

UMFELDORIENTIERTE FARBSEGMENTIERUNG AUF BASIS EINES FUZZY-BASIERTEN FARBBELEHRUNGSTOOLS

Bernhard Frei Martin Hund Markus Schnitzlein

Océ Document Technologies GmbH

Abteilung DE / Scanner-Technik und Bildverarbeitung

Max-Stromeyerstr. 116

78467 Konstanz

www.odt-oce.com

email: markus.schnitzlein@odt-oce.com

Abstract

Die Bildaufbereitung zur optimierten Verarbeitung von Dokumenten zur Zeichenerkennung und inhaltlichen Analyse setzt eine Separation des Bildes in Nutz- und Hintergrundinformation voraus. Nachfolgend wird ein umfeldorientiertes, 2-dimensionales Farbsegmentierungsverfahren vorgestellt, das eine farb- und strukturbasierte Trennung von Hintergrund- und Nutzinformationen ermöglicht. Das spezifische Vorwissen über die farbliche Gestaltung der Dokumente kann zusammen mit den dokumentabhängigen Farbausprägungen mit Hilfe eines fuzzy-basierten Belehrungstools verknüpft werden. Diese Belehrung bildet die Grundlage für die Farbsegmentierung, die sich aufgrund der Auswertung der Pixelumfeldes einer $(2n+1, 2n+1)$ -Matrix, an Farbveränderung innerhalb eines Dokumentes adaptieren kann.

1. EINLEITUNG

1.1 Probleme der Dokumentgestaltung

Dokumente bilden die Grundlage der meisten Geschäftsprozesse in der Verwaltung sowohl in der Industrie, als auch im behördlichen Bereich. Etwa 70%-80% aller Dokumente werden heute noch in Papierform weitergeleitet und abgelegt.

Zur inhaltlichen digitalen Weiterverarbeitung von Dokumenten werden Methoden der Bildverarbeitung und -analyse, sowie der ICR (Intelligent Character Recognition) benötigt.

Gerade auf stark strukturierten Vorlagen (Formulare) und semi-strukturierten Dokumenten (Rechnungen, Lieferscheine) sind die farblichen Gestaltungsformen vielfältig. Die Separation der Nutzinformation vor einem nicht eindeutig beschreibbaren Hintergrund ist schwierig, bildet aber die Voraussetzung für eine zuverlässige Zeichenerkennung.

Aufgrund der großen Bandbreite der Dokumentgestaltung und der erheblichen Varianz der drucktechnischen Ergebnisse bei der Vorlagenherstellung müssen auch auf strukturell definierten Dokumenten intelligente und adaptive Bildverarbeitungsverfahren eingesetzt werden, um eine für die automatische Weiterverarbeitung zufriedenstellende Zeichenerkennungsrate zu erhalten. Die nachfolgend beschriebene farbadaptive Segmentierung soll am Beispiel von problematischen und wenig definierten Formularen ein Bildverarbeitungsverfahren erläutern, das trotz großer farbmischer Ähnlichkeit von Nutz- und Hintergrundfarben eine zuverlässige Farbsegmentierung erlaubt und sich dabei robust gegenüber Bildstörungen verhält.

1.2 Drucktechnische Varianz und scannerseitige Erfassungsartefakte

Bei der farborientierten Analyse von Dokumenten müssen die typischen Artefakte, die bei der drucktechnischen Erzeugung der Vorlagen und bei der digitalen Erfassung auftreten können, berücksichtigt werden:

- Veränderungen der Farbsättigung über die Dokumentbreite
- Verwendung unterschiedlicher Druckfarben durch nicht eindeutige Farbdefinition
- Auflösung von Rasterdruckpunkten
- Registration-Fehler bei Farbsensoren
- Farbrauschen
- Verwendung von Scanner ohne ICC-Profilierung

2. UMFELDORIENTIERTE FARBSEGMENTIERUNG

Im Bereich der automatischen Zeichenerkennung wird in der Regel vor dem eigentlichen Erkennungsprozess eine Trennung zwischen Nutz- und Störinformation durchgeführt. Im einfachsten Fall wird durch eine Binarisierung die dunkle Schrift vom hellen Hintergrund separiert. Bei farbiger Bedruckung des Hintergrundes gestaltet sich die Trennung erheblich schwieriger. Bei Formularen sind z.B. Ausfüllhinweise farbig gedruckt und der auszufüllende Bereich durch einen farbigen Rahmen markiert. Die Farbe, die zum Ausfüllen der Formulare verwendet wird, ist in der Regel nicht definiert und ist zur besseren visuellen Lesbarkeit meist dunkel. Bei der herkömmlichen Farbsegmentierung ist es in vielen Fällen nicht möglich, allein aufgrund der Farbinformation eines Bildpunktes über die Zuordnung zu Vordergrund oder Hintergrund zu entscheiden, da die oben beschriebenen Farbartefakte (Rasterdruck, der in Einzelpunkten sichtbar wird; Farbrauschen; Kompressionsartefakte) zur punktwisen Ähnlichkeit oder gar Identität der Farbwerte zwischen Vorder- und Hintergrund führen.

Mit der umfeldbezogenen Farbsegmentierung wird ein Verfahren definiert, bei dem durch eine farbliche Betrachtung des Umfeldes auf die Zugehörigkeit eines Bildpunktes zu Schrift oder Hintergrund geschlossen werden kann. Die umfeldbezogene Farbsegmentierung ermöglicht auch dann eine saubere Trennung von Vorder- und Hintergrund, wenn teilweise die gleichen Mischfarben bei Schrift und Hintergrundbedruckung auftreten.

Zur primären Unterscheidung der Nutz- und Störinformation werden „sicher zugeordnete“ Farben beschrieben, die dann als Kristallisationspunkte für eine umfeldorientiertes (2-dimensionales) Farbsegmentierungsverfahren dienen.

Diese Spezifikation der Farben erfolgt anhand der Terme Farbe, Helligkeit und Sättigung.

Um die Handhabung der Belehrung zu vereinfachen, können auf einem Dokument Bereiche ausgewählt werden, die als Histogramm ausgewertet werden und in dem die relevanten Kristallisationsfarben gefunden werden können.

2.1 Farbbelehrungsverfahren

Die Belehrung der Farbtypen „Vordergrund“ oder/und „Hintergrund“ durch Auswahl von Farbreferenzen oder durch Selektion von realen Farben auf einem Trainingsdokument gestaltet sind als äußerst schwierig und wenig effizient. Weder

das Vergleichen der idealen Auswahlfarbe mit Farbreferenzen noch die Auswahl von Einzelfarbpixeln im Image sind dem Bediener (auch nicht einem geübten Bediener) zumutbar und bergen darüber hinaus viele Fehlerquellen.

Zur Beschreibung von Dokumentfarben (Vorder-/ Hintergrund) wird nachfolgend ein Verfahren gezeigt, das - ausgehend von der menschlichen Wahrnehmung – im Farbraum „Lu*v*“ arbeitet. Die Formulierung der Farbcharakterisierung wird durch die Begriffe Farbe, Helligkeit und Sättigung vorgenommen.

Die Belehrung der Hintergrund- (Vordergrund-)Farben wird auf der Basis einer globalen Histogrammanalyse eines Referenzdokumentes durchgeführt. Gestützt durch dieses Histogramm wird eine einfache Farbbeschreibung erreicht, die sich auf die Begriffe Farbe, Helligkeit und Sättigung stützt. Durch diese Belehrung werden die „sicher zugeordneten“ Farben beschrieben, die dann als Kristallisationspunkte für eine umfeldorientiertes (2-dimensionales) Farbsegmentierungsverfahren dienen.

Für die drei wichtigen Farbbeschreibungparameter wird jeweils ein Histogramm dargestellt, das auf einem ausgewählten Dokumentbereich erzeugt wurde. Durch Grenzmarken lassen sich in dem selektierten Bereich die relevanten Bereiche für die zu beschreibenden Farben markieren. Die schrägen Begrenzungslinien, die trapezförmige Bereiche abgrenzen, beschreiben die Stärke der Zugehörigkeit der einzelnen (Farb-, Helligkeits-, und Sättigungswerte) zu der Klasse der zu beschreibenden Farbe.

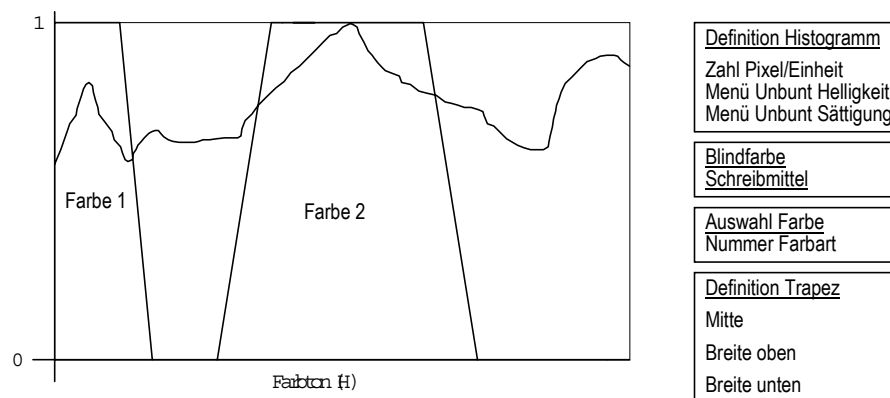


Abb.1: Beispiel für das Definitionsfenster „Farbton“ zur Trennung von Vorder- und Hintergrundbereich

In gleicher Weise wird mit jedem der drei Beschreibungsfelder verfahren. Im farbigen Bildanhang finden Sie die Darstellung aller drei Beschreibungsfenster.

Auf Basis dieser „ungenauen“ Beschreibung der Farbcharakteristika Helligkeit, Farbton und Sättigung wird nun ein „Fuzzy“-basierter Ansatz verfolgt, um die Zugehörigkeitsfunktion eines Pixels zu der zu beschreibenden Klasse zu berechnen.

Für das Verfahren wurde folgender Forderungskatalog erarbeitet:

1. Es muss mit Stichprobenmaterial belehrbar sein. Im einfachsten Fall liegt ein Referenzdokument vor, dessen farbliche Komposition eine eindeutige Abbildung der Farbe auf die nachfolgenden Verarbeitungsschritte ergibt. Dies sind im einfachsten Fall eine Zuordnung von Farbe auf die Elemente eines Bitonalbildes, also die Entscheidung ob schwarz oder weiss. Das zu entwickelnde Verfahren muss automatisch oder durch interaktive Probenentnahme diese Zuordnung leisten.

In die so aus Stichprobenmaterial gewonnene Zuordnung von Farbe zu Merkmal müssen noch die beim Erfassungsprozeß typisch auftretenden Streuungen durch die Eigenschaften der Erfassungsgeräte und des Belegmaterials eingebracht werden können.

2. Die Zuordnung von Farbe zu Merkmal und/oder Verarbeitungsschritt muss frei definierbar sein. So kann gewünscht sein, generell alle z.B. „roten“ Merkmale eines Dokumentes einer besonderen Verarbeitung zuzuführen. Wie mehr oder weniger allgemein dieses Rot sein kann, muss durch den Benutzer festgelegt werden können. Als anschauliche und benutzerfreundliche Größen bieten sich die Verwendung von Helligkeit, Farbton und Sättigung als Parameter an.
3. Ist aber eine Farbe definiert gegeben muss das System die Eingabe in absoluten farbmtrischen Größen wie X,Y,Z oder Lab unterstützen. Daraus ergibt sich die Notwendigkeit einem geräteunabhängigen Farbraum zu arbeiten.
4. Ebenfalls müssen die in der Drucktechnik benutzten Farbsysteme wie Pantone für die Farbdefinition Verwendung finden können.
5. Es muss auch eine Mischung dieser Art von Farbfestlegungen möglich sein. So kann durchaus eine Hintergrundfarbe im Pantonesystem, Schriftfarben aber allgemeiner nach 2. definiert werden.

Zur Umsetzung dieses Forderungskataloges wurde ein Graphical-User-Interface (GUI) definiert, das dem Benutzer die oben beschriebenen Leistungen bietet.

Für die Definition eines freien Farbbereiches wurde ein Fuzzy-Ansatz gewählt, der mit den linguistischen Variablen Helligkeit, Sättigung und Farbton arbeitet.

Das Arbeiten in diesem Farbraum bietet sich an. Das Problem, das hierbei aber zu lösen war, ist folgendes:

Um der Forderung in Punkt 3 gerecht zu werden müssen diese Variablen in einem standardisierten Farbraum definiert werden.

Die allgemein verwendeten Farbräume HSI, HSL und HSV, die mit diesen Variablen arbeiten, sind geräteabhängig und haben somit keinen farbnormierten Bezug. Damit wird ein Ansatz verhindert, der alle Forderungen vereinigt.

Diese Problematik wurde durch folgenden Ansatz einer Lösung zugeführt:

Helligkeitsachse : $L=L^*$

dafür wird die L-Komponente, des nach CIE genormten Farbraumes verwendet

Farbton: $H = \arctan [(v^* / u^*)$

v^* und u^* sind Komponenten des $L u^*v^*$ Farbraumes nach CIE

Farbsättigung: $s = [(u' - u'n)^2 + (v' - v'n)^2]^{0.5}$

mit $u'n$ und $v'n$ wird auf den Weißpunkt normiert

sie ergeben sich durch Umrechnung aus dem CIE XYZ Farbraum

Die Verwendung dieser Variablen erlaubt einen intuitiven Umgang bei der Definition von Farbbereichen.

Für die Fuzzydefinition der Zugehörigkeit werden diese Größen auf Fuzzyzugehörigkeitsfunktionen mit dem Wertebereich 0..1 abgebildet.

Damit ist das gewünschte Ziel, den intuitiven Umgang bei der Farbbereichsdefinition mit einer farbmétrisch exakten Basis zu verbinden, erreicht worden.

Dieses Verfahren gibt uns jetzt die Möglichkeit entweder mit dieser Information direkt die Farbsegmentierung anzustoßen – hierbei wird aufgrund der Zugehörigkeit der einzelnen Bildpunkte das Farbbild in ein 8-bit Grauwertbild überführt und anschließend binarisiert – oder man nutzt die jetzt beschriebenen Farben als Kristallisationsfarben für die adaptive strukturorientierte Farbsegmentierung.

2.2 Adaptive strukturorientierte Farbsegmentierung

Bei der herkömmlichen Farbfilterung ist es in vielen Fällen nicht möglich, allein aufgrund der Farbinformation eines Bildpunktes über die Zuordnung zu Vordergrund oder Hintergrund zu entscheiden. Abbildung 3 dient hier zur Verdeutlichung.

Die Papierfarbe hat eine leichte grünliche Färbung, die herauszufilternde Farbe ist durch die grüne Bedruckung gegeben. Für die ICR-Verarbeitung soll lediglich die graue Bedruckung herangezogen werden. Um die grüne Bedruckung zu unterdrücken, müssen die vorhandenen grünen Bildpunkte der Bedruckung aufgrund ihrer Farbe als Hintergrund definiert werden. Damit diese Farbfilterung auch auf verschiedenen Druckqualitäten dieses Formulars funktioniert, muß unter Berücksichtigung der Schwankungen der Druckqualität und Papierfarbe ein entsprechender Toleranzbereich in den Farbfilter eingerechnet werden. Der Bereich der zu filternden Farben muß um diesen Toleranzbereich vergrößert werden.

Teilweise ergeben sich jedoch exakt die gleichen Farben durch Farbmischung der grauen Bedruckung mit der Papierfarbe, so daß bei der eindimensionalen Farbfilterung zwangsläufig Bildpunkte die zur grauen Schrift gehören, dem Hintergrund zugeordnet werden.

Abbildung 4 zeigt dieses Problem: Obwohl bei der eindimensionalen Farbfilterung nicht einmal alle grünen Bereiche unterdrückt wurden, werden die grauen Schriftzeichen schon stark ausgedünnt. Dies verhindert den sinnvollen Einsatz einer automatischen Zeichenerkennung.

Mit der umfeldbezogenen Farbfilterung wird ein Verfahren definiert, bei dem durch eine farbliche Betrachtung des Umfeldes auf die Zugehörigkeit eines Bildpunktes zu Schrift oder Hintergrund geschlossen werden kann. Die umfeldbezogene Farbfilterung ermöglicht auch dann eine saubere Trennung von Vorder- und Hintergrund, wenn, wie oben beschrieben, teilweise die gleichen Mischfarben bei Schrift und Hintergrundbedruckung auftreten.

Bei der nachfolgenden Beschreibung wird angenommen, daß bei einem Bild diejenigen Bildpunkte gefunden werden sollen, die zum Hintergrund gehören und deshalb unterdrückt werden sollen. Das Verfahren kann jedoch analog auch so definiert werden, daß diejenigen Punkte gesucht werden sollen, die zur Schrift gehören und demzufolge nicht unterdrückt werden dürfen.

Definitionen:

Jede Farbe eines beliebigen Farbraums (z.B. Grau, RGB, HLS, YCrCb, Lab, YIQ, CMYK, ...) wird einem der drei folgenden Bereiche zugeordnet.

Bereich 1:

Alle Farben, die mit Sicherheit zur Hintergrundbedruckung gehören. Mischfarben zwischen Hintergrundbedruckung und Schriftfarbe gehören nicht zum Bereich 1.

Bereich 2:

Alle Farben, die nach einem gegebenen Toleranzmaß (z.B. Sättigung $\pm 10\%$, Helligkeit $\pm 5\%$) um die Farben von Bereich 1 liegen.

Bereich 0:

Alle Farben, die nicht zum Bereich 1 und nicht zum Bereich 2 gehören.

Für den Filterungsalgorithmus wird ein matrixartig aufgebautes Abtastfenster (Suchmatrix) mit einer zur Bildgröße vergleichsweise geringen Spalten- und Zeilenanzahl schrittweise in Zeilen- und Spaltenrichtung über den zu betrachtenden Bildbereich geführt. Dabei soll jeder Bildpunkt des zu verarbeitenden Bildes einmal zum Mittelpunkt des matrixartigen Abtastfensters werden.

Nachfolgend soll über das Mittelpunktpixel die Entscheidung getroffen werden, ob es als Hintergrundpixel unterdrückt oder als Vordergrundpixel (Schrift) bestehen bleiben soll. Für die Entscheidung werden drei Regeln herangezogen:

1. Ist das Mittelpunktpixel Element von Bereich 1, so wird das Mittelpunktpixel als Hintergrund interpretiert und damit unterdrückt.
2. Ist das Mittelpunktpixel Element von Bereich 2 und mindestens einer der unmittelbaren Nachbarpixel (es gibt 8 direkte Nachbarpixel) Element von Bereich 1, so wird das Mittelpunktpixel als Hintergrund interpretiert und damit unterdrückt.
3. Ist das Mittelpunktpixel Element von Bereich 2 und mindestens eines der Pixel im Abtastfenster Element von Bereich 1 (= Zielpixel), dann wird das Mittelpunktpixel als Hintergrundpixel interpretiert, wenn alle Pixel, die auf einer Verbindungslinie von Mittelpunktpixel zu Zielpixel liegen, entweder zum Bereich 2 oder 1 gehören. Es genügt, wenn eine Verbindungslinie innerhalb der Suchmatrix die geforderten Eigenschaften erfüllt.

Die maximale Länge der Verbindungslinie, die Form der zulässigen Verbindungslinien und Größe der Suchmatrix ermöglichen vielfältige Anpassungen an die Problemstellung.

Das Verfahren kann als Vorverarbeitungsstufe zu einem Binarisierungsalgorithmus eingesetzt werden, dabei kann z.B. der zu unterdrückende Bildpunkt durch einen bekannten Hintergrundwert ersetzt werden. Das Verfahren eignet sich durch seinen Klassifikationsmechanismus (Vordergrund \Leftrightarrow Hintergrund) jedoch auch direkt zu einer Binarisierung auf farbigen Bildern.

3. ZUSAMMENFASSUNG

Die Untersuchung verschiedener Ansätze zur adaptiven Farbfilterung zeigt, dass das - an die menschliche Wahrnehmung adaptierte Verfahren - eine intuitive und damit auch eine zuverlässige Handhabung der Farbbelehrung erlaubt. Die Vorteile dieser Methode, die durch das Verfahren der 2-dimensionalen Farbsegmentierung ergänzt wurde, wurden auf kritischen Beispieldokumenten erprobt und zeigen eine deutlich bessere Wirkung, als die bislang angewendeten Verfahren.

Eine hardwarenahe Implementierung der Verfahren wird zur Realisierung vorgeschlagen, da in vielen dokumentverarbeitenden Prozessen Echtzeitanforderungen gestellt werden, um eine zur Bildgewinnung synchrone Verarbeitung zu ermöglichen.

4. BILDANHANG

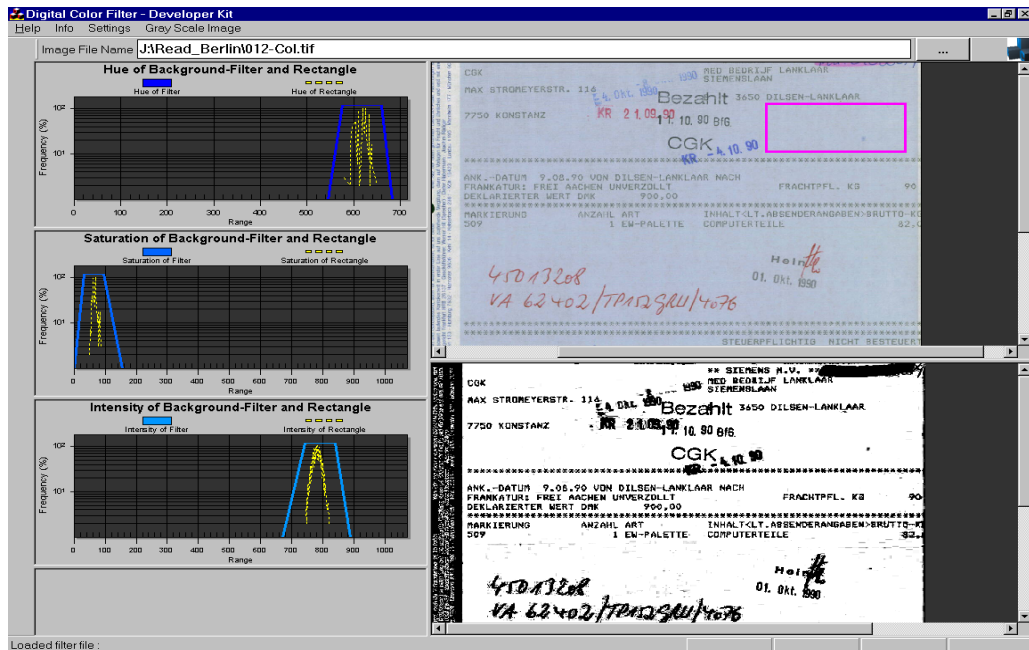


Abb.2: GUI zur Beherrschung von Farben auf Dokumenten

Beispiel für das Unterdrücken des Hintergrundes durch die umfeldbezogene Farbfilterung

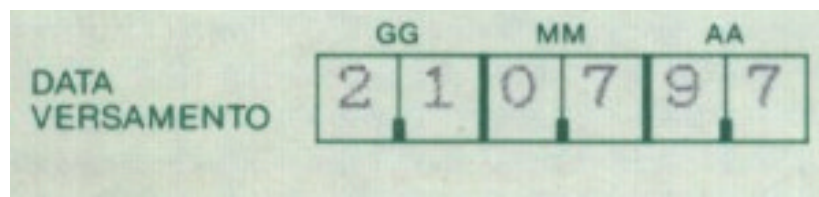


Abb. 3: Original

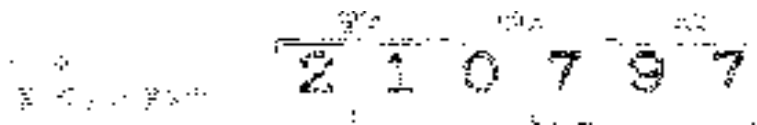


Abb. 4: Binarisierungsergebnis bei herkömmlicher Farbfilterung

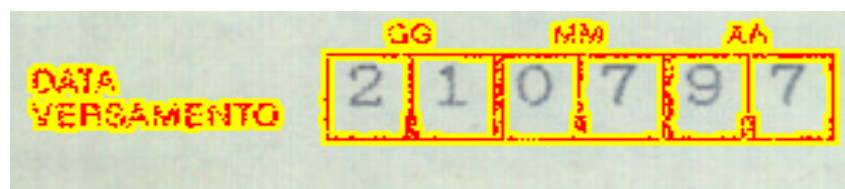


Abb. 5: Markierung der Suchbereiche bei der umfeldbezogenen Farbfilterung (Rot = Bereich 1, Gelb = durch umfeldbezogene Farbfilterung gefundene Pixel aus Bereich 2)

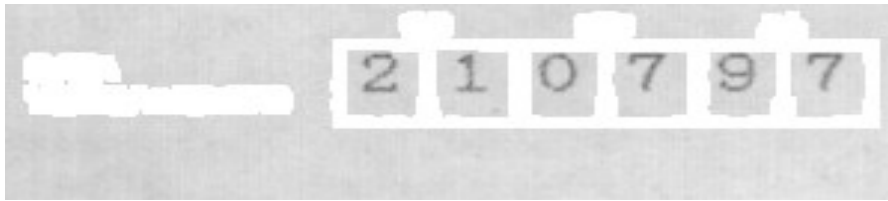


Abb. 6: Unterdrückung der Hintergrundbedruckung bei der umfeldbezogenen Farbfilterung

2 1 0 7 9 7

Abb. 7: Ergebnis der umfeldbezogenen Farbfilterung nach einer Binarisierung

Beispiel für das Herauslösen der Schrift durch die umfeldbezogene Farbfilterung

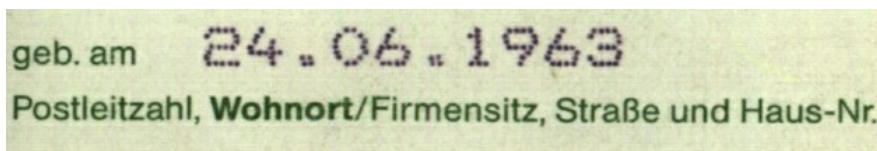


Abbildung 8: Original

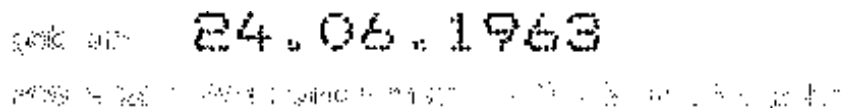


Abb. 9: Ergebnis der herkömmlichen Farbfilterung nach einer Binarisierung

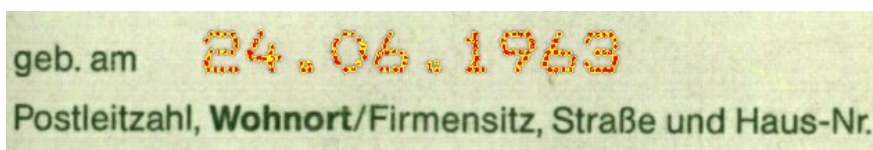


Abb. 10: Markierung der Schrift bei umfeldbezogener Farbfilterung (Rot = Bereich 1, Gelb = durch umfeldbezogene Farbfilterung gefundene Pixel aus Bereich2)

24.06.1963

Abb. 11: Ergebnis der umfeldbezogenen Farbfilterung nach einer Binarisierung