

# Statistikbasierter Navigationstunnel für die roboterassistierte Sinus-Chirurgie

J. Kluth<sup>1</sup>, M. Rilk<sup>1</sup>, R. Westphal<sup>1</sup>, F. Wahl<sup>1</sup>, F. Bootz<sup>2</sup>, K.W. Eichhorn<sup>2</sup>

<sup>1</sup> TU Braunschweig, Institut für Robotik und Prozessinformatik, Germany

<sup>2</sup> Universität Bonn, Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/Chirurgie, Germany

Kontakt: jkl@rob.cs.tu-bs.de

## Abstract:

In diesem Artikel wird eine Möglichkeit der Verwendung statistischer Daten für einen Bahnplanungsalgorithmus in der roboterassistierten Nasenendoskopie beschrieben. Das Ziel einer solchen Roboterassistenz ist die Unterstützung des Operateurs, indem der Roboter das Endoskop als ‚dritte Hand‘ führt. Dabei muss der Roboter den Bewegungen des menschgeführten Instruments in intuitiver Weise folgen und gleichzeitig Risikobereiche im geometrisch hochkomplexen Nasenraum meiden.

Der hier beschriebene Ansatz für die benötigte Bahnplanung benutzt multivariate Normalverteilungen, um vorher aufgezeichnete statistische Daten komprimiert zu repräsentieren. Das Ergebnis ist ein Navigationstunnel, der den Frei-raum beschreibt. Mithilfe eines Optimierungsverfahrens wird dieser Navigationstunnel dann für eine Bahnplanung genutzt.

In der Evaluation wird die Methode in einer Simulation getestet und qualitativ mit einer existierenden Navigationslösung verglichen.

Schlüsselworte: Robotik, Nasenendoskopie, FESS, Bahnplanung, Statistik, Navigation, multivariate Normalverteilung

## 1 Problem

In der roboterassistierten Endoskopie der Nasenhöhlen (FESS) unterstützen Roboter den Operateur durch eine (semi-) automatische Führung des Endoskops. Die Nutzung eines Roboters für diese Aufgabe ermöglicht es dem Operateur, beide Hände für die Arbeit mit Instrumenten zu verwenden, anstatt eine Hand mit der Endoskopführung zu blockieren. Die optimale Führung eines Endoskops ist ein komplexes Problem, bei dem zahlreiche Randbedingungen zu berücksichtigen sind. So wird das Endoskop in einem anatomischen Bereich mit einer hochkomplexen geometrischen Struktur bewegt. Dieser Bereich ist umgeben von Risikostrukturen deren Verletzung zum Teil lebensbedrohliche Komplikationen bedeuten können. Die Einhaltung der Arbeitsraumgrenzen und das Vermeiden hoher Kontaktkräfte sind daher zwingend erforderlich und muss in der Bahnplanung berücksichtigt werden. Unter Einhaltung dieser Randbedingungen muss das Endoskop durch den Roboter jederzeit so positioniert werden, dass der Operateur zum einen eine optimale Sicht auf die Instrumente hat und zum anderen ein maximaler Bewegungsfreiraum für die Instrumente verbleibt.

Für diese Anforderungen bietet sich eine geschwindigkeitsbasierte Robotersteuerung [1] an. Dabei werden für jeden Bahnpunkt neue Geschwindigkeiten für den Roboter generiert. Hierfür können keine klassischen Bahnplanungsalgorithmen angewendet werden, da der jeweils nächste Schritt nicht nur vom Bahnplanungsalgorithmus beeinflusst wird, sondern auch anderweitig beschränkt wird. Beispielsweise durch Kollisionstests in vorhandenen CT-Daten oder durch einen Kraft-Momenten-Sensor, der in Kontaktsituationen die maximal erlaubte Geschwindigkeit begrenzt. Weiterhin ist das Ziel der Bahn beweglich, da das anvisierte Instrument ständig vom menschlichen Operateur bewegt wird.

Diese Einschränkungen zeigen, dass es nicht möglich ist, eine vollständige Trajektorie im Vorhinein zu berechnen. Stattdessen wird eine Navigationsfunktion  $f$  definiert, die jeweils den gewünschten Geschwindigkeitsvektor  $v$  für eine gegebene Zielpose  $target$  und eine aktuelle Pose  $x$  liefert:  $v = f(x, target)$

Das Ergebnis  $v$  wird zusammen mit den o.g. Beschränkungen zu einer endgültigen Geschwindigkeit verrechnet. In jedem Zeitschritt bewegt sich der Roboter damit in Richtung Zielpose unter Einhaltung von schützenden Bedingungen.

In dieser Veröffentlichung werden statistische Daten für die Konstruktion der Funktion  $f$  benutzt. Diese können durch einen Anwender des Systems bereitgestellt werden, indem das Endoskop manuell innerhalb der Nase bewegt wird, während ein Tracking-System die Position und Orientierung des Endoskops kontinuierlich aufzeichnet. Für diesen Lernprozess können einerseits Aufzeichnungen einer kompletten Operation und andererseits Aufzeichnungen eines bestimmten Vorgangs verwendet werden. Diese Art des Lernansatzes ist in der Literatur auch bekannt als „Programming by Demonstration“ [2].

## 2 Methoden

Ziel ist es, einen Raum zu konstruieren, in dem es möglich ist, von einem gegebenen Punkt zu einem anderen navigieren zu können. Um diesen Raum zu beschreiben, wird eine tunnelähnliche Struktur benutzt. Die aufgezeichneten Daten von den manuellen Endoskopbewegungen werden verwendet, um die Größe des Tunnels in allen Dimensionen zu bestimmen. Die Bahnplanung soll nicht direkt auf den aufgezeichneten Daten erfolgen, da mit wachsender Datenmenge die Verarbeitungszeit zu stark ansteigen würde.

Stattdessen wird ein auf multivariaten Normalverteilungen basierendes Modell für die Komprimierung der Rohdaten benutzt. Diese, auch „Gaussians“ genannten, Normalverteilungen beschreiben die Wahrscheinlichkeitsdichteverteilung von Sample-Punkten. Mit einer solchen Verteilung können aufgezeichneten Sample-Punkte in einer kompakten Form repräsentiert werden. Ein Gaussian besteht aus einem Vektor, der die Mittelwerte der Daten in sämtlichen Dimensionen enthält, und aus einer Kovarianzmatrix, die die Abweichungen der Daten von den Mittelwerten beschreibt. Die Idee, Gaussians zu benutzen, um eine Art Tunnel zu konstruieren, ist von der Verwendung von sogenannten Gaussian Mixture Models inspiriert, die bereits in der Vergangenheit bei „Programming by Demonstration“ verwendet wurden [3].

### Aufbau

Es wird davon ausgegangen, dass eine Menge von 6D-Endoskop-Posen existiert, die jeweils aus drei Koordinaten für die Position und drei Werten für die Orientierung (RPY-Notation) bestehen. Für die Navigation ist die Orientierung um die Endoskop-Achse allerdings nicht relevant, da sie nur die Ausrichtung des Bildes bestimmt. Daher kann die entsprechende Komponente entfernt werden, d.h. es müssen nur noch fünf Freiheitsgrade repräsentiert werden. Der entstehende Raum ähnelt einem Vektorfeld mit den Endoskopblickrichtungen als Vektoren an den Positionskomponenten.

Durch eine stereografische Projektion werden die beiden rotatorischen Komponenten auf ein Maß abgebildet, das sich besser für die Distanzberechnung eignet. Dabei wird ein Parameter für den Projektionskugelradius festgelegt, der den Einfluss der Rotation auf das Distanzmaß steuert; empirisch wurde ein Radius von 20mm ermittelt.

Nach der Konvertierung aller Poses in diesen 5D-Raum wird die Hauptachse in den aufgezeichneten Daten ermittelt, um entlang dieser Achse den Tunnel auszurichten. Hierfür wird eine passende homogene Transformation gesucht, die das ursprüngliche Koordinatensystem so rotiert, dass die neue x-Achse in Richtung der Hauptachse zeigt. In dem beschriebenen experimentellen Aufbau erfolgte dies zunächst manuell, aber automatische Lösungen sind denkbar. Eine normale Hauptkomponentenanalyse (PCA) empfiehlt sich nicht, da hier die Freiheitsgrade für Position und Rotation vermischt werden, was eine spätere visuelle Analyse der Ergebnisse erschwert. Eine PCA lediglich auf den Positionsdaten wäre eine mögliche Alternative für die Zukunft, jedoch ist der Einfluss der Krümmung des Freiraums zu prüfen.

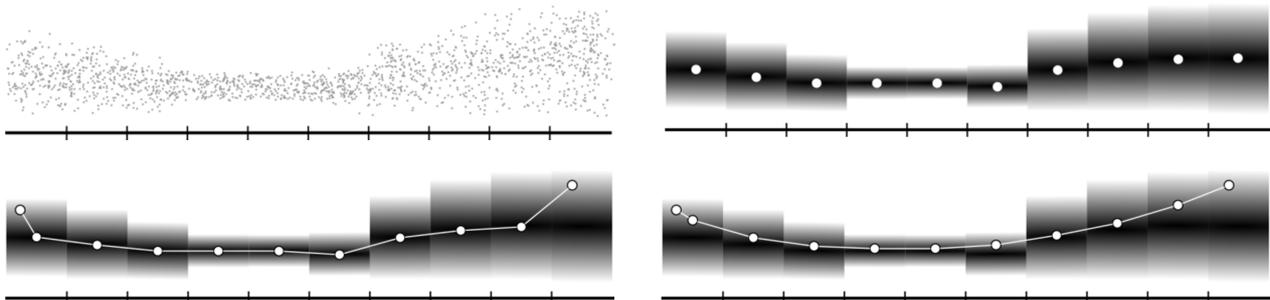
Im nächsten Schritt wird der Bereich auf der neuen x-Achse in gleich große Zellen unterteilt. In jeder dieser Zellen wird eine multivariate Normalverteilung erstellt, die die Verteilung der verbleibenden vier Komponenten beschreibt. Dafür könnte man lediglich genau die Datenpunkte verwenden die innerhalb der Zelle liegen. Um aber eine Filterung der Datenpunkte zu erreichen und einen „glatteren“ Tunnel zu bekommen, werden sämtliche Datenpunkte benutzt, die jeweils unterschiedlich gewichtet werden. Für die Gewichtung wird eine Normalverteilung verwendet, diesmal eindimensional. Abhängig von der ersten Komponente werden alle Datenpunkte für jede Zelle einzeln gewichtet.

Das Resultat ist eine Menge von Tunnelsegmenten, die jeweils durch einen Abschnitt der Hauptkomponente (x-Achse) und einer vierdimensionalen Normalverteilung, basierend auf den verbleibenden vier Komponenten, beschrieben sind. Der Aufbau des Tunnels geschieht offline vor der eigentlichen Nutzung und dauert lediglich wenige Sekunden auf einem normalen Büro-PC. Der nächste Abschnitt ist der Nutzung dieser Datenstruktur gewidmet.

### Nutzung

Nach der Offline-Erstellung des Tunnels, kann dieser für eine Online-Navigation benutzt werden. Der Tunnel repräsentiert den Raum, in dem sich das Endoskop bewegen darf. Mithilfe der Segmente können Stützpunkte generiert und ein Pfad von einer Start- zu einer Zielpose geplant werden.

Die einzelnen Stützpunkte werden mit dem Mittelwert der Normalverteilung eines jeden Segments initialisiert und anschließend optimiert. Die Optimierung hat zum Ziel, die „Reisezeit“ von Startpose zur Zielpose zu minimieren, wäh-



**Abb. 1: Schematische Darstellung:** links oben: Menge an Datenpunkten; rechts oben: Berechnete Normalverteilungen mit Mittelwerten; links unten: Initialer Pfad mit Stützpunkten; rechts unten: Finaler Pfad mit Stützpunkten

rend gleichzeitig sichergestellt wird, dass der Verlauf des Pfades innerhalb des Tunnels bleibt. Dies wird erreicht, indem die Größe der Optimierungsschritte kleiner wird, umso weiter die Stützpunkte vom ursprünglichen Mittelwert entfernt werden. Dadurch können die Stützpunkte sich nicht beliebig weit von den Mittelwerten entfernen.

Die Reisezeit bestimmt sich aus dem Abstand der Stützpunkte und der erlaubten Geschwindigkeit. Letztere bestimmt sich aus dem Abstand der Stützpunkte zum Mittelwert. Dieses Vorgehen nimmt an, dass die Mittelwerte der Tunnelsegmente der sichersten Position entsprechen und dadurch dort die höchsten Geschwindigkeiten erlaubt sind. Bei zunehmendem Abstand zum Mittelwert wird angenommen, dass die Chance, auf unerwartete Hindernisse zu stoßen, größer ist und dementsprechend langsamer gefahren werden muss. Das übergeordnete Steuerungssystem für den Roboter wird beispielsweise die Geschwindigkeit reduzieren, sobald bekannte Hindernisse in der Nähe des Endoskops sind oder unerwartete hohe Kräfte und Drehmomente von den Sensoren erfasst werden. Daher ist es von der Geschwindigkeit her am besten sich nahe den Mittelwerten zu bewegen.

In Abb. 1 ist der Prozess schematisch dargestellt. Das erste Bild zeigt eine beispielhafte Datenmenge aus zweidimensionalen Datenpunkten. Im nächsten Bild wurden die Normalverteilungen erzeugt. Es ist zu beachten, dass diese Normalverteilungen hier lediglich eindimensional sind, da die erste Dimension zur Einteilung der Segmente verwendet wurde. Die weißen Punkte repräsentieren jeweils den Mittelwert der Normalverteilung, während die Verlaufstreifen die Wahrscheinlichkeitsdichte symbolisieren. Ein dunkler Farbton kennzeichnet eine höhere Wahrscheinlichkeitsdichte. Im dritten Bild werden die Mittelwerte als Stützpunkte zwischen einem Startpunkt und einem Zielpunkt benutzt. Das letzte Bild stellt die Verschiebung der Stützpunkte im Zuge der Optimierung dar und repräsentiert den finalen Pfad.

In der Praxis ist es nicht wünschenswert, darauf warten zu müssen, dass die Optimierung vollständig abgeschlossen ist. Stattdessen kann direkt damit begonnen werden, den Pfad abzufahren, da die initiale Lösung auf jeden Fall eine sichere Lösung ist. Dadurch kann eine gute Reaktionszeit des Systems erreicht werden, während trotzdem noch ein nahezu optimaler Pfad abgefahren wird.

Im Folgenden soll der entwickelte Lösungsansatz formal beschrieben werden: Die Distanz zwischen zwei Posen ist die

Euklidische Distanz zwischen zwei Vektoren:  $\text{dist}(x, y) = \sqrt{\sum_{i=1}^5 (x_i - y_i)^2}$

Dabei sind  $x$  und  $y$  zwei 5D-Posen. Zusätzlich wird eine Geschwindigkeitsschätzung benötigt. Dabei handelt es sich nicht um eine reale Geschwindigkeitsangabe, sondern um eine Relation zu einer maximal möglichen Geschwindigkeit. Daher wird die relative Wahrscheinlichkeitsdichtefunktion der Normalverteilungen benutzt, um eine Pseudogeschwindigkeit zu berechnen:  $\text{vel}(x, \text{seg}) = \frac{P(x; \text{seg})}{P(\text{mean}_s; \text{seg})}$

Dabei ist  $x$  wiederum eine Pose und  $\text{seg}$  ein Segment des Tunnels.  $P(x; \text{seg})$  sei die Wahrscheinlichkeitsdichtefunktion an der Stelle  $x$  der zugehörigen multivariaten Normalverteilung von  $\text{seg}$ .  $\text{mean}_s$  sei der entsprechende Mittelwert.

Mit diesen beiden Teilen kann die Reisezeit anhand der klassischen Formel für Geschwindigkeit formal beschreiben werden:  $v = \frac{s}{t}$ , also  $t = \frac{1}{v} s$ . Dabei ist zu beachten, dass es sich hierbei nur um eine Schätzung handelt, da in der Realität noch andere Faktoren wie z.B. die Beschleunigung hinzukommen. Als Resultat erhält man:

$$\text{time}(x, y) = \frac{1}{\frac{1}{2} \cdot \text{vel}(x, \text{seg}_x) + \frac{1}{2} \cdot \text{vel}(y, \text{seg}_y)} * \text{dist}(x, y)$$

Dabei sind  $\text{seg}_x$  und  $\text{seg}_y$  die Segmente, in denen jeweils  $x$  bzw.  $y$  liegt. Hier wird eine Mischung aus beiden Geschwindigkeitskomponenten benutzt, da die betrachtete Strecke jeweils zur Hälfte in den betroffenen Segmenten liegt.

Für die Gesamtstrecke wird eine Summe über die Terme vom Start- bis zur Zielpose gebildet:

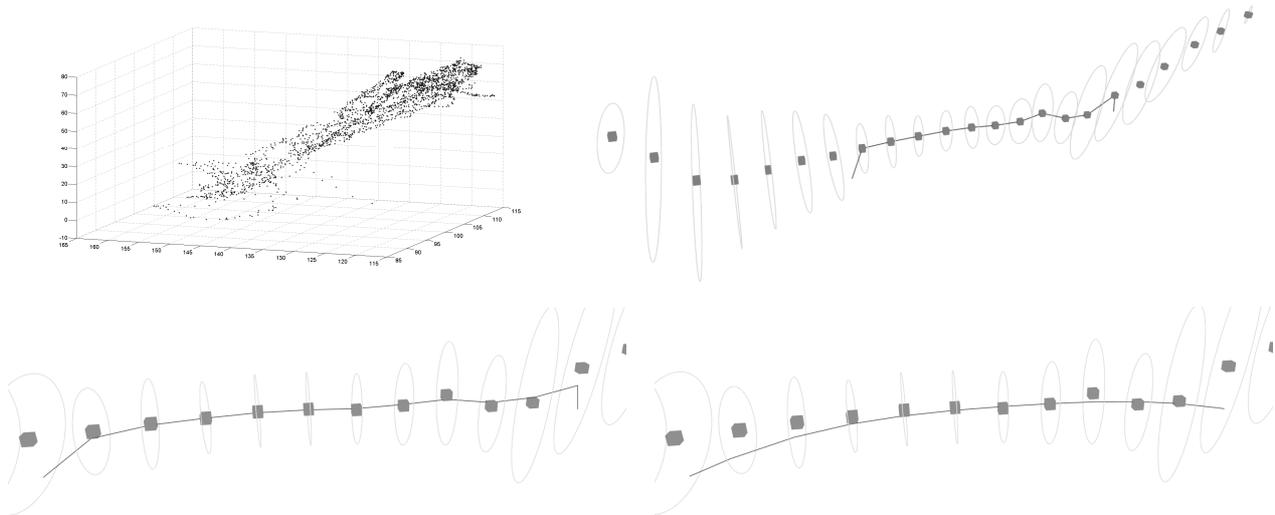
$$\text{totalTime} = \text{time}(\text{start}, x_1) + \text{time}(x_1, x_2) + \dots + \text{time}(x_{n-1}, x_n) + \text{time}(x_n, \text{target})$$

Das Ziel ist es, diese Summe durch Änderungen der  $x_i$  zu minimieren. Mithilfe der dadurch bestimmten Bahn kann mit der oben erwähnten Navigationsfunktion  $f$  eine neue Geschwindigkeit für den nächsten Zeitschritt errechnet werden.

### 3 Ergebnisse

Für eine Testdatenmenge wurde ein Endoskop etwa fünf Minuten von einem Nicht-Mediziner in einem Nasenmodell bewegt und die Posen aufgezeichnet. Ein erfahrener Operateur würde das Endoskop wahrscheinlich auf eine andere Art und Weise bewegen, was in einen anderen Tunnel resultieren würde. Die erste Grafik in Abb. 2 zeigt die aufgenommenen Rohdaten (nur Positionsanteil), die ca. 2700 5D-Datenpunkte umfassen. Eine offensichtliche Hauptrichtung ist sichtbar, und wurde für die folgende Verarbeitung manuell festgelegt. In dem Datensatz befindet sich der Eingang des Nasenmodells im linken unteren Bereich. Sichtbar ist, dass außerhalb des Nasenraums eine hohe Varianz der Daten vorhanden ist, gefolgt von einem engeren Bereich, der dann wieder in einem Bereich höherer Varianz mündet.

Die Abbildung rechts oben zeigt die initiale Lösung ohne Optimierung in einer Simulationsumgebung. Dargestellt ist eine Projektion der Bahn in eine 2D-Ebene. In der unteren Zeile sind zwei vergrößerte Ausschnitte des Tunnels zu sehen, die jeweils verschiedene Optimierungszustände darstellen. Man kann sehen, dass die Bahn zunehmend kürzer wird, aber innerhalb der Tunnelgrenzen verbleibt.



**Abb. 2: Daten in Simulationsumgebung:** links oben: Rohdaten; rechts oben: Gesamter Tunnel, initialer Pfad; links unten: Ausschnitt des Tunnels, während der Optimierung; rechts unten: nach der Optimierung. Legende: Kovarianzen der Tunnelsegmente jeweils als Ellipsen, Mittelwerte als Quader, Linie als geplante Bahn

Während der Optimierung werden also Schritt für Schritt die Stützpunkte verschoben bis ein Gleichgewicht zwischen erreichbarer Geschwindigkeit und günstiger Lage entsteht. Die Bahn stellt jeweils die Lage der Endoskopspitze dar. Die Orientierung des Endoskops wird ebenfalls mitgesteuert, ist hier aber nicht dargestellt.

Als Zielpose können beliebige Datenquellen dienen. Dies ist zum Beispiel eine gute Beobachtungspose für die aktuelle Instrumentenpose, die man über ein Trackingsystem erhalten kann. Weiterhin kann auch manuell eine Zielpose vorgegeben werden, um das Endoskop ohne eine Instrument im Naseninnenraum führen zu können.

## 4 Diskussion

Für den beschriebenen Lösungsvorschlag wird angenommen, dass der Operateur das Endoskop die meiste Zeit im Zentrum des Tunnels benutzt. Wenn das Instrument vom Zentrum des Tunnels fortbewegt wird, folgt das Endoskop erst einer Bewegung in Richtung der Tunnelmitte, bevor es die endgültige Pose ansteuert. Dieses Verhalten kann zu unerwarteten Bewegungen führen. Dieses Verhalten zeigt ein simplerer Ansatz, der aktuell ebenfalls für das erwähnte Nasendoskopie-Robotersystem implementiert ist, nicht. Dieser simplere Ansatz versucht dauerhaft eine festgelegte Lage für das Endoskop bezüglich des Instruments anzusteuern. Nachteil dabei ist, dass das Verfahren zwingend ein Instrument voraussetzt und damit keine alleinige Navigation des Endoskops ermöglicht. Weiterhin sind keine komplexen Bahnen, wie z.B. Naseneinfahrt bei bereits eingeführtem Instrument möglich. Mit der in diesem Beitrag vorgestellten Methode ist beides möglich.

Im Rahmen unserer zukünftigen Arbeiten muss ein invarianter Raum für die Beschreibung des Naseninnenraums geschaffen werden, damit dieses Verfahren sinnvoll in der Praxis eingesetzt werden kann. Aktuell müssten für jeden Patienten manuell Trainingsdaten erzeugt werden, da die statistischen Daten jeweils nur genau für einen Nasenraum passend sind. Mit einer invarianten Beschreibung der Daten und einem automatisierten Verfahren zur Übertragung dieser Daten auf einen konkreten Patienten mittels CT/MRT-Daten kann der vorgestellte Ansatz praktische Relevanz erreichen. Für den Patienten bieten sich durch den Einsatz eines endoskopführenden Roboters mit einer intelligenten – wie der hier vorgestellten – Bewegungsplanung erhebliche Vorteile: Das Infektionsrisiko kann durch eine geringe Anzahl an Instrumentenwechseln reduziert werden, die Operationszeit kann verkürzt werden und kritische Grenzstrukturen wie etwa zur Schädelbasis können zuverlässig respektiert werden.

## 5 Referenzen

- [1] M. Rilk, F. M. Wahl, K. W. G. Eichhorn, I. Wagner und F. Bootz: Path Planning for Robot-Guided Endoscopes in Deformable Environments, Advances in Robotics Research, Springer Berlin Heidelberg, 2009
- [2] A. Billard, S. Calinon: Robot Programming by Demonstration, Springer Handbook of Robotics, Chapter 59, Springer, 2008
- [3] S. Calinon, F. Guenter, A. Billard: On Learning, Representing, and Generalizing a Task in a Humanoid Robot, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2007