

Identifikation von Zellen in intaktem Gewebe

Segmentierung und Klassifikation von Zellkernen in 3D Volumendaten mittels voxel-basierter Grauwertinvarianten

Janis Fehr¹, Catharina Sauer^{1,2}, Haymo Kurz²,
Olaf Ronneberger¹ und Hans Burkhardt¹

¹Institut für Informatik, Lehrstuhl für Mustererkennung und Bildverarbeitung,

²Institut für Anatomie und Zellbiologie,
Albert-Ludwigs-Universität Freiburg
Email: fehr@informatik.uni-freiburg.de

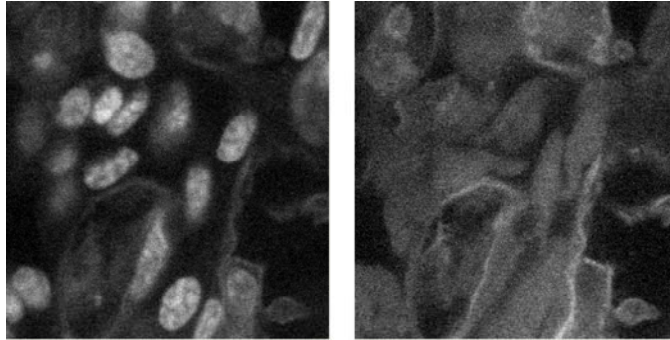
Zusammenfassung. Wir präsentieren erste Ergebnisse eines Ansatzes zur vollautomatischen Segmentierung und Klassifikation von Zellen (Zellkernen) in 3D Volumendaten. Durch Einsatz voxel-weiser invarianter Merkmale und Support Vektor Maschinen wird die Klassenzugehörigkeit jedes Voxels anhand von Trainingsbeispielen gelernt. Dies ermöglicht eine simultane Segmentierung und Klassifikation von Objekten selbst in stark verrauschten und kontrastschwachen konfokalen Laserscanning-Mikroskop Aufnahmen.

1 Einleitung

Die Untersuchung von Zellen in natürlichem Kontext ist für viele biologische und medizinische Fragestellungen von entscheidender Bedeutung. Die Analyse unterschiedlichster Gewebsverbände mittels konfokaler Laserscanning-Mikroskopie (LSM) hat sich auf vielen Feldern etabliert, dennoch stehen kaum Werkzeuge zur automatischen Segmentierung und Klassifikation von Strukturen in 3D LSM Aufnahmen zur Verfügung. Gewebsaufnahmen mit ihrer teilweise sehr dichten Packung, oft schlechtem Signal-Rauschverhältnis und unklaren Objektkanten machen einer Segmentierung mittels bekannter Verfahren wie z.B. Schwellwert- oder Regiongrowingverfahren nahezu unmöglich. Zudem verlangt die Anwendung oftmals eine schnelle und einfache Adaption an unterschiedliche Bedingungen (z.B. andere Färbeverfahren, Mikroskopie-Parameter, etc) und verschiedene Zelltypen.

Im Bereich der Segmentierung und Klassifikation von 2D Daten existiert eine Vielzahl von Verfahren welche aber in der Regel nicht auf 3D Daten erweiterbar sind. Bei den wenigen in der Literatur beschriebenen Algorithmen für 3D LSM-Daten handelt es sich um hoch spezialisierte Lösungen wie z.B. Atlas- und Matching-basierte Ansätze, welche nur mit großem Aufwand auf neue Strukturen (z.B. andere Zell- und Gewebstypen) angepasst werden können. Biologische Verfahren setzen spezifische Marker, fluoreszierende Reporter-Proteine und Immun-Fluoreszenz-Methoden bei der Präparation der Gewebeproben ein, um

Abb. 1. xy-Schnitt einer typischen Zweikanalprobe: links der erste Kanal mit YoPro Färbung stellt im wesentlichen die Zellkernstruktur da. Rechts der zweite Kanal in dem bei dieser Probe mit Cy3 u.a. Zellmembranen abgebildet werden.



teils markante, teils subtile Unterschiede zu erfassen. Die Nachteile dieser Verfahren liegen in der zeitaufwändigen Präparation und der begrenzten Anzahl simultan messbarer Parameter.

In einem neuen sehr allgemeinen Ansatz soll mit möglichst wenig Vorarbeiten und ohne Apriori-Wissen die simultane Segmentierung und Klassifikation von anisotropen 3D Objekten anhand von Trainingsbeispielen gelernt werden. Die Arbeit ist wie folgt strukturiert: Abschnitt 2 gibt einen Überblick über die verwendeten Präparate und die daraus erzeugte Datenbank. Die voxel-weisen Grauwertinvarianten werden in Abschnitt 3 kurz vorgestellt. Abschnitt 4 zeigt die Ergebnisse der Versuche die in 5 diskutiert werden.

2 Methoden

Gewebsproben aus der Chorioallantois-Membran des Hühnereis wurden mit YoPro (Färbung der Kerntextur) und mit Cy3 oder Alexa 546 gekoppelten spezifischen Markern gefärbt [1]. Mit dem LSM wurden dann zwei Kanäle YoPro (509nm Emission) und Cy3/Alexa 546 (570nm) aufgenommen (Abb 1).

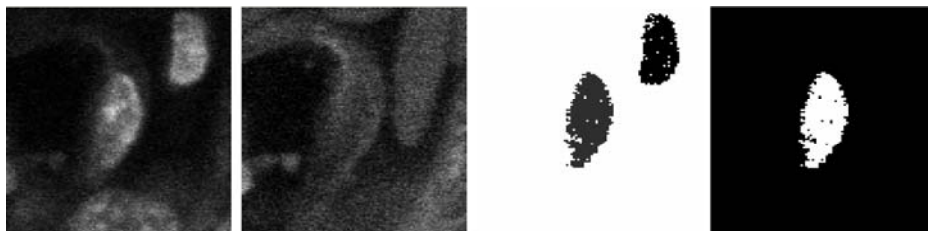
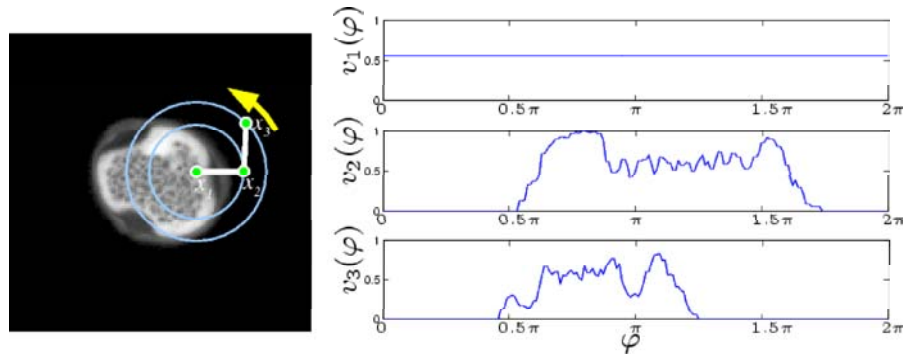


Abb. 2. Datenbank-Element einer Endothelzelle mit Umgebung aus Abb. 1, xy-Schnitt (von links nach rechts): Ausschnitt YoPro Kanal, Ausschnitt Cy3 Kanal, Label der Zellkerne im Ausschnitt, Maske des Endothelzellkerns.

Abb. 3. Bei Verwendung von 3-Punkt-Kernen auf einen kontinuierlichen 2D Bild und unter Betrachtung der Rotationsgruppe, ergibt die Abtastung durch die Kern-Punkte \mathbf{x}_1 , \mathbf{x}_2 und \mathbf{x}_3 , eindimensionale Funktionen $v_1(\varphi)$, $v_2(\varphi)$ und $v_3(\varphi)$.



2.1 Zell-Datenbank

Aus 50 Aufnahmen unterschiedlicher Gewebeproben wurde eine Datenbank mit insgesamt 229 Zellkernen in 4 Klassen erstellt (30 Epithelzellen, 27 Fibroblasten und 172 Erythrozyten sowie Hintergrund). Dabei wurden die Zellkerne, unterstützt durch ein Region-Growing-Verfahren, manuell segmentiert und gelabelt. Abb. 2 zeigt die Datenbankenrepräsentation eines Zellkerns. Auf dieser werden dann voxel-weise 3D rotationsinvariante Grauwertmerkmale mittels Haar-Integration berechnet (siehe Abschnitt 3) und zum Trainieren einer Supportvektor-Maschine (SVM) [2] verwendet.

3 Voxel-Weise Grauwertinvarianten

Voxel-weise Grauwertinvarianten (GSI) [3][4][5] sind 3D-Rotationsinvariante Merkmale welche für jedes Voxel mittels Haar-Integration über die Rotationsgruppe eines Voxels und seiner Umgebung berechnet werden. Formel (1) zeigt die allgemeine Berechnung der GSI.

$$T[f](\mathbf{X}) := \int_G f(g\mathbf{X}) dg \quad (1)$$

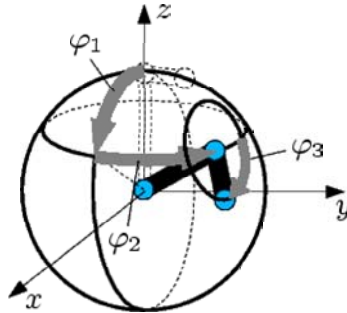
G : Transformationsgruppe (Rotation)
 g : ein Element der Transformationsgruppe
 f : nichtlineare Kernfunktion
 \mathbf{X} : n -dim, Multikanaldaten
 $g\mathbf{X}$: transformierte n -dim Daten

Aus Gründen der Berechnungskomplexität beschränken sich unsere Versuche auf die handhabbare Untermenge der "Separierbaren 3-Punkt Kerne" [5] der Form:

$$f(\mathbf{X}) = f_a(\mathbf{X}(\mathbf{0})) \cdot f_b(\mathbf{X}(\mathbf{q}_1)) \cdot f_c(\mathbf{X}(\mathbf{q}_2))$$

f_a, f_b, f_c : beliebige Funktion auf den Grauwerten
 \mathbf{q}_i : Radius

Abb. 4. Parameterisierung der 3D Rotation mit $\lambda = (\varphi_1, \varphi_2, \varphi_3)^T$



Für 3D Volumendaten (siehe Abb. 4) lässt sich das Haar-Integral (1) für 3-Punkt Merkmale mittels Faltung auf einer Entwicklung in Spherical-Harmonics für alle Voxel gleichzeitig sehr schnell berechnen [5]. Um die unendlich große Menge der möglichen Kernfunktionen weiter einzuzugrenzen, werden aus allen Kombinationen praktisch relevanter Radien und nichtlinearer Abbildungen mittels Maximum Marginal Diversity [6] die Merkmale mit den besten Separationseigenschaften ausgewählt. Auf diese Weise wird so für jeden Voxel ein Merkmalsvektor generiert.

4 Ergebnisse

Die Versuche wurden mittels der Leave-One-Out-Methode durchgeführt, dabei wurden immer alle Merkmalsvektoren eines Datenbankeintrages entnommen. Auf den restlichen Daten wurde eine SVM trainiert gegen welche die entnommenen Merkmale dann klassifiziert werden konnten. Tab. (1) zeigt die Erkennungsrate der Klassifikation jeweils voxel-weise und pro Zellen, dabei wird eine Zelle als korrekt angenommen, wenn mindestens 60% der Voxel korrekt klassifiziert sind.

Die schlechteren Erkennungsraten von Endothelzellen und Fibroblasten im Vergleich mit Erythrozyten und Hintergrund erklärt sich wohl durch die deutlich geringere Anzahl an Trainingsbeispielen die für diese beiden Klassen zur Verfügung standen.

5 Diskussion

Unsere Ergebnisse zeigen, dass erstmals die simultane Segmentierung und Klassifikation von Zellkernen anhand Voxel-weißer Invariantenberechnung mit einer guten Qualität realisiert werden konnte[4].

Tabelle 1. Erkennungsrate der Klassifikation

Zelltyp	Erkennungsrate / Voxel	Erkennungsrate / Zelle
Erythrozyten	90,4%	95,3%
Endothelzellen	81,2%	84,6%
Fibroblasten	78,3%	79,8%
Hintergrund	94,1 %	-

Dieser fundamental neue Ansatz in der Mustererkennung in 3-D-Datensätzen sollte neue, vollautomatisierte Anwendungen für die Analyse komplexer biologischer Gewebs-Architekturen ermöglichen. Die Grauwert-Invarianten haben sich auch in verrauschten Ein-Kanal-Daten als robuste und schnell berechenbare Merkmale erwiesen. Durch Hinzunahme weiterer Kanäle lässt sich erforderlichenfalls die Klassifikationsqualität nochmal verbessern. Die weitere Entwicklung wird sich besonders auf das quantitativ anspruchsvolle Datenmanagement (> 100000 Merkmalsvektoren, Rekombination der klassifizierten Voxel zu segmentierten Objekten, Quantitative Analyse auf Objekt-Ebene) zu konzentrieren haben.

Literaturverzeichnis

1. Kurz H, Papoutsis M, Wilting J, Christ B. Pericytes in experimental MDA-MB231 tumor angiogenesis. *Histochem Cell Biol* 2002;p. 117:527–534.
2. Vapnik VN. The nature of statistical learning theory. Springer; 1995.
3. Ronneberger O, Fehr J, Burkhardt H. Voxel-Wise Gray Scale Invariants for Simultaneous Segmentation and Classification. In: *Procs DAGM*. vol. 3663 of LNCS. Springer; 2005.
4. Fehr J, Ronneberger O, Kurz H, Burkhardt H. Self-Learning Segmentation and Classification of Cell-Nuclei in 3D Volumetric Data using Voxel-Wise Gray Scale Invariants. In: *Procs DAGM*. vol. 3663 of LNCS. Springer; 2005.
5. Ronneberger O, Fehr J, Burkhardt H. Voxel-Wise Gray Scale Invariants for Simultaneous Segmentation and Classification – Theory and Application to Cell-Nuclei in 3D Volumetric Data. IIF-LMB, University of Freiburg, Germany; 2005.
6. Vasconcelos N. Feature Selection by Maximum Marginal Diversity: optimality and implications for visual recognition. In: *Procs IEEE Conf on Computer Vision and Pattern Recognition*, Madison, USA; 2003.