Stereo Light Microscope Calibration for 3D Submicron Vision

Gaudenz Danuser
Communication Technology Lab, Image Science Group
Swiss Federal Institute of Technology, ETH
CH-8092 Zürich
email: danuser@vision.ee.ethz.ch
Commision V, Working Group 2

KEY WORDS: Three-dimensional, Calibration, Bundle, Adjustment, Stereoscopic, Light Microscopy

ABSTRACT

The rapid development in micro- and nanotechnology has introduced new challenges to photogrammetry and computer vision. Microscopic imagery has to be seriously investigated by these communities. This paper represents a first step in this direction. It addresses the calibration of Stereo Light Microscopes. Various new aspects have to be investigated when trying to develop computer vision and photogrammetric approaches for this type of sensor. Normally, in macroscopic imagery the contrast and image quality can be controlled to some degree. Due to poor contrast and worse quality of the microscopic imagery the reliable and precise localization of image coordinates becomes a much more difficult task. A high level vision algorithm followed by a modified Least Squares Template Matching adequately solves these problems. In addition, a new imaging model for Stereo Light Microscopy is introduced. It is based on the weak perspective situation in microscopy and includes a new distortion term describing the non paraxial imaging of a Stereo Light Microscope. The results of various calibration runs demonstrate the suitability of this new approach. The new model is explicitly compared with the performance of standard imaging models used in computer vision and photogrammetry, when applying them to microscopy. Relative accuracies of laterally $1:1000$ and vertically $2:100$ are obtained in the object space. In contrast, the relative accuracy in the image space reaches $2:10000$. The discrepancy between the relative accuracies in object and image space results from the limited precision of the underlying calibration standard.

1 INTRODUCTION

The development of technologies to handle and manufacture complex systems on the micro- and nanometer scale is a grand challenge for science and engineering. This development is accompanied by a quickly growing demand to precisely measure three dimensional micro structures. For many applications (quality control, object description, manipulation of micro-structures, etc.), it is of utmost importance to gain geometrical information from a relatively large field of view. Thus, a new challenge for the photogrammetry and machine vision communities is to extend their techniques for extracting accurate spatial data to make them applicable to microscope imagery.

Quantitative techniques for stereoscopic measurements have largely been limited to the Stereo Scanning Electron Microscope (SSEM). The idea of introducing photogrammetry to the SSEM came soon after the development of the first instruments [Maune 1973] and it is still in progress [Gleichmann *et al.* 1994]. The state of the art in photogrammetric treatment of SSEM imagery is summarized in [Ghosh 1989].

In contrast to the SSEM, stereoscopic measurements using Stereo Light Microscopes (SLM) have not been deeply investigated. The limited depth of field in a SLM is most likely the key factor that discouraged intensive research in SLM imagery. To circumvent this problem, the optical community developed other techniques to reconstruct a 3D micro space: confocal microscopy, optical sectioning, laser confocal microscopy, etc. [Taylor *et al.* 1992]. Another difficulty in quantitative stereo light microscopy is to obtain images of sufficiently high contrast. Specular reflections and specimen which may be transparent further diminish the image quality. Above all, the low Numerical Aperture (N.A.) limits the sensor resolution to about $2\,\mu$m. However, SLM provide considerable advantages over the microscope types mentioned above. Real-time observation of a 3D scene is impossible with all the other microscopes as their data acquisition relies on a scanning procedure. Furthermore, with a SLM nearly every object, e.g. living cells, can be observed. In addition, the low N.A. results in a quite large working distance between the specimen and the front lens. This may become important when observing objects with a relatively large vertical extension or if they have to be handled while viewing at them.

This paper addresses the problem of 3D measurements with SLM. Studies of the same topic can hardly be found in the literature. To the best of my knowledge, the only relevant paper has been presented by [Kim *et al.* 1990]. It describes a method for 3D reconstruction using a SLM. The authors demonstrate their approach on biological objects. Their imaging model is based on an orthographic projection and they operate with the viewing angles given in the technical specifications of the manufacturer. Neither an initial calibration nor an a posteriori analysis of the 3D accuracy obtained by their approach has been carried out. They emphasize the limitation of their approach and propose that a refined imaging model, as well as powerful statistical tools, may improve the quality of the results. It is exactly the purpose of my work to focus on the description of microscopic imaging with a photogrammetric approach.

A first paper about this topic has been presented by [Danuser and Kübler 1995]. A thorough numerical analysis allowed us to define the set up for a novel Bundle Adjustment. Since then, the proposed calibration procedure has been successfully implemented on a fully operational micro vision system. Compared to the first paper, this one addresses more the practical aspects of microscope calibration. It discusses the system hardware (section 2) and the key problems of the algorithm, namely the image coordinate acquisition (section 4) and the parameter estimation in the Bundle Adjustment (sec-

tion 3). Finally, in section 5, I discuss results from various calibration runs.

2 SYSTEM HARDWARE AND OVERVIEW

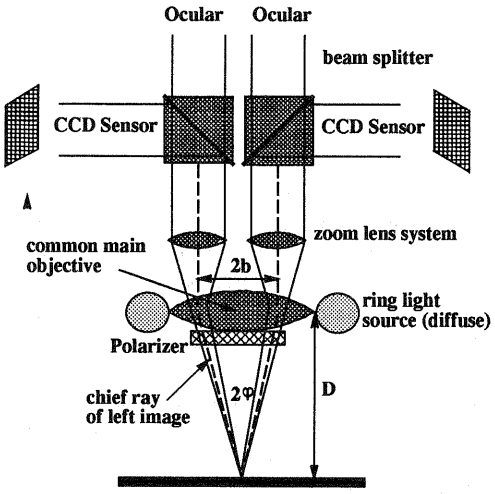

Figure 1: Cross section of the common main objective lens (CMO) type stereo microscope.

In our project, the SLM to be calibrated is integrated into a nano robot system [Codourey *et al.* 1995]. The stereo effect of the microscope is generated by refraction in a common main objective (CMO) lens. A scheme of the fundamental optical subsystems of this microscope type is shown in figure 1. The specific advantages of the CMO stereo image formation for our robot system with respect to two separate objective lenses (Greenough system) are discussed in [Danuser and Kübler 1995]. For detailed information the reader can refer to [Richardson 1991]. The relatively low N.A. of 0.136 results in enough space ($D = 6.3\,\mathrm{cm}$) for the robot manipulation tools. The serious disadvantage of the low N.A. lens is its poor resolution of $2.4\,\mu\mathrm{m}$. Fortunately, for the 2D case we have demonstrated in [Mazza *et al.* 1995] that as long as the target object itself is larger than the optical resolution, motion and deformation of micro-structures can be analyzed down to a measurement limit of at least $50\,\mathrm{nm}$.

The digital imagery is collected by two monochrome CCD cameras with 756x582 elements. The crucial point in microscopy is the illumination. The optimum setup for our applications consists of a ring light source enclosing the CMO lens with a supplementary diffuser. Before entering the optical system, the light waves are polarized. This strongly reduces the appearance of glancing spots in the imagery. The latter problem also requires to keep the illumination power as low as possible. Thus, to get images of sufficient brightness, cameras with optional frame integration are used.

For the geometric calibration with Bundle Adjustment we employ planar gratings. The required 3D point distribution is then generated by lifting the grating within the depth of focus. The gratings used for simulations in [Danuser and Kübler 1995] turned out to be inappropriate for the real system. A pattern of squares is now employed (see figures 3-B and 3-F). The image contrast results from the different reflectivity of Silicon (dark) and Siliconoxyd (bright). Such gratings are manufactured with photo lithography. The accuracy of patterns produced by this process is limited by the mechanical tolerance of the mask as well as by the quality of the lithography itself. The original photo lithographic mask has been produced with an optical pattern generator which yields an accuracy of about $3\,\mu\mathrm{m}$. It will be demonstrated in section 5 that this accuracy is the limiting factor of my calibration.

Calibration runs with simulated data proved that for sufficient determinability of all the parameters, images of four views of the grating rotated 90° between each view have to be acquired. In our setup the lifting and rotation of the grating are carried out by the robot system. However, for my calibration procedure it is sufficient to have approximate rotations. In particular, the mathematical model of the Bundle Adjustment takes into account that the rotation is not concentric with respect to the optical axis. Thus, all the needed motions of the grating could be achieved with off the shelf micrometer tables.

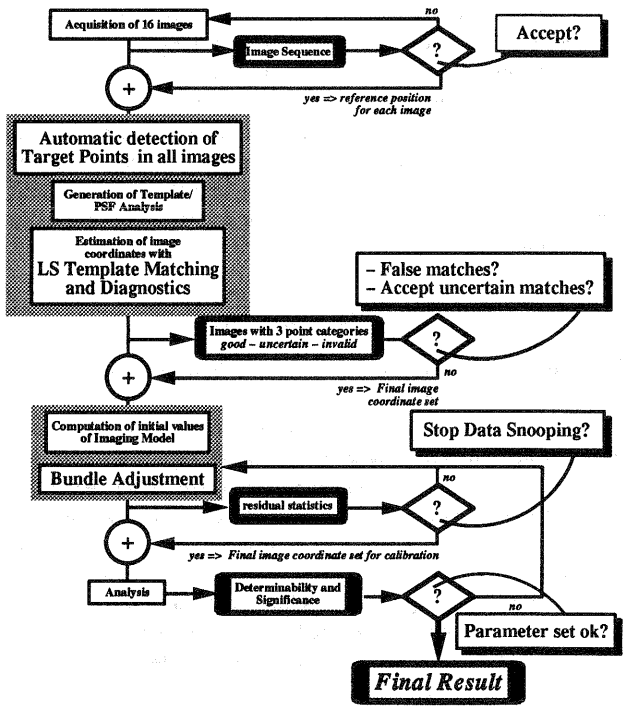

Figure 2: Scheme of processes, data flow and user interaction in the calibration procedure.

Figure 2 schematically shows the various processes, the data flow and the user interaction of the calibration procedure. The left column contains the computation steps that run completely automatically. The kernel of the algorithm consists of two blocks outlined in light gray boxes. The first block is responsible for the detection and for measuring the precise location of target points in all the images. The processes of the second block compute the parameters of the imaging model. Both blocks are based on the properties of SLM imagery and therefore require a completely new photogrammetric implementation. The image coordinate acquisition is described in section 4. The mathematical model of the Bundle Adjustment is briefly outlined in section 3.

User interaction, symbolized by question marks in figure 2, is reduced to four stages. The dark boxes in the center of the flow chart represent the data interfaced to the user. The only required user interaction is to accept or reject the images, i.e., the user has to decide whether the illumination and

the sharpness are appropriate. The three optional user interactions (accepting uncertain matches, supervision of data snooping and refinement of the parameter set) may be valuable to improve the quality of the final results. They can be omitted in situations where fast and automatic calibration is more important than results of highest quality.

3 MATHEMATICAL FRAMEWORK

To provide the context for sections 4 and 5 the mathematical framework is outlined very briefly. An extended version of this discussion can be found in [Danuser and Kübler 1995] and for the derivations I refer to [Danuser 1995].

The Bundle Adjustment involves four types of observation equations. The relation between the image coordinate observations and the 3D position of the target points on the calibration standard is introduced with

$$\begin{aligned}(\xi^{\text{view}}_{i,\Omega} - \xi^{\text{view}}_{0}) \cdot [1 - \sum_{k=1}^{\infty} \vartheta^k] + e_{\xi^{\text{view}}_{i,\Omega}} \\ = \; M^{\text{view}} R_{\varphi^{\text{view}},\omega^{\text{view}},\kappa^{\text{view}}} \cdot (S_\Omega \cdot \tilde{x}_i + x_{0,\Omega}) \\ - \; \delta_\xi(K_1^{\text{view}}, K_2^{\text{view}}, P_1^{\text{view}}, P_2^{\text{view}}, E, a^{\text{view}}, s^{\text{view}}) \quad (1)\end{aligned}$$

The index i runs over all target points. Ω is the rotation angle of the grid with respect to the object coordinate system. My calibration procedure is based on four rotation angles, each approximately 90° apart. ξ is the observed image coordinate vector. Its index *view* specifies whether the point is observed in the left or right image. The corresponding 3D position $\tilde{x}_i$ is defined with respect to a coordinate system rotating together with the calibration grid. Note that through the formulation of $(S_\Omega \cdot \tilde{x}_i + x_{0,\Omega})$ a non concentric rotation of the calibration grid in the superior fixed object coordinate system is introduced. The translation between the grid coordinate system and the object system is defined by $x_{0,\Omega=0} = 0$.

Equation (1) incorporates the weak perspective situation in microscopy. The nonlinearity of perspective imaging therefore disappears and a perspective distortion term $\sum_{k=1}^{\infty} \vartheta^k$ is introduced, instead. Its magnitude depends on the lateral and vertical position of the target point with respect to a perspective distortion free point. The latter's image coordinates are given by ξ_0^{view}. The expansion of the distortion series is stopped when $\vartheta^{2k} \cdot (\xi^{\text{view}}_{i,\Omega} - \xi^{\text{view}}_{0}) < \text{diag}(Q_{\xi\xi})$ [1] which holds in most cases for $k = 2$. This approach guarantees that the inaccuracy of the weak perspective model becomes smaller than the accuracy of the image coordinates.

M^{view} symbolizes the magnifications of the left and right images, respectively. R^{view} is an orthonormal matrix describing the orientation of the image coordinate system with respect to the object coordinate system. The choice of the primary, secondary and tertiary angle has been adapted to the microscopic situation retaining the technical meaning for $\varphi^{\text{view}}, \omega^{\text{view}}, \kappa^{\text{view}}$ in standard photogrammetry.

The distortion function δ_ξ contains the well known radial and decentering distortion for each view. Its coefficients are — according to the technical terms in photogrammetry — $K_1^{\text{view}}, K_2^{\text{view}}, P_1^{\text{view}}, P_2^{\text{view}}$. The terms s^{view} and a^{view} compensate for scale and shear distortion in the image coordinate frames. The CMO distortion function $\delta_\xi(E)$ has been derived to particularly compensate for geometric deformations originating from the non paraxial imaging (see figure 1). Even in the case of perfect lenses, non paraxial optics leads to image deformations. To get highly accurate vertical positions, the introduction of this term is of utmost importance, as theoretically demonstrated in [Danuser and Kübler 1995] and empirically verified in section 5. A remarkable behavior of this distortion type is that the very same parameter E is involved in the distortion model of the left and the right view. Therefore, in contrast to macroscopic photogrammetry, the two image spaces of the stereo rig are closely related to each other. In particular, this renders a separate calibration of the left and the right imaging function impossible.

The coordinate residuals $e_{\xi^{\text{view}}_{i,\Omega}}$ may origin either from noise in the image coordinate measurements, from unmodeled systematic errors or from the position errors of the control points on the photo lithographic grating. An adequate error model of the control point precision has to be introduced:

$$\ell_{\tilde{x}^i} + e_{\ell_{\tilde{x}^i}} = \tilde{x}^i \;\text{ with }\; Q_{\ell_{\tilde{x}}\ell_{\tilde{x}}} \quad (2)$$

To get reliable results the relation between $Q_{\xi\xi}$ and $Q_{\ell_{\tilde{x}}\ell_{\tilde{x}}}$ must be very thoroughly determined. This can be achieved by dividing the target points into a set of control points and a set of check points. The empirical RMS of the check points contains valuable information to define $Q_{\ell_{\tilde{x}}\ell_{\tilde{x}}}$.

The full parameter vector p to be estimated in the Bundle Adjustment includes p',

$$\begin{aligned}p' \;=\; [&\varphi^{\text{view}}, \omega^{\text{view}}, \kappa^{\text{view}}, M^{\text{view}}, \\ &K_1^{\text{view}}, K_2^{\text{view}}, P_1^{\text{view}}, P_2^{\text{view}}, E, a^{\text{view}}, s^{\text{view}}]^T\end{aligned}$$

all the grid rotation angles Ω, the corresponding translation vectors $x_{0,\Omega}$ and all the coordinates $\tilde{x}^i$ of the target points on the calibration grid.

The observation equations (3) and (4) allow me to influence the estimation of the parameter sub-vector p'.

$$\ell_{p'} + e_{\ell_{p'}} \;=\; p' \;\text{ with }\; Q_{\ell_{p'}\ell_{p'}} \quad (3)$$

$$e_c \;=\; Bp' \;\text{ with }\; Q_{cc} \quad (4)$$

With the equations of type (3), I can gradually turn on and off the estimation of a certain parameter. With those of type (4), I can specify similarities between parameters of the same type in the left and right view, where B is the corresponding similarity matrix.

A Least Squares approach using equations (1) to (4) is implemented as in classical Bundle Adjustment. Data Snooping is employed to efficiently detect and eliminate outliers in the image coordinate observations. Each individual image coordinate residual is compared with the corresponding diagonal element of the residuals covariance matrix Q_{vv}. This is a well known standard procedure for blunder detection in the photogrammetric Bundle Adjustment. Due to the limited depth of field the estimation process suffers from a considerable numerical weakness. Simulations in [Danuser and Kübler 1995] turned out that gross errors in the image coordinates can cause serious numerical oscillations in the parameter space. Such oscillations indirectly affect the quality of blunder detection, too. To solve this problem, as many erroneous image coordinate observations as possible must be detected before entering the Bundle Adjustment. An expensive but fully automatic algorithm has been implemented to achieve this demanding task. The next section describes the most important steps in this procedure.

[1] $Q_{\xi\xi}$ is the cofactor matrix of the image coordinates.

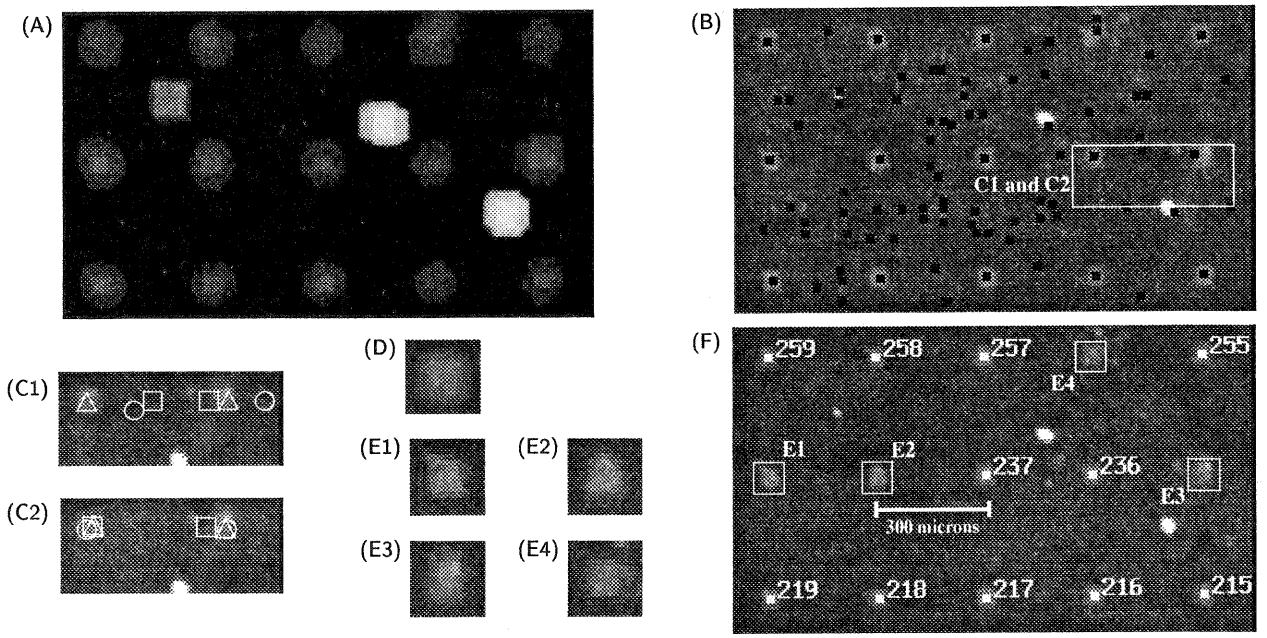

Figure 3: Image coordinate acquisition demonstrated with a 300x100 Pixel border region of a full view from the calibration standard. (A) continuous response field of the interest operator. (B) the set $\mathcal{S}$ of local maxima overlaid on the original image. (C1) and (C2) two iterations of the ICP. $\bigcirc \in \hat{\mathcal{P}}_k$, $\triangle \in \hat{\mathcal{P}}_{k+1}$,$\square \in \mathcal{P}_k$ extracted from $\mathcal{S}$ through a closest point search. The inaccuracy of the affinity A_k lets the algorithm pick out the wrong local maximum for the left point in (C1). However, the update of the affinity, also controlled by many other image points, results in a better location of the points in $\hat{\mathcal{P}}_{k+1}$. $\triangle$ of (C1) becomes $\bigcirc$ of (C2). After a second update of the affinity the $\triangle$ moves to the correct place in (C2) even for the right point where the local maximum $\square$ is slightly misplaced. (D) Template generated by analyzing the PSF of the image. (F) Result of the LSM completed with diagnostics. Points without label are rejected points; (E1) to (E4) detail views of rejected points.

4 IMAGE COORDINATE ACQUISITION

The calibration procedure is run on 16 images, each containing more than 100 target points. As mentioned in section 3, outliers in the image coordinates have to be reliably excluded in an early stage of the procedure. Furthermore, an efficient calibration is only possible with a fully automatic target point detection and location. This becomes a demanding task with light microscope imagery. Not working in a clean room environment, small particles and dust stick on the standard. In addition, scratches originating from the manufacturing and the handling of the grating appear. Both, the detection and the location of the points are affected by such damages.

The target points can be detected by an interest operator. Normally, these operators use a threshold on their local response. An application of this technique to microscope images will fail even if the threshold is determined data driven. Dust and scratches produce much higher image contrast and therefore also higher interest operator responses than target points (see figure 3-A and 3-B). Generally, it is impossible to predict the frequency distribution of the interest operator responses for good target points and of those for corrupt image features of high contrast. In my approach, I first compute a continuous response E of an interest operator (figure 3-A). In a second step its local maxima are detected (figure 3-B). Thus, a set $\mathcal{S}$ of image points is obtained

$$\mathcal{S} = \{\text{local maxima}\,(E)\} \tag{5}$$

Note that the strongest responses do not necessarily appear at the target point positions. I have experimented with several implementations of interest operators. Mainly because of its fastness and its close relation to the normal equations of the Least Squares Template Matching (LSM) the [Förstner and Gülch 1987] operator is used.

The second step of the detection algorithm has to partition $\mathcal{S}$ in two subsets $\mathcal{P}$ and $\overline{\mathcal{P}}$

$$\mathcal{S} = \mathcal{P} \cup \overline{\mathcal{P}} \tag{6}$$

where $\mathcal{P}$ contains those local maxima that correspond to a target point. A priori knowledge about the target point distribution on the calibration standard is introduced. Each point in $\mathcal{P}$ is a unique image point of the known target point set $\mathcal{O}$. The relation between the image point coordinates ξ_i of points in $\mathcal{P}$ and the object coordinates $\tilde{x}_i$ in $\mathcal{O}$ is given by an unknown, point dependent transformation T_i.

$$T_i : \ \tilde{x}_i \longmapsto \xi_i \tag{7}$$

If $\mathcal{O}$ is a set of points all laying on a plane and when the distortion terms in (1) are neglected, a global transformation T for all points can be described by an affinity A. Therefore, simultaneous to the set partitioning, the parameters of the unknown affinity have to be estimated. A method based on an *Iterative Closest Point Search (ICP)* has been implemented to solve this problem. Starting with an approximate transformation A_0 and a subset $\mathcal{O}_0$ a first point set $\hat{\mathcal{P}}_0$ can be computed applying (7). $\mathcal{O}_0$ contains only points mapped close to the image center since those are least affected by errors in A_0. Searching for each point in $\hat{\mathcal{P}}_0$ the closest point in $\mathcal{S}$, a set $\mathcal{P}_0$ is obtained. The unique correspondence between $\hat{\mathcal{P}}_0$ and $\mathcal{P}_0$ now allows me to estimate a better transformation A_1 and its quality $\hat{\sigma}_1$. The further iterations for the computation of A_k and the extraction of $\mathcal{P}_k$ are analogous. The

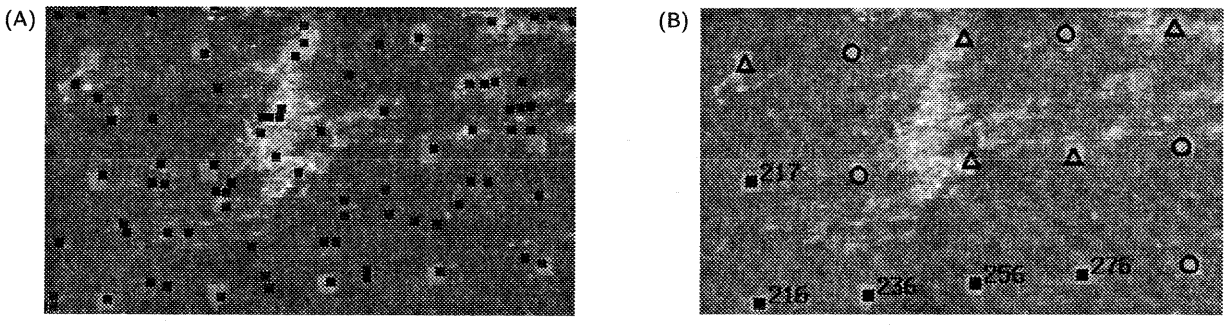

Figure 4: Example of target point detection and location in the case of a rotated and widely damaged grating. (A) local maxima of the interest operator. (B) result of LSM. Points labeled with △ are fully rejected, points with a ○ are declared as uncertain and may be manually accepted, points with a number and a filled □ are automatically accepted.

better A_k, the more can $\hat{\mathcal{P}}_k$ be extended towards the image border. And vice versa, the more correct points are contained in $\mathcal{P}_k$, the more precise A_{k+1} can be estimated. A sample for two iteration steps is given in figures 3-C1 and 3-C2. The iteration is stopped if one of the two criteria (8) or (9) is fulfilled.

$$Pr[\hat{\sigma}_{k-1} > \hat{\sigma}_k] > 95\% \quad (8)$$

$$\mathcal{P}_k \cap \mathcal{P}_{k-1} = \mathcal{P}_k \quad (9)$$

The criterion (8) means that false point correspondences between $\hat{\mathcal{P}}_k$ and $\mathcal{P}_k$ have been established and thus A_k become worse than A_{k-1}. In this case A_{k-1} is presumed to be the best known transformation $\hat{A}$. The criterion (9) means that no new points in $\mathcal{S}$ can be identified as target points and therefore A_k is the optimal transformation $\hat{A}$. Transforming the whole set $\mathcal{O}$ with $\hat{A}$ results in a coordinate set $\mathcal{Q}$ which can be used as initializations for LSM. Note that $\mathcal{Q}$ and $\mathcal{P}$ may differ from each other. The poor contrast of the target point pattern in the image does not guarantee a corresponding local maximum in E for all target points. However, $\mathcal{Q}$ contains the full grating independent of the successful detection of each target point by the interest operator.

A priori knowledge about the shape of the target points and an analysis of the Point Spread Function (PSF) for each image allows me to generate artificial templates (see figure 3-D). This procedure is necessary because the sharpness differs for each of the 16 images due to the small depth of field. The definitive image coordinates are computed applying LSM. A chain of statistical tests (global model test, joint significance of all parameters and parameter subsets, exclusion of non-determinable parameters, parameter testing in the eigenspace) forms a powerful diagnostic tool. Its main purpose is to detect false matches due to target point patterns corrupted by manufacturing damages, scratches or dust (see figures 3-E1 to 3-E4). The mean precision of the point location turned out to be 0.03 pixel.

The high quality of this image coordinate acquisition scheme can be verified in the data snooping procedure of the Bundle Adjustment. My experience from many calibration runs is that less than 0.05% of image points still have to be excluded. Another example in figure 4 demonstrates that the algorithm works well even if the grating is rotated 7° out of the expected orientation (figure 4-B) and if only a few target points result in a well centered local maximum of the interest operator response field (figure 4-A).

5 RESULTS

As in any calibration the suitability and the quality of the underlying standard is the key for a successful procedure. The currently available grating has a target point spacing of $300\,\mu$m (see figure 3-F). The estimation of the imaging model requires at least seven by seven target points within the field of view. Therefore, using this grating, the microscope can only be calibrated on a medium zoom level where the field of view exceeds 2.5×2.5 mm. For higher magnification levels, a grating with $100\,\mu$m spacing is necessary. A new type of grating will be used in the future[2]. The most important modifications of the new gratings are:

- Various zoom level dependent point features on one single wafer.
- Improved contrast by etching Aluminium layers instead of Siliconoxyd
- The mask is produced by a laser pattern generator which is the very best currently available in micro electronic manufacturing. Thus, a point precision all over the standard of 80 nm is expected (compared to the $3\,\mu$m of the mask used for the tests below).

However, even the relatively inaccurate standard allows me to demonstrate the great potential of the new algorithm. The resolution of the microscope is $4\,\mu$m on the zoom level used for the presented tests. The depth of field measures about $330\,\mu$m. This must be compared with the 3.5×3.5 mm field of view. Thus the relative depth of field can be defined as about $1 : 10$. Note that macroscopic close range photogrammetry never is obliged to work with such asymmetric point fields.

Table 1 contains results of calibration runs that investigate various a priori stochastic models and six different imaging models. The a priori model mainly consists of three parameters. The a priori standard deviation (Stdev) of the image coordinates is described by $\sigma_{\xi\eta}$ [3][4]. The Stdev of the lateral and the vertical components of the control point position are determined by σ_{XY} and σ_Z, respectively. From these values the cofactor matrices $Q_{\xi\xi}$ and $Q_{\ell_{\tilde{x}}\ell_{\tilde{x}}}$ are gained. The lateral Stdev is obtained from the specifications of the grating while

[2] Due to some delivery problems of the manufacturer, they are not yet available.

[3] When writing ξ and η separately the components of an image coordinate vector ξ are meant.

[4] For simplicity only one value for both components is shown. In the software σ_ξ and σ_η can be defined, independently.

V	A Priori Model			Imaging Model								Empirical Accuracy						
													Image		Chk Pts		Ctl Pts	
	[Pix]	[μm]		PT	Distortion δ_ξ								[Pix]		[μm]		[μm]	
	$\sigma_{\xi\eta}$	σ_{XY}	σ_Z		E	a	s	K_1	K_2	P_1	P_2	$\hat{\sigma}_0$	μ_ξ	μ_η	μ_{XY}	μ_Z	μ_{XY}	μ_Z
0	0.17	3.0	8.0	-	-	-	-	-	-	-	-	2.05	1.3	1.6	6.8	30.1	8.4	26.1
1	0.03	3.0	5.0	x	x	x	-	x	x	x	x	0.13	0.10	0.07	3.5	16.1	3.8	16.7
2	0.12	3.0	5.0	x	x	x	-	x	x	x	x	0.14	0.10	0.08	3.5	12.0	4.0	9.6
3				x	-	x	x	-	-	-	-	0.17	0.11	0.08	4.0	15.8	3.7	18.9
4				x	-	x	x	x	x	x	x	0.15	0.10	0.08	3.8	10.4	3.5	12.3
5	0.17	3.0	8.0	x	x	x	-	-	-	-	-	0.20	0.12	0.15	3.8	**6.8**	3.5	**6.7**
6				x	x	x	x	-	-	-	-	0.16	0.10	0.08	3.8	10.4	3.6	12.4
7				x	x	x	-	-	-	x	x	0.15	0.10	0.08	3.9	10.5	3.6	12.5
8				**x**	**x**	**x**	-	**x**	**x**	-	-	**0.19**	**0.11**	**0.14**	**4.3**	**6.5**	**3.8**	**6.8**

Table 1: Calibration versions using various estimation and imaging models.

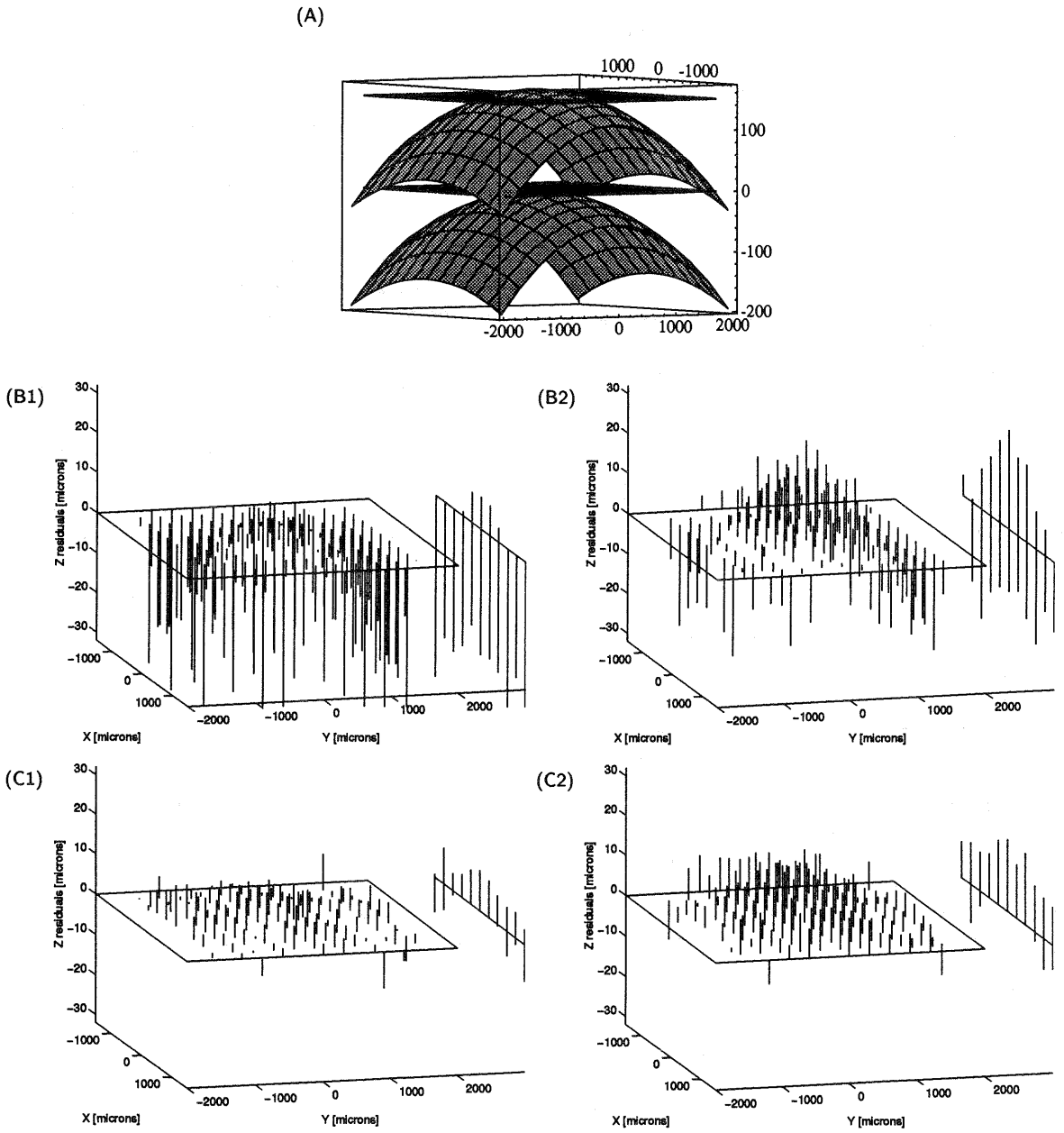

Figure 5: (A) Theoretically predicted deformation of two parallel planes in the object space when not compensating CMO distortion. (B) Vertical residuals of version 3 in table 1 not compensating for CMO distortion and (C) those of version 8 compensating for CMO distortion. (?1) residual field at the lower border, (?2) at the upper border of the depth of field. The vectors at the right side of the drawings are the overall projection of the residuals along the Y axes of the object space.

	Φ [o]	σ_ϕ [o]	M^{left} [ ]	$\sigma_{M^{\text{left}}}$ [ ]	M^{right} [ ]	$\sigma_{M^{\text{right}}}$ [ ]	E $[10^{-10}]$	σ_E $[10^{-10}]$	μ_{XY} $[\mu m]$	μ_Z $[\mu m]$	$\hat{\sigma}_0$ $[Pix]$
1	18.733	0.052	1.8953	0.0014	1.9214	0.0014	-1.5537	0.0092	3.57	5.60	0.17
2	18.640	0.053	1.8956	0.0014	1.9217	0.0014	-1.5429	0.0093	3.65	6.58	0.17
$\lvert\Delta\rvert$	0.093	0.074	0.0003	0.0020	0.0003	0.0020	0.0108	0.0132			
$\lvert\Delta\rvert/\sigma_{\lvert\Delta\rvert}$	**1.25**		**0.15**		**0.15**		**0.82**				

Table 2: Consistency test for some parameters: the empirical accuracies are shown for the check points

	Φ [o]	σ_ϕ [o]	M^{left} [ ]	$\sigma_{M^{\text{left}}}$ [ ]	M^{right} [ ]	$\sigma_{M^{\text{right}}}$ [ ]	E $[10^{-10}]$	σ_E $[10^{-10}]$	μ_{XY} $[\mu m]$	μ_Z $[\mu m]$	$\hat{\sigma}_0$ $[Pix]$
1	17.800	0.057	1.8935	0.0017	1.9185	0.0017	-1.4450	0.0090	3.81	6.77	0.19
2	17.311	0.044	1.9002	0.0013	1.9256	0.0014	-1.3839	0.0079	3.51	5.82	0.16
$\lvert\Delta\rvert$	0.489	0.087	0.0067	0.0021	0.0071	0.0022	0.061	0.020			
$\lvert\Delta\rvert/\sigma_{\lvert\Delta\rvert}$	**6.79**		**3.19**		**3.23**		**3.09**				

Table 3: System repeatability over 2 month: the empirical accuracies are shown for the check points

the vertical one is currently taken from the specifications of the micrometer table carrying out the vertical motion. Its relative position for motions up to $500\,\mu m$ is specified with 1%. In the future, using more precise gratings the vertical position has to be measured by a laser interferometer in order to get a homogeneous 3D precision distribution.
The imaging model described by equation (1) is partitioned into the weak perspective transformation (PT) and the distortion function δ_ξ. The quality of the results is judged by investigating empirical accuracies in the image space and in the object space.

In version 0 no parameters except the magnification M^{view} are estimated. They are fixed to the values provided by the manufacturer. The strikingly improved results of the other versions clearly demonstrate the necessity of a thorough calibration. Versions 1 and 2 demonstrate the sensitivity of the estimation procedure to a disproportion of $Q_{\xi\xi}$ and $Q_{\ell_{\tilde{x}}\ell_{\tilde{x}}}$. In version 1, $\sigma_{\xi\eta}$ is received from LSM results. These by far too optimistic values are corrected for a better Bundle Adjustment in version 2. For versions 3 to 8, $\sigma_{\xi\eta}$ and σ_Z are further adapted to the corresponding empirical accuracies.
Version 3 to 8 consider various combinations of distortion terms. The imaging model of version 3 principally corresponds to the standard camera model often used in computer vision ([Faugeras 1993]). A perspective transformation (in my case modified to the weak perspective situation) is completed by terms that compensate for shear and scale distortion in the image space. In version 4 radial and decentering distortion terms well known in photogrammetry are additionally introduced. Both models considerably improve the overall accuracy with respect to version 0. The finer distortion model of version 4 provide significantly increased vertical accuracy. However, version 5 clearly demonstrates that the vertical accuracy is mainly affected by the non paraxial imaging. The vertical accuracy is further improved by a factor of nearly two when the new term E is introduced.

Figure 5 contrasts the the vertical residuals of versions 3 and 8. Not correcting for CMO distortion results in a spherical deformation of the object space 5-A. This behavior has been theoretically predicted in [Danuser and Kübler 1995] and now is perfectly verified on real data (figures 5-B1 and 5-B2). Figures 5-C1 and 5-C2 demonstrate that through the new CMO term E the vertical residuals are reduced to values below the vertical precision limit of the target point field. The remaining systematic errors at the upper border of the depth of field (figure 5-C2) may origin from different scales of the vertical axis (defined by the micrometer table) and the lateral axes (defined by the grating). This is the major draw back of using planar gratings for the 3D calibration. On the micrometer scale it is very delicate to tune the metric systems of various dimensions.
Unfortunately, the term E causes reduced determinability for the scale and the decentering distortion terms. It is of utmost importance to detect this weakness as the introduction of s^{view} (version 6), as well as of P_1^{view} and P_2^{view} (version 7) worsens the performance of the system in the object space. Without check point analysis, the determinability is tested by the contribution of each distortion term to the trace of the cofactor matrix Q_{pp}. A large relative contribution of a certain parameter indicates its high correlations to other parameters and therefore its weak determinability. Simulations in [Danuser and Kübler 1995] let me expect this problem, but it has turned out to be even more serious. The relative contributions of P_1^{view} and P_2^{view} reach values of 8% which are not tolerable. On the other hand the empirical accuracy in the image space is improved when estimating scale and decentering distortion terms. Versions 6 and 7 show that particularly the η-direction (perpendicular to the epipolar direction) is positively affected by s^{view} and $P_{1/2}^{\text{view}}$. Through the high correlations between the rotation angles φ^{view} and the CMO term E the latter becomes overestimated. As E affects the ξ- and the η-component of the function $\delta_\xi(E)$ the η-compensation is slightly too large. This overestimation now is partially picked up by the terms s^{view} and $P_{1/2}^{\text{view}}$ but with the demonstrated negative side effect in the object space. Therefore, I propose to use the imaging model of version 8. Version 6 or 7 can be reevaluated if the application uses only measurements in the center of the field of view. In this case the relative depth of field is increased which stabilizes the estimation process.

In most versions, the empirical accuracies of the check points are very similar to those of the control points. Furthermore, in the most accurate versions 5 and 8, they correspond very well to the specified a priori Stdev. Hence, the imaging model would allow me to derive more precise positions when using a more precise calibration standard. Note that not the optical resolution is responsible for the precision limit. For the highest zoom level (optical resolution $2.3\,\mu m$) sub-micron ac-

curacies can be expected when employing the more precise standards. Relative measurements have been carried out for lateral positions ([Mazza *et al.* 1995]). The results — also verified by measurements using other sensor types — have demonstrated that super resolution down to $50\,\mathrm{nm}$ can be achieved with the proposed approach.

Two other tests investigating numerical weaknesses in the Bundle Adjustment are shown in table 2 and 3. The self-consistency of four numerically weak parameters, which are the parallactic angle Φ[5], the magnification M^{view} and the CMO term E, is tested in table 2 . In the self consistency test check points and control points of the very same data set are exchanged. None of the normalized differences $|\Delta|/\sigma_\Delta$ are statistically significant. Notice that σ_Δ is computed from the Bundle Adjustment's normal equations. Therefore one can state a satisfactory self-consistency of the parameters.
The same procedure is carried out in table 3. This time two data sets acquired with a time difference of 2 month are compared. It is obvious that the parallactic angle Φ suffers from relatively weak repeatability. Of course, not only numerical oscillations but also mechanical instabilities of the system hardware and inhomogeneities of the calibration standard (the target points used for calibration are not the same in both runs) lead to differences. A part of the differences Δ_Φ can be explained by the correlation between Φ and E (worst case: 85%). High correlations occur also between various other parameters. Thus, the whole parameter set must be considered as a unity. Repeatability tests can not concern one single term of the whole imaging model but must be analyzed through the accuracy performed in the object space. The latter is adequate also for the long term repeatability.

6 CONCLUSION

A new imaging model for Stereo Light Microscopes has been theoretically derived in [Danuser and Kübler 1995] and is now successfully implemented on a fully operating micro vision system. On a medium zoom level (optical resolution $4\,\mu\mathrm{m}$) empirical accuracies of laterally $3.8\,\mu\mathrm{m}$ and vertically $6.5\,\mu\mathrm{m}$ are obtained. This corresponds to a relative accuracy of $1:1000$ (lateral) and $2:100$ (vertical), respectively. In the image space the relative accuracy is $2:10000$. The discrepancy between object space and image space accuracy is caused by the limited precision of the underlying calibration standard currently available. Better standards will be used in the future, hopefully allowing me to measure with a precision of some tens of nanometers on the highest magnification level. Good results for lateral position measurements support this expectation.
One of the outstanding features of the imaging model is the computation of distortions originating from non paraxial optics. The application of this model is not limited to microscopy but may as well improve the performance of macroscopic systems relying on non paraxial optics.

REFERENCES

[Codourey *et al.* 1995] A. Codourey, W. Zesch, R. Büchi, and R. Siegwart. A Robot System for Automated Handling in Micro-World. In *IROS 95 Conference on Intelligent Robots and Systems*, volume 3, pages 185 – 190. IEEE/RSI, Aug. 1995.

[Danuser and Kübler 1995] G. Danuser and O. Kübler. Calibration of CMO-stereo-microscopes in a micro robot system. *IAPRS - International Archives of Photogrammetry and Remote Sensing*, 30/5W1:345 – 353, 1995.

[Danuser 1995] G. Danuser. A photogrammetric model for CMO Stereo Light Microscopes. Technical Report 167, Image Science, ETH Zurich, 1995.

[Faugeras 1993] O. Faugeras. *Three-Dimensional Computer Vision*. The MIT Press, 1993.

[Förstner and Gülch 1987] W. Förstner and E. Gülch. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In *Fast Processing of Photogrammetric Data*. ISPRS, 1987.

[Ghosh 1989] S.K. Ghosh. Electron Microscopy: Systems and Applications. In H.M. Karara, editor, *Non-Topographic Photogrammetry*, chapter 13, pages 187 – 201. American Society for Photogrammetry and Remote Sensing, 1989.

[Gleichmann *et al.* 1994] A. Gleichmann, M. J. Köhler, M. Hemmleb, and J. Albertz. Photogrammetric determination of topography of microstructures by scanning electron microscope. In C.J. Cogswell and K. Carlsson, editors, *Three-Dimensional Microscopy: Image Acquisition and Processing*, volume 2184, pages 254 – 265. SPIE, 1994.

[Kim *et al.* 1990] N. H. Kim, A. C. Bovik, and S. J. Aggarwal. Shape Description of Biological Objects via Stereo Light Microscopy. *IEEE Transactions on Systems, Man and Cybernetics*, 20:475 – 489, 1990.

[Maune 1973] D.F. Maune. *Photogrammetric Self-Calibration of a Scanning Electron Microscope*. PhD thesis, The Ohio State University, 1973.

[Mazza *et al.* 1995] E. Mazza, G. Danuser, and J. Dual. Lightoptical deformation measurements in microbars with nanometer resolution. *Microsystem Technologies - Sensors - Actuators - System Integration*, to appear, 1995.

[Richardson 1991] J. H. Richardson. *Handbook for the Light Microscope*. Noyes Publications, 1991.

[Taylor *et al.* 1992] D. L. Taylor, M. Nederlof, F. Lanni, and Waggoner A. S. The New Vision of Light Microscopy. *American Scientist*, 80:322 – 335, July 1992.

[5] $\Phi = \varphi^{\mathrm{right}} - \varphi^{\mathrm{left}}$