**34th Annual International Conference of the IEEE EMBS**
**San Diego, California USA, 28 August - 1 September, 2012**

# Automatic 3D reconstruction of quasi-planar stereo Scanning Electron Microscopy (SEM) images*

S. Roy, J. Meunier, *Member, IEEE*, A.M. Marian, F. Vidal, I. Brunette and S. Costantino

*Abstract*— **Scanning Electron Microscopy (SEM) is widely used in science to characterize the surface roughness of materials. Three-dimensional information can be obtained with SEM based on stereovision techniques. A stereo pair is typically obtained by tilting the sample by a few degrees. In this paper we present a fully automated method for 3D reconstruction from a SEM stereo pair without any particular constraint. Results are presented for corneal stromal surfaces.**

## I. INTRODUCTION

Scanning Electron Microscopy (SEM) is a powerful technology to assess the topography of material. It uses a particular type of electron microscope that scans a sample with a beam of electrons in a raster pattern to reveal surface smoothness and other properties. SEM is used in many fields ranging from biology to materials science and offers very high resolution on the order of nanometers.

Although some three-dimensional information appears in a single SEM image, more sophisticated techniques are needed to recover accurately the 3D roughness of a surface. One of these techniques relies on stereovision to achieve this goal. Two images (stereo pair) are taken with a small tilt angle difference. When observed with a stereo viewer a real 3D image shows up. Several researchers have investigated stereovision techniques (e.g. [1-5]). Currently, the extraction of 3D data using SEM stereo techniques is carried out most of the time semi-automatically [1-4]. For instance, the matching of corresponding points in the stereo pair must be performed manually. This is a very fastidious task for the operator. Automatic methods exist but require severe image acquisition constraints. For instance Ponz et al. [5] impose that the image center coincides with the eucentric point i.e. the point where the tilt axis intercepts the optical axis. Unfortunately this is not easy to do because precise alignment is required to get satisfactory 3D reconstructions.

To address these problems, in this paper we present a novel methodology for fully automatic 3D reconstruction of a SEM stereo pair without any constraints.

## II. METHODOLOGY

### A. Stereo pair acquisition

A SEM stereoscopic pair is acquired simply by tilting the specimen a few degrees and capturing approximately the same region of interest. The rotation angle depends on the relief of the material; the flatter the topography is, the larger the rotation angle must be. Some trials and errors might be necessary if no prior knowledge of the roughness of the material is known. The stereo reconstruction algorithm requires that, when observed from two different angles, the structures in the sample become displaced since the magnitude of this displacement is used to assign a surface elevation.

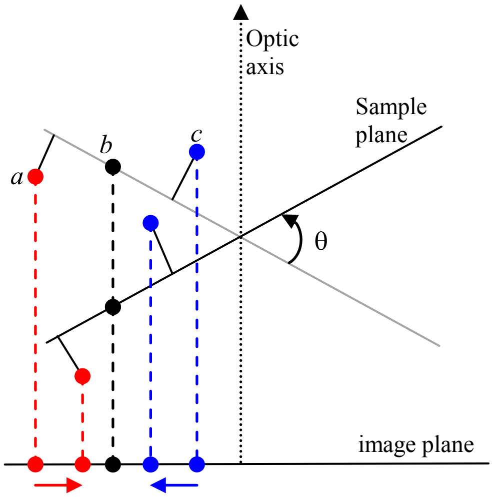

Figure 1. The reference plane concept.

### B. Reference plane

To assess the sample topography, we need a reference place corresponding to the zero height level of the specimen. For that purpose, a set of similar points is automatically identified in both stereo-images and the coordinates of these matching points are extracted. These correspondences are used to find a 2D affine transformation T, which models geometrically a tilt movement of the microscope stage.

---

*Research supported by The Canadian Institutes of Health Research (CIHR), the FRSQ Research in Vision Network, the Canadian National Institute for the Blind (CNIB), and the Fond de Recherche en Ophtalmologie de l'Université de Montréal.

S. Roy is with the Department of Computer Science and Operations Research, Université de Montréal, Montreal (QC), H3C3J7 Canada (e-mail: roys@iro.umontreal.ca).

J. Meunier is with the Biomedical Engineering Institute and Department of Computer Science and Operations Research, Université de Montréal, Montreal (QC), H3C3J7 Canada (corresponding author e-mail: meunier@iro.umontreal.ca).

A.M. Marian, I. Brunette and S. Costantino are with the Maisonneuve-Rosemont Hospital Research Center and Department of Ophthalmology, Université de Montréal, Montreal (QC), Canada.

F. Vidal is with the Institut National de la Recherche Scientifique – Énergie, Matériaux et Télécommunications, Varennes (QC), Canada.

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

By registering the stereo-images with the transformation T we obtain a reference plane (height=0) passing through the specimen for the subsequent measurement of heights. Fig. 1 represents graphically the sample reference plane concept. The projection on the image plane of a point *b* that is on the sample reference plane does not move from one image to the other (after tilting). A point *a* under the reference plane (negative height) moves to the right while a point *c* above the plane moves to the left. This is the amplitude of this displacement, known as disparity that is used to assign a surface height. Notice that we assume in this example that the reference plane was computed from several points in addition to the point *a, b* and *c* illustrated here.

Notice that the transformation T is Identity (no transformation) only when the specimen reference plane is tilted *exactly* at symmetrically opposite angles with respect to the optical axis as shown in Fig. 1, but this is rarely the case in practice.

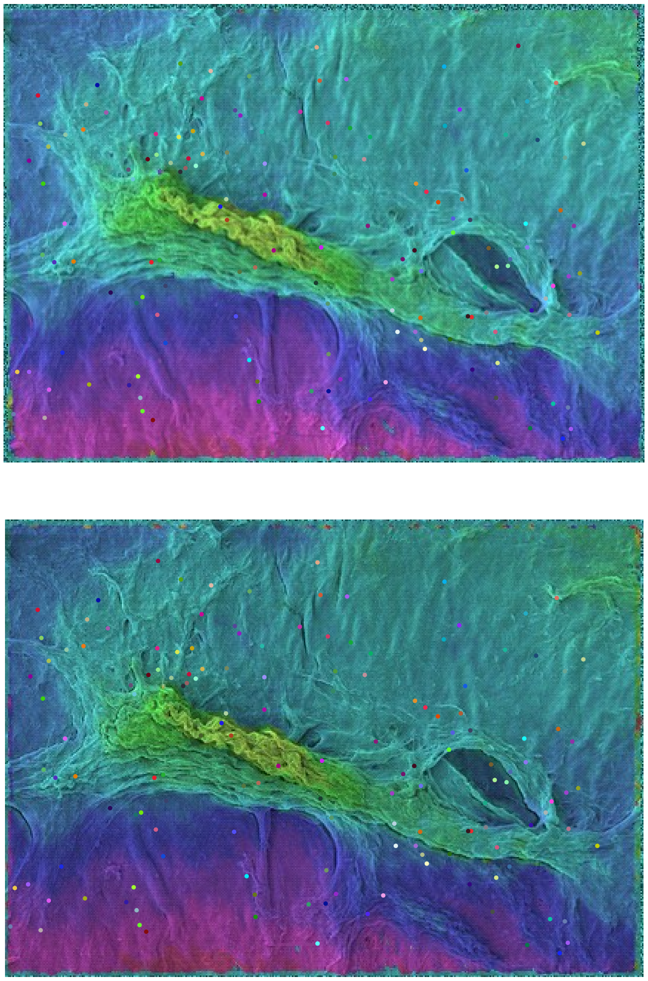

Figure 2. SURF keypoints of a SEM sample stereo pair (original images are monochrome, here false colors indicates depth).

Similar points are extracted with the SURF (Speeded Up Robust Feature) algorithm. SURF is a robust landmark point detector and descriptor [6] that was inspired by another well-known descriptor: the Scale-Invariant Feature Transform (SIFT) descriptor. However SURF is several times faster and more robust. The descriptor is based on Harr wavelets and makes an efficient use of integral images. One important advantage of these descriptors is that they are invariant to image scale, translation and rotation, a necessary feature for our work since the specimen is tilted between the two image captures. Point matching between the stereo pair is based on the descriptors and done through a Euclidean-distance based nearest neighbor approach. SURF (and SIFT) is used in computer vision tasks such as object recognition, panorama stitching or 3D reconstruction.

Fig. 2 shows SURF keypoints displayed with different colors (the corresponding points have the same color) for a SEM sample stereo pair. Notice that most keypoints are easily recognizable in both images.

These corresponding SURF keypoints in both stereo-images are used to find the best 2D affine transformation T with the Random Sample Consensus Method (RANSAC) [9]. RANSAC is a robust iterative fitting method that discards erroneous matched points (outliers) to determine the best geometric transformation.

### C. *Dense disparity map*

When comparing the registered stereo images, since the original sample was not planar, any depth deviation (from the reference plane) shows up as a residual displacement (or disparity) of any corresponding pair of pixels (see Fig. 1). For SEM, the imaging geometry can be modeled as an orthographic projection of specimen points onto the image plane with negligible error [7]. This ensures that all disparities are along the same direction. To ease computation, the images are rotated to make these displacements horizontal in order to establish a dense stereo matching of all pixels. This rotation is automatically obtained from the direction of displacements of SURF keypoints.

The horizontal stereo matching step is done with a robust and efficient algorithm based on semi-global matching [8]. The cost function used in this study is very simple since the brightness change of corresponding pixels is negligible for the small tilt involved in stereo SEM. In this study we used a sum of squared differences (SSD) over a 7 x 7 pixel window. Since it relies on dynamic programming, the matching algorithm uses an occlusion parameter which penalizes large depth discontinuities and, in practice, smoothes the disparity map and the corresponding reconstructed surface [8].

This step yields a disparity map i.e. the differences between the horizontal coordinates of the matching points in the stereo-images.

### D. *Depth (Height) map*

The horizontal disparity map is finally converted into heights according to the acquisition parameters: tilt angle, magnification, pixel size, with simple trigonometric equations. From Fig. 1, height $h$ and disparity $d$ in microns are related by the following equation: $d = 2h \sin(\theta/2)$. Therefore, the height $h$ (in microns) of a point whose disparity is $d$ (in pixels) is:

$$h = \frac{d \cdot p}{2 \sin\left(\frac{\theta}{2}\right)}$$

where $\theta$ is the total tilt angle and $p$ the pixel size in sample units (e.g. microns). The latter can be obtained from the scale provided by the SEM system or a calibration object (e.g. a microscopic grid) used for that purpose. Missing or hidden pixels can be estimated by interpolation.

Notice that this height computation is valid even if the specimen is not tilted exactly at symmetrically opposite angles with respect to the optical axis because it is relative to a reference plane passing through the specimen, which is implicitly computed with the registration of the stereo-images using transformation T (section II.B).

## III. EXPERIMENTAL RESULTS

The 3D height map of each surface using stereo pairs of SEM images were computed with an algorithm developed using Mathematica (Wolfram, Champaign, IL) [10].

### A. *Dataset*

A validation test was first performed using as reference sample a microscopic grid imaged at 1000X (magnification) with the scanning electron microscope. The reconstructed surface was then compared with manufacturer specifications (10 μm step width and 2 μm step height).

Several (17) lamellar cornea stroma (inner layer) cuts performed with a microkeratome or a femtosecond laser on human corneas unsuitable for transplantation and obtained from the local eye bank (Banque d'Yeux du Québec, Montreal, QC, Canada) were also used for this study. Samples were coated with a 20-nm layer of gold and also imaged at 1000X.

### B. *3D reconstruction*

All stereoscopic pairs were captured for 3D reconstruction of their surfaces at -3 and +3 degrees with respect to the optical axis.

Figure 3 shows one image of the reference grid at 1000X and a 2D horizontal profile of the reconstruction of its surface. This shows the precision of our methodology with the correct assessment of the 2 μm step height provided by the manufacturer.

We also successfully performed 3D reconstructions of the 17 samples (Fig. 4). A SEM image of one sample (Fig. 4A) and its corresponding elevation pseudo-color map overlaid on the original SEM image (Fig. 4B) are shown as an example.

A 3D rendering of the surface using the same color scale as in Fig. 4B is shown in Fig. 5. The X axis (Y axis) is the horizontal (vertical) axis of the sample. The topography displays a range of elevations of ±9 μm around the sample reference plane with less-than-a-micron height resolution.

By comparing Fig. 4A and Fig. 5 it becomes evident that this sample presents large scale height variations that could not be appreciated using a single SEM image (Fig. 4A).

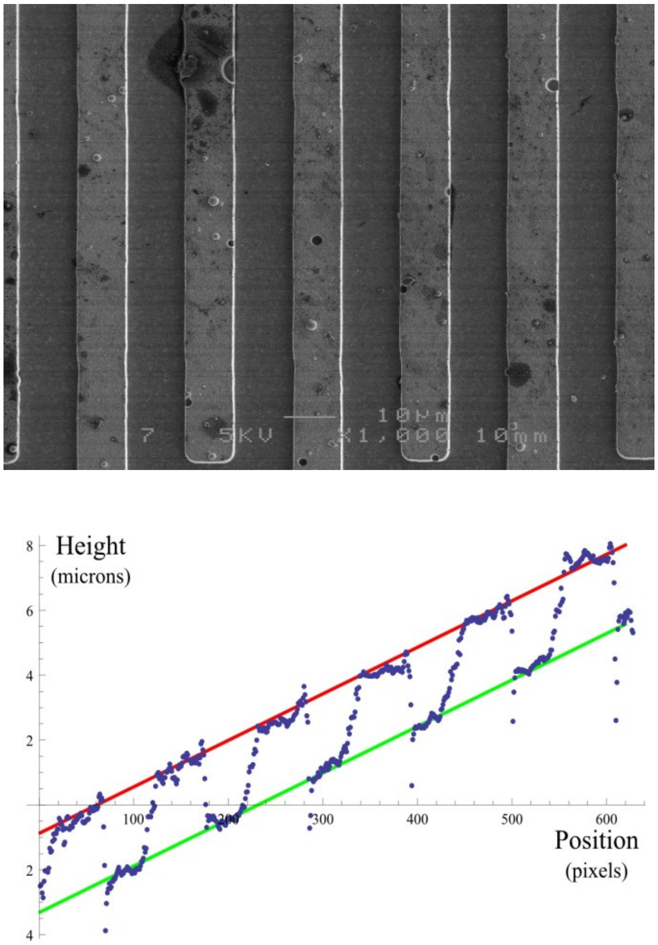

Figure 3. Validation test with a reference grid. One of the two SEM stereo images required for the reconstruction (top). 2D horizontal profile of a 3D reconstruction of the grid showing the correct (manufacturer) step height of 2 μm (bottom).

## IV. CONCLUSION

The stereo methodology presented here for SEM sample 3D reconstruction is fully automated. The operator has only to tilt the sample by a known angle and take two pictures, and then our methodology takes care automatically of the 3D reconstruction. No manual intervention (e.g. point matching) or severe image acquisition constraints (e.g. special positioning of the eucentric point) is needed. In practice, the specimen is tilted approximately at symmetrically opposite angles with respect to the optical axis but this is not mandatory. The total tilting angles typically range from 2° to 12° according to surface characteristics. Smooth surfaces require larger angles than rough surfaces to show measurable disparities. Large angles will therefore amplify disparities and show more subtle height variations, however, when elevation changes are large as for rough surface, low angles are preferred to avoid occlusions.

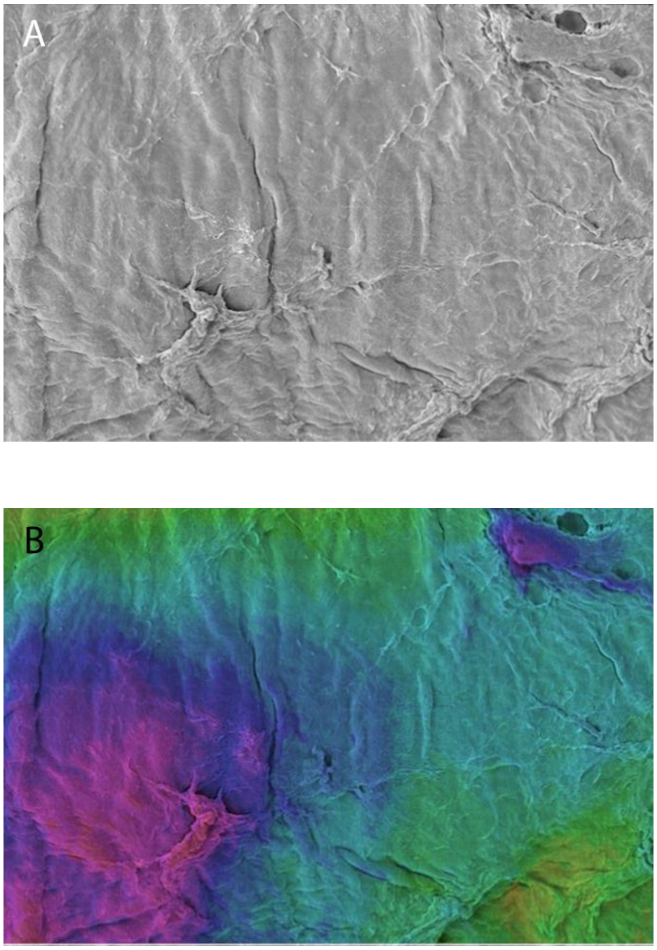

Figure 4. Representative example of SEM image 3D reconstruction. (A) One of the two stereo images required for the reconstruction. (B) Color map obtained after the 3D reconstruction overlying the original 2D image.

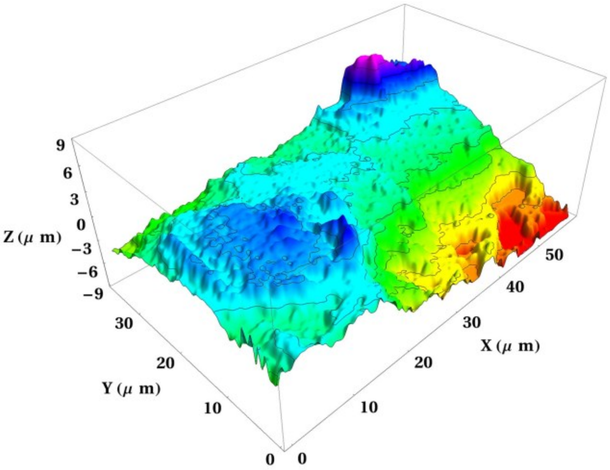

Figure 5. 3D rendering of the surface obtained after reconstruction.

Finally it is important to note that the real 3D topography of a sample surface is not the same as the one perceived from a single standard SEM image (Fig. 4A vs. Fig. 5). In a 2D image, small structures with small height variations (e.g. cracks, wrinkles) are more visible and evident while large height fluctuations are somewhat masked by them. Therefore 3D stereo reconstructions provide complementary information that is useful to assess surface roughness of materials. For instance, the standard deviation of the height, computed from the elevation map, can be used as a 3D roughness coefficient. In the future we intend to study the clinical utility of such 3D reconstructions to assess the quality of corneal lamellar dissections.

## REFERENCES

[1] GS Lane “The application of stereographic techniques to the scanning electron microscope. Journal of Physics E. Scientific Instruments 2:565, 1969

[2] JE Hilliard “Quantitative analysis of scanning electron micrographs. Journal of Microscopy, February, p. 45, 1972.

[3] G Piazzesi “Photogrammetry with the scanning electron microscope”, Journal of Physics E. Scientific Instruments 6:392, 1973

[4] A Kayaalp, AR Rao, R Jain, “Scanning Electron Microscope-Based Stereo Analysis”, Machine Vision and Applications 3:231-246, 1990

[5] E Ponz, JL Ladaga, RD Bonetto, “Measuring surface topography with scanning electron microscopy. I. EZElmage: A program to obtain 3D surface data”, Microsc Microanal 12, 170-177, 2006.

[6] H Bay, A Ess, T Tuytelaars, L Van Gool "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008

[7] PGT Howell “A theoretical approach to the errors in SEM photogrammetry”, Scanning 1:118-124, 1978.

[8] H Hirschmuller, “Accurate and efficient stereo processing by semi-global matching and mutual information”, Proceeding IEEE CVPR, 807-814, 2005

[9] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". Comm. of the ACM 24 (6): 381–395, 1981

[10] S Roy, J Meunier, AM Marian, F Vidal, I Brunette, S Costantino, Automatic 3D reconstruction of quasi-planar stereo SEM images. Technical report #1361, DIRO, Université de Montréal, 2012.