

Diss. ETH No. 12191

**Quantitative Stereo Vision for the Stereo Light
Microscope:**
An Attempt to Provide Control Feedback for a Nanorobot
System

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Technical Sciences

presented by
GAUDENZ DANUSER
Dipl. Verm. Ing. ETH
ETH - Swiss Federal Institute of Technology
born 15th of February, 1969
citizen of Mastrils and Pontresina(GR)

accepted on the recommendation of
Prof. Dr. Olaf Kübler, examiner
Prof. Dr. Hans Tiziani, co-examiner
Dr. Markus Stricker, co-examiner

April 1997

Abstract

Quantitative microscopy is a prerequisite to fulfill the sensing tasks that arise with the demand to not only visually observe microscopic specimen but to measure and manipulate structures at the microscopic scale. With the term *quantitative microscope*, sensor systems are addressed which consist of a microscope equipped with devices for capturing digital images and with tools for processing the image data in terms of extracting a geometric representation of the observed scene. This thesis deals with the implementation of a quantitative *stereo light microscope* dedicated for the survey of generic and dynamic scenes.

The considered microscope system is used predominantly, but not exclusively, for the visual control of a robot aimed at the manipulation of μm -sized objects with sub-micron precision. Although the research is focused on this application, much of the material is generic and can be adapted for any type of measurement task at this scale. Particular emphasis is put on the discussion of algorithms for visual inference which are robust with respect to the often difficult contrast situation occurring in light microscopy.

The first part of the thesis reports on investigations of the *resolution* and *accuracy* performance of the stereo light microscope. The discussion on resolution shows that image processing has a strong impact on the total resolution performance of the system. In general, this result is not new. However, the contribution of this work is to investigate the impact of image processing not just on the signal level but on a more symbolic level as it is typically performed in the field of computer vision. This kind of visual inference is needed to obtain a geometric representation of the observed scene. On the one hand, the step of feature extraction bears substantial limits on resolution. On the other hand, the concepts of computer vision allow for effectively including a priori knowledge about the observed scene, which yields an improved resolution performance.

The conceptual material is validated in practical experiments. Applications to micro-manipulation and quality control in the microscopic domain are reported where benefit is derived from the high resolution performance of the implemented computer vision system integrated into the stereo light microscope. The microscope used for all experiments has a numerical aperture of $\text{NA} = 0.136$. The empirically determined resolution in localization amounts to 20 nm and the resolution in separability achieved is better than 1/3 of the Rayleigh limit in cases use can be made of the knowledge about the scene dynamics and the structure of the objects.

The accuracy performance of the microscope depends on the quality of the ge-

ometric imaging model which describes the positional relation between an object feature and its corresponding image feature. The parameters of the imaging model are determined in a calibration procedure using photogrammetric techniques. The calibration framework allows for an efficient testing of the parameter estimates. Different types of imaging models are evaluated by analyzing the stability of the parameter estimates involved and their respective accuracy performance. The accuracies achieved amount to 700 nm laterally and 1.6 μm axially in the overall volume of the working space and to 250 nm laterally and 600 nm axially within a neighborhood of 100 μm^3 .

In the second part of the thesis, the geometric reconstruction of objects from stereo and the application of the framework to the task of tracking and controlling a micro-pipette mounted on one of the robot arms is addressed. The focus is directed towards the problem of tracking the pipette in the near-field of the object, where the patterns of the pipette and the approached specimen interfere. This situation bears difficulties for segmenting the pipette from the object. Nevertheless, it is demonstrated that rigorous multi-cue processing, *i.e.*, the combination of various visual cues, such as feature evidence, motion evidence, and stereo, together with prior knowledge about the 3D structure of the pipette and the imaging model, yields high accuracy and separability.

Kurzfassung

Die quantitative Mikroskopie bietet wesentliche Hilfsmittel zur Lösung der verschiedenen Sensorik-Aufgaben, welche mit der Vermessung und Manipulation von Mikro-Strukturen verbunden sind. Ein quantitatives Mikroskop besteht aus einem Mikroskopsystem, Video-Kameras und einer Einheit zum Digitalisieren der Bilder. In diesen Punkten unterscheidet sich das quantitative Mikroskop nicht von einem herkömmlichen Videomikroskop. Das entscheidende zusätzliche Element in einem System für *quantitative* Mikroskopie ist die Bildverarbeitung, wobei Algorithmen zur Extraktion einer geometrischen Beschreibung des beobachteten Arbeitsraums im Vordergrund stehen. Die vorliegende Arbeit beschäftigt sich mit einem speziellen Typ quantitativer Mikroskopie – der quantitativen *Stereo-Lichtmikroskopie* – welche zur Beobachtung und Vermessung von prinzipiell beliebigen mikroskopischen Szenen verwendet wird.

Dieses Mikroskop wird vorwiegend, aber nicht ausschliesslich, für die visuelle Steuerung eines Roboters eingesetzt. Dieser Roboter dient zur Manipulation von Objekten von einigen zehn Mikrometer Grösse, wobei eine Manipulationsgenauigkeit von einigen hundert Nanometer angestrebt wird. Obwohl sich die Forschungsarbeit hauptsächlich auf diese spezielle Steuerungsaufgabe konzentriert, sind viele der vorgestellten Verfahren direkt für allgemeinere Messaufgaben in der Mikroskopie anwendbar. Die Arbeit beschäftigt sich insbesondere mit generischen Algorithmen zur Extraktion visueller Information aus Licht-Mikroskopbildern, welche sehr oft ungünstige Kontrasteigenschaften aufweisen.

Der erste Teil der Arbeit behandelt Fragen der Auflösung und Genauigkeit in der quantitativen Stereo-Lichtmikroskopie. Die Diskussion unterstreicht den entscheidenden Einfluss der digitalen Bildverarbeitung auf das gesamte Auflösungsvermögen eines Mikroskops. Diese Aussage an sich ist nicht neu. Der wesentliche Beitrag dieser Arbeit liegt in der Untersuchung des Auflösungsverhaltens von Algorithmen zur Extraktion visueller Daten auf einer abstrakten Informationsebene. Bezüglich der Auflösungsfragen werden in der Literatur sehr oft nur Methoden der reinen Signalverarbeitung beachtet, obwohl solche für die Extraktion einer geometrischen Repräsentation der beobachteten Szene nicht ausreichen. Zur Lösung dieser Aufgabe müssen Bildmerkmale in einem grösseren Zusammenhang detektiert und lokalisiert werden, was eine symbolische Beschreibung der betrachteten Strukturen mit Methoden der Computer Vision fordert. Um das Auflösungsvermögen eines quantitativen Mikroskops vollständig zu behandeln, müssen deshalb nebst den bekannten optischen Auflösungsgrenzen alle Schritte der Merkmalsextraktion – von der Si-

gnalfilterung bis zur Gruppierung und Beschreibung von komplexen Bildstrukturen – bezüglich dieser Fragestellung untersucht werden. Verschiedene Prozeduren zur Merkmalsextraktion führen zu einer wesentlichen Begrenzung des Auflösungsvermögens. Gleichzeitig erlauben aber die Methoden der Computer Vision ein effizientes Ausnutzen von Vorinformation bezüglich der beobachteten Objekte und der Abbildungseigenschaften des Mikroskops, wodurch dessen Auflösung verbessert werden kann.

Praktische Beispiele erläutern die konzeptionell geführte Diskussion. Es werden Messaufgaben im Zusammenhang mit der Manipulation und Qualitätskontrolle von Mikro-Objekten vorgestellt, zu deren Lösung die Methoden der Computer Vision konsequent ausgenutzt werden. In einem Stereo-Lichtmikroskop mit einer numerischen Apertur $NA=0.136$ wird empirisch eine Bewegungsauflösung von 20 nm und eine Separabilitätsauflösung von besser als $1/3$ der Rayleigh-Grenze nachgewiesen. Letzteres lässt sich durch Vorwissen über die Dynamik und die geometrische Struktur der Objekte erreichen.

Die Genauigkeit des Mikroskops ergibt sich aus den Genauigkeiten des Abbildungsmodells und der geometrischen Beschreibung von Korrespondenzen zwischen Objekt- und Bildmerkmalen. Die Parameter des Abbildungsmodells werden mithilfe der Kalibrierung des Mikroskops bestimmt. Dabei werden Techniken aus der Photogrammetrie eingesetzt. Das vorgestellte Kalibrierungsverfahren erlaubt ein effizientes Testen der geschätzten Parameter. Dadurch wird es möglich, verschiedene Abbildungsmodelle in Bezug auf deren Stabilität und Genauigkeit zu analysieren. Über das gesamte Volumen des Arbeitsraums ergibt sich eine Messgenauigkeit von lateral 700 nm und axial $1.6 \mu\text{m}$. Die Nachbarschaftsgenauigkeit in einem Würfel von $100 \mu\text{m}^3$ beträgt lateral 250 nm und axial 600 nm.

Der zweite Teil der Arbeit beschäftigt sich mit der Stereo-Rekonstruktion von Objekten. Die Problemstellungen sind durch die konkrete Anwendung der 3D Daten für das visuelle Tracking und Steuern einer auf dem dreiarmligen Roboter montierten Mikro-Pipette definiert. Die Diskussion konzentriert sich vorwiegend auf das Tracking der Pipettenspitze im Nahfeld eines Objektes, wobei hier das *Nahfeld* die Zone bezeichnet, in der die Bilder von Werkzeug und Objekt interferieren. Dadurch ergeben sich Schwierigkeiten in der Segmentierung der Pipette. Dank der simultanen Verarbeitung visueller Information mit unterschiedlichem Inhalt (z.B. Merkmals- und Bewegungsevidenzen, Stereoinformation und Vorwissen über die 3D Geometrie der Pipette sowie das Abbildungsmodell des Mikroskops) ist es trotz der schwierigen Kontrastverhältnisse möglich, die Spitze mit hoher Genauigkeit und Separabilität in Bezug auf das angefahrene Objekt zu lokalisieren.