Bildverarbeitung in der Endoskopie des Bauchraums

Vogt, F. [1,*] , Klimowicz, C.[1], Paulus, D.[1], Hohenberger, W.[2], Niemann, H.[1], Schick, C. H.[2]

[1] Lehrstuhl für Mustererkennung
Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen
Email: {vogt,paulus,niemann}@informatik.uni–erlangen.de
{crklimow}@immd5.informatik.uni–erlangen.de

[2] Chirurgische Universitätsklinik
Krankenhausstr. 12, 91058 Erlangen
Email: schick@chirurgie-erlangen.de

Zusammenfassung. Derzeit werden bei endoskopischen Operationen, unter der Verwendung von Spezialfarbkameras, die unverarbeiteten Bilder aus dem Körper direkt auf einem Fernsehbildschirm dargestellt. Diese „direkte“ Darstellung der Bilder auf dem Fernsehbildschirm ist mit Einschränkungen der Darstellungsqualität verbunden. Vor allem Glanzlichter, Verzerrungen, Farbfehler und Schwebepartikel beeinträchtigen die Bildqualität. Es werden verschiedene Verfahren zur Verbesserung der Bildqualität vorgestellt und evaluiert.

1 Einleitung

Die Entwicklung der chirurgischen Technik verläuft hin zu den die Patienten weniger traumatisierenden, so genannten „minimal invasiven“ Operationen. Am Beispiel der Gallenblasenentfernung hat sich diese Entwicklung bereits soweit vollzogen, dass sie als „laparoskopische Cholezystektomie“ den alten Eingriff mittels großem Bauchschnitt weitgehend abgelöst hat.

Am Beispiel des laparoskopischen Eingriffs wird im Folgenden das Prinzip der minimal invasiven Chirurgie erläutert. Der Zugang in die Bauchhöhle erfolgt durch einen etwa 2 cm langen Schnitt unterhalb des Nabels. In die Bauchhöhlenöffnung wird ein Schleusenrohr, ein so genannter Trokar, eingeführt, über das zunächst CO_2-Gas bis zu einem vordefinierten Druck (10-14 mm Hg) in die Bauchhöhle eingeblasen wird.

Durch den Trokar kann nun das optische System, welches aus einer circa 30 cm langen und 1 cm durchmessenden Optik mit Kameraaufsatz besteht, in den Bauchraum geschoben werden. Um die zur Operationsdurchführung benötigten Instrumente in den Bauchraum einführen zu können, werden drei weitere Tokare mit 0,5 bis 1,2 cm Durchmesser in die Bauchdecke eingesetzt. Mittels langer, schlanker Instrumente wird die Gallenblase im rechten Oberbauch am Leberunterrand aufgesucht und entfernt. Dabei kommt ein elektrischer Gewebeschneider zum Einsatz.

2 Medizinische Problematik

Während der Operation werden die von der Kameraoptik gelieferten Bilder unverarbeitet, direkt auf dem Fernsehbildschirm dargestellt. Die Bildqualität wird vor allem

* Teile dieser Arbeit wurden von der DFG unterstützt im Teilprojekt C2 des SFB 603.

durch Glanzlichter, Verzerrungen, Farbfehler und Schwebepartikel beeinträchtigt (siehe Abb. 2, 3 und 4). Aus Sicht des Arztes wäre es wünschenswert alle genannten Einschränkungen zu verringern.

Die Farbfehler entstehen durch vorwiegend rötliche Farben (z.B. Blut und Gewebe) im Bild. Durch Reflexionen der Lichtquelle des Endoskops an feuchten Organ-Oberflächen entstehen so genannte Glanzlichter (siehe Abb. 3). Sie sind gekennzeichnet durch starke Kontraste im Bild, welche die Qualität des Bildes in ihrer Umgebung beeinträchtigen. Durch Verbrennen von Gewebe (beim Präparieren) während der Operation entstehen störende Schwebepartikel (siehe Abb. 4). Wenn zuviele Schwebepartikel vorhanden sind ist derzeit ein kompletter Austausch des Gases im Bauch notwendig, um die Operation fortsetzen zu können. Dies hat einen Zeitverlust zur Folge, der möglichst gering ausfallen sollte. Erschwerend zu obigen Bildstörungen kommen Verzerrungen des Bildes, verursacht durch die eingesetzte Kameraoptik, hinzu, wie sie in der Aufnahme des Kalibriermusters in Abb. 2 zu sehen sind. Vor allem die kleine Brennweite (Weitwinkel) spielt hierbei eine große Rolle.

3 Methoden

Aus der Bildverarbeitung sind Techniken zur Bildverbesserung seit Jahren bekannt. Es stellt sich die Frage, welche dieser Methoden bei endoskopischen Bildern sinnvoll eingesetzt werden können und dem Arzt eine verbesserte Sicht bieten.

Das dichromatische Modell [1] lässt sich auch für die Erkennung von Glanzlichtern bestimmter Oberflächentypen anwenden [2, 3, 4]. Die Voraussetzungen für dieses Modell sind im Falle von endoskopischen Bildern nicht erfüllt. Dieses Reflexionsmodell wurde jedoch erfolgreich auf Gewebebildern angewandt, wie in [2] für Bilder des Rachenraums und in [5] für Gesichtsbilder gezeigt wurde. Nach dem dichromatischen Reflexionsmodell setzt sich die spektrale Verteilung des reflektierenen Lichts Y aus einer Linearkombination aus spiegelnder (interface) L_{I} und diffuser (body) Reflexion L_{B} zusammen:

$$Y(\boldsymbol{\theta}, \lambda) = c_{\mathrm{I}}(\boldsymbol{\theta}) \cdot L_{\mathrm{I}}(\lambda) + c_{\mathrm{B}}(\boldsymbol{\theta}) \cdot L_{\mathrm{B}}(\lambda)$$

$$L_{\mathrm{I}}(\lambda) = S_{\mathrm{I}}(\lambda) \cdot E(\lambda), L_{\mathrm{B}}(\lambda) = S_{\mathrm{B}}(\lambda) \cdot E(\lambda)$$

Die beiden Farbkomponenten L_{I} und L_{B} sind nur von der Wellenlänge λ des einfallenden Lichts abhängig. $S_{\mathrm{I}}(\lambda)$ sowie $S_{\mathrm{B}}(\lambda)$ sind Reflexionseigenschaften. Mit $E(\lambda)$ wird das Spektrum der Lichtquelle beschrieben. Die Gewichstfaktoren c_{I} und c_{B} werden durch die Aufnahmegeometrie $\boldsymbol{\theta}$ bestimmt; im Vektor $\boldsymbol{\theta}$ werden alle Winkel zusammegefasst. Ein Lichtstrahl welcher das Objekt trifft wird bereits an der Oberfläche reflektiert. Ein Teil davon dringt unter die Oberfläche, wird teils absorbiert, teils gestreut und verläßt ohne Vorzugsrichtung (diffus) das Objekt (siehe Abb. 1).

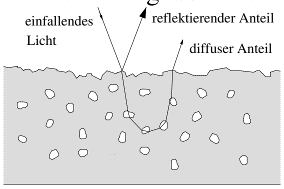
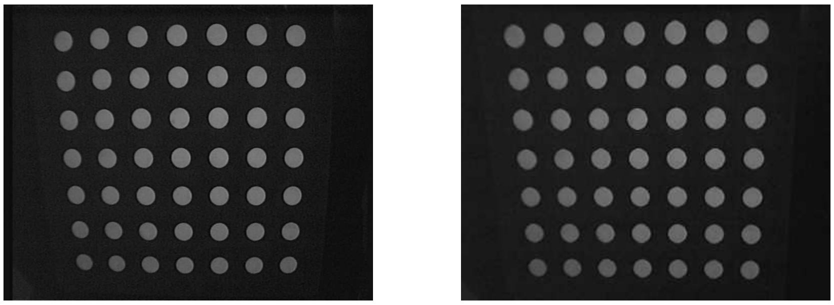


Abb.1.: Reflexionsmodell

Aufgrund des Modells werden Unterräume im RGB Farbraum definiert, deren Schnittgerade die Berechnung der Beleuchtungsfarbe ermöglicht.

Abb.2. Kalibriermuster original (links) und entzerrt (rechts)



Durch Kalibrierung der intrinsischen Kameraparameter können Verzerrungsparameter ermittelt werden [6]. Der Abbildungsprozess von unverzerrten Bildkoordinaten $(X_\mathrm{u}\ Y_\mathrm{u})^\mathrm{T}$ auf Rechnerkoordinaten $(X_\mathrm{f}\ Y_\mathrm{f})^\mathrm{T}$ wird in zwei Schritten durchgeführt. Bei der Tranformation von $(X_\mathrm{u}\ Y_\mathrm{u})^\mathrm{T}$ in $(X_\mathrm{f}\ Y_\mathrm{f})^\mathrm{T}$ werden radiale Verzerrungen durch ein Polynom $1 + \sum_{i=1}^{n} \kappa_\mathrm{i} \cdot R^{2i}$ mit $R = \sqrt{{X_\mathrm{u}}^2 + {Y_\mathrm{u}}^2}$ modelliert, wobei hier nur ein Polynom ersten Grades betrachtet wird, da sich gezeigt hat, dass dies in den meisten Fällen ausreichend ist:

$$(X_\mathrm{d}\ Y_\mathrm{d})^\mathrm{T} = (1 + \kappa R^2)(X_\mathrm{u}\ Y_\mathrm{u})^\mathrm{T} \tag{1}$$

Im zweiten Schritt werden die verzerrten Bildkoordinaten in die vom Rechner verwendeten Pixelkoordinaten umgerechnet. Die verzerrten Bildkoordinaten $(X_\mathrm{d}\ Y_\mathrm{d})^\mathrm{T}$, welche die reale Position des Bildpunktes auf dem Kamerasensor darstellen, können auch als Sensorkoordinaten bezeichnet werden. Der Urspung der Sensorkoordinaten wird durch den Schnittpunkt $(C_\mathrm{x}\ C_\mathrm{y})^\mathrm{T}$ der optischen Achse des Objektivs mit der Sensorebene festgelegt. Die Transformation für den zweiten Schritt wird wie folgt definiert:

$$\begin{pmatrix} X_\mathrm{f} \\ Y_\mathrm{f} \end{pmatrix} = \begin{pmatrix} \frac{s_\mathrm{x}}{{d_\mathrm{x}}'} X_\mathrm{d} \\ \frac{1}{{d_\mathrm{y}}'} Y_\mathrm{d} \end{pmatrix} + \begin{pmatrix} C_\mathrm{x} \\ C_\mathrm{y} \end{pmatrix} \quad . \tag{2}$$

Dabei bezeichnen ${d_\mathrm{x}}'$ und ${d_\mathrm{y}}'$ in Sensorkoordinaten die Abstände benachbarter Bildpixel in x- bzw. y-Richtung, wie sie im Bildspeicher nach der Abtastung (durch den Framegrabber) vorliegen. Ungenauigkeiten in der Abtastung durch den Framegrabber werden durch den Faktor s_x repräsentiert (gebräuchliche Annahme $s_\mathrm{x} = 1,0$).

Durch Aufnahme eines Kalibriermusters werden die Parameter κ, C_x und C_y ermittelt. Anschließend wird die Verzerrung invertiert.

Zur Farbnormierung bieten sich diejenigen Verfahren an, die eine Korrektur so durchführen, dass der Eindruck auf den Betrachter natürlich wirkt. Farbkonstanzverfahren sind wegen der aufwändigen Kalibrierung nicht möglich. Statt dessen wird eine affine Transformation des Farbraums durchgeführt. Hierzu wird wie in [7] beschrieben von dem gegebenen Bild zunächst die Farbkovarianzmatrix $\boldsymbol{C}$ ermittelt. Eine Ermittlung der Eigenwerte und Eigenvektoren dieser 3×3 Matrix schließt sich an. Der zum größten Eigenwert gehörige Eigenvektor gibt die Richtung der Farbhauptachse an. Eine Rotation des Farbraums wird nun mit einer Matrix $\boldsymbol{R}$ so durchgeführt, dass diese

Abb.3. Ausgangsbild (li.), Farbnormierung (mi.), Glanzlichterkennung (re.)

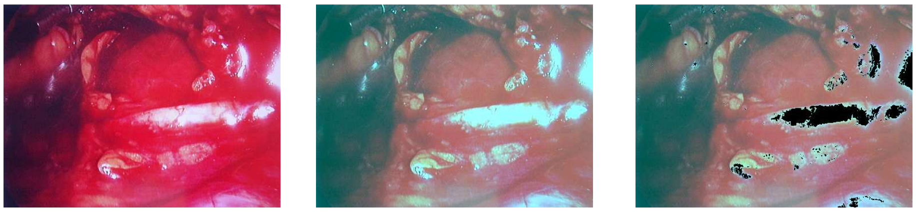
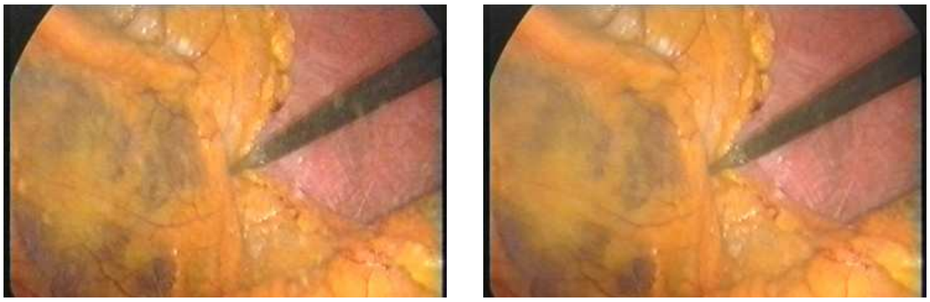


Abb.4. Originalbild (links) und nach zeitlicher Medianfilterung der Größe 3 (rechts)



Hauptachse auf die Unbuntachse des RGB Raums abgebildet wird und gleichzeitig die beiden anderen Achsen möglichst wenig verdreht werden.

Medianfilter (örtlich und zeitlich) dienen zur Reduktion von Punktstörungen. Bei Farbbildern werden in einer (örtlichen oder zeitlichen) Umgebung des betrachteten Farbpixels $\boldsymbol{x}$ Pixel ausgewählt und nach der euklidischen Länge des Farbvektors sortiert. Dem betrachteten Pixel wird der Median der sortierten Folge zugewiesen. Durch Gaußfilterung (Tiefpass) des Farbbildes wird hochfrequentes Rauschen reduziert. Dabei werden die einzelnen Farbkanäle mit diskreten Gaußmasken der Größe 3 x 3 gefaltet.

4 Experimente

Die in Abs. 3 beschriebenen Methoden werden nun auf laparoskopische Bilder der Größe 360 x 288 angewandt. Abb. 2 zeigt den Unterschied vor und nach der Entzerrung eines Kamerabildes, wobei der Verzerrungsparameter $\kappa = 0.006$ ermittelt wurde. Die Entzerrung eines Bildes dauerte 0,25 sec. Ergebnisse der Farbnormierungs-Experimente sind zu sehen in Abb. 3. Die zugehörigen Matrizen C und R sind:

$$C = \begin{pmatrix} 3516.00 & 1120.85 & 1334.25 \\ 1120.85 & 1860.67 & 1886.28 \\ 1334.25 & 1886.28 & 2022.03 \end{pmatrix} \quad R = \begin{pmatrix} 0.99 & -0.09 & -0.07 \\ 0.10 & 0.99 & 0.02 \\ 0.07 & -0.02 & 0.99 \end{pmatrix} \tag{3}$$

Die Farbnormierung eines Bildes dauerte 0,85 sec. Das Ergebnis einer zeitlichen Medianfilterung ist in Abb. 4 zu sehen. Ergebnisse der Glanzlichterkennung sind zu sehen in Abb. 3; Glanzlichtpixel sind schwarz dargestellt. Nach der Detektion können die gefundenen Pixel geeignet ersetzt werden. Zur Glanzlichterkennung wurden 1,10 sec pro Bild benötigt.

Die objektive Evaluation der Ergebnisse ist schwierig zu gestalten. Zunächst ist das Ziel, eine subjektive Verbesserung der Bildqualität zu erreichen, wobei die Beurteilung von dem medizinischen Personal zu treffen ist. Für 40 Bilder wurde daher ein Vergleich durchgeführt, in dem einem Arzt die Originale und verarbeiteten Einzelbilder in zufälliger Reihenfolge und Anordnung gezeigt wurden. In den meisten Fällen spiegelte der Farbeindruck des normierten Bilds eher die Realität einer offenen Operation wider. Der verfälschte Farbeindruck durch das Endoskop war jedoch gewohnter. Von 10 zeitlich gefilterten Bildern wurden in allen Fällen die gefilterten Bilder als besser eingestuft, da die Reduktion von Schwebepartikeln eine deutliche Bildverbesserung darstellte. Die räumliche Filterung dagegen führte zu verwaschenen Details.

Für einen weitergehenden Vergleich werden derzeit normierte Bildsequenzen vorbereitet, die mehreren Ärzten zur Evaluation vorgelegt werden sollen.

5 Zusammenfassung und Ausblick

Das Projekt, in dem unser System eingesetzt wird, befindet sich im Anfangsstadium. Wie jedoch jetzt schon gezeigt werden konnte, lässt sich durch geeignete Bildverarbeitung die Bildqualität verbessern. Auch hat sich gezeigt, dass nicht alle Verarbeitungsschritte, z.B. Gauß- oder Medinafilter, bedingungslos anwendbar sind.

References

1. S. A. Shafer. Using color to separate reflection components. *COLOR research and application*, 10(4):210–218, 1985.
2. C. Palm, T. Lehmann, and K. Spitzer. Bestimmung der Lichtquellenfarbe bei der Endoskopie makrotexturierter Oberflächen des Kehlkopfs. In K.-H. Franke, editor, *5. Workshop Farbbildverarbeitung*, pages 3–10, Ilmenau, 1999. Schriftenreihe des Zentrums für Bild- und Signalverarbeitung e.V. Ilmenau.
3. Th. Gevers and H. M. G. Stokman. Classifying color transitions into shadow-geometry, illumination highlight or material edges. In *Proceedings of the International Conference on Image Processing (ICIP)*, pages I:521–524, Vancouver, BC, September 2000. IEEE Computer Society Press.
4. K. Schlüns and A. Koschan. Global and local highlight analysis in color images. In CGIP00 [8], pages 147–151.
5. M. Störring, Granum E., and H. J. Andersen. Estimation of the illuminant color using highlights from human skin. In CGIP00 [8], pages 45–50.
6. R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, Ra-3(3):323–344, August 1987.
7. D. Paulus, L. Csink, and H. Niemann. Color cluster rotation. In *Proceedings of the International Conference on Image Processing (ICIP)*, Chicago, October 1998. IEEE Computer Society Press.
8. *First International Conference on Color in Graphics and Image Processing*, St. Etienne, 2000.