SHAPE FROM FOCUS: FULLY AUTOMATED 3D RECONSTRUCTION AND VISUALIZATION OF MICROSCOPIC OBJECTS

Markus NIEDEROEST[a], Jana NIEDEROEST[b], Jirí ŠCUCKA[c]

[a]markus.niederoest@freesurf.ch
[b]Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland
jana@geod.baug.ethz.ch
[c]Institute of Geonics, Academy of Sciences of the Czech Republic, Ostrava, Czech Republic
scucka@ugn.cas.cz

KEY WORDS: Geomaterials, Multifocus Imagery, Image Processing, 3D Reconstruction, Visualization

ABSTRACT
Optical microscopy as an important technique in the petrological and mineralogical research enables the observation of highly magnified objects and material structures. However, in case of a non-planar preparation the acquisition of sharp images is impossible due to the small focus depth of an optical microscope. The commonly used method to resolve this problem is the acquisition of a sequence of monoscopic images with a constant change of the focal length between consecutive images. The final sharp image is then calculated from this sequence by combining the sharp regions of the raw images. In a previous paper (Niederoest et al, 2003) we described a simple method to reconstruct the composed image and, as our main focus, we presented innovative procedures for generation of interactive 3D models from the multifocus image sequence. Since at that time the procedures did not yet work fully automatically - the generation of a binary mask for exclusion of the background from the reconstruction process was still done in a manual mode - we have automated this step in our recent work. In this paper we report on all stages of the processing chain, present our easy-to-use software Shape From Focus and validate our results.

1 INTRODUCTION

The generation of high quality microscopic images and 3D models of spatial objects is a task of great concern in the study of geomaterials. Scanning electron microscopy provides a very detailed and sharp 3D representation of the studied object on the computer screen and thus it is a suitable means for reconstructing various mineralogical and petrological samples. However, scanning electron microscopy is very expensive and time-consuming, which prompts many researches to resort to the much more accessible technique of optical microscopy. The major difficulty associated with using an optical microscope for 3D visualization is its small focus depth that makes the acquisition of sharp images of three-dimensional objects impossible. Thus, to obtain a sharp 2D sample representation as well as to recover its 3D structure, methods of image processing must be applied.

We developed a software system that combines multifocus optical microscopy and digital image processing for automated generation of interactive textured 3D models. The idea of such a monocular depth recovery goes back to the early eighties. The experimental work of (Pentland, 1982) showed that two images taken with different apertures or with different focal lengths are a source of depth information. In the following years the methods were improved and became more general (Ens et al, 1993, Kaufhold et al, 1998, Baba et al, 2002). Surprisingly, these techniques remain rather a computer vision domain; researchers in the application fields such as biology and geology mostly employ stereomicroscopy to obtain 3D views of studied objects (Greenberg et al, 1997). The most recent study of the two- and three-dimensional processing of multifocus images – an approach similar to the one we use – can be found in (Martišek, 2002). As there is a lack of operational software offering a 3D information of optical microscope outputs, we see our main contribution in providing an easy-to-use fully automated system including several innovations in terms of interactive 3D visualization.

2 3D VISUALIZATION IN THE STUDY OF GEOMATERIALS

Conventional optical microscopes provide magnified and sharp *two-dimensional* images of *planar* objects. For the microscopic observation in the petrological and mineralogical research, two types of particularly prepared planar preparations are usually used, so called polished and thin sections, which enable visualization and measurement of various structure and texture characteristics of the studied material. However, there exist many geomaterials where the creation of such preparations is not possible, because this procedure would cause the distortion or loss of the image information. In addition, it is also necessary to visualize and analyze *spatial* objects such as microcrystals, loose material particles, microfossils etc. The high quality microscopic images as well as 3D models of such objects can provide essential information for many applications. There are three commonly used methods for 3D visualization in the study of geomaterials nowadays; and all of them face some drawbacks in terms of applicability and quality of the three-dimensional result. In the following we list the main features of these methods.

1. **Scanning Electron Microscopy (SEM):** The Scanning Electron Microscope creates highly magnified grey scale images of studied objects. Samples must be carefully prepared to withstand the vacuum inside the microscope and to conduct electricity. The result is a very detailed, complex and interactive 3D sample representation. The drawback of this method is – except of its high costs – a time-consuming preparation of the samples. In addition, it is not possible to visualize orientation and quality of cracks and inclusions inside the studied material.
2. **Stereo Optical Microscopy:** The stereo optical microscope allows viewing and measuring objects in stereo mode. However, this effect is only temporary and cannot be preserved in the same quality outside the microscope. There exist several systems for so called direct-view 3D microscopy using conventional lenses, e.g. a method of multiple oblique sample illumination or a dual oblique viewing. For the latter, a novel 3D head for a microscope replaces the standard trinocular head and converts the 2D image into a 3D image (Greenberg et al, 1997).
3. **Optical Microscopy:** As mentioned before, the small focus depth of an optical microscope causes the effect that a non-planar sample cannot be displayed sharply in a full extent. Therefore the following procedure of image reconstruction can be applied: A sequence of monoscopic images of the studied object (e.g. grain) is scanned step by step with changing focal length. Then a digital image processing system selects only the sharp fractions, the so called "optical cut", from every image of the sequence. The final high quality grain image is obtained by the composition of extracted optical cuts (Fig. 1). To generate a spatial model of the object, the depth is estimated according to focusing distance of a particular image of the sequence (Ens et al, 1993, Kaufhold et al, 1998, Baba et al, 2002, Martišek, 2002).

For practical reasons (availability of an appropriate system), we decided to use the monoscopic optical microscopy in combination with digital image processing. The question is *how* to find the optimal image fractions for the 2D and 3D reconstruction procedure. Most of the existing techniques are based on modeling the optical system's Point Spread Function and removal of out-of-focus blur. This deconvolution method usually requires a calibration procedure or an object containing clear edges (Kaufhold et al, 1998, Asada et al, 2001). Other methods are semi-automated, such as (Kaufhold et al, 1998). The advantage of our approach to 2D (Section 3) and 3D (Section 4) reconstruction is that almost no pre-knowledge is required and the procedures work fully automatically. The result of our spatial reconstruction is an interactive textured 3D model of the sample containing information that could not be provided merely by the optical system of the conventional microscope.

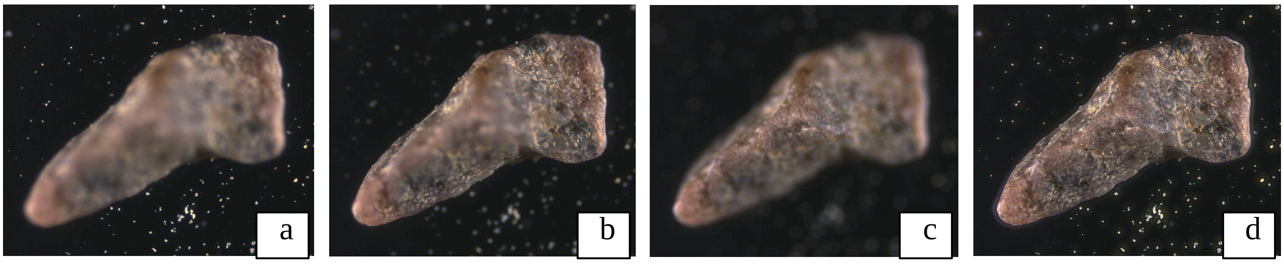

Fig. 1: The procedure of image reconstruction. (a)-(c) three of 24 images of a grain of sand taken through an optical microscope with varying focusing distance, (d) the composed image.

3 MULTIFOCUS IMAGE SEQUENCE AND THE COMPOSED IMAGE

Due to the small focal length of an optical microscope, it is not possible to get a sharp representation of a 3D object with just one image. Therefore a sequence of monoscopic images of the studied object is scanned step by step with a constant change of the focussing distance between consecutive images. The composed image is then calculated from this sequence by combining the sharp regions (optical cuts) of the raw images. In our approach we use local variance as criterion for sharpness. The result is a focused representation of the complete object.

In our system, the color $RGB_{x,y}$ of a pixel at x,y position in the composed image is determined in the following manner: For each of the raw images, the grey value variance of the *n x n* neighbourhood at position x,y is calculated. For the variance calculation the color values of the pixels are transformed to grey scale. Variance $v_{xy}(R)$ of raw image R at x, y is then given by

$$v_{xy}(R) = \frac{1}{n^2} \sum_{i=x-k}^{x+k} \sum_{j=y-k}^{y+k} \left(GV_{ij} - \overline{GV} \right)^2 \qquad (1)$$

where	$v_{xy}(R)$	Variance of the R-th raw image at x, y
	n	Patchsize selected by the user before processing the data
	k	(n-1) / 2, calculated from the patchsize
	x,y	Image coordinates of current pixel
	GV_{ij}	Grey value at image coordinates i, j
	$\overline{GV}$	Mean grey value in *n x n* current pixel neighbourhood

Now the raw image R_{max} with the highest variance $v_{xy}(R)$ is determined. The color RGB_{xy} in the composed image is taken from the color raw image R_{max} at x,y.

The result of the multifocus image reconstruction is depicted in Fig. 1d. Our procedure is simple: we did not invest in finding more sophisticated solutions on purpose. The reason was that at the Institute of Geonics Ostrava, which is involved in the geological part of our project, the complex image analysis system LUCIA DI is available. This system, developed by the Czech company Laboratory Imaging Ltd., Prague, supports multifocus image acquisition in combination with an optical microscope, generation of the composed image and offers many other analysis functions.

Therefore we rather focused on the topics described in the following, as they are not part of the LUCIA DI software package.

4 3D RECONSTRUCTION USING A MULTIFOCUS IMAGE SEQUENCE

We developed two methods for 3D object reconstruction from a multifocus image sequence. The first one, which we call the method with maximum variance, can be found in (Niederoest et al, 2003). In this paper we only present the second, more accurate method that uses a parable fit to a sample of variances.

In the first step, a binary mask that separates the fore and the background of the image data is generated automatically (Fig. 2). This serves to eliminate errors in the background of the reconstructed object.

After choosing one image of the acquired sequence, a non-critical threshold is used for the separation of the scene to foreground and background. The resulting image contains a large white area (the object to be reconstructed) and a number of small white areas (dust, small objects in the background). The area of interest is determined by a connectivity analysis, and all non-maxima areas are rejected. In a postprocessing step, small black areas in the foreground are removed by dilation with the patch size previously set by the user.

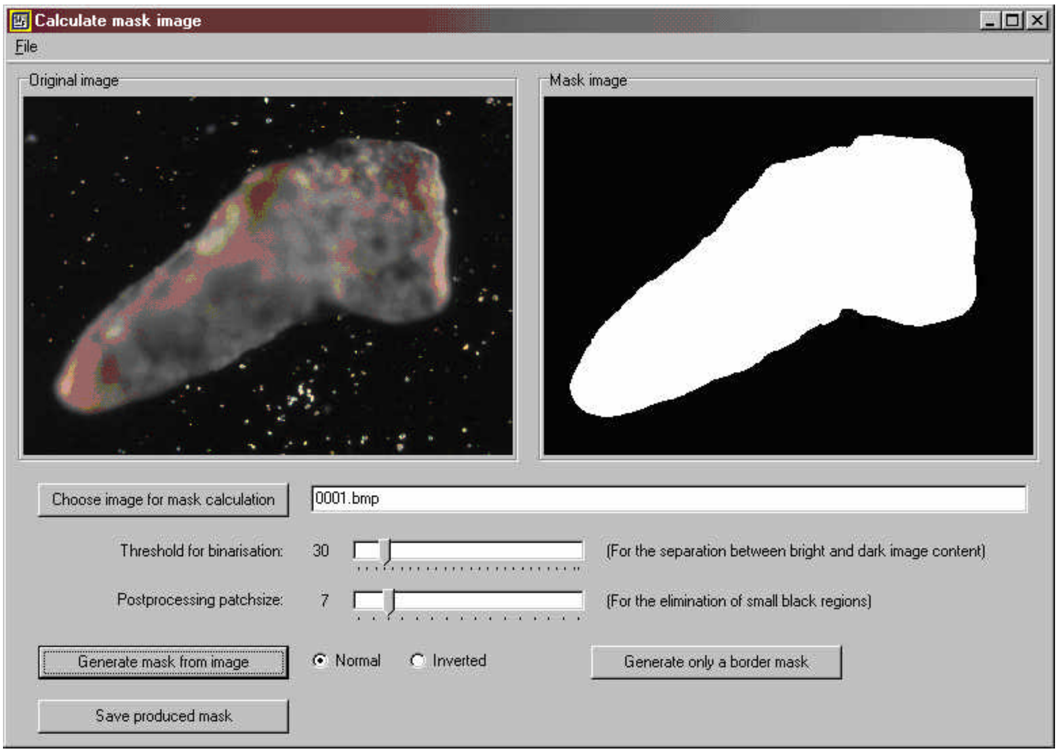

Fig. 2: Tool for automatic generation of a binary mask to exclude the background from the reconstruction process.

In the following, the content of the used raw images is transformed to grey scale and a patchsize *n x n* is selected by the user. The raw images have to be ordered according to the focussing distance and an index *i* is assigned to each of them. The number of raw images is *N*. Image 1 is the one with the longest focussing distance, image *N* the one with the shortest focussing distance. For data volume optimization in the resulting height model, only each *t-th* image point is processed. The thin out factor *t* is user defined and is typically assigned with a value of 4. For each point x,y where x and y are multiples of *t*, the variance $v_{xy}(R)$ of raw image *R* at x,y is calculated from Equation 1.

Now the raw image R_{max} with the highest variance $v_{xy}(R)$ is determined (see Equation 1). The height of the currently processed point x, y is given by the index $i(R_{max})$ of that particular raw image, multiplied with a height factor:

$$h_{xy} = i(R_{\max}) \cdot hf \tag{2}$$

where h_{xy} Height of the object at coordinates x,y
$i(R_{max})$ Index of the raw image with highest variance
hf Height factor

The height factor is used to preserve the correct ratio between the planar extensions and the height of the object. Taking into account that a thin out factor *t* was used for data reduction, the height factor *hf* is calculated from the focusing distance step *fds* and the resolution of the image. The resolution *res* of the image (metric size of one pixel) is given by the image dimensions n_x, n_y in pixel and the metric dimensions d_x ,d_y of the sensor in microns:

$$hf = \frac{fds}{res \cdot t}, \qquad \text{with} \qquad res = \frac{d_x}{n_x} = \frac{d_y}{n_y} \tag{3}$$

Now a parable is optimally fit to a graph of variances at x,y, using 7 observation points:

$$v_{xy} = a \cdot h_{xy}{}^2 + b \cdot h_{xy} + c \tag{4}$$

where v_{xy} Variance of the current raw image with the highest variance at x,y (from Equation 1)
h_{xy} Height of the point (from Equation 2)
a, b, c Parameters to be estimated

The equation system for the least squares fitting can now be formulated with the given 7 points (Fig. 3). After calculating the unknown parameters, the abscissa value of the apex is calculated.

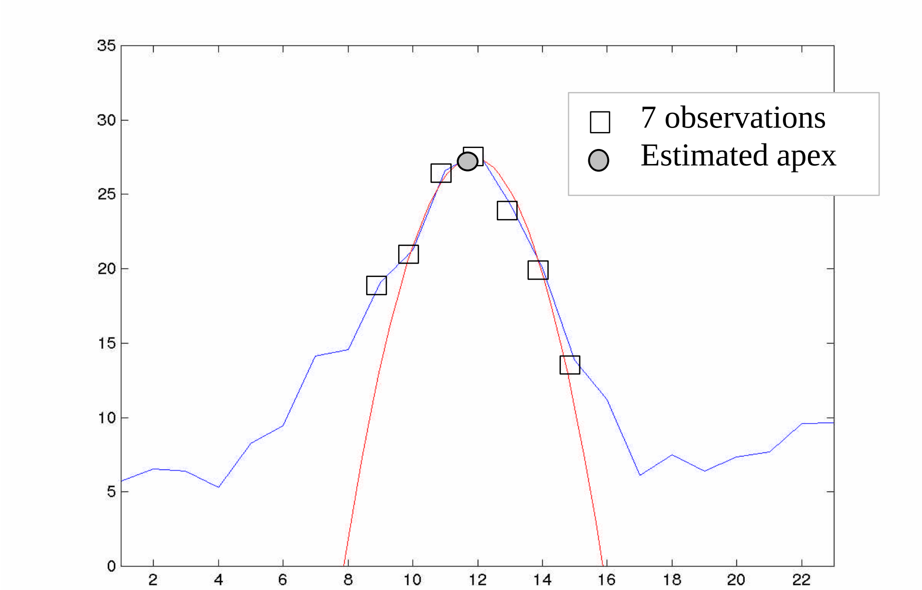

Fig. 3: Graph of grey value variance through a sample of raw images at the same pixel position x,y. The parable is fit to the point with maximum variance and the 3 closest samples on the left and the right side, taken from the adjacent raw images.

The x-coordinate of the apex is now taken as height value h_{xy} for the resulting height model. The postprocessing using the local median and the local average leads to a visual improvement of the object representation.

5 VISUALIZATION

The resulting height model is saved in the commonly known VRML2 format (Virtual Reality Modelling Language), where the composed image calculated in Section 3 is used as texture. An example of a VRML model of the reconstructed grain of sand with the free browser plugin Cosmoplayer is depicted in Fig. 4. We also developed an own 3D viewer (Fig. 5), which enables navigation in reconstructed 3D models, generation of videos in AVI format and creation of stereo anaglyph images (Fig. 6). It is possible to visualize wireframe models, too. An integrated part of our viewer is a novel tool for measurements in the 3D model. A 3D measuring mark placed onto the model surface enables measurement of 3D coordinates (Fig. 7).

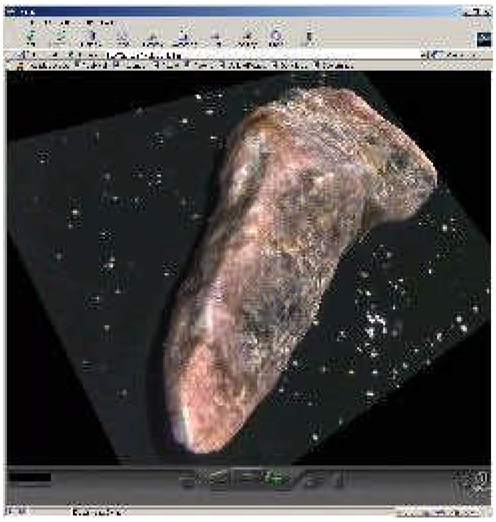

Fig. 4: The reconstructed grain of sand (object size in plane about 4 x 2 mm) with Cosmoplayer.

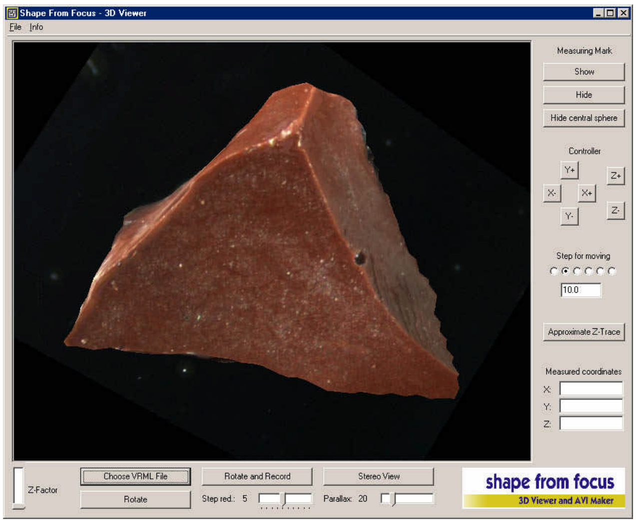

Fig. 5: The reconstructed copper slag grain (object size in plane about 1.5 x 1.5 mm) in the viewer of our software Shape From Focus.

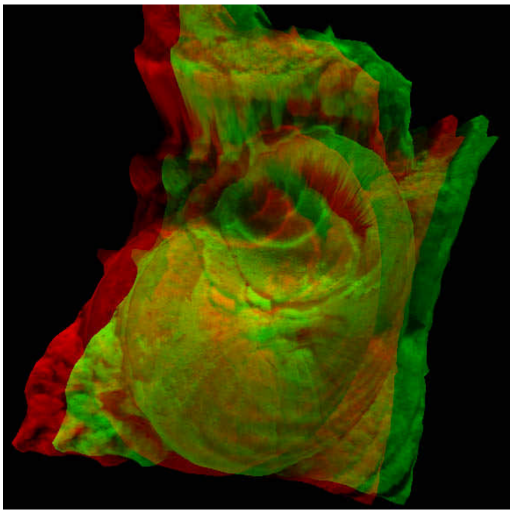

Fig. 6: Stereo anaglyph view of the gastropod fossil (object size in plane about 3 x 4 mm) to be viewed with red/green anaglyph glasses.

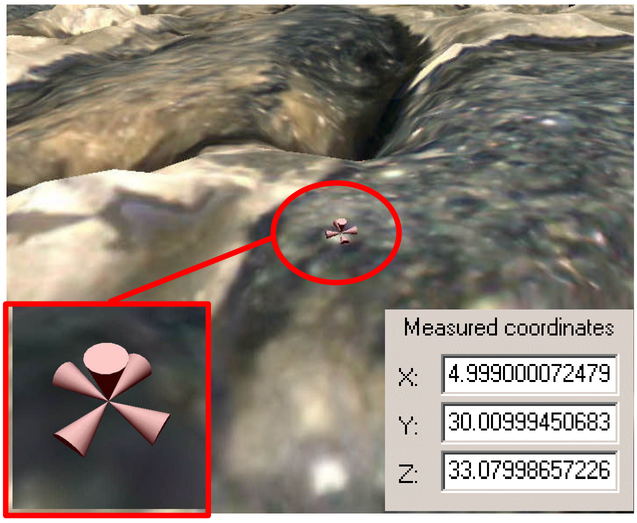

Fig. 7: 3D measuring mark in Shape from Focus. In the background is the 3D model of a crushed slate hardened with polyurethane resin.

6 VALIDATION OF RESULTS

For the evaluation of our implemented methods a tilted thin section, i.e. a perfectly planar cut was scanned in a sequence of 20 images with a focusing distance step of 51 microns. The object was reconstructed with the previously described method, using a 9x9 patch size for the variance (Fig. 8a). For the postprocessing a 9x9 median filter was applied, followed by calculation of the local average in a 5x5 neighbourhood. Surface profiles of the tilted thin section should be straight lines, so the following procedure was used to estimate the height error of a reconstructed object: From the 16 created profiles the standard deviation of a single height was estimated with linear regression (Fig. 8b).

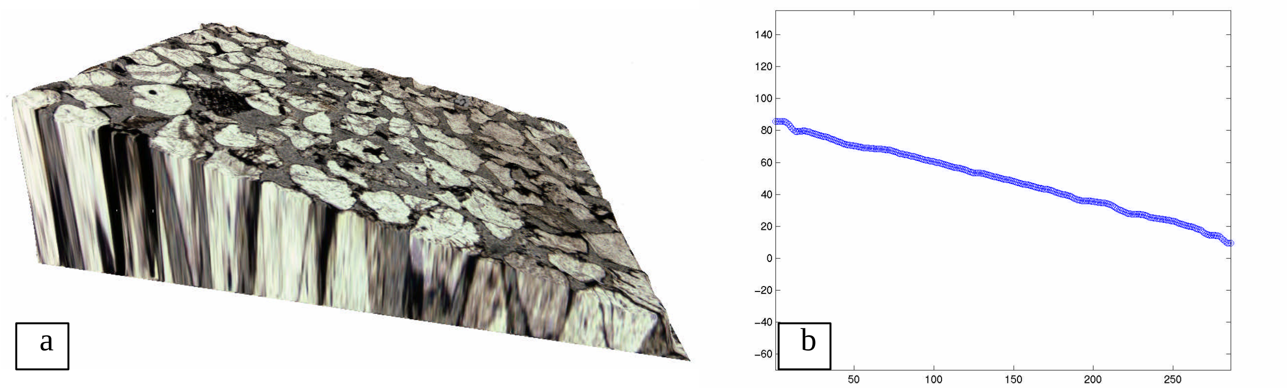

Fig. 8: (a) The 3D model of the test object (tilted thin section of sandstone) that was used for the estimation of the height accuracy, (b) Profile of the test object after the reconstruction with the parable method.

The resulted mean standard deviation from 16 profiles is 9.3 microns. Considering the object size of 4179 x 3297 x 972 microns, the accuracy is not excellent, but certainly sufficient for visualisation purposes. The accuracy of the results could be improved, if the focussing distance step between two adjacent raw images was reduced. This would result in a larger number of raw images which is not critical due to the fast and automated processing.

7 THE SOFTWARE

The presented algorithms were implemented as a stand-alone software package called Shape From Focus, running on Microsoft Windows. The module for the calculation of composed images and for the 3D reconstruction was programmed in Borland Delphi 6.0, whereas for the viewer Delphi 6.0 was used in combination with GLScene, an OpenGL based 3D library for Delphi that is distributed under the Mozilla Public License (GLScene, 2003).

Tests were done on a computer with a 1.7 GHz Pentium 4 processor and 512 MB RAM in order to estimate the time for different processing steps. The calculation of the composed image from 23 raw images takes around 1 minute. This gives an average processing time of less than 3 seconds per raw image. The 3D reconstruction requires for the same input data about 20 seconds. The postprocessing of the resulting 3D model with the local median or the local average is done in a second.

The numbers above show that even on an older computer it won’t take longer than 5 minutes to produce an interactive 3D model or an anaglyph stereo view of a microscopic object form a multifocal image sequence in a fully automated mode.

8 CONCLUSIONS

Our system 'Shape From Focus' represents an effective way of fully automated reconstruction and 3D visualization of microscopic objects. Its main value in opposite to all commonly used procedures is that it can produce interactive, texture-mapped and durable 3D models for visualization purposes within a minimal time span and with very little operator interference. Although the principle of the applied methods has been known for many years, we see our main contribution in the development of a robust and user-friendly software system with the stress on innovative visualization tools.

ACKNOWLEDGEMENT

The part of this work done at the Institute of Geonics in Ostrava was financially supported by the Grant Agency of the Academy of Sciences of the Czech Republic, No. A 3086201.

REFERENCES

Asada, N., Baba, M., Oda, A., 2001. Depth from Blur by Zooming. Proceedings of the Vision Interface Annual Conference, Ottawa, Canada.

Baba, M., Asada, N., Migita, T., 2002. A Thin Lens Based Camera Model for depth Estimation from Defocus and Translation by Zooming. Proceedings of the 15th International Conference on Vision Interface, Calgary, Canada.

Ens, J., Lawrence, P., 1993. An Investigation of Methods for Determining Depth from Focus. IEEE Transactions on Patern Analysis and Machine Intelligence. Vol. 15, No. 2, pp. 97-108.

GLScene, 2003. GLScene – OpenGL Solution for Delphi. Founded and previously developed by M. Lischke, currently developed and maintained by E. Grange. www.glscene.org, last visited on 20 January 2003.

Greenberg, G. L., Boyde A., 1997. Direct View 3-D Microscopy Using Conventional Lenses. *European Microscopy and Analysis*, March, pp. 5-7.

Kaufhold, J., Karl, W. C., 1998. A Nonparametric Defocus-based Approach to Reconstructing Thin 3D Structures in Optical Sectioning Microscopy. Proceedings of the International Conference on Image Processing, Part II, pp. 953-957.

Martišek, D., 2002. The Two-Dimensional and Three-Dimensional Processing of Images Provided by Conventional Microscopes. *Scanning*, Vol. 24, 284-295, pp. 284-295.

Niederöst, M., Niederöst, J., Šcucka, J., 2003. Automatic 3D reconstruction and visualization of microscopic objects from a monoscopic multifocus image sequence International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV-5/W10.

Pentland, A. P., 1982. Depth of Scene from Depth of Field. Proceedings of the Image Understanding Workshop, pp. 253-259.