

## Bewegungsanalyse für fiberskopische Superposition\*

Winter C<sup>2</sup>, Rupp S<sup>1</sup>, Wittenberg T<sup>1</sup>

<sup>1</sup>Abteilung für Bildverarbeitung und Medizintechnik, Fraunhofer-Institut für Integrierte Schaltungen, Deutschland

<sup>2</sup>Lehrstuhl für Informations- und Kommunikationstechnik, Friedrich-Alexander Universität Erlangen-Nürnberg, Deutschland

wnt@like.e-technik.uni-erlangen.de

\* Diese Arbeit wurde im Rahmen des Sonderforschungsbereichs 603 (TP A7) und dem Graduiertenkolleg 244 von der Deutschen Forschungsgemeinschaft gefördert.

**Einleitung und Fragestellung** Der Einsatz faseroptischer Abbildungssysteme ermöglicht eine im wörtlichen Sinne "flexible" und minimal-invasive Begutachtung schwer zugänglicher Hohlräume. Insbesondere in der medizinischen Diagnostik, Behandlung und Forschung spielen deshalb biegsame Endoskope (sog. Fiberskope) eine tragende Rolle. Trotz ihrer Vorzüge stehen sie linsenoptischen, starren Endoskopen in der Qualität der übertragenen Bilder deutlich nach. Gerade hochwertige Fiberskope mit Bildleitern aus homogen angeordneten Glasfasern zeichnen sich konstruktionsbedingt durch eine relativ geringe Anzahl an Einzelfasern aus (Größenordnung: 3k – 15k), die zudem durch die Ummantelung eine kontraststarke Wabenstruktur aufweisen und daher einen adäquaten Einsatz für die digitale Weiterverarbeitung verhindern.

Neben gängigen passiven Methoden zur Glättung von gerasterten Bilddaten [1] tragen erste Forschungsarbeiten [2] dazu bei, Verfahren zur aktiven Steigerung von Qualität und Auflösung bewegter Bildsequenzen durch Überlagerung (*Superposition*) für den Anwendungsbereich der flexiblen Endoskopie zu nutzen. Dies ähnelt dem Verhalten des menschlichen Auges beim Arbeiten mit Fiberskopen, nämlich kleine Seitwärtsbewegungen zu nutzen, um einen detailreicheren Eindruck zu gewinnen.

Entscheidend für den Detail- und damit Informationsgewinn durch *Superposition* ist die Präzision der ermittelten Bewegungen (*Tracking*) innerhalb der Bildsequenz. Die vorliegende Arbeit beschreibt die Methode und Randbedingungen von *Superposition* im Anwendungsbereich fiberskopischer Systeme und stellt eine Methode zur systematischen Evaluierung dar, welche die Anwendbarkeit quantitativ belegt.

**Material und Methoden** Wir gliedern den Prozess einer aktiven Auflösungssteigerung im Ortsraum für die Fiberskopie in fünf Abschnitte. (1) Die Bildgebung: Eine Szene wird von der distalen Optik eines Fiberskops erfasst und über den flexiblen Bildleiter auf eine am Okular montierte Industriekamera abgebildet, wo sie die beschriebene wabenförmig Struktur aufweist. (2) Die Wabenreduktion: Als Vorbereitung für Schritt (3) müssen verdeckte Bildbereiche aus der Helligkeitsverteilung der umliegenden Faserzentren (siehe Abb. 2) rekonstruiert werden, um den Bewegungsdetektor nicht durch künstlich harte Kanten im Bild negativ zu beeinflussen. Dies kann spektral im Frequenzraum [3] oder örtlich, beispielsweise barizentrisch erfolgen, wobei jeweils die Auflösung informationstechnisch nicht beeinflusst wird. (3) Die Bewegungsdetektion: Aus der Sequenz wird ein Basisbild gewählt, in dem durch ein geeignetes Verfahren Merkmale in definiertem Mindestabstand  $d_{min}$  detektiert werden. Aufgabe eines *Trackers* ist es nun, die Verschiebungen dieser Merkmale in benachbarten Bildern Subfaser-genau zu schätzen. (4) Die *Superposition*: Ein kontinuierliches Gitter, das an sog. Stützstellen die Intensitäten der Faserzentren des Basisbilds enthält werden nun an Stellen sicher georteter Verschiebung die Intensitäten der nächstgelegenen Faserzentren in den korrespondierenden Aufnahmen ergänzt. Der Versatz zu den bereits enthaltenen Stützstellen wird aus den invertierten Bewegungen bestimmt. (5) Die Nachbearbeitung: Das Basisbild kann nun in höherer Auflösung aus dem ergänzten Gitter kartesisch abgetastet werden. Auf die Erläuterung weiterer Schritte zur abschließenden Bildaufbereitung wird hier verzichtet.

Um die Genauigkeit und damit die Eignung der erarbeiteten Algorithmen in Schritt (3) systematisch zu evaluieren, werden für das Bewegungsmodell der abgebildeten Szene Referenzdaten (*Ground-Truth*) benötigt. Selbst kalibrierte endoskopische Aufnahmesysteme bieten für die vorliegende Anforderung jedoch keine ausreichende Genauigkeit, um aus realen Szenenbewegungen metrische *Ground-Truth*-Daten ableiten zu können. Deshalb werden die Schritte (1) bis (2) durch mehrere Syntheseschritte ersetzt. Dadurch ist es möglich, Szenenabbildungen durch ein dediziertes Bewegungsmodell zu transformieren, bevor sie als Ersatz für die fiberskopische Abbildung wabenförmig degradiert und anschließend mit Schritt (2) gefiltert oder interpoliert werden.

Unabhängig vom eingesetzten Verfahren zur Bewegungsdetektion liegen nun für jedes Bildpaar Soll- und Ist-Werte für die Verschiebungen vor. Es kann gezeigt werden, dass eine ermittelte Verschiebung innerhalb eines Bildpaares genau dann einen Informationsgewinn bewirkt, falls dieser geringer ist, als der mittleren Abstands zweier Fasern. Als Maß für die Güte der Bewegungsdetektion wird daher eine Statistik über die Abweichungen der detektierten Verschiebung zu den jeweiligen Referenzen des Bewegungsmodells geführt und ausgewertet. Zur Visualisierung der Ergebnisse dient ein Diagramm, das für jede Referenzverschiebung die zwei Dimensionen für Mittelwert und Varianz der ermittelten realen Bewegungen durch eine Ellipse veranschaulicht. Der Auflösungsgewinn im resultierenden Bild nach Abschluss von Schritt (5) kann im Fall der Synthese (Modellierte Degradierung und Interpolation) durch Vergleich mit dem Originalbild über gängige Differenzmaße quantifiziert werden.

**Ergebnisse** Für die Auswertung des Verfahrens werden verschiedene Bildsequenzen aus industriellem und medizinischem Umfeld verwendet. Am eindrücklichsten für das menschliche Auge sind Bildausschnitte mit lesbarer Schrift oder ähnlichen Strukturen. Ein reales Videobild wird vor der fiberskopischen Degradierung und Interpolation mäanderförmig an 24 verschiedene Positionen verschoben. Bei diesem trivialen Bewegungsmodell korrelieren alle Bewegungen innerhalb eines Bildes mit der gleichen Referenz-Verschiebung und häufen sich daher um die mit 'x' gekennzeichneten Positionen in Abb. 1. Die Bewegungsanalyse wird für die Auswertung im Diagramm o.B.d.A. mit dem Pyramiden-basierten Ansatz von Kanade-Lucas durchgeführt. Dabei spielen neben Gütekriterien der relevanten Merkmale auch deren Maximalanzahl eine Rolle, die hier auf 5000 gesetzt ist, sowie der Mindestabstands  $d_{min}$  für neu lokalisierte Merkmale, der typischer Weise in der Größenordnung des mittleren Abstands  $d_f$  zweier Fasern in der Referenzabbildung gewählt wird. Mittelwert und Standardabweichung der detektierten Bewegungen bilden für jede Referenz-Verschiebung die Kenngrößen einer Ellipse, die eine Abschätzung der Zuverlässigkeit vermittelt. Liegt eine Abweichung innerhalb der Größenordnung des Kreises im Nullpunkt der Grafik mit dem Radius  $d_f/2$ , so trägt er bei der *Superposition* zum Informationsgewinn bei (*Inlier*). Der obere Wert jedes Zahlentupels steht für die Anzahl dieser *Inlier*, der untere für die Gesamtanzahl der gültigen Verschiebungsvektoren. Der relativ niedrige Wert von durchschnittlich 3500 Merkmalen kommt von der spezifischen Vorselektion von Merkmalen auf Grund der Apertur in endoskopischen Aufnahmen. Der Prozentsatz informationstheoretisch gewinnbringender Merkmale liegt im veranschaulichten Fall bei 96.8%.

Die positiven Werte der quantitativen Auswertung werden durch visuelle Ergebnisse in Abb. 2 und 3 bestätigt. Hier sind stark vergrößerte Ausschnitte zweier realer Aufnahmen zu sehen, die mit dem vorgestellten Verfahren aufbereitet wurden. Jeweils links ist der Ausschnitt des Basisbilds aus der Originalsequenz zu sehen. Daneben schließt sich das degradierte und interpolierte Einzelbild an. Abb. 2 zeigt als drittes Bild das superpolierte Ergebnis unter Verwendung von 8 Nachbarbildern und rechts das Ergebnis unter Verwendung von 24 Nachbarbildern. Zu letzterem zeigt Abb. 1 die Visualisierung der Bewegungsanalyse. Das superpolierte Ergebnis rechts in Abb. 3

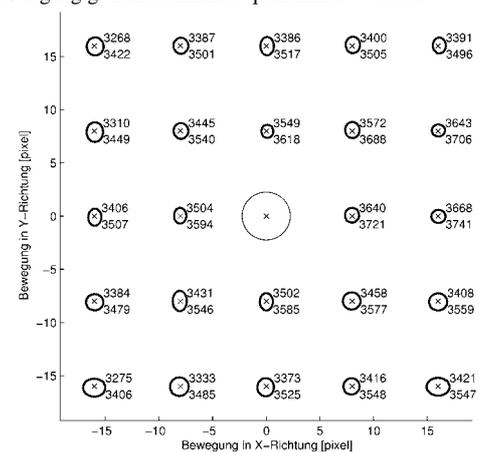


Abb. 1 Visualisierung einer Bewegungsanalyse



Abb. 2 Vergrößerter Ausschnitt einer faseroptischen Aufnahme (v.l.n.r.): Weiß-Bild, Interpolation einer einzelnen Aufnahme, Superposition mit 8 Bildern, Superposition mit 24 Bildern



Abb. 3 Vergrößerter Ausschnitt einer faseroptischen Aufnahme (v.l.n.r.): Original, Interpolation einer einzelnen Aufnahme, Superposition mit 15 Bildern

entsteht durch Überlagerung von 15 Einzelaufnahmen zum Basisbild.

**Diskussion** Die Auswertung der anwendungsspezifisch definierten Gütemaße bestätigt die prinzipielle Anwendbarkeit der *Superposition* für faseroptische Abbildungssysteme und gibt die Möglichkeit einer automatisierten Beurteilung des *Tracking*-Verfahrens und der Optimierung dessen Parameter. Es kann quantitativ untersucht werden, wie erfolgreich verschiedene *Tracking*-Algorithmen auf Bildmaterial fiberskopischer Abbildungssysteme eingesetzt werden können, welche Parameter zu einer zuverlässigen Weiterverfolgung von Merkmalen über bestimmte Distanzen führen und welchen Nutzen ihre Verwendung im Rahmen der *Superposition* darstellt. Erste visuelle Ergebnisse aus realen und synthetisierten Bilddaten belegen ebenso eine deutliche Zunahme von Details in strukturierten Bildbereichen, beispielsweise in Abb. 2 der Kantenbereiche oder in Abb. 3 der lesbaren Schrift, was sich bei synthetischer Bildgebung wiederum im Vergleich mit dem Originalbild über gängige Differenzmaße analysieren lässt.

Eine derartige Auflösungssteigerung eröffnet für die rechnergestützte Weiterverarbeitung von fiberskopischem Bildmaterial neue interessante Anwendungsgebiete, die bislang wegen ihrer hohen Anforderung an die Auflösung bislang starren Endoskopen vorenthalten war und insbesondere für bestehende faseroptische Systeme einen interessanten Mehrwert bieten.

## Literatur

- [1] Dickens M M et al. Removal of optical fiber interference in color micro-endoscopic images. Proc's 11th IEEE Symposium on Computer-Based Medical Systems 1998 : 246–251.
- [2] Lertrattanapanich S, Bose N. High resolution image formation from low resolution frames using delaunay triangulation. IEEE Transactions on Image Process 2002; 11 : 1427-1441.
- [3] Winter C, et al.: Automatic adaptive removal of fiberscopic comb-structure by spectral masking. Proc's Computer Assisted Radiology and Surgery, Amsterdam 2005 : 1283.