

# Ein automatisiertes Verfahren zum Image-Mosaicing bei endoskopischen Videoaufnahmen

Beate Breiderhoff<sup>1</sup>, Wolfgang Konen<sup>1</sup>, Martin Scholz<sup>2</sup>

<sup>1</sup>Institut für Informatik, Fachhochschule Köln, 51643 Gummersbach

<sup>2</sup>Neurochirurgische Universitätsklinik, Knappschafts Krankenhaus, 44892 Bochum  
breider@gm.fh-koeln.de

**Zusammenfassung.** In diesem Beitrag beschreiben wir ein automatisiertes Verfahren zur Registrierung und anschließender Panoramaerstellung bei endoskopischen Videoaufnahmen. Ein bereits existierendes Verfahren [1] wird um die Einbeziehung der endoskoptypischen Maske erweitert und um einen hierarchischen Ansatz ergänzt. Die Minimierung der Grauwertdifferenzen der zu registrierenden Bilder erfolgt gemäß Levenberg-Marquardt-Algorithmus. Das Verfahren unterziehen wir ersten, exemplarischen Tests mit Einzelbildern. Die erweiterten Methoden zeigen eine gute Eignung für Endoskopbilder. Eine signifikante Steigerung der Geschwindigkeit wird durch den Einsatz von Pyramiden erreicht. Eine vorher durchgeführte Glättung der Bilder durch Gaußfilterung erhöht die Robustheit der Registrierung.

## 1 Einleitung

Die Erzeugung eines Mosaikbildes aus endoskopischen Bildern stellt einen bislang unbearbeiteten Teil der medizinischen Bildverarbeitung dar. Bisher beschränkt sich die Anwendung der digitalen Bildverarbeitung in der Endoskopie auf die Dokumentation und die digitale Bildoptimierung. Neben diesen Möglichkeiten wäre eine virtuelle Bildausschnittsvergrößerung für die Bestimmung der anatomischen Position während einer Operation hilfreich. Dies gilt in besonderem Maße in der Neuroendoskopie [2].

Auf dem Markt vorhandene Navigationssysteme können durch Einspielung von CT- oder MRI-Bildern bereits eine Navigation und verbesserte Orientierung ermöglichen. Systeme, die zur Navigation nur die endoskopische Bildinformation nutzen, existieren aber noch nicht. Techniken zur Mosaikerstellung bei digitalen Bildern sind seit langem bekannt (z. B. [1, 3, 4]). Die wenigsten arbeiten vollautomatisch und nur ein Teil arbeitet unter real-time Bedingungen. Zur Überblendung endoskopischer Videoaufnahmen gibt es eine interessante Arbeit von Wald et al. [5]. Wichtige Wege zur Qualitätsverbesserung der Bilder z. B. durch Reduzierung von Farbfehlern oder Ersetzung von Glanzlichtern zeigen Vogt et al. [6].

Endoskoptaufnahmen weisen Eigenschaften auf, die sowohl eine Registrierung als auch eine Überblendung der Bilder erschweren. In Abschnitt 2 werden diese Problemfelder genauer beleuchtet. Abschnitt 3 beschreibt unser erweitertes Verfahren im Detail. Der vierte Abschnitt fasst die Ergebnisse erster Experimente zusammen.

## 2 Problemfelder

Endoskopbilder weisen, bedingt durch die eingesetzten Optiken, besonders starke *Verzeichnungen* auf. Die *Kamerabewegung* ist freihändig, d.h. sie unterliegt keinen Einschränkungen, die eine Registrierung erleichtern könnten. Hinzu kommt, dass sich die Kamera nahe zur nicht regelmäßigen Oberfläche befindet, und die Strukturen im Bild sich bewegen, d.h. die Szenen sind nicht immer statisch. Durch die Bewegung der Kamera kann es zu schwerwiegenden *Beleuchtungsunterschieden* kommen, da sich gleichzeitig mit der Kamera die integrierte Beleuchtungsquelle bewegt, und sich die Beleuchtung einer Szene ändert. Wandernde Glanzlichter können z. B. die Folge sein. Zusätzlich erschweren Schwebepartikel, Rauchentwicklung und Blutungen die Sicht und somit die räumliche Orientierung. Verfahren zur Bildverbesserung müssen als Vorverarbeitung eingesetzt werden. Das Auffinden markanter Stellen im Bild, sogenannter Features, ist in den *strukturschwachen* Endoskopbildern erschwert. Befinden sich Objekte unterschiedlicher *Tiefe* im Bild, so existiert im Allgemeinen keine geschlossene, stetige 2D-Abbildung, die die Ansichten ineinander überführt.

*Ergonomische Aspekte* im Hinblick auf eine immer ausgeprägtere Mensch/Maschine Interaktion werden in der Endoskopie bisher nur unzureichend berücksichtigt. Ergebnisbilder müssen ergonomischen Test durch Experten unterzogen werden, damit die für den Operateur optimale ergonomische Konfiguration ermittelt werden kann. Eine Mindestanforderung ist, dass das Mosaicing *automatisiert* arbeiten muss, da der Arzt während der Operation keine Passpunkte setzen kann. Zudem muss eine Verwirrung oder Ablenkung des Operateurs durch inkonsistente Informationen oder zu lange Rechenzeiten vermieden werden.

## 3 Methoden

Die meisten Registrierungsalgorithmen lassen sich grob in intensitätsbasierte Verfahren und featurebasierte Verfahren aufteilen. Featurebasierte Verfahren konzentrieren sich auf markante Stellen im Bild (z.B. Eckpunkte). Aufgrund der Schwierigkeit, Eckpunkte in den strukturschwachen Endoskopiebildern zu finden, und der Schwierigkeit eines automatischen Feature-Matching kommt diese Klasse von Verfahren nicht zum Einsatz.

Das ausgewählte Verfahren nutzt die Intensitätswerte und basiert auf dem Verfahren nach Szeliski [1]. Es bietet den Vorteil der automatischen Transformationsuche. Gegenüber Szeliski kommen folgende Aspekte hinzu:

- Um das Verfahren erstmalig auf Endoskopbilder anzuwenden ist eine Einbeziehung der kreisförmigen Maske notwendig.
- Zur Verbesserung der Genauigkeit und Geschwindigkeit der lokalen Minimumsuche, wird das Verfahren um einen hierarchischen Ansatz erweitert.

Nach Szeliski kann man die Registrierungsparameter berechnen durch Minimierung der Differenz der Grauwerte sich entsprechender Pixel in den Bildern  $I$  und  $I'$ .

$$\chi^2 = \sum_i (I'(x'_i, y'_i) - I(x_i, y_i))^2 \quad (1)$$

Um das Minimum zu suchen kommt die iterative, nichtlineare Least-Square-Optimierung nach Levenberg-Marquardt zum Einsatz [7]. Aufgrund der speziellen endoskopischen Maske sind sowohl bei der Transformation als auch bei Filterung und Gradientenberechnung Randeffekte zu beachten, für Details hierzu s. [8].

Zur Überblendung wird zunächst ein einfaches Verfahren angewendet, bei dem ein Parameter  $t$  die lineare Mischung von  $I$  und  $I'$  in der Überlappregion definiert:

$$g = (1-t)I(x, y) + t I'(x', y') \quad (2)$$

Szeliski [1] schlägt vor, die Bildpunkte eines Quellbildes  $I$  umso schwächer zu gewichten, je näher sie am Bildrand liegen. Diese Gewichtung verhindert sichtbare Artefakte an den Bildübergängen im Mosaikbild. Ein spezielles Verfahren zur fließenden Überblendung bei Endoskopbildern stellen Wald et al. [6] vor.

## 4 Experimente

In einer ersten Testreihe soll die Geschwindigkeitssteigerung durch die beschriebene AuflösungsPyramide untersucht werden. Dazu werden zwei 8-bit Grauwert-Bilder der Größe 512 x 512 mit einer kreisförmigen Maske verwendet. Das Basisbild wird mit Hilfe der entwickelten Software transformiert, um ein leicht gedrehtes und skaliertes Bild zu erhalten (z. B. achseneinheitliche Vergrößerung um Faktor 1.1 und Drehung um Winkel von 10 Grad). Die Transformationsklasse für die Registrierung hat sechs Freiheitsgrade (affine Abbildung). Die Parameter der genauen Registrierung sind, da das veränderte Bild selbst erstellt wurde, bekannt. In einer Reihe von fünf Registrierungen wird die Höhe der Pyramide variiert. Der *Abstand* (durchschnittliche Distanz, um die die Pixel bewegt werden müssen) des idealen Parametervektors zum Startvektor wird im 2. und 3. Durchlauf vergrößert. Zur Performancesteigerung wird über einen Sample-Faktor die Anzahl der in das Optimierungsmodell eingehenden Pixel unter 5000 gebracht.

Einen Auszug aus den Ergebnissen auf gastroendoskopischen Bildern zeigt Tabelle 1, die Ergebnisse aller Durchläufe sind in Abb. 1 zusammengefasst. Man erkennt, dass gerade bei größerem Abstand die Rechenzeit ohne Pyramide (nur Ebene 1) explodiert. Weitere Ergebnisse bei optischer Überprüfung der Ergebnisbilder: Die Resultate der ersten Versuche bis hin zu drei Pyramidenstufen sind sehr zufriedenstellend. Beim Durchlauf mit vier Ebenen in der Pyramide sind kleinere Ghosting-Artefakte zu sehen, die in der geringeren Genauigkeit der Registrierung begründet liegen. Die Grauwerte beider Bilder gehen hier zu gleichen Teilen in den Überlappungsbereich ein. Das ermöglicht eine gute optische Kontrolle der Registrierungsgüte. Ab einer Höhe der AuflösungsPyramide von fünf Ebenen ist das Ergebnis nicht mehr ausreichend (siehe Abb. 2, rechts). Durch Verwendung der Pyramide lässt sich die Rechenzeit stark verringern. Die richtige Höhe der AuflösungsPyramide ist aber entscheidend. Eine zwei- oder dreistufige Pyramide liefert in unserem Fall das beste Ergebnis, eine vierstufige resultiert dagegen in nur geringfügig oder gar nicht verbesserter Rechenzeit und einem Registrierungsergebnis, das keine „glatten“ Mosaik mehr zur Folge hat.

Der Einsatz eines Gaußfilters zur Glättung der Pyramiden zeigte gute Ergebnisse, denn dies verbessert die Signaldarstellung auf den höheren Pyramidenebenen. Die Bilder aus der vorherigen Testreihe wurden vor der Berechnung der nächsten Ebene einer Filterung mit einem Gaußfilter vom Radius 5 unterzogen. Wie Abb. 3 zeigt, ist bei großem Abstand der initialen Parameterschätzung von den tatsächlichen Parameterwerten ohne Filterung eine gute Registrierung oft nicht mehr möglich.

## 5 Fazit und Ausblick

Die grundsätzliche Anwendbarkeit des Verfahrens auf Endoskopbilder und problemlose Einbeziehung der Maske konnte gezeigt werden. Der pyramidale Ansatz bietet erhebliche Verbesserung hinsichtlich der Geschwindigkeit und eine Erhöhung der Genauigkeit bei größeren Abständen (Tab. 1, rechts). Eine zusätzliche Filterung verbessert die Registrierungsergebnisse. Für einen Vergleich des Verfahrens mit anderen Methoden und geplante Weiterentwicklungen sei auf [8] verwiesen.

## 6 Literatur

1. Szeliski R: Image Mosaicing for Tele-Reality, Technical Report, Cambridge Research Laboratory, 1994
2. Konen W, Scholz M, Tombrock S: The VN-project: endoscopic image processing for neurosurgery, Computer Aided Surgery 3(3), 1998
3. Peleg S, Herman J: Panoramic mosaics by manifold projection, In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 338-343, 1997
4. Sawhney H, Hsu S, Kumar R: Robust video mosaicing through topology inference and local to global alignment, Procs ECCV, Springer, 1998.
5. Wald D, Reeff M, Székely G, Cattin P, Paulus D: Fließende Überblendung von Endoskopiebildern für die Erstellung eines Mosaiks, Procs BVM 2005, 287-291, 2005
6. Vogt F, Klimowicz C, Paulus D, Hohenberger W, Niemann H, Schick C: Bildverarbeitung in der Endoskopie des Bauchraums, Procs BVM 2001, 320-324, 2001
7. Press W, Teukolsky S, Vetterling T, Flannery B: Numerical Recipes in C, The Art of Scientific Computing, 2nd ed. University Press, Cambridge, 1992
8. Konen W, Breiderhoff B, Scholz M: Real-time image mosaic for endoscopic video sequences, submitted to Procs BVM 2007

Tabelle 1: links: 1. Durchlauf, Abstand = 1.660, rechts: 3. Durchlauf, Abstand =16.580. Es wird die Anzahl der Ebenen in der Pyramide variiert (1= ohne Pyramide)

Ebenen	Zeit(s)	RMSE	Ebenen	Zeit(s)	RMSE
1	28	6.229	1	57	13.339
2	12	7.714	2	11	2.910
3	4	6.985	3	3	4.956
4	3	11.871	4	3	7.135
5	3	20.670	5	2	10.767

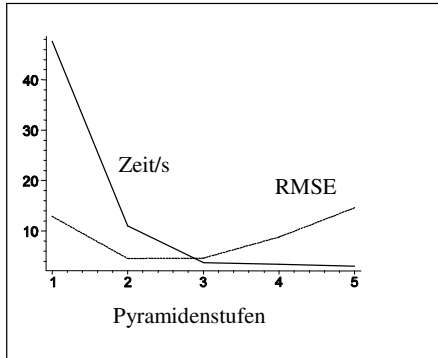


Abb.1: Durchschnittlicher Registrierungsfehler (RMSE = mittlere Abweichung der Grauwerte pro Pixel am Ende einer Registrierung) und durchschnittliche Rechenzeit gemittelt über 4 Testreihen mit unterschiedlichem Abstand.

Abb. 2: Ergebnisse des 1. Testdurchlaufes (kleiner Abstand), links: Höhe der Pyramide = 3, rechts: =5. Rechts sind deutliche Ghosting-Artefakte zu erkennen.

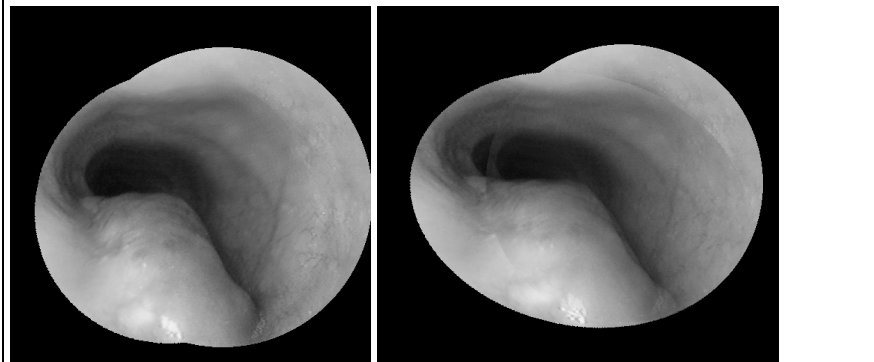


Abb. 3: Ergebnisse der Gaußfilterung im 4. Durchlauf (großer Abstand). Links ohne Filterung, rechts mit Filterung (Höhe der Pyramide jeweils 3). Links ist keine Registrierung möglich, während sie rechts fehlerfrei gelingt.

