

Partikelfilter-basiertes Tracking chirurgischer Instrumente in Endoskopbildern

Markus E. Rilk, Simon Winkelbach und Friedrich M. Wahl

Institut für Robotik und Prozessinformatik,
Technische Universität Braunschweig, 38106 Braunschweig
Email: m.rilk@tu-bs.de

Zusammenfassung. Das hier vorgestellte Verfahren erkennt und verfolgt chirurgische Instrumente in Endoskopbildern ohne Hilfe von speziellen Markern oder farblichen Unterschieden zwischen Instrument und Hintergrund. Die ermittelte 3d-Position einer bestimmten Instrumentenspitze verwendet ein Roboter, um automatisch das Endoskop nachzuführen. Die Erkennung selbst ist aufgeteilt in zwei Phasen. In der Initialisierungsphase ruht das Endoskop, so dass das zu verfolgende Instrument aufgrund seiner Eigenbewegung gefunden werden kann. In der Trackingphase wird das Instrument mit Hilfe eines Partikelfilters verfolgt. Die Partikel beschreiben dabei in jedem Zeitschritt verschiedene Lagehypothesen des Instrumentes mit unterschiedlichen Wahrscheinlichkeiten.

1 Einleitung

Der Hauptkritikpunkt an der Video-Endoskopie ist, dass der Chirurg das Endoskop meist selbst halten muss. Dies ist auf Dauer ermüdend und dem Chirurg fehlt oftmals eine freie Hand, um beispielsweise weitere Instrumente zu bedienen. Versuche, einem Operationsassistenten die Führung des Endoskops zu überlassen, waren wenig ermutigend. Neben den ebenfalls auftretenden Ermüdungserscheinungen (zittriges Endoskopbild), sah der Chirurg oft nicht die von ihm gewünschte Region. Im hier vorgestellten Ansatz wird die Führung des Endoskops einem Roboter übergeben. Dieser kann das Endoskop nicht nur präzise in einer ganz bestimmten Position halten, sondern auch durch intuitive Bewegungen die Instrumente des Chirurgen im Endoskopbild zentrieren.

Die hierfür benötigte Position der Instrumentenspitze kann man u.a. mit Hilfe der Endoskopbilder gewinnen, vorausgesetzt das Instrument ist in ihnen sichtbar. Ein einfacher Ansatz besteht darin, das Instrument mit einem Marker auszustatten [1, 2, 3]. In jedem Bild wird dieser Marker gesucht, segmentiert und dessen Mittelpunkt bestimmt. Die Nachteile dieses Ansatzes sind die erhöhten Kosten, da jedes zu erkennende Instrument mit einem Marker auszustatten ist, eine wenig robuste Erkennung des Markers bei Verdeckungen oder Verschmutzungen und eine eventuell erhöhte Infektionsgefahr für den Patienten durch den Marker. Eine andere Vorgehensweise, die ohne Marker auskommt, beruht auf der Ausnutzung der Farbunterschiede zwischen Instrument und Hintergrund zur Segmentierung des Instrumentes [4, 5]. Unsere Untersuchungen haben ergeben,

dass die Güte dieser Verfahren sehr stark vom Material sowie der Farbe der Instrumente und des Hintergrunds abhängig ist. Ist das Instrument - wie in den meisten Fällen - metallisch, so kann es aufgrund seiner Reflexionseigenschaften die Farbe der näheren Umgebung oder einen gräulichen Ton annehmen. In beiden Fällen ist eine Segmentierung anhand der Farbinformation sehr schwierig. Ein Trackingalgorithmus, der weder einen speziellen Marker noch eine farbliche Unterscheidbarkeit benötigt, basiert auf der Feststellung, dass die bei endoskopischen Operationen verwendeten Instrumente typischerweise einen langen zylindrischen Schaft aufweisen [6]. Eine Seite dieses Schaftes versuchen die Autoren mit Hilfe einer Hough-Transformation für Geraden zu finden und suchen dann entlang der wahrscheinlichsten Gerade die Instrumentenspitze. Diese Grundidee erweitern wir, indem wir zur Erhöhung der Robustheit und Genauigkeit beide Seiten des Schaftes betrachten und zudem mit Hilfe eines Partikelfilters mehrere Lagehypothesen gleichzeitig über die Zeit verfolgen.

2 Methoden

Die Instrumente modellieren wir durch ein 3d-Zylindermodell mit bekanntem instrumentenspezifischen Radius. Zur Beschreibung der Instrumentenlage bzgl. des Kamerakoordinatensystems sind insgesamt fünf Parameter notwendig, die im Folgenden auch als Partikel (Lagehypothese) $\mathbf{s} = (x, y, z, \alpha, \beta)$ zusammengefasst werden. Die x-, y-, und z-Werte geben die Translation zwischen Brennpunkt der Kamera und Spitze des Instrumentes an. Die rotatorischen Freiheitsgrade sind durch den Winkel α für die Drehung des Zylindermodells um die z-Achse der Kamera und β für die Rotation um die y-Achse festgelegt. Aufgrund der Rotationssymmetrie kann der dritte rotatorische Freiheitsgrad unberücksichtigt bleiben.

Mit Hilfe dieses Modells haben wir ein neues Verfahren entwickelt, das aus einer Initialisierungsphase und einer Trackingphase besteht. Die Initialisierungsphase ist vor dem Tracking notwendig sowie wenn der Trackingalgorithmus das Instrument verliert. In dieser Phase ruht der Roboter und somit auch das Endoskop. Das Instrument kann nun leicht im Bild segmentiert werden, sobald es sich vor dem nahezu statischen Hintergrund bewegt. Sind gültige Lagehypothesen enthalten, so werden diese an die Trackingphase übergeben.

In der Trackingphase kommt der Condensation-Partikelfilter zum Einsatz [7]. Die Grundidee ist, den Zustand eines dynamischen Systems (in diesem Fall die sich über die Zeit ändernde Lage des Instrumentes) mit Hilfe einer allgemeinen Wahrscheinlichkeitsdichte zu beschreiben. Da dies analytisch schwer berechenbar ist, wird die Wahrscheinlichkeitsdichte mit Hilfe von gewichteten Partikeln diskretisiert. Für diese Partikel ist in jedem Zeitschritt (etwa 25 pro Sekunde) eine Neuverteilung, Vorhersage und eine Neugewichtung notwendig, weil die Wahrscheinlichkeitsdichte sich über die Zeit ändern kann. Falls der beste Partikel ein gewisses Mindestgewicht unterschreitet, ist im nächsten Zeitschritt die Initialisierungsphase erneut durchzuführen. Ansonsten wird die Position der Instrumentenspitze für die Robotersteuerung ermittelt.

Die Neuverteilung ist vergleichbar mit dem Evolutionsprinzip. Gute Partikel (mit einem hohen Gewicht) erzeugen viele Nachkommen und schlechte Partikel wenige bzw. gar keine. Die alten Partikel werden danach verworfen, um die Anzahl der Partikel in jedem Zeitschritt konstant (bei $N = 1000$) zu halten. Hierfür haben wir direkt den Algorithmus implementiert, der auch in [7] vorgeschlagen wurde. Für die Vorhersage der neuen Partikelpositionen haben wir ein einfaches Bewegungsmodell gewählt. Der j -te Partikel $\mathbf{s}_{j,t} = \mathbf{s}_{j,t-1} + \mathbf{w}_{j,t}$ in Zeitschritt t ergibt sich aus dem j -ten Partikel in Zeitschritt $t - 1$ und dem normalverteilten Zufallsvektor $\mathbf{w}_{j,t}$. Dabei sind in der Implementierung die Streuungen $(r, r, r, 5^\circ, 3^\circ)$ gewählt worden, wobei r dem Radius des Instrumentes entspricht. Nach der Vorhersage erfolgt jetzt eine Neugewichtung der Partikel mit Hilfe einer Qualitätsfunktion. Hierfür wird für jeden Partikel die Silhouette des Zylindermodells im Endoskopbild durch ein Viereck approximiert. Entlang des Vierecks wird an den M Messpunkten, die einen konstanten Abstand voneinander haben, die Übereinstimmung mit dem Endoskopbild mittels Gradient Direction Matching (GDM) [8] geprüft. Bei GDM bestimmt man den durchschnittlichen Fehlerwinkel $\bar{\theta}$ zwischen den Normalenvektoren der Silhouette und den Gradienten des Bildes an den jeweiligen Messpunkten. Mit Hilfe von Normalverteilungen kann nun die Qualitätsfunktion

$$p(\mathbf{z}|\mathbf{x}) = k \exp\left(-\frac{\bar{\theta}^2}{2\sigma_{\bar{\theta}}^2}\right) \exp\left(-\frac{\lambda^2}{2\sigma_{\lambda}^2}\right) \quad (1)$$

definiert werden, wobei k ein Normalisierungsfaktor ist und $\sigma_{\bar{\theta}} = 0.14$ sowie $\sigma_{\lambda} = 0.4$ experimentell bestimmt wurden. Wir haben zu dem eigentlichen GDM den Korrekturterm $\lambda = (M_{max} - M) / M_{max}$ hinzugefügt. Dieser ist notwendig, da aufgrund des einfachen Viereckmodells Lagehypothesen, die entlang der medialen Achse in Richtung Instrumentenende verschoben sind, die gleiche Gewichtung bekämen oder sogar präferiert würden.

Für die Positionsbestimmung könnte man einfach den translatorischen Teil $T(\mathbf{s}_j) = (x_j, y_j, z_j)$ des Partikels \mathbf{s}_j mit dem höchsten Gewicht π_j nehmen. Dies würde allerdings zu starken Fluktuationen der Positionswerte führen, weshalb wir eine Mittelung über alle gewichteten Partikel verwendet haben:

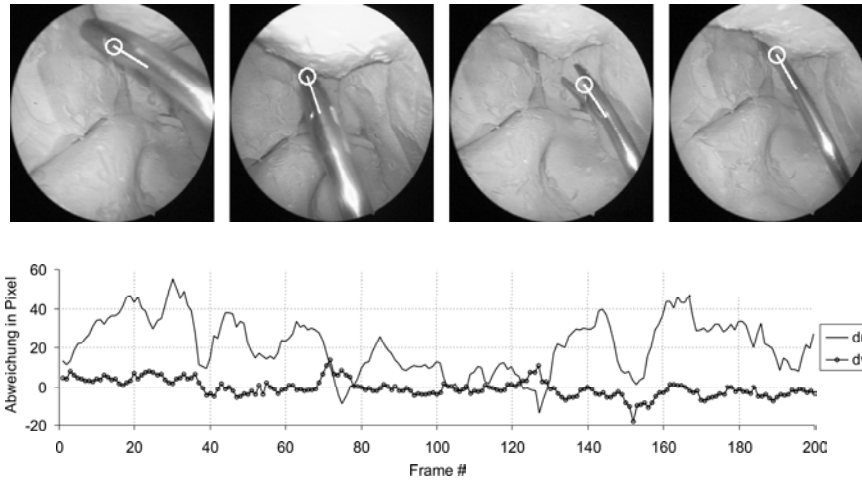
$$(x, y, z) = \frac{1}{\sum_{i=1}^N \pi_i} \sum_{i=1}^N \pi_i T(\mathbf{s}_i). \quad (2)$$

Dies lässt sich noch leicht und effizient erweitern, indem man mit Hilfe des k-means-Clustering die Partikel zuvor gruppiert. Dadurch kann man auch eine höhere Robustheit bei Verwendung mehrerer Instrumente erreichen.

3 Ergebnisse

Die Ergebnisse einer der verwendeten Videosequenzen sind in Abb. 1 dargestellt (siehe auch www.cs.tu-bs.de/rob/mri). Es hat sich gezeigt, dass die Genauigkeit

Abb. 1. Testvideo mit einer Auflösung von 360x288 Pixel und einer durchschnittlichen Breite des Instrumentes von 20 Pixel: (oben) Frame 20, 80, 140 und 200 von links nach rechts, (unten) Darstellung des Fehlers entlang (du) und orthogonal (dv) zur medialen Achse des Instrumentes in Pixel.



orthogonal zur medialen Achse des Instrumentes und die Orientierung sehr gut für die Robotersteuerung geeignet sind. Lediglich die Varianz des Fehlers entlang der medialen Achse ist noch etwas zu stark. Hier besteht jedoch die Möglichkeit in einem separaten Verfahren die Lösung entlang der Achse durch eine eindimensionale Suche noch entscheidend zu verbessern. In Tabelle 1 sind die Versuchsergebnisse am Roboter dargestellt (Durchmesser der Instrumente etwa 3.5 mm). Hier zeigt sich, dass die Tiefeninformation gegenüber den anderen Richtungen ungenau, aber noch durchaus brauchbar ist, und dass die Genauigkeit der Position aufgrund der Trägheit der Partikel abhängig von der Geschwindigkeit des Instrumentes im Bild ist. Die Nachzieheffekte sind allerdings durchaus erwünscht, damit die Bewegungen des Roboters nicht zu ruckartig werden. Die Analyse der Robustheit hat sehr positive Ergebnisse geliefert. Das Instrument konnte in Sequenzen mit teilweise über 2000 Bildern erfolgreich verfolgt werden. Nur bei sehr schnellen ruckartigen Bewegungen ist das Instrument komplett verloren gegangen, so dass eine erneute Initialisierungsphase nötig geworden ist.

4 Diskussion und Ausblick

Da die Instrumentenspitze nur ungefähr in der Bildmitte zu halten ist, bedarf es keiner sehr hohen Genauigkeit. Entscheidender ist, dass die Erkennung robust und stabil über lange Zeit hinweg erfolgt und ein Verlust des Instrumentes sofort erkannt wird. Das vorgestellte Verfahren liefert unter diesen Gesichtspunkten gute Ergebnisse, wobei die Genauigkeit entlang der medialen Achse noch zu

Tabelle 1. Ungenauigkeit der 3d-Position bzgl. des Koordinatensystems der Kamera in mm für verschiedene Geschwindigkeiten des Instrumentes.

	10 mm/s			40 mm/s			160 mm/s		
	dx	dy	dz	dx	dy	dz	dx	dy	dz
Maximum	5.98	5.53	13.28	17.19	15.20	15.03	17.51	19.03	26.92
Durchschnitt	1.78	1.04	8.46	3.39	2.87	9.05	4.45	3.58	10.23
Std.-Abweichung	1.13	0.92	2.02	3.65	2.92	2.73	4.36	3.47	3.94

verbessern ist. Das verwendete 3d-Zylindermodell ist eine gute Approximation der meisten Instrumente. In einigen Fällen könnte aber eine Modellierung mit Hilfe von B-Splines (z.B. für gekrümmte Instrumente) sinnvoller sein. Zudem ist die verwendete Positionsvorhersage noch sehr einfach und kann voraussichtlich mit Hilfe einer stochastischen a priori Analyse der Instrumentenbewegungen in realen Operationen verbessert werden. Ebenfalls Gegenstand weiterer Untersuchungen ist die Verwendung von Motion Flow zum Finden und Bewerten von Hypothesen. Insgesamt ist das Verfahren äußerst erfolgsversprechend und bietet hohes Potential für den klinischen Einsatz.

Literaturverzeichnis

1. Wei GQ, Arbter K, Hirzinger G. Automatic Tracking of Laparoscopic Instruments by Color Coding. *Lecture Notes in Computer Science* 1997;1205:357–366.
2. Zhang X, Payandeh S. Application of Visual Tracking for Robot-Assisted Laparoscopic Surgery. *Journal of Robotic Systems* 2002;19(7):315–328.
3. Matthes J, Gröll L, Jäkel J. Optimale Führung von Endoskopen mit redundanter Kinematik. *Automatisierungstechnik* 2001;49(4):168–176.
4. Boisvert J, Cheriet F, Grimard G. Segmentation of Laparoscopic Images for Computer Assisted Surgery. In: *13th Scandinavian Conference on Image Analysis*; 2003. p. 587–594.
5. Uecker DR, Lee C, Wang YF, Wang Y. Automated Instrument Tracking in Robotically-Assisted Laparoscopic Surgery. *Journal of Image Guided Surgery* 308-325;1(6):1995.
6. Climent J, Marés P. Automatic Instrument Localization in Laparoscopic Surgery. *Electronic Letters on Computer Vision and Image Analysis* 2004;4(1):21–31.
7. Isard M, Blake A. CONDENSATION – Conditional Density Propagation for Visual Tracking. *International Journal of Computer Vision* 1998;29(1):5–28.
8. Lichtenauer J, Reinders MJT, Hendriks EA. Influence of the Observation Likelihood Function on Particle Filtering Performance in Tracking Applications. In: *International Conf. on Automatic Face and Gesture Recognition*; 2004. p. 767–772.