Registration of Microscopic Range Images Obtained by Shape from Focus

Reinhard Danzl and Konrad Karner
VRVis Research Center for Virtual Reality and Visualization
email: danzl@vrvis.at,karner@vrvis.at, http://www.vrvis.at/ar2

Abstract:
We present a shape from focus algorithm to reconstruct range images of microscopic objects from different directions using a hardware setup with a rotatable imaging tube and show how the orientation between the range images can be calculated using a variant of the iterative closest point algorithm. We first describe the special imaging geometry of our setup and the changes that have to be made to algorithms where the imaging tube is perpendicular to the object stage. Then we describe how the range images are distorted if the orientation of the imaging tube is not known exactly and propose that proper registration can be achieved if we do not use the metric coordinates of the points but the sensor coordinates after projection into the image stacks. Our method allows to obtain sharp stereoscopic images that can be used for shape from stereo and to obtain range images from multiple directions that can be fused to a refined 3D model. The importance of proper modelling of the imaging geometry and the distortion effects is evaluated on synthetic and real image stacks.

1 Introduction

Shape from focus (SfF) [7] has become one of the most important methods for the reconstruction of microscopic objects in the last decade. In contrast to other methods such as shape from stereo that suffer from the small depth of field of light microscopes [3], SfF exploits the small depth of field by taking a whole stack of images with varying distance between the objective and the specimen and searching for the image planes where a point appears sharpest.
Traditionally the imaging tube is perpendicular to the motorized object stage which results in image stacks where a point of the specimen is located at the same position in each image. Although these methods perform well if a lot of texture is present they face difficulties in regions with homogenous grey levels and steep edges.
In order to overcome the problems of traditional focusing methods we introduce a new microscope setup where the imaging tube can be tilted in an eucentric way which allows to obtain multiple image stacks of a specimen from different directions. Before the 3D information of the different stacks can be fused or stereoscopic methods can be applied two problems have to be solved in advance which will be discussed in this paper. The first is to model the imaging

geometry of the setup where the imaging tube is not perpendicular to the object stage and to adapt traditional SfF algorithms in order to obtain range images for each image stack. The second is to estimate the exterior orientation parameters between the different stacks since the rough orientation that can be obtained from an angular scale on the microscope is not sufficient for further processing.

The application of classical shape from focus methods to tilted imaging tubes has not been tried before. The only method for tilted image sensors we know is by [5] who proposed an online measurement algorithm by searching for sharp contours in the different images. We describe a generalization of traditional SfF methods that uses all images of the stack during computation and thus allows better interpolation between the stacks.

The estimation of the orientation for the different stacks is done using a variant of the iterative closest point [1] algorithm which has become the most prominent method for the alignment of different range images. In contrast to previous applications we have to cope with range images that contain a systematic error since the SfF procedure requires the correct unknown orientation to produce satisfying results. This systematic error does not allow correct registration if the initial range images are used. Therefore we model the projection of the 3D points to their corresponding pixels in the image stacks including a special distortion matrix and perform the registration using these stack coordinates which allows registration of wrong initial range images.

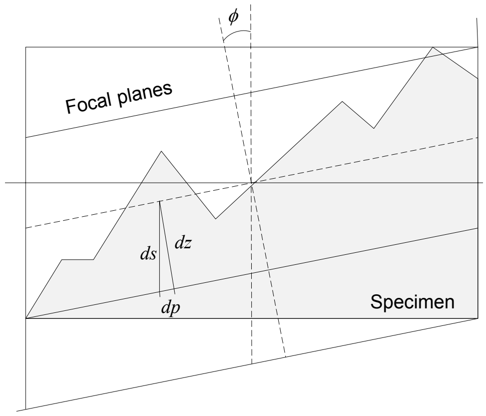
In Section 2 we propose our hardware setup and explain how range images can be calculated using tilted imaging tubes. Then we describe the imaging geometry of the system and model the projection of object points into the image stacks (Section 3) which will be used in Section 4 that describes the registration of the different range images. The performance of the algorithm is evaluated in Section 5 both on synthetic and real image stacks.

2 Shape from Focus using Tilted Imaging Tubes

Fig. 1a shows the hardware setup of our system containing a motorized xyz-stage, an imaging tube that can be tilted in an eucentric way so that the specimen stays focused before and after the rotation, a cold light source which focuses light via a beam splitter towards the specimen, a color camera of 1280*1024 pixels and an angular scale from which the tilt angle of the tube can be roughly estimated.

Our shape from focus method is quite similar to other methods that operate on image stacks with non-tilted tubes [7]. First we calculate a focus measure at each point of the image stack [6]. Then we estimate the height for each point of the specimen by searching for the largest focus measure along the different image planes. In contrast to traditional approaches the position of an object point is not the same for all images but is shifted between two planes

(a)



(b)

Figure 1: (a) Hardware setup of the microscope using a rotatable imaging tube (b) Schematic view of a specimen and the position of the focal planes for a tilted imaging tube and different distances between the objective and the specimen

by the amount

$$dpx = +\frac{ds}{dx}\sin\phi \quad (1)$$

$$dpy = -\frac{ds}{dy}\cos\phi\sin\omega \quad (2)$$

where ds is the displacement of the object stage between two images (Fig. 1b), (dx, dy) is the pixel-size and ϕ, ω and κ are the three rotation angles of the imaging tube.

3 Range Image Distortion and Projection Model

Since the orientation values are read from the angular scale they are rough estimations only which results in a wrong estimated shift between the image planes and two types of errors in the range images. First there are more outliers since the pixels that are used during the search for the focus maximum do not correspond to the same object point throughout the stack but are slightly shifted. This problem can be reduced by discarding wrong points during the registration algorithm as described in Section 4. The second problem is more severe since all reconstructed points contain a systematic distortion that results in large errors when the range images are registered without additional consideration. In order to remove this distortion we model the projection of the reconstructed points into the different image stacks and only use the projected coordinates during the registration.

Actually we are interested in how an object point $X = (X_1, X_2, X_3)$ in world coordinates is projected to a point $x = (x_1, x_2, x_3)$ in a stack, where (x_1, x_2) are the pixel coordinates of the corresponding range image reconstructed by SfF and x_3 is the image plane number in the stack. So, instead of aligning the object points $\{X_i\}$ and $\{X_j\}$ of two range images in world

coordinates, we align the projected points $\{x_i\}$ and $\{x_j\}$.
The whole projection is modelled using a projection matrix P that transforms homogenous world coordinates $\tilde{X}$ to homologue stack coordinates $\tilde{x}$ and can be split up into three matrices

$$\tilde{x} = P\tilde{X} = (A * D * E)\tilde{X} \tag{3}$$

E contains the parameters of the unknown exterior orientation and consists of a rotation matrix R that depends on the three Euler angles ω, φ, κ and a translation vector $t = (t_x, t_y, t_z)$.

$$E = \begin{pmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \quad \mathit{with}\ \mathbf{0}^T = (0,0,0) \tag{4}$$

A is the matrix that contains the interior orientation parameters and transforms from metric coordinates to sensor coordinates

$$A = \begin{pmatrix} \frac{1}{dx} & 0 & 0 & \frac{cols}{2} \\ 0 & -\frac{1}{dy} & 0 & \frac{rows}{2} \\ 0 & 0 & \frac{1}{dz} & \frac{planes}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

where dx and dy are the horizontal and vertical size of a pixel in mm. $dz = ds \cos\phi \cos\omega$ is the distance between two image planes in mm and can be derived from ds which is the vertical distance between the corresponding positions of the xyz-stage. We use an affine projection instead of a perspective one since this is more suitable for our microscope objectives and the projection into a stack of images.
The crucial matrix in the projection is D which models the distortion between the coordinates of a range image reconstructed with the correct orientation parameters and one with the wrong parameters (ϕ', ω') that were initially used:

$$D = \begin{pmatrix} 1 & 0 & s_x & -s_x \frac{planes}{2} dz \\ 0 & 1 & -s_y & s_y \frac{planes}{2} dz \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6}$$

where

$$s_x = \frac{1}{\cos\phi \cos\omega}(\sin\phi - \sin\phi') \tag{7}$$

$$s_y = \frac{1}{\cos\phi \cos\omega}(\sin\omega \cos\phi - \sin\omega' \cos\phi') \tag{8}$$

Since a stack contains image planes that are shifted along the x- and y-direction one has to determine which image plane shall be used as reference frame for the range image. As can be seen in the distortion matrix D we have chosen the middle plane of the image stack since it is the plane where most part of the specimen is visible. Another aspect to mention is that the

z-value is not changed by the transformation D which is obvious since a wrong shift does not affect the z-value but only the x- and y-coordinates.
The interior orientation parameters dx and dy are estimated in advance using a planar calibration target, while the stage displacement ds can be derived accurately from the stage sensors. Thus the only parameters that have to be estimated during the registration process are the rotation and translation parameters of the matrix E.

4 Registration

In order to fuse the different range image stacks we need to refine the three rotation and the three translation parameters of the exterior orientation matrix E. Since photometric differences in the images between two stacks do not lead to reliable homologue points we only use the 3D information of the initial range images for the estimation.
Early approaches of this registration task mainly used pre-computed features for alignment [4] but since features were often rather sparse this approach had inherent problems with fine registration. The proposal of the Iterative Closest Point Approach (ICP) developed by Besl and McKay [1] and Chen and Medioni [2] reduced this problem and is the basis of almost all fine registration algorithms. It basically iteratively pairs each point in the first set to its closest neighbor in the second set and computes the transformation that minimizes the error between the sets. Since this approach was very successful many different variants of the original algorithms followed and tried to improve them mainly in terms of speed and robustness.
Based on a recent survey by [9] we decided on an ICP variant that iteratively performs the following steps:

1. Transformation of the points of the first range image using the matrix P

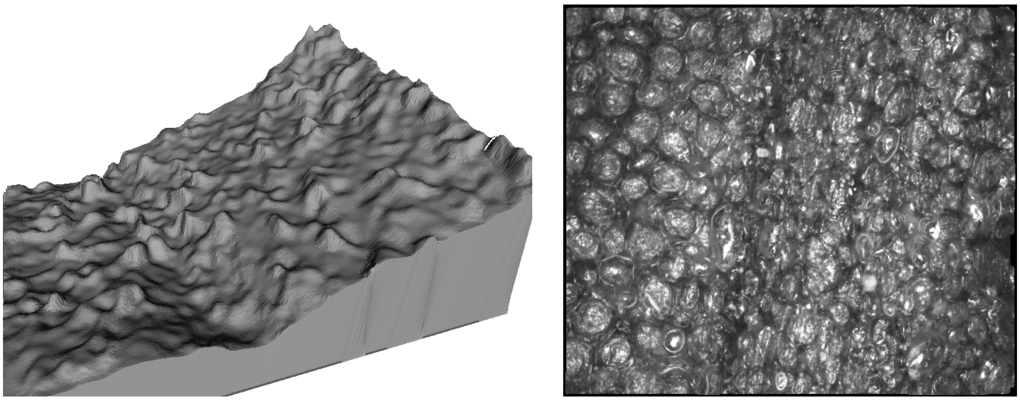
2. Projection of the points onto the second range image using the range direction of the second range image in order to find the corresponding points.

3. Sorting of all distances between the corresponding points and discarding a certain percentage of points with the largest distances.

4. Calculating the tangent plane of the surface at the corresponding point.

5. Refinement of the orientation parameters by minimizing the sum of the distances of the points and the corresponding tangent planes using a conjugate gradient algorithm.

In contrast to the algorithm by [8] which is closest to our approach we do not search for the closest neighbor in the second mesh but perform simple projection as described above. The main difference however is that we use the whole projection matrix P (Section 3) instead



(a) (b)

Figure 2: (a) Ground truth and (b) texture of the synthetic object

of the exterior orientation matrix alone. Another important issue is that we do not use the distance between corresponding points (point-point-metric) to model the alignment error as initially proposed by [1] but the point-plane metric. Together with the correct modelling of the distortion this is the most important part of the algorithm as shown in the next section.

5 Results

In this section we provide three experiments that evaluate the performance of our registration algorithm. The first experiment compares the results between the point-point and the point-plane metric (Section 4), the second evaluates the improvements that are achieved using our projection model instead of the raw range images and the third is an experiment on real image stacks obtained with the new hardware setup.

In the first experiment we tried to register two image stacks of a synthetic object (Fig. 2) once using the point-point metric [1] and once using the point-plane metric [2]. The first image stack was created with the imaging tube perpendicular to the object stage, the second with the exterior orientation parameters $(Rx, Ry, Rz) = (0°, 10°, 0°)$ for the rotation and $(Tx, Ty, Tz) = (0, 0, 0)$ pixel for the translation. In order to evaluate only the effect of the used metric both range images were reconstructed using the correct orientation and then the registration was started using a wrong orientation estimate. When we used the point-point metric the errors for the orientation were up to $0.05°$ and the errors for the translation up to 0.5 pixel as shown in Fig. 3. In the case of the point-plane metric the errors dropped below $0.006°$ and 0.06 pixel which shows that the point-plane metric is preferable especially when registering range images that allow easy calculation of surface tangent planes.

For the second experiment we used the same object as in experiment 1 and registered the image stacks once without modelling the distortion effect due to the wrong initial orientation and once with our new registration procedure using the projected coordinates. The first

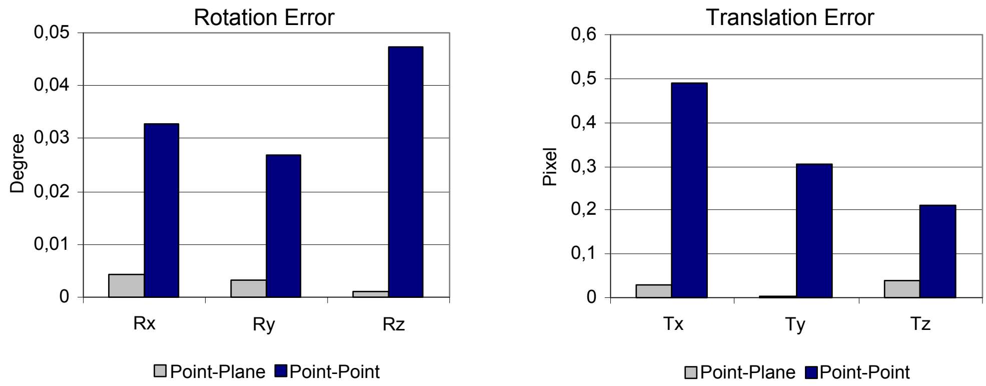


Figure 3: Experiment 1: Absolute rotation and translation errors for registration using the point-point and the point-plane metric

Errors	Rx	Ry	Rz	Tx	Ty	Tz
Correction	0.0051	0.0043	0,0104	0,03	0	0,0644
No Correction	0,0560	0,1784	0,2483	0,559	0,908	0,1588

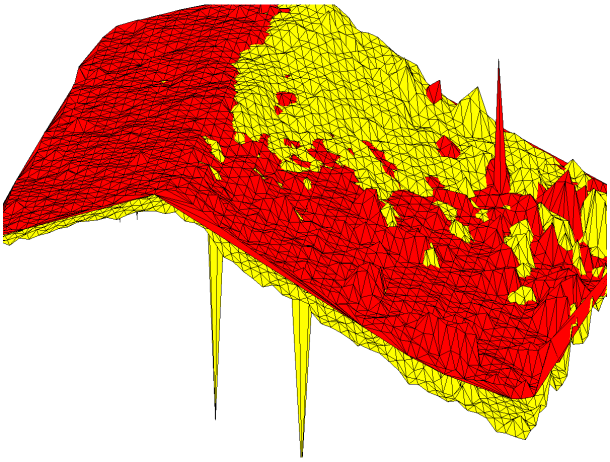
Table 1: Experiment 2: Absolute rotation and translation errors for registration once with and once without correction of the special range image distortion

image stack was created as in experiment 1 while the rotation parameters of the second were $(-1°, 11°, 2°)$. Both registrations were performed using the point-plane metric and outlier rejection of 10 percent. According to typical orientation errors in the real setup we chose an initial orientation of $(0°, 10°, 0°)$ for the rotation and $(20, -20, 10)$ pixel for the translation. When the distortion was not correctly modelled (by removing the distortion matrix D in the projection process) the rotation errors were up to 0.25° and the translation errors up to 0.9 pixel as shown in Table 1. However when the registration was performed using the projected coordinates the resulting rotation errors were 0.01° at most and the translation errors were below 0.07 pixel which shows that a good distortion model is crucial for proper alignment.

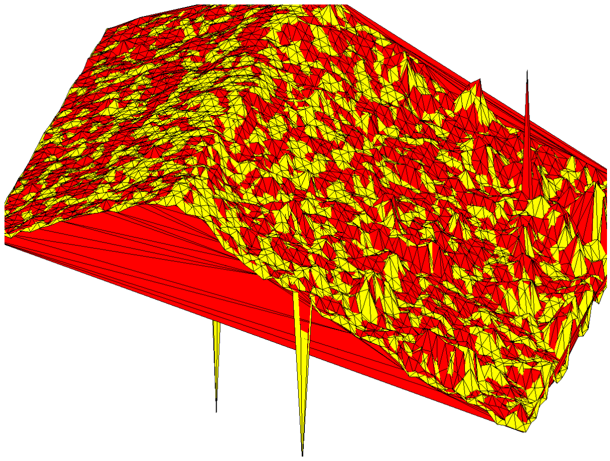
The third experiment was performed using real image stacks of an eraser obtained with our microscope setup. One of the stacks was obtained using the rough orientation $(0°, 0°, 0°)$, the second with the orientation $(0°, 10°, 0°)$. Although the orientation parameters from the angular scale of the microscope provided a good rough registration as shown in Fig. 4a, the final result after registration with our new algorithm (Fig. 4b) was much better and demonstrates the performance of the algorithm.

6 Conclusions

We have presented a new hardware setup for light microscopes that allows to reconstruct range images of objects from different directions using shape from focus. We have shown that good registration of these range images is only possible when the projection of reconstructed points into the different image stacks is properly modelled. Additionally the use of a point-plane



(a)



(b)

Figure 4: Registration of real image stacks: a) Before Registration b) After Registration

metric and significant outlier rejection is crucial for proper alignment of the range images. Future work will focus on fusion of the different range images based on the orientation results achieved by the proposed algorithm.

References

[1] P. J. Besl and N.D. McKay. A method for registration of 3d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–259, 1992.

[2] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10:145–155, 1992.

[3] G. Danuser. Photogrammetric calibration of a stereo light microscope. *Journal of Microscopy*, 193:62–83, 1999.

[4] O. D. Faugeras and M. Hebert. The representation, recognition, and locating of 3-d objects. *Intern. Journal of Robotic Research*, 5:27–52, 1986.

[5] A. Ishii. 3-d shape measurement using a focused-section method. *Proc. of the Int. Conf. on Pattern Recognition*, 4:828–832, 2000.

[6] E. Krotkov. Focusing. *Int. Journal on Computer Vision*, 1:223–237, 1987.

[7] S. K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:824–831, 1994.

[8] K. Pulli. Multiview registration for large data sets. *Second Int. Conf. on 3D Digital Imaging and Modeling*, pages 160–168, 1999.

[9] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. *Proc. Third Intern. Conf. on 3D Digital Imaging and Modeling*, pages 145–152, 2001.