

Bildverarbeitung für ein motorisiertes Lichtmikroskop zur automatischen Lymphozytenidentifikation

Michael Beller¹, Rainer Stotzka¹, Hartmut Gemmeke¹
Karl-Friedrich Weibezahn², Gudrun Knedlitschek²

¹Institut für Prozessdatenverarbeitung und Elektronik
Forschungszentrum Karlsruhe, 76344 Eggenstein, Germany

²Institut für Medizintechnik und Biophysik
Forschungszentrum Karlsruhe, 76344 Eggenstein, Germany
Email: Michael.Beller@ipe.fzk.de

Zusammenfassung. Zur Diagnose potentiell maligner Erkrankungen können die Größenverhältnisse von Lymphozyten aus arteriellem und venösem Blut herangezogen werden. Die Bestimmung dieser Größen wird bisher manuell durchgeführt. Es wird ein Detektionssystem weiterentwickelt, das die Selektion der Lymphozyten und ihre Auswertung objektiv gewährleisten soll. Dafür wird ein motorisiertes Lichtmikroskop zusammen mit einer CCD-Kamera benutzt. Bei 20facher Vergrößerung wird in einem Blutausschnitt zunächst nach angefärbten Zellen gesucht, die danach mit 100facher Vergrößerung angefahren und genau vermessen werden. Zuletzt werden diese Zellen in Lymphozyten und Nicht-Lymphozyten klassifiziert.

1 Problemstellung und medizinischer Hintergrund

Tumorzellen sind Körperzellen, die nicht mehr einer normalen Wachstumsregulierung unterliegen. Die Zelle verändert jedoch häufig nicht nur Form und Größe, sondern meist auch ihr Verhalten und ihre Oberflächencharakteristik und ist so prinzipiell durch das Immunsystem zu erkennen. Hinweise auf die Anwesenheit von Tumorzellen im Körper sollten sich aus Veränderungen bestimmter Komponenten des Immunsystems ergeben.

Dieser Hypothese entsprechend ergaben Beobachtungen an einer ukrainischen Klinik, dass die Größenverhältnisse von Lymphozyten aus Kapillarblut (überwiegend arteriell) zu solchen aus venösem Blut Rückschlüsse auf eine potentielle maligne Erkrankung gestatten. Über die Ursachen, die hier zugrunde liegen, ist noch wenig bekannt. Diese vergleichende Methode könnte den zur Zeit verwendeten bildgebenden Verfahren (Strahlenbelastung bei einigen Verfahren erheblich), histologischen Untersuchungen an Biopsiematerial und Verfahren mit Tumormarkern überlegen sein, da sie lediglich eine Blutentnahme erfordert und zudem letztgenannte in der Unterscheidung in gut- und bösartig oft unsicher sind.

Zur wissenschaftlichen Analyse dieser Beobachtung ist die Untersuchung einer

großen Zahl von Blutaussstrichen notwendig. Da ein einigermaßen geschultes Auge etwa ein bis fünf Tage benötigt, um die Probenträger für einen Patienten vollständig zu analysieren [1], wurde ein automatisches System entwickelt, das auf einem Probenträger angefärbte Objekte segmentiert und vermisst.

2 Stand der Forschung

Die zur Zeit verwendeten Verfahren der Zellzählung und Zellgrößenbestimmung sind deren manuelle Bestimmung, das Coulter-Prinzip und die Flusszytometrie. Die manuelle Bestimmung ist aufgrund der Kosten- und Zeitintensivität sowie der Interpretationsfehler durch den Untersuchenden nicht geeignet [2].

Beim Coulter-Verfahren und der Flusszytometrie werden die Zellen in einem isotonischen Medium aufgenommen. Dies führt zu Volumenänderungen der Zellen [1]. Während des Passierens des elektrischen Messfelds bei Einsatz der Coulter-Methode können die Zellen zudem verformt werden, was zu einer Veränderung der Größenverhältnisse führt [2]. Diese Verhältnisse sind jedoch für die Interpretation wichtig. Weiterhin kann mit beiden Verfahren die gewünschte Größenauflösung von $0,5 \mu\text{m}$ nicht erreicht werden.

Es existieren Systeme, die Blutzellen erkennen und solche, die Blutzellen vermessen. Es gibt auch Systeme, die beide Funktionen vereinen, jedoch müssen diese konstant überwacht werden und klassifizieren häufig falsch. Dazu wurde in [2] die Analyse von Hundebloodproben herangezogen: „Denn bei einer Differenzierung von 100 Leukozyten (...) betrug die Zahl der nicht erkannten und manuell nachzudifferenzierenden Zellen durchschnittlich 18,8 % und maximal sogar 50,7 %.“. Als Alternative werden in [2] halbautomatische Systeme genannt, die das Orten und Fokussieren der Zelle selbst übernehmen und die eigentliche Identifikation dem Untersucher überlassen. Dies ist aus o.g. Gründen nicht akzeptabel.

Unser Ziel ist daher die Entwicklung eines Systems, das nahezu vollautomatisch weiße Blutkörperchen aufnimmt, zählt, segmentiert, vermisst und klassifiziert. Die benötigte Zeit wird gegenüber einer manuellen Vermessung wesentlich verkürzt werden. Durch den Einsatz von herkömmlichen Geräten werden Aufwand und Kosten verringert. Zusätzlich werden spezielle Nachteile der bisherigen automatischen Methoden zur Zellzählung und Identifikation umgangen. Eine breit angelegte Studie zur Verifikation der Beobachtungen ist somit möglich.

3 Materialien und Methoden

Ein Standard-Lichtmikroskop (Axioskop, Zeiss) mit einem xyz-Tisch (Märzhäuser) wird mit einer CCD-Kamera (SIS ColorView Camera) gekoppelt. Die Optik und das Mikroskop selbst sind manuell optimal eingestellt. Die Auflösung der Kamera beträgt 1280×1024 Pixel bei einer Pixelgröße von $0,32 \mu\text{m} \times 0,32 \mu\text{m}$ für die 20fache bzw. $0,06 \mu\text{m} \times 0,06 \mu\text{m}$ für die 100fache Vergrößerung.

Die Kamera ist an einen handelsüblichen PC angeschlossen. Bei der verwendeten Software handelt es sich um AnalySIS® der Firma SIS [3]. Dieses Programm ermöglicht die Steuerung des Mikroskops, die Verwendung der

CCD-Kamera und die Bildverarbeitung. Innerhalb von AnalySIS existiert ein Derivat der Programmiersprache C, welches um verschiedene Bildverarbeitungs-routinen erweitert wurde und die Erstellung eigener Verarbeitungsfunktionen gestattet. Das vorgestellte Verfahren wurde hierin implementiert.

Die verwendeten Blutproben wurden auf Objektträgern ausgestrichen und gemäß der Methode nach Pappenheim [4] angefärbt. Die Erythrozyten wurden durch eine Säurebehandlung zerstört, da sie den Segmentierungsprozess stören. Dies hat keinen Einfluß auf die Größenverteilung der Lymphozyten.

Der Objektträger wird bei 20facher Vergrößerung mäanderförmig abgetastet. Wird ein angefärbtes Objekt gefunden, wird dessen Position gespeichert. Bei 100facher Vergrößerung werden dann die abgespeicherten Positionen angefahren, die Objekte aufgenommen und analysiert. Ihre Merkmale werden extrahiert und die Objekte als Lymphozyt/Nicht-Lymphozyt klassifiziert.

3.1 Segmentierung

Das Auffinden eines angefärbten Objektes bei 20facher Vergrößerung erfolgt durch eine Suche nach Objekten, die in der Größenordnung der gesuchten Zellen liegen und eine ähnliche Form haben.

Bei 100facher Vergrößerung werden die Charakteristika der Zellen untersucht. Es wird zuerst das RGB-Bild in ein Grauwertbild transformiert. Danach wird derjenige Schwellwert für das trimodale Histogramm gesucht, der die Zelle vom Hintergrund separiert. Das Histogramm ist trimodal, da das Bild aus Zellkern, Zellplasma und Hintergrund besteht. Mittels einer Region-Of-Interest wird der Hintergrund ausgewählt und dessen Mittelwert μ und dessen Standardabweichung σ bestimmt. Der Schwellwert wird bei $\mu - 3\sigma$ gesetzt, wodurch 99,73% der Hintergrundwerte berücksichtigt sind [5].

Für das nun bimodale Histogramm existiert ein Schwellwert, der Kern und Plasma gut trennt. Ein gängiges Verfahren ist die Methode von Otsu [6]. Hierbei wird der Schwellwert dort gesetzt, wo die Summe der Varianzen der entstehenden Gruppen minimal wird. Die Erkennung des Zellkerns ist nötig, da sich Lymphozyten von z.B. Granulozyten nicht in der Größe sondern in der Form des Kerns unterscheiden. Somit sind nun 2 Objekte für die weitere Analyse vorhanden.

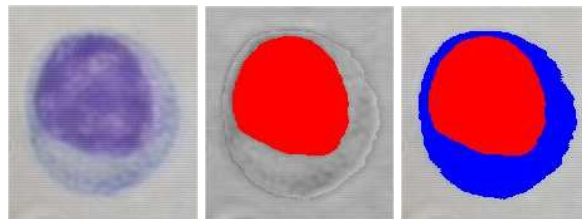


Abb. 1. Von links nach recht: Nach Pappenheim angefärbter Lymphozyt. Segmentierter Zellkern (rot). Segmentiertes Plasma (blau) und Kern (rot).

3.2 Merkmalsextraktion und Klassifikation

Nach der Segmentierung werden 24 Merkmale der detektierten Objekte extrahiert. Der Merkmalsvektor besteht unter anderem aus den folgenden Elementen: (konvexe) Fläche, (konvexer) Umfang, Equivalent Circle Diameter (ECD), Objektschwerpunkt in x und y , minimaler, mittlerer und maximaler Grauwert des Objektes, Standardabweichung σ , Varianz σ^2 , mittlerer und maximaler Durchmesser, maximale Objektausdehnung in x und y , Orientierung und Zirkularität. Die Klassifikation in Lymphozyt/Nicht-Lymphozyt erfolgt mit gängigen Methoden [7]: Bayes, k-Nearest Neighbour (kNN) und einem Klassifikationsbaum. Die Merkmale bilden einen 48-dimensionalen Raum, der sich aus den jeweils 24 Merkmalen von Kern und Plasma zusammensetzt. Bayes nimmt an, dass die Merkmale multivariat normalverteilt sind.

4 Ergebnisse

Die Ergebnisse, die die Klassifikation nach vollständiger Kreuzvalidierung [8] für 190 untersuchte Zellen liefert, finden sich in Tab. 1.

Das Augenmerk ist auf die Lymphozyten gerichtet. Handelt es sich bei der betrachteten Zelle um einen Nicht-Lymphozyten, der als Lymphozyt klassifiziert wurde, bezeichnet man dies als falsch-positiv (f.-pos.). Falsch-negativ (f.-neg.) ist äquivalent dazu ein Lymphozyt, der als Nicht-Lymphozyt klassifiziert wurde.

Klassifikator	wahr	falsch	Fehlerrate	f.-pos.	f.-neg.	f.-pos.-Rate	f.-neg.-Rate
Bayes	181	9	4,7%	1	8	0,5%	4,2%
kNN, {k=1,3}	185	5	2,6%	3	2	1,6%	1,1%
K-Baum	181	9	4,7%	4	5	2,1%	2,6%

Tabelle 1. Ergebnisse der Klassifikation. f.-pos. = falsch positiv, f.-neg. = falsch negativ. Für Bayes gilt, dass die a priori Wahrscheinlichkeiten beider Klassen gleich sind.

5 Diskussion

Das vorgestellte Verfahren zur automatischen Lymphozytenidentifikation ermöglicht es, Größenhistogramme von Lymphozyten anhand von Blutaussstrichen zu erstellen. Diese können zur Diagnose einer malignen Erkrankung herangezogen werden. Unter idealen optischen Voraussetzungen liegt die Schwierigkeit in der Segmentierung von Kern und Plasma, da der Kontrast oft gering ist. Modifikationen der Färbemethoden könnten hier möglicherweise zu weiteren Verbesserungen führen. Zur Positionsermittlung hingegen muss nur ein angefärbtes Objekt gefunden werden.

Die mit dem Verfahren von Otsu erzielten Ergebnisse zur Trennung von Zellkern und Zytoplasma sind einer manuellen Segmentierung durchaus ebenbürtig. Zur Bewertung wird der *Equivalent Circle Diameter (ECD)* verwendet. Dieser Wert wird zu einer diagnostischen Aussage herangezogen [1]. Die Güte des automatischen Vorgangs bestimmt sich durch den Vergleich zur manuellen Segmentierung. Wurden Zellen manuell segmentiert, betragen die Abweichungen der subjektiven Beurteilungen des ECD durch mehrere Personen bis zu $0,8 \mu\text{m}$. Somit ergibt der Vergleich nur eine Abschätzung der tatsächlichen Übereinstimmung. Als „gut segmentiert“ werden diejenigen automatisch segmentierten Zellen bezeichnet, deren ECD weniger als $0,5 \mu\text{m}$ – also unterhalb der geforderten Auflösung – von den manuell segmentierten abweicht. Nach diesem Kriterium werden 88 % der Lymphozyten und 92 % ihrer Kerne „gut segmentiert“. Lässt man eine Abweichung von $1 \mu\text{m}$ zu, so sind es 95 % resp. 98 %. Auch ein Vergleich der Größenhistogramme zeigt keine nennenswerten Unterschiede.

Aufgrund des vielfach schwachen Kontrasts ist eine perfekte automatische Segmentierung ohne a priori Wissen über die Zellen nicht möglich. Ein geschulter Mensch besitzt dieses Wissen und kann fehlende Bildinformation kompensieren – ein Algorithmus, der nur auf Intensitäten basiert, jedoch nicht.

Die Klassifikation liefert mit Fehlerraten zwischen 2,6 % und 4,7 % gute Ergebnisse. Die interessante falsch-positiv-Rate liegt mit 0,5 % bis 2,1 % unter dem Wert von 5,5 %, der in [2] genannt wird. Die falsch-negativ-Rate liegt mit 1,1 % bis 4,2 % ebenfalls unter der Literaturangabe von 12 %.

Das automatisierte Verfahren benötigt etwa 10 Stunden, was gegenüber einer manuellen Analyse deutlich Zeit spart. Eine vollständige Automatisierung ist aufgrund des Wechsels der Vergrößerung mit konventionellen Geräten jedoch nicht möglich (Vollautomatische Mikroskope sind inzwischen auf dem Markt). Zu untersuchen bleibt, ob und inwiefern die Fehlerrate mit sinnvollem Aufwand verbessert werden kann. Möglicherweise ließe sich die Klassifikation durch Segmentierungen mit regionen- oder texturorientierten Verfahren verbessern.

Literatur

1. Andreas Berting. Automatische Lymphozytenidentifikation. Master's thesis, Fachhochschule Gießen-Friedberg, 2000.
2. Marion Püsch. *Hämatologiesystem ADVIA 120, Softwareadaption und Evaluation bei den Tierarten Schaf und Ziege*. PhD thesis, Universität Gießen, 2002.
3. SIS GmbH. <http://www.soft-imaging.de>.
4. B. Romeis. *Mikroskopische Technik*. Oldenbourg-Verlag, 1968.
5. Lothar Sachs. *Angewandte Statistik*. Springer, 9th edition, 1999.
6. Al Bovik. *Handbook of Image and Video Processing*. Academic Press, 2000.
7. Casimir A. Kulikowski Sholom M. Weiss. *Computer Systems that learn*. Morgan Kaufmann, 1991.
8. D. Michie, D. J. Spiegelhalter, C. C. Taylor. *Machine Learning, Neural and Statistical Classification*. Ellis Horwood, 1994.