automatisch erfolgen. Bild 3 zeigt einen Ausschnitt aus einer Karte, bei der nach einer automatischen Reduktion auf 32 Farben die Möglichkeit einer weiteren Separation in vorgegebenen Klassen mit vorgegebenen Paletten gezeigt wird. Da die Farbwerte und ihre Verteilungen bei verschiedenen Karten recht unterschiedlich sein können, wurden Verfahren zur automatischen Anpassung der Paletten entwickelt, die automatisch Veränderungen von Clustern und Verteilungen erkennen.

**Filter**, die die Umgebung von Pixeln einbeziehen, sind, um sinnvolle Regionen bilden zu können, unabdingbar. Die beinahe fraktalen Abbildungen der Randbereiche der Regionen führen zu stark gestörten Kanten und einer Vielzahl unerwünschter Regionen. Für die technisch sinnvolle Filterung wurde ein SNN-Filter für den HSV-Raum entwickelt und optimiert, der aber wegen der benötigten Rechenzeit nur für kleinere Formate angewendet werden kann. Allgemein verwendet wird ein Modalfilter nach dem Klassifizieren. Durch ge-

schickte Optimierung bei der Berechnung des häufigsten Wertes im Fenster, die sich an Verfahren zur Berechnung von Rangordnungen anlehnen, können auch bei größeren Fenstern noch hohe Verarbeitungsgeschwindigkeiten erreicht werden. Schon mit dem Modalfilter mit der Fenstergröße  $7 * 7$  wurde die Anzahl der kleinen Störungen drastisch verringert, wobei die relevanten Objekte kaum beeinträchtigt wurden.

Nach der Klassifizierung wird sofort auf eine **Regionsbildung** und listenförmige Beschreibung übergegangen. Dabei werden grundsätzliche Mängel der traditionellen Pixelverarbeitung überwunden:

- verfahrensbedingter Zeitaufwand
- schwieriges oder unmögliches Einbringen von globalen Informationen wie Know-how, Vermutungen oder Kriterien für interessierende Regionen in die Verarbeitungsschritte
- notwendige Rechnung mit 2-dimensionalen "Punkten" (Bild 1a)
- notwendige Zuordnung der Kontur zu ausgewählten Objekten (Bild 1b)
- Verletzung des Jordanischen Kurvensatzes (Bild 1b)

Bei dem gewählten Modell (Bild 2) wird ein Pixel dann durch vier 0-dimensionale Ecken, vier 1-dimensionale Kanten und eine 2-dimensionale Fläche (siehe Bild 2a) und Konturen durch eine geordnete Folge von Pixelkanten (Bild 2b) beschrieben. Das Eckenmodell beseitigt einige grundsätzliche Mängel des Pixelmodells. Alle Objekte können durch das gleiche Modell abgebildet werden. Damit kann die Breite der Konturlinie als Grenzwert gegen Null angenommen und jede Region durch die ihr zugeordnete Kontur zu begrenzt werden. Für eine eindeutige Beschreibung von farbigen Dokumenten ist dies unbedingt erforderlich, da sich bis zu 4 gleichberechtigte Objekte in einem Punkt berühren können.

Dieser Eckencode sowie das darauf aufbauende Verfahren wurde 1990 von MANDLER und OBERLÄNDER zunächst für binäre, später auch für mehrwertige Labelbilder [1] eingeführt. Graphikon hat sich seit seiner Gründung darauf konzentriert [3], [4], den Ansatz der Regionsbildung über Eckencodes auf die Analyse von Farbbildern zu erweitern und zum Teil auf Arbeiten aus der Akademie der Wissenschaften zurückgegriffen [2]. Der Eckencode wird durch einen in Scanreihenfolge über das Bild laufenden  $2 * 2$ -Filter abgeleitet, indem aus der Nachbarschaft einer von neun möglichen Codewerten entsprechend Bild 2c gebildet wird.

Handelt es sich um eine relevante Ecke entsprechend Bild 2c, dann und nur dann wird der Code zusammen mit den Labelwerten der benachbarten Regionen und den Spaltennummern abgespeichert. Ergänzend dazu werden die Zeilenenden durch ein spezielles Codemuster gekennzeichnet. Im Ergebnis ist nach einem einzigen Durchlauf durch diese Matrix das ursprüngliche digitale Abbild des Dokumentes in eine nach Koordinaten sortierte Liste von Codes von relevanten Ecken sortiert.

Die **Ableitung der Regionen** erfolgt auf Basis des in [1] beschriebenen Algorithmus in einem einzigen Durchlauf (single-pass) durch die Folge der Eckencodes durch eine von Graphikon modifizierte und optimierte Version.

Das Verfahren liefert für jede Region eine Konturbeschreibung in Form von verketteten Eckpunktkoordinaten und geometrische Merkmale wie Fläche, Umfang und umschreibendes Rechteck. Die Konturbeschreibung jeder Region enthält immer genau eine Außenkontur und je nach Anzahl der Löcher  $n$  Innenkonturen (vgl. Bild 4a und 4b, Eckencode siehe Bild 4c). Eine Ausnahme bildet nur die zur Vermeidung topologischer Widersprüche

zusätzlich eingeführte "Weltregion", die nur eine das Dokument umschließende Innenkontur umfaßt (siehe außen umlaufende Kontur in Bild 4b).

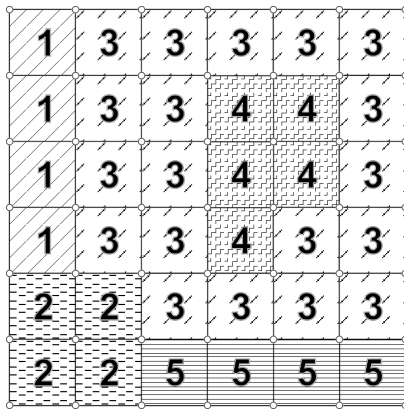


Bild 4a: Labelung

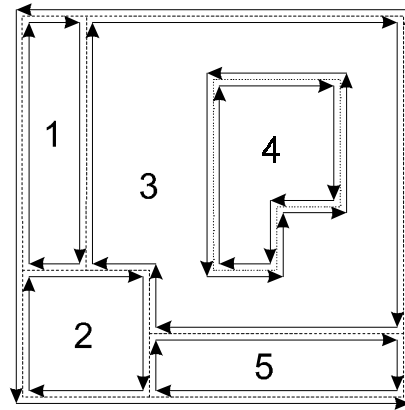


Bild 4b: Regionen und ihre Relationen

Code	Pos.	Symbol	Beschreibung
0x20000001	0,0	CC_UL	0  ,1 neu
0x70010013	1,0	CC_D	3 neu
0x30060030	6,0	CC_UR	
0xf0000000		eol 0	
0x20033334	3,1	CC_UL	3  ,4 neu
0x30053343	5,1	CC_UR	
0xf0010000		eol 1	
.			
.			
0x40000200	0,6	CC_LL	
0x90022500	2,6	CC_U	2 fertig
0x50065000	6,6	CC_LR	5,0   fertig
0xf0060000		eol 6	

Bild 4c: Codierung

Bild 4: Beispiel mit Kodierung und Listenstruktur

Gleichzeitig werden folgende topologische Beziehungen zwischen den Regionen erfaßt: Vater-Sohn, Bruderschaft und Halbbruderschaft. Eine Region A ist Vater von B und B Sohn von A, wenn A die kleinste Region ist, die B vollständig einschließt oder wenn jeder Weg von B in die Weltregion durch A führt. Zwei Regionen B und C sind echte Brüder, wenn sie eine gemeinsame Vaterregion haben und zum selben Loch der Vaterregion gehören. Sie sind Halbbrüder, wenn sie eine gemeinsame Vaterregion haben, aber sich in unterschiedlichen Löchern befinden. (In Bild 4b sind die Regionen 1, 2, 3 und 5 Brüder, wobei die "Weltregion" ihr gemeinsamer Vater ist. Die Region 4 ist Sohn der Region 3.)

Der Gesamtaufwand zur Berechnung der Regionen ist zwar abhängig von der Menge der Eckencodes, jedoch sind dafür keinerlei Iterationen notwendig. Damit ist die Grundlage gegeben, auch die Verarbeitung der Eckencodes schritthaltend mit dem Scanvorgang auszuführen. Gleichzeitig können wesentliche Merkmale zur Beschreibung der Regionen, wie ihre Fläche, ihr Umfang und das sie umschreibende Rechteck abgeleitet werden. An-

hand dieser Beschreibungen der Dokumente können unerwünschte Objekte aus den Listen einfach entfernt werden.

Im Ergebnis ist eine **regionenbezogene Liste**, die das Dokument beschreibt, entstanden. Die Daten sind auf relevante und globale Informationen reduziert. Liegen die Beschreibungen der Regionen, ihrer Relationen zueinander sowie geometrische Merkmale vor, ist damit eine weiterführende symbolische aufgabenabhängige Auswertung in einem Hostrechner (PC) in Echtzeit möglich, da jetzt die Daten drastisch reduziert sind.

An einem kleinen ungefilterten Ausschnitt wird das Ergebnis im Bild 5b der Regionen- und Zusammenhangsanalyse zur Verifizierung der Regionenlisten dargestellt. Dabei wurde zur Verdeutlichung jeder Region eine Farbe zugewiesen, die nicht in einer benachbarten Region auftritt. Die Regionen werden anhand von automatisch bei der Regionsbildung mit abgeleiteten Konturmerkmalen klassifiziert und

blau	dargestellt:	interessierende Schriftkandidaten
grün	dargestellt:	unerwünschte strichförmige Regionen
rot	dargestellt:	unerwünschte kleine kompakte Regionen
schwarz	dargestellt:	Hintergrund

### 3. Nachbearbeitung

Durch die Regionenbildung sind wesentliche Merkmale und Attribute (Farbwerte) der Regionen sowie Hierarchien und Nachbarschaften bekannt. Aus der Spezifik der Karten lassen sich auch die erwarteten Farbüberlagerungen und Merkmale der Objekte ableiten. Beispiele hierfür sind Abbildungen von Überlagerungen durch Straßen verschiedener Ordnungen oder Texte und Höhenlinien, die über verschiedene Objekte wie Wiesen, Wald oder Felder gedruckt sind.

Unter Nutzung dieser Kenntnisse erfolgen automatische Analysen, vor allem an Regionsgrenzen, die auch nach der Klassifizierung als separate Regionen abgebildet sein können. In kritischen Bereichen erfolgt ein Rückgriff auf die 24-Bit-Abbildungen im HSV-Raum.

Zur Weiterverarbeitung kann jede Farbebene vektorisiert werden, wobei von den "dünnen" Linien die Mittelachsen und von den Flächen die Konturen abgeleitet und an den Übergangsbereichen geeignet miteinander verbunden werden [6].

### **Literatur**

- [1] Mandler, Eberhard; Oberländer, Matthias F.: One-pass Encoding of Connected Components in Multi-Valued Images, International Conference on Pattern Recognition, Atlantic City, 1990
- [2] Kovalevski, Wladimir; Strukturen der Bildträger und Bilder in AUTBILD 86/2, Friedrich-Schiller- Uni Jena, 1986
- [3] Schlegel, Jan: Diplomarbeit "Robustheit und Invarianz bei der Beschreibung von Regionen", Technische Universität Dresden, 1992
- [4] Patentschrift DE 4419395, Graphikon GmbH: Verfahren und Einrichtung zur Analyse und Verarbeitung von Farbbildern, 1994
- [5] Rehrmann, V.: Stabile echtzeitfähige Farbbildauswertung. Koblenzer Schriften zur Informatik, 1994
- [6] Handbuch VECTOR 4.0, 1999

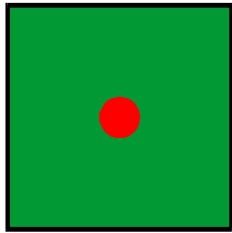


Bild 1a: Pixel

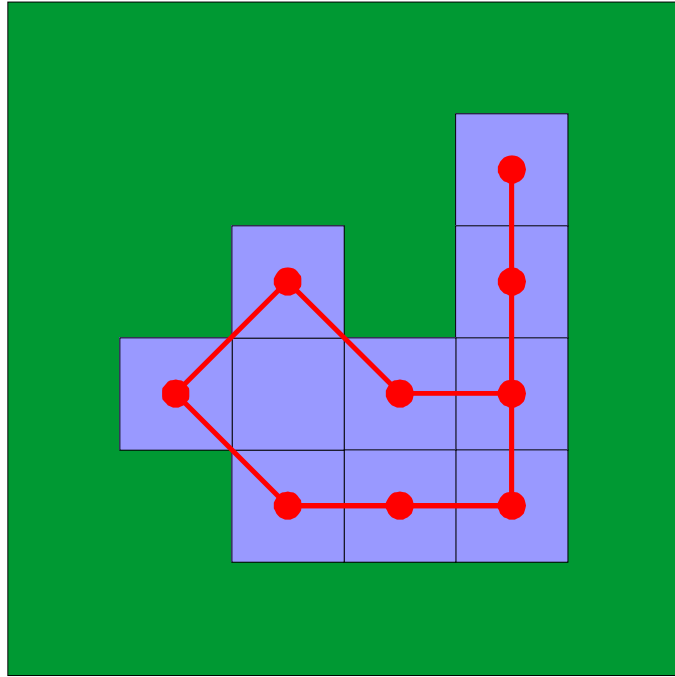


Bild 1b: Region und ihre Konturenpixel

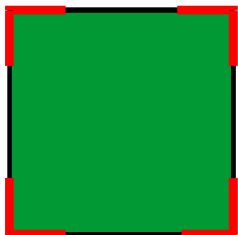


Bild 2a: Eckenmodell eines Pixels

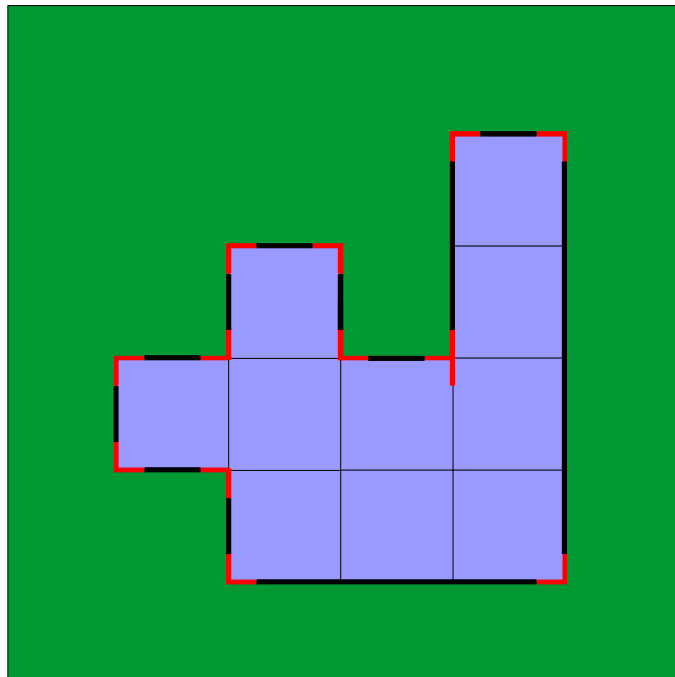
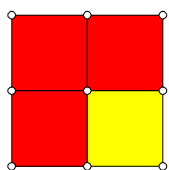
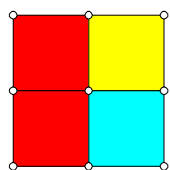


Bild 2b: Regionen und ihre Kontur

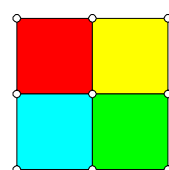


CC\_UL  
rechts oben  
1

...



CC\_R  
T nach rechts  
8



CC\_4  
4 Ecken  
9

Bild 2c: Aufbau und Klassifikation des Eckencodes

Bild 2: Eckenmodell

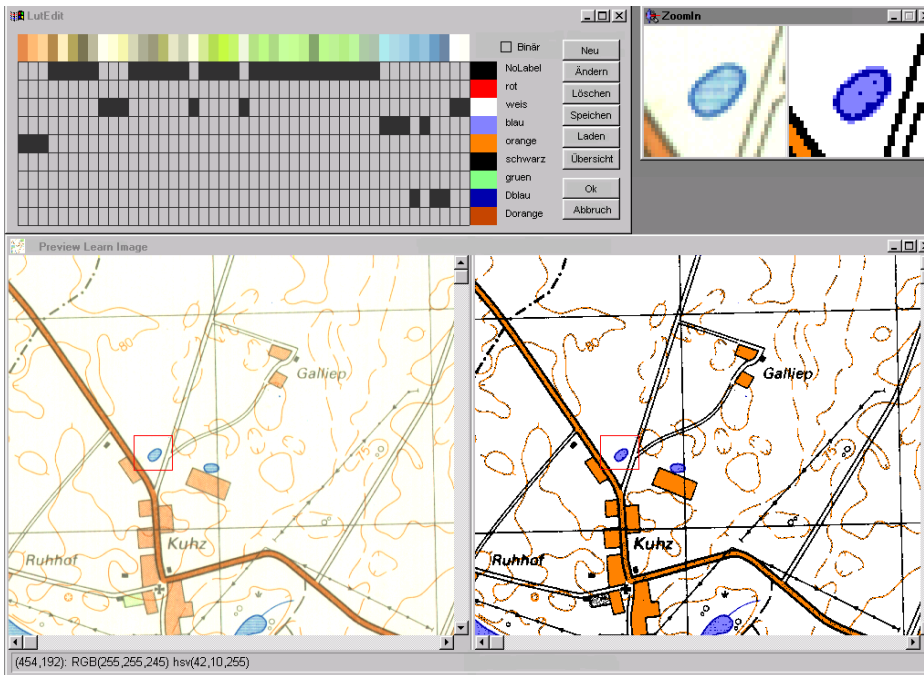


Bild 3: Farbseparation über Paletten



Bild 5a) gescannter ec-Scheck

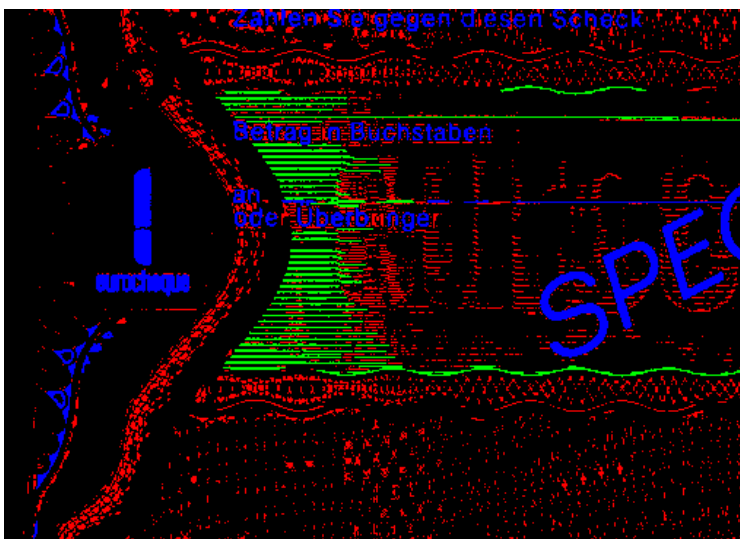


Bild 5b) Klassifikation in Regionen