

Echtzeitfähige Extraktion scharfer Standbilder in der Video-Koloskopie

Sebastian Gross^{1,2}, Martin Schink¹, Thomas Stehle¹, Alexander Behrens¹,
Jens Tischendorf², Christian Trautwein² und Til Aach¹

¹Lehrstuhl für Bildverarbeitung, RWTH Aachen University

²Medizinische Klinik III, Universitätsklinikum Aachen

sebastian.gross@lfb.rwth-aachen.de

Kurzfassung. Die endoskopische Untersuchung des Darms (Koloskopie) ist ein wichtiges Verfahren zur Diagnose und Therapie vieler Erkrankungen. Während der Koloskopie werden u. a. zu dokumentarischen Zwecken oder als Grundlage für die computergestützte Diagnose und Therapie Standaufnahmen erzeugt. Aufgrund der Peristaltik des Darms und der eingeschränkten Kontrolle über das Koloskop ist es schwierig, Standbilder aufzunehmen, die keine Bewegungs- oder Focus-bedingte Unschärfe aufweisen. Eine Verbesserung der Situation kann erreicht werden, indem statt einer einzelnen Aufnahme eine Sequenz von Bildern gewonnen wird, aus denen in Echtzeit mit Hilfe von schnellen Schärfemaßen zur Bewertung der Bildschärfe das detailreichste Bild gewählt wird. Wir stellen drei Schärfemaße vor und vergleichen ihre Ergebnisse auf sechs verschiedenen Bildersequenzen aus der Koloskopie.

1 Einleitung

Die Endoskopie ist als minimalinvasives Verfahren zur Diagnose und Therapie einer großen Anzahl an Krankheiten im klinischen Alltag von großer Bedeutung. Für die Diagnose von Dickdarmkrebs ist die Koloskopie (Darmspiegelung) das Standardverfahren, welches in Deutschland ab dem 55. Lebensjahr von den Krankenkassen alle zehn Jahre empfohlen und finanziert wird. Darmkrebs ist hierzulande bei beiden Geschlechtern die zweithäufigste Krebserkrankung und es werden im Jahr ca. 73.000 Neuerkrankungen registriert [1].

Während der Untersuchung werden digitale Momentaufnahmen zur Dokumentation und ggf. zur computergestützten Diagnose und Therapie [2] erstellt. Die Aussagekraft dieser Aufnahmen hängt nicht nur von der Qualität des bildgebenden Systems, sondern auch von Geschicklichkeit und Erfahrung des Untersuchers ab. Während einer Koloskopie bewegt sich der Darm durch die Peristaltik unentwegt, was die Aufnahme scharfer Einzelbilder erschwert. Ein Verfahren, welches vom Arzt gestartet, selbständig die schärfste Aufnahme aus einer frei wählbaren Anzahl von Folgebildern wählt, könnte hier verhindern, dass der Untersucher die Aufnahme mehrfach wiederholen muss und gleichzeitig die Qualität der Aufnahmen steigern. Aus diesen Gründen untersuchen wir verschiedene Schärfemaße im Echtzeiteinsatz auf Videos aus der Koloskopie.

Der Abschnitt 2 beschäftigt sich mit der Schärfe von Bildern und deren Messung. Hier werden verschiedene Schärfemaße und die durchgeführten Untersuchungen vorgestellt. In Abschnitt 3 werden die Resultate der Untersuchungen und die benötigten Rechenzeiten dargelegt. Im abschließenden Abschnitt 4 werden Schlussfolgerungen gezogen und ein Ausblick auf zukünftige Arbeiten gegeben.

2 Material und Methoden

Schärfe bezeichnet die Eigenschaft eines Bildes, Details an Kanten und Ecken genau darstellen zu können. Dies ist mit Hilfe lokal hoher Ortsfrequenzen möglich. Es gibt Verfahren, welche die Schärfe in Bildern bewerten. Die Farbwerte werden hierzu in Graustufen umgewandelt. Drei echtzeitfähige Verfahren werden hier vorgestellt und miteinander in der Anwendung auf Dickdarmpolypen unter NBI-Beleuchtung [3] verglichen, da diese Beleuchtung die diagnostisch wichtigen Blutgefäße [4] auf den Polypen hervorhebt. Im Folgenden werden die Schärfemaße Varianz, Sum Modulus Difference (SMD) und Perceptual Sharpness Metric (PSM) beschrieben.

2.1 Varianz

Im Jahr 1982 haben Erasmus et al. [5] vorgeschlagen, die über das gesamte Bild gemessene Varianz als Referenz für die Schärfe eines Bildes zu nutzen. Die Berechnung dieses Schärfemaßes erfolgt mit

$$S_{\text{Varianz}} = \frac{1}{N \cdot M} \sum_{i=1}^M \sum_{j=1}^N \left(x_{i,j} - \frac{1}{N \cdot M} \sum_{i=1}^M \sum_{j=1}^N x_{i,j} \right)^2 \quad (1)$$

wobei N der Bildbreite in Pixeln und M der Bildhöhe in Pixeln entspricht. Die Grauwerte der einzelnen Bildpunkte (i, j) werden mit $x_{i,j}$ bezeichnet.

2.2 Sum Modulus Difference

Sum Modulus Difference (SMD) [6] bewertet die partiellen Ableitungen nach x und y mit Hilfe der Amplitudendifferenzen von horizontalen und vertikalen Nachbarpixeln und summiert diese betragsmäßig auf. Das Schärfemaß wird anhand der Vorschrift

$$S_{\text{SMD}} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (|x_{i,j} - x_{i-1,j}| + |x_{i,j} - x_{i,j-1}|) \quad (2)$$

berechnet.

2.3 Perceptual Sharpness Metric

Yang et al. [7] schlugen 2006 ein Schärfemaß namens Perceptual Sharpness Metric (PSM) vor. Das Bild wird überlappungsfrei in 8×8 -Pixel große Blöcke zerlegt. Auf diese Blöcke wird die Diskrete Kosinus-Transformation (DCT) wie in

$$DCT_{m,n}(k_1, k_2) = \sum_{i=0}^7 \sum_{j=0}^7 x_{i,j} \cos \left[\frac{\pi}{8} \left(i + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{8} \left(j + \frac{1}{2} \right) k_2 \right] \quad (3)$$

angewendet, wobei $k_1 = 0..7$ und $k_2 = 0..7$ die Indices für die Koeffizienten der DCT des transformierten 8×8 -Blockes mit der Position (m, n) darstellen. Die Koeffizienten der ersten Zeile und ersten Spalte des transformierten 8×8 -Blockes werden mit Hilfe der Sharpness Sensitivity Function $SSF(d) = d^{0.269} (-3.533 + 3.533d) \exp(-0.548d)$, welche auf Studien über die Empfindlichkeit des menschlichen Auges basiert [7], gewichtet und nach der Vorschrift

$$\tilde{S}_{m,n} = \sum_{d=1}^8 SSF(d) [DCT_{m,n}(0, d-1) + DCT_{m,n}(d-1, 0)] \quad (4)$$

aufsummiert. Abschließend wird die Summe

$$S_{PSM} = \frac{1}{M \cdot N} \sum_{m=1}^{M/8} \sum_{n=1}^{N/8} \tilde{S}_{m,n} \quad (5)$$

über alle Blöcke gebildet, wobei das Gesamtergebnis S_{PSM} auf die Bildbreite N in Pixeln und die Bildhöhe M in Pixeln normiert wurde.

2.4 Versuchsaufbau

Um die vorgestellten Verfahren zu vergleichen, haben wir sechs verschiedene Sequenzen aus endoskopischen Videos verwendet. Die Sequenzen zeigen verschiedene Dickdarmpolypen unter NBI-Beleuchtung. Die Videoausschnitte wurden zur einfacheren Handhabbarkeit in Halbbilder der Größe 1440×540 zerlegt. Die einzelnen Sequenzen sind zwischen 50 und 100 Bilder lang. Wie in der Koloskopie üblich zeigen viele Bilder Lichtreflexionen an der Darmoberfläche. Diese Glanzlichter wurden durch einen von Stehle im Jahr 2006 vorgestellten Algorithmus [8] lokalisiert und während der Berechnung der Schärfemaße ausgespart.

Verwendet wurde für die Testreihen ein Intel Core2 Duo E6550 mit 1 GB DDR2 RAM und NVIDIA GeForce 8400GS. Die Verfahren haben ihre Berechnungen in Echtzeit auf den zentralen Bildbereichen (576×468) ausgeführt, die Bilder wurden zur Dokumentation auf der Festplatte gespeichert und das beste Bild ausgewählt. Im Routineeinsatz reicht die Speicherung des derzeit besten Bildes im Speicher. Zu den drei Ergebnisbildern der Schärfemaße wurden scharfe Bilder aus den Serien herausgesucht und beigemischt. Diese Gruppe von Bildern für jede Serie wurde im Anschluss sieben in der Bildverarbeitung tätigen

Tabelle 1. Die Tabelle zeigt den prozentualen Anteil p der gewonnenen Entscheidungen für jedes Verfahren und jede Serie. In der vorletzten Spalte ist der Mittelwert jedes Verfahrens über alle Serien aufgeführt, in der letzten Spalte erscheint die Varianz.

	Serie 1	Serie 2	Serie 3	Serie 4	Serie 5	Serie 6	Mittelwert	Varianz
p_{Varianz}	0,64	0,43	0	0,14	0,45	0,49	0,36	0,057
p_{SMD}	0,40	0,76	0,60	0,66	0,68	0,12	0,54	0,056
p_{PSM}	0,52	0,03	0,45	0,19	0,04	0,32	0,26	0,043

Tabelle 2. Vergleich der durchschnittlichen Rechenzeiten T pro Bild in ms für die verschiedenen Serien.

	Serie 1	Serie 2	Serie 3	Serie 4	Serie 5	Serie 6	Mittelwert	Varianz
T_{Varianz}	3,77	4,06	3,92	3,65	3,78	3,92	3,85	0,02
T_{SMD}	12,07	12,04	12,50	13,60	12,78	11,73	12,45	0,45
T_{PSM}	19,21	19,84	20,28	21,37	19,68	20,05	20,07	0,54

Betrachtern vorgelegt. Die Aufgabe bestand darin, im direkten Vergleich der Bilder das jeweils schärfere zu bestimmen. Falls der Betrachter beide Bilder für gleich scharf hielt, wurde die Wertung geteilt. Die Gesamtsumme der von einem Verfahren erreichten Wertungen innerhalb einer Serie wurde dann auf die Anzahl der durchgeführten Vergleiche pro Bild normiert. Der Mittelwert aller Serien wurde dann genutzt, um das insgesamt beste Verfahren zu bestimmen.

3 Ergebnisse

Die drei Verfahren haben in den sechs Sequenzen jeweils das vermeintlich beste Bild ausgewählt. In der Abb. 1 sind beispielhaft die aus der Serie 6 gewählten Bilder für die drei Verfahren dargestellt. In Tabelle 1 sind die Wertungen aus der Studie zusammen gefasst. Die beste Gesamtbewertung durch die Betrachter erzielte das Verfahren SMD. Ebenfalls gemessen wurden die Rechenzeiten der einzelnen Verfahren, welche in Tabelle 2 dargestellt sind. Die niedrigste Rechenzeit T weist das Verfahren Varianz auf.



Abb. 1. Beispielhaft zeigt die Grafik die gefundenen Resultate von Varianz (Frame 6, links), SMD (Frame 17, Mitte) und PSM (Frame 60, rechts).

4 Diskussion

Das Ergebnis der Studie zeigt, dass die Bilder, die vom Verfahren SMD bestimmt wurden, von den Betrachtern als schärfer empfunden wurden als die Bilder der anderen Verfahren. Die Rechenzeiten T aller Verfahren zeigen deutlich, dass ihre Berechnung in Echtzeit möglich ist. Bis zu 50 Bilder pro Sekunde sind selbst mit PSM möglich. Allerdings schneidet dieses Verfahren erheblich schlechter ab als SMD, welches auch deutlich geringere Rechenzeiten ausweist. Wenn die Bilder als Grundlage für weiterführende Analysen [2] in Echtzeit genutzt werden, kann die Varianz Rechenzeit sparen. Im Falle der Gewinnung von Standaufnahmen senkt die Anwendung der Technik den Arbeitsaufwand für den untersuchenden Arzt, da weniger untaugliche Aufnahmen entstehen. Gleichzeitig steigert sich die Qualität der Aufnahmen für Dokumentation und Weiterverarbeitung.

In weiteren Untersuchungen sollte das Verfahren im Rahmen einer klinischen Studie getestet werden. Auf diesem Wege könnte die Qualität der gewonnenen Bilder direkt mit den Ergebnissen der herkömmlichen Standbilder verglichen werden. Weiterhin sollte untersucht werden, ob es möglich ist, zusätzliche Informationen über den Bildinhalt aus den 1-2 Sekunden Video zu extrahieren und in einem scharfen Standbild zu vereinen. Ein möglicher Ansatz wäre Extended Depth of Field (EDoF) [9].

Danksagung. Das Projekt wurde aus Mitteln der Exzellenzinitiative des Bundes und der Länder gefördert.

Literaturverzeichnis

1. Gesundheitsberichterstattung des Bundes. Krebs in Deutschland 2003-2004 - Häufigkeiten und Trends. 8th ed. Robert Koch Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V.; 2008.
2. Stehle T, Gross S, Behrens A, et al. Classification of colon polyps in NBI endoscopy using vascularization features. *Proc SPIE*. 2009;6918:12.
3. Gono K, Obi T, Yamaguchi M, et al. Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J Biomed Opt*. 2004;9(3):568–77.
4. Tischendorf JJW, Wasmuth HE, Koch A, et al. Value of magnifying chromoendoscopy and narrow band imaging (NBI) in classifying colorectal polyps: A prospective controlled study. *Endoscopy*. 2007;39(12):1092–6.
5. Erasmus SJ, Smith KCA. An automatic focusing and astigmatism correction system for the SEM and CTEM. *J Microsc*. 1982;127:185–199.
6. Brenner JF, Dew BS, Horton JB, et al. An automated microscope for cytologic research. *J Histochem Cytochem*. 1976;24(1):100–11.
7. Yang KC, Guest CC, Das P. Perceptual sharpness metric (PSM) for compressed video. In: *Proc. IEEE ICME*; 2006. p. 777–80.
8. Stehle T. Removal of specular reflections in endoscopic images. *Acta Polytechn J Adv Eng*. 2006;46(4):32–6.
9. Bradley AP, Bamford PC. A one-pass extended depth of field algorithm based on the over-complete discrete wavelet transform. In: *Proc Image Vis Comput*; 2004. p. 279–84.