Automatische 3D-Rekonstruktion aus endoskopischen Bildfolgen

Thorsten Thormählen, Hellward Broszio, Peter N. Meier[1]

Universität Hannover, Laboratorium für Informationstechnologie,
Abteilung Systemtechnik, Schneiderberg 32, 30167 Hannover
[1]Medizinische Hochschule Hannover, Abteilung Gastroenterologie
und Hepatologie, Carl-Neuberg-Str.1, 30623 Hannover
Email: thormae@tnt.uni-hannover.de

Zusammenfassung. In diesem Beitrag wird ein automatisches Verfahren zur Erstellung eines 3D-Oberflächenmodells aus endoskopischen Bildfolgen vorgestellt. Die Aufnahme der Bildfolgen erfolgt mit handelsüblichen Video-Endoskopen. Zunächst schätzt das Verfahren die Bewegung der Endoskopkamera aus korrespondierenden Bildmerkmalen, die in der Bildfolge detektiert werden. Anschließend werden die Raumkoordinaten dieser korrespondierenden Merkmalspunkte bestimmt. Die Auswertung der Information aus vielen Bildern der Sequenz erhöht die Genauigkeit der Raumkoordinaten. Durch die Raumkoordinaten wird ein Dreiecksnetz gelegt, das die Oberfläche der aufgenommenen Szene approximiert. Mit der Texturierung des Dreiecksnetzes aus Bildern der Sequenz entsteht ein realistisches 3D-Oberflächenmodell.

1 Einleitung

Bei der Endoskopie erfolgt die Diagnose und Intervention durch die Auswertung der Bildfolge einer Videokamera. Die Bildfolge der Kamera bietet keine Tiefeninformationen. Daher fehlt das gewohnte stereoskopische Sehen, wodurch die Navigation von Instrumenten erschwert wird. Das Erkennen und Bewerten pathologischer Strukturen und die Beurteilung ihrer räumlichen Ausdehnung ist daher nur durch große Erfahrung des Untersuchers möglich und beruht auf subjektiver Einschätzung. Mit Hilfe von 3D-Oberflächenmodellen der umliegenden Magen- bzw. Darmwand können endoskopische Befunde objektiver und reproduzierbarer gestaltet werden. In diesem Beitrag wird daher ein automatisches Verfahren zur Erstellung eines 3D-Oberflächenmodells aus endoskopischen Bildfolgen vorgestellt.

Das Verfahren schätzt die Bewegung der Endoskopkamera und verwendet die geschätzte Kamerabewegung zur Rekonstruktion der beobachteten Szene. Da die Bewegung der Kamera nicht kontrolliert gesteuert oder zuverlässig gemessen werden kann, muss sie aus der Bildfolge geschätzt werden. Verfahren, die 3D-Modelle auf diese Weise aus Bildfolgen generieren, werden in der Computer Vision Structure-from-Motion genannt. Der Vorteil des Einsatzes eines

Structure-from-Motion-Verfahrens in der Endoskopie liegt darin, dass ausschließlich Bildfolgen handelsüblicher Video-Endoskope verwendet werden. Die meisten anderen Ansätze zur 3D-Rekonstruktion in der Endoskopie benötigen ein spezielles Endoskop oder verwenden ein zusätzliches Messgerät, das in den Arbeitskanal des Endoskops geschoben wird [1].

Das hier vorgestellte Verfahren besteht aus zwei Teilen. Der erste Teil, eine sehr robuste Schätzung der Kamerabewegung, wird im nächsten Abschnitt beschrieben. Der zweite Teil besteht aus der Rekonstruktion der 3D-Oberfläche bei bekannter Kamerabewegung. Diese wird in Abschnitt 3 beschrieben. In Abschnitt 4 werden Ergebnisse des Verfahrens präsentiert. Der Beitrag endet mit Schlussfolgerungen und einem Ausblick.

2 Robuste Schätzung der Kamerabewegung

Die Schätzung der Kameraparameter aus endoskopischen Bildfolgen erfolgt in drei Schritten: Detektion von Merkmalspunkten, Korrespondenzanalyse und die robuste Schätzung der Kamerabewegung. In Abbildung 1 ist das Ablaufdiagramm des Schätzverfahrens dargestellt.

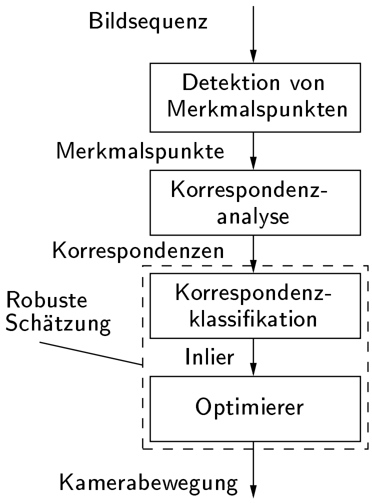

Abb. 1. Schritte zur Schätzung der Kamerabewegung aus Bildfolgen.

Die Detektion von Merkmalspunkten erfolgt mit Subpelgenauigkeit durch den Algorithmus von Harris [2], der zuverlässige Ergebnisse liefert. In der anschließenden Korrespondenzanalyse werden die gleichen Merkmalspunkte in aufeinanderfolgenden Bildern bestimmt. In jeweils zwei aufeinanderfolgenden Bildern werden Merkmalspunkte mit der höchsten normierten Kreuzkorrelation zu einem Paar zusammengefasst. Die normierte Kreuzkorrelation wird über ein 15×15 pel großes Fenster berechnet, das die Merkmalspunkte umgibt. Ein Paar von Merkmalspunkten wird im Folgenden Korrespondenz genannt. Die Korrespondenz beschreibt die Verschiebung der Projektion der zugehörigen 3D-Koordinate in aufeinanderfolgenden Bildern in Abhängigkeit von der Kamerabewegung. Daher kann die gesuchte Kamerabewegung aus den Korrespondenzen geschätzt werden. Bei der Detektion von Merkmalspunkten und bei der Korrespondenzanalyse treten in der praktischen medizinischen Anwendung immer Fehler auf. Um eine robuste Schätzung zu erhalten, wird daher zunächst eine Korrespondenzklassifikation durchgeführt. Mit Hilfe der Korrespondenzklassifikation werden Fehlzuordnungen, genannt Outlier, gefunden und eliminiert. Hierzu wird ein „Random Sampling“ Algorithmus verwendet. Die verbleibenden guten Korrespondenzen, genannt Inlier, werden vom Optimierer ausgewertet, um die Kamerabewegung durch Maximierung einer Gütefunktion zu bestimmen. Mathematische Details zu diesem Verfahren sind in den Veröffentlichungen [3] und [4] beschrieben.

3 3D-Rekonstruktion

Die Kenntnis der Kamerabewegung bzw. -position ermöglicht eine Rekonstruktion der 3D-Oberfläche der umliegenden Magen- bzw. Darmwand. Dazu werden Sichtlinien durch alle Merkmalspunkte konstruiert, die bei der Schätzung der Kamerabewegung als Inlier klassifiziert wurden. Die Berechnung des Schnittpunktes von Sichtlinien korrespondierender Merkmalspunkte bestimmt die Raumkoordinate der Merkmalspunkte. In Abbildung 2 ist dieses Verfahren veranschaulicht. Das Ergebnis ist eine 3D-Punktwolke, deren Raumkoordinaten alle auf der zu rekonstruierenden Oberfläche liegen. Um eine geschlossene Oberfläche zu erhalten, werden die Raumkoordinaten durch Dreiecke verknüpft. Die Verknüpfung von 3D-Punkte einer Punktwolke zu einem Dreiecksnetz ist mehrdeutig. Daher wird die Triangulation der Merkmalspunkte in der Bildebene durchgeführt und auf die zugehörigen Raumkoordinaten übertragen. Für die gewonnenen Dreiecke des Netzes wird aus der endoskopischen Bildfolge eine Textur berechnet. Unter Verwendung der bekannten Kamerapositionen werden dazu die Dreiecke der Oberfläche in jedes Bild der Folge projiziert. Die projizierten Dreiecke beranden den Bildausschnitt, der ihre Textur beinhaltet. Die Textur wird aus demjenigen Bild der Sequenz gewonnen, in dem der Bildausschnitt die größte Fläche besitzt. Details zur Texturierung eines 3D-Modells aus Bildfolgen sind in [5] veröffentlicht.

4 Ergebnisse

Das Verfahren wurde auf eine Bildfolge angewendet, die bei einer echten endoskopischen Untersuchung aufgezeichnet wurde. In der Bildfolge ist ein Polyp im Dickdarm zu sehen. Mit Hilfe des erläuterten Verfahrens konnte ein texturiertes 3D-Oberflächenmodell des Polypen erzeugt werden (siehe Abbildung 3). Durch Mittelung der Raumkoordinaten über eine größere Anzahl von Bildern und Eliminierung von Merkmalspunkten auf Reflexionen konnte im Vergleich zu früheren Ergebnissen in [4] ein exakteres Oberflächenmodell generiert werden.

5 Schlussfolgerung und Ausblick

Das vorgestellte automatisches Schätzverfahren kann zur Erzeugung eines 3D-Oberflächenmodells aus endoskopischen Bildfolgen verwendet werden. Computerprogramme, die das gewonnene 3D-Oberflächenmodell visualisieren, bieten neue Möglichkeiten zur Diagnose, Operationsplanung und Dokumentation. In zukünftigen Arbeiten soll durch Vergleich mit CT- oder MRT-Daten eine Analyse des Schätzfehlers durchgeführt werden.

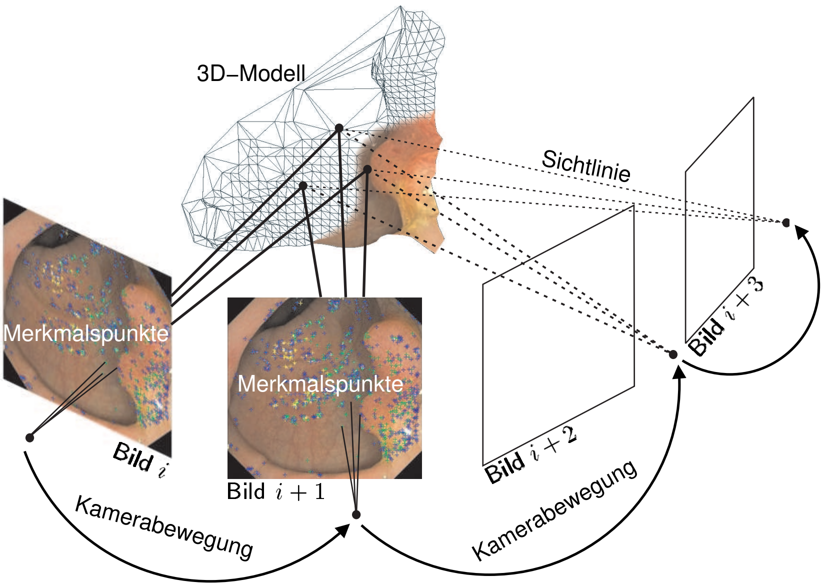

Abb. 2. Erzeugung eines 3D-Modells bei bekannter Kamerabewegung.

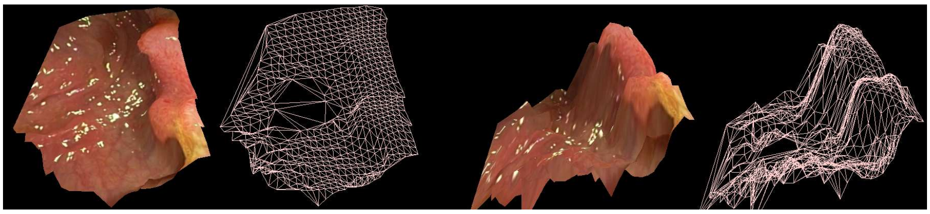

Abb. 3. Rekonstruiertes 3D-Oberflächenmodell eines Polypen im Dickdarm

Literatur

1. Schubert M, Müller A: Evaluation of endoscopic images for 3-dimensional measurement of hollow organs. Biomedical Engineering. 43(1); p 32-3; 1998.
2. Harris C, Stephens M: A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference. p 147-51; 1988.
3. Broszio H, Grau O: Robust Estimation of Camera Parameters Pan, Tilt and Zoom for Integration of Virtual Objects into Video Sequences. International Workshop on Synthetic-Natural Hybrid Coding and Three-Dimensional (3D) Imaging. Santorini, Greece; September 15-17; 1999.
4. Thormählen T, Broszio H, Meier PN: 3D Endoscopy. Falk Symposium No 124. Medical Imaging in Gastroenterology and Hepatology. September 28-29; 2001. Kluwer 2002.
5. Niem W, Broszio H: Mapping texture from multiple camera views onto 3D-object models for computer animation. International Workshop on Stereoscopic and Three-Dimensional Imaging. Santorini, Greece; September 6-8; 1995.