

# GPGPU-basierte Echtzeitdetektion von Nanoobjekten mittels Plasmonen-unterstützter Mikroskopie

Frank Weichert<sup>1</sup>, Constantin Timm<sup>2</sup>, Marcel Gaspar<sup>1</sup>, Alexander Zybin<sup>3</sup>,  
Evgeny L. Gurevich<sup>3</sup>, Heinrich Müller<sup>1</sup>, Peter Marwedel<sup>2</sup>

<sup>1</sup>Lehrstuhl für Graphische Systeme, Technische Universität Dortmund

<sup>2</sup>Lehrstuhl für Eingebettete Systeme, Technische Universität Dortmund

<sup>3</sup>ISAS - Institut for Analytical Science, Dortmund

frank.weichert@tu-dortmund.de

**Kurzfassung.** Die Verfügbarkeit echtzeitfähiger und mobiler Biosensoren gewinnt durch die zunehmende Verbreitung viraler Infektionen zunehmend an Bedeutung. Im Gegensatz zu Virusdetektionsmethoden wie beispielsweise ELISA erlaubt die neuartige Plasmonen-unterstützte Mikroskopie von Nanoobjekten, Proben innerhalb von wenigen Minuten auf Viren analysieren zu können. Die Herausforderung für ein, auf dieser Analysemethoden beruhendes In-situ-Virusdetektionssystem, besteht in der Echtzeitverarbeitung von extrem hohen Datenmengen. Hier setzt die vorliegende Arbeit an, welche eine hoch parallele GPU-basierte Verarbeitungspipeline zur echtzeitfähigen Virusdetektion vorstellt. Durch die konsequente Ausnutzung der GPGPU-Fähigkeiten von Grafikkarten kann auf teure Spezialhardware verzichtet werden, um eine echtzeitkonforme Beschleunigung notwendiger Bildverarbeitungs- und Bildanalysealgorithmen bereitzustellen, die auch den Anforderungen an ein eingebettetes Virusdetektionssystem gerecht wird.

## 1 Einleitung

Der Einsatz ubiquitär verfügbarer echtzeitfähiger Biosensoren wird vor dem Hintergrund eines Anstiegs epidemisch auftretender viraler Infektionen zunehmend relevant [1]. Mit der neuartigen, am Leibniz-Institut für Analytische Wissenschaften entwickelte PAMONO-Technik (engl. Plasmon-assisted Microscopy of Nano-Size Objects) [2] steht eine Methodik zur Verfügung, welche vergleichbar zu etablierten Verfahren, wie z.B. ELISA die gestellten Anforderungen an eine zuverlässige Virusdetektion erfüllt, aber zudem die Umsetzung an ein echtzeitfähiges portables System ermöglicht. Daher könnte dieses System auch außerhalb von Speziallaboren, direkt vor Ort, z.B. an Flughäfen eingesetzt werden, da es auf keine aufwendige Infrastruktur angewiesen ist. Prinzipiell basiert der PAMONO-Sensor [2] auf der Erkennung von markierungsfreien biomolekularen Bindungsreaktionen an einer Goldoberfläche, in einer mit einer CCD-Kamera aufgenommenen Bildserie. Zur Detektion von Bindungsereignissen wird der Effekt ausgenutzt, dass polarisierendes Licht (Laser), welches gebündelt über ein

Prisma auf eine Metallschicht trifft, reflektiert wird und dieses zu einer Anregung der Oberflächenplasmonen innerhalb der Metallschicht führt. Eine Virusbindung verändert dabei die reflektierte Intensität [2].

Eine CPU-basierte Analyse mit hoher Detektionsrate für PAMONO-Daten wurde in [3] vorgestellt. Die automatische Verarbeitung dieser Bilderserien stellt aber mit einer Datenrate von circa 50 Megabyte/s im Hinblick auf das verarbeitende System besondere Anforderungen, da selbst aktuelle Multicore-CPU's mit einer durchschnittlichen Verarbeitungszeit zwischen 70 – 100 ms pro Frame für die vorliegende Detektionsaufgabe nicht mehr konform zu einer Echtzeitdetektion sind. Eine effiziente Beschleunigungsmöglichkeit für parallele Algorithmen bieten Grafikkarten. Diese werden zunehmend auch im medizinischen Umfeld eingesetzt, beispielsweise bei der Verarbeitung von Ultraschallbildern, da sie im Gegensatz zu Spezialprozessoren [4] einen flexibleren Entwurf von Algorithmen unterstützen. Ausgehend von dieser einleitenden Darstellung thematisiert der Abschnitt 2, die echtzeitfähige Analyse und Verarbeitung der PANOMO-Daten mittels hoch parallelisierter Detektionsalgorithmen und der Abschnitt 3 die Entwurfsraumexploration mit verschiedenen Grafikkartenplattformen. Zur Bewertung der Nachhaltigkeit der Umsetzung erfolgt diese ausgehend von der Bewertung der Erkennungsrate der Verarbeitungspipeline (Abschn. 3).

## 2 Material und Methoden

Die Erkennung von Nanoobjekten beruht konzeptionell auf drei algorithmischen, echtzeitfähigen Basisschritten (Abb. 1). Methodisch wird diese parallele Verarbeitung und Klassifikation des Eingabebilderstroms über zwei Konzepte erreicht (Abb. 2). Im Hinblick auf die Vorverarbeitung und Klassifikation mittels Zeitreihenanalyse werden die Zeitreihen pro Bildposition durch die Prozessoren der Grafikkarte gleichzeitig analysiert (Abb. 2a). Zur Aggregation der klassifizierten Pixel zu (Virus-)Objekten erfolgt eine Fraktionierung der Bildebene in parallel zu verarbeitende Frames (Abb. 2b). Die Thread-basierte Verarbeitung der PAMONO-Daten geschieht dabei komplett auf der Grafikkarte, um Verzögerungen durch Speichertransfers zwischen der Grafikkarte und dem Analyse-rechner auszuschließen.

Der initiale Schritt zur Verarbeitung der PAMONO-Daten ist eine Bildverbesserung in Form einer Hintergrundbereinigung und Rauschreduktion. In diesem Schritt wird insbesondere der zeitlich nicht bzw. langsam variierende Bildanteil durch Subtraktion eines Hintergrundes entfernt, um Partikelanhaftungen vi-

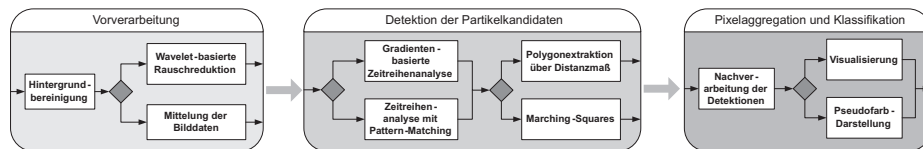
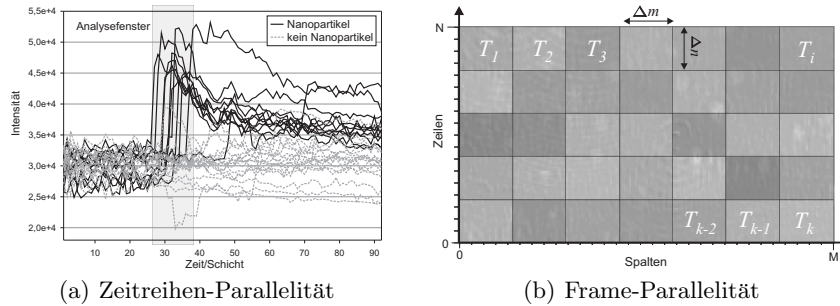


Abb. 1. Schematische Darstellung der GPGPU-Verarbeitungspipeline.

suell und maschinell erfassbarer zu machen. Weiterhin findet in diesem Schritt eine Glättung der zeitlichen Intensitätsvariation (Rauschreduktion) mittels Wavelets statt. Die Algorithmen der Vorverarbeitung können bis auf die einzelne Pixelposition parallelisiert werden, d.h. bei  $M \times N$ -großen Bildern in der aufgenommenen Bilderserie, werden  $M \cdot N$  parallele Threads (Zeitreihen) ausgeführt (Abb. 2a), die automatisch auf die verschiedenen Rechenkerne der Grafikkarte verteilt werden.

Im Anschluss an die Vorverarbeitung der Daten erfolgt die Differenzierung jedes Pixels aufgrund seiner zeitlichen Intensitätsvariation (Abb. 2a) in die Klassen (Virus-)Partikel (positiv) und Hintergrund (negativ). Ein Virus bzw. Nanoobjekt (schwarze Kurven) manifestiert sich durch einen im Verhältnis zu den Hintergrundstrukturen (graue Kurven) prägnanten Sprung im Intensitätsprofil. Die Pixelklassifikation in diesem Schritt basiert auf einem kombinierten Ansatz aus einer gleitenden Sprungdetektion im Sinne eines gleitenden Mittelwertprozesses und auf der Berechnung der Ähnlichkeit eines Zeitreihenabschnittes zu einer musterhaften Sprungfunktion. Die Algorithmen dieser Pipelinestufen werden analog zur Vorverarbeitung parallelisiert.

Der letzte Schritt aggregiert positiv klassifizierte Pixel (Abb. 3a, schwarz markierte Pixel) in zusammenhängende Bereiche bzw. Segmente und klassifiziert jedes dieser Segmente unter Bewertung von Formfaktoren wie beispielsweise Kreishaftigkeit. Die Granularität der Parallelisierung erfolgt in dieser Stufe der Verarbeitungspipeline unter Berücksichtigung der Zerlegung der Frames in gleichgroße Teilbilder (Abb. 2b), die in parallelen Threads verarbeitet werden. Falls sich Segmente über mehrere Teilbilder erstrecken, werden die einzelnen Berechnungen synchronisiert und die Verarbeitung in einem Thread fortgesetzt. Die als Nanoobjekt (Virus) klassifizierten Bildbereiche sind in der Abbildung 3a durch weiße Polygone visualisiert.



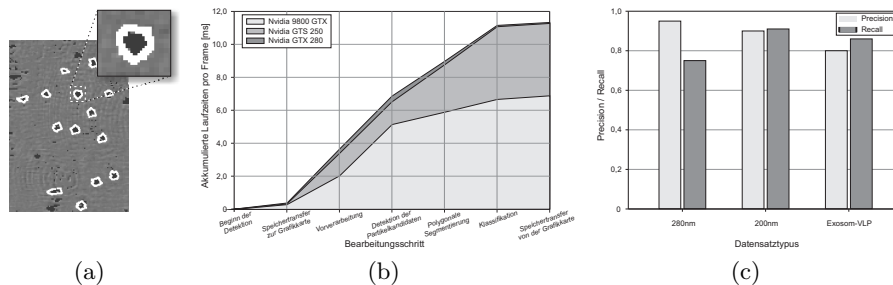
**Abb. 2.** Exemplarische Darstellung der verschiedenen Ebenen der parallelen Verarbeitung eines Bildes der Größe  $M \times N$ : (a) Prinzip der parallelen Analyse der  $M \cdot N$  Zeitreihen und (b) parallele Verarbeitung von  $k$  Teilbildern der Größe  $\Delta m \times \Delta n$ .

### 3 Ergebnisse

Die Evaluierung erfolgte ausgehend von einem PC-Komplettsystem (Intel Core 2 Duo 2.8GHz, 4GB, Nvidia GTS 250) und einer per Firewire verbundenen CCD-Kamera (Kappa DX 4 – 1020). Als Testdatenfundus standen drei exemplarische Typen von Datensätzen (inkl. manueller Segmentierung) mit 4316 Bildern zur Verfügung. Zwei Typen von Datensätzen mit einer Auflösung von  $1000 \times 566$  Pixeln respektive  $1000 \times 367$  Pixeln beruhen auf synthetischen Partikeln der Größe  $200nm$  bzw.  $280nm$ , für den dritten Typus mit einer Auflösung von  $1000 \times 295$  Pixeln wurden virusähnliche Partikel (Virus-like Particles) verwendet. Unabhängig vom betrachteten Datensatz konnte aus der Evaluierung abgeleitet werden, dass die Verarbeitungsgeschwindigkeit der Pipeline für die maximale Datenrate der Kamera (30 fps bei  $800 \times 600$  Pixeln) ausreichend ist und damit eine Echtzeit-konforme Auswertung prinzipiell garantiert werden kann.

Initial wurde die Erkennungsgüte der automatisch detektierten Strukturen (Abb. 3a) mit den Referenzsegmentierungen innerhalb einer ROC-Analyse bewertet (Abb. 3b): Sofern eine Überschneidung zwischen den extrahierten und den manuell bestimmten Polygonen vorlag, wurden diese als True-Positives bezeichnet, nicht geschnittene Polygone aus der automatischen Bestimmung als False-Negatives und nicht geschnittene Polygone aus der Referenzsegmentierung als False-Negatives. Sowohl die synthetischen als auch die realkonformen Nanoobjekte werden in ausreichender Güte erkannt. Dass die Identifikation der synthetischen Strukturen eine geringere Ausprägung falsch positiver Strukturen aufweist, liegt an der bedingt besseren Ausprägung in der Darstellung.

Zur näheren Analyse der Skalierbarkeit und Performanz der Verarbeitungspipeline wurde zudem eine Variation der Grafikkartentypen (Nvidia GTS 250 und 9800 GTX: 128 Rechenkerne; Nvidia GTX 280: 240 Rechenkerne) vorgenommen. Gemäß der Auswertung – akkumulierte Laufzeit in Millisekunden pro Verarbeitungsschritt der Pipeline – kann eine Skalierbarkeit der Geschwindigkeit mit der Anzahl der Rechenkerne festgestellt werden (Abb. 3c). Aus der Aufschlüsselung der einzelnen Laufzeiten ist erkennbar, dass die Detektion der



**Abb. 3.** (a) Darstellung der Partikelkandidaten (schwarz) und der zu Segmenten aggregierten Pixel (weiß); Statistischer Vergleich zur (b) Klassifikationsbewertung für drei exemplarische Datensätze und (c) zur Laufzeit der verschiedenen Pipelineschritte in Abhängigkeit zur Grafikkarte.

Partikelkandidaten einen prägnanten Anteil der Bearbeitungszeit innerhalb der Pipelineverarbeitung aufweist (2 – 3 ms). Die Gesamtlaufzeit pro Frame variiert zwischen 7 ms (Nvidia GTX 280) und 11 ms, d.h. es können zwischen 90 und 142 Frames pro Sekunde verarbeitet werden – bei der CPU-basierten Implementierung nur ca. 10 Frames pro Sekunde. Die Auswertung der GPU-basierten Detektion zeigt daher, dass die Kamera (maximale Übertragungsrate der Kamera: 30 fps bei  $800 \times 600$  Pixeln) momentan der Flaschenhals der Anwendung ist, die GPU-basierte Detektion hingegen auch effizient genug ist, um einen höheren Datendurchsatz zu ermöglichen.

## 4 Diskussion

Ausgehend von der Motivation, eine Echtzeit-konforme Detektion von Nanoobjekten (insbesondere Viren) für die neuartige PAMONO-Technik zur Verfügung zu stellen, wurden in dieser Arbeit GPGPU-basierte Verarbeitungsalgorithmen innerhalb einer hoch-parallelen Pipeline zur Anwendung gebracht. Es konnte gezeigt werden, dass die Verarbeitungspipeline bis zu 142 Bilder pro Sekunde verarbeiten kann und damit die Verarbeitung im Moment sogar oberhalb der maximalen Datenrate der aktuell verwendeten Kamera liegt. Durch die integrierte Kombination von zur parallelen Verarbeitung ausgerichteten Algorithmen, u.a. zur Wavelet-basierten Rauschreduktion und merkmalsbasierten Klassifikation auf Zeitreihen, kann trotz der performanten Verarbeitung weiterhin eine hohe Erkennungsgüte von Nanoobjekten erreicht werden. Zusammenfassend ist festzuhalten, dass das vorliegende Verfahren die Voraussetzungen an ein In-situ-Virusdetektionssystem erfüllt. Trotz der subjektiv recht guten Ergebnisse besteht die Notwendigkeit, die vorliegenden Algorithmen weiter zu optimieren, insbesondere im Hinblick auf eine weitergehende Elimination von Artefakten in den Bildfolgen und einer frühzeitigen Einschränkung des Suchbereichs. Zudem sollen verstärkt hybride Verfahren zwischen Template-basierten und Partitions-basierten Ansätzen sowie der Verwendung von Strukturparametern von Partitions-basierten Verfahren integriert werden. Im Hinblick auf ein angestrebtes portables Virusdetektionssystem ist eine Adaptierung an die speziellen Anforderungen mobiler Grafikchips (z.B. Nvidia ION) geplant.

## Literaturverzeichnis

1. Erickson D, et al. Nanobiosensors: optofluidic, electrical and mechanical approaches to biomolecular detection at the nanoscale. *Microfluid Nanofluidics*. 2008;4(1):33–52.
2. Zybin A, et al. Real-time detection of single immobilized nanoparticles by surface plasmon resonance imaging. *Plasmonics*. 2010;5:31–5.
3. Weichert F, et al. Signal analysis and classification for surface plasmon assisted microscopy of nanoobjects. *Sens Actuators B Chem*. 2010;151:281–90.
4. Dasika G, et al. MEDICS: ultra-portable processing for medical image reconstruction. In: *Proc PACT*. ACM; 2010. p. 181–92.