

## 3D Stereo Reconstruction of SEM Images

J. C. Henao-Londoño<sup>1</sup>, J. C. Riaño-Rojas<sup>2</sup>, J. B. Gómez-Mendoza<sup>3</sup> & E. Restrepo-Parra<sup>1</sup>

<sup>1</sup> Department of Physics and Chemistry, Universidad Nacional de Colombia, Manizales, Colombia

<sup>2</sup> Department of Mathematics and Statistics, Universidad Nacional de Colombia, Manizales, Colombia

<sup>3</sup> Department of Electric, Electronic and Computer Engineering, Universidad Nacional de Colombia, Manizales, Caldas, Colombia

Correspondence: E. Restrepo-Parra, Department of Physics and Chemistry, Universidad Nacional de Colombia, Manizales, Colombia. E-mail: erestrepopa@unal.edu.co

Received: Sep. 19, 2018

Accepted: October 8, 2018

Online Published: November 16, 2018

<https://doi.org/10.5539/mas.v12n12p57>

URL: <https://doi.org/10.5539/mas.v12n12p57>

### Abstract

In this work is proposed a new fully automated methodology using computer vision and dynamic programming to obtain a 3D reconstruction model of surfaces using scanning electron microscope (SEM) images based on stereovision. The horizontal stereo matching step is done with a robust and efficient algorithm based on semi-global matching. The cost function used in this study is very simple since the brightness and contrast change of corresponding pixels is negligible for the small tilt involved in stereo SEM. It is used a sum of absolute differences (SAD) over a variable pixel size window. Since it relies on dynamic programming, the matching algorithm uses an occlusion parameter which penalizes large depth discontinuities and, in practice, smooths the disparity map and the corresponding reconstructed surface. This step yields a disparity map, i.e. the differences between the horizontal coordinates of the matching points in the stereo images. The horizontal disparity map is finally converted into heights according to the SEM acquisition parameters: tilt angle, image magnification and pixel size. A validation test was first performed using as reference a microscopic grid with manufacturer specifications. Finally, with the 3D model are proposed some applications in materials science as roughness parameters estimation and wear measurements.

**Keywords:** SEM, 3D reconstruction, stereo vision, dynamic, programming

### 1. Introduction

Stereo matching has been one of the fundamental and most widely studied problems in computer vision (Scharstein & Szeliski. 2002, pp. 7-42). It has been successfully used in a variety of applications, including robot navigation, space missions, etc. (Tingbo, et. al. 2012, pp. 908-921). Stereo matching is the process of taking two or more images from different viewpoints simultaneously and estimating a 3D model of the scene by finding matching pixels in the images and converting their 2D positions into 3D depths (Tingbo, et. al. 2012, pp. 908-921) (Richard Szeliski. 2010).

The basic principle involved in the recovery of depth using passive imaging is triangulation. The triangulation needs to be achieved with the help of only the existing ambient illumination. Hence a correspondence need to be established between the features from two images that correspond to some physical feature in space. The major steps involved in the process of stereopsis are preprocessing, establishing correspondence, and recovering depth (Umesh, et. al. 1989, pp. 1489-1510). The solution may be represented as a disparity field specifying the positional differences of corresponding feature points relative to the image coordinate systems. The distance to the scene may then be computed from the disparity field, and from knowledge on the transformation relating the two image coordinate systems (Soren. 1990, pp. 309-315). A main problem in stereo analysis is to detect disparities over a large range of values. To reduce the search space, it is commonly assumed that the epipolar line geometry is known. This knowledge reduces the search space to a line segment (Soren. 1990, pp. 309-315).

The Scanning Electron Microscope (SEM) is a tool in a broad range of scientific and engineering applications (Johnson. 1996). The SEM is useful to observe and characterize different kind of materials on a nanometer (nm) to micrometer ( $\mu\text{m}$ ) scale (Joseph, et. al. 2003). The SEM can provide information on surface topography, crystalline structure, chemical composition and electrical behavior of the top 1  $\mu\text{m}$  or so of specimen (Shweta & Rama. 2013, pp. 105-114) (Bogner, et. al. 2007, pp. 390-401). The capability of the SEM to obtain three-

dimensional-like images of the surfaces is much appreciated (Mahovic, et. al. 2008, pp. 3449-3458) (Risovi, et. al. 2008, pp. 3063-3070). This characteristic makes the SEM very popular in a wide variety of media, from scientific journals to popular magazines to the movies (Joseph. et.al. 2003).

All SEM are composed of an electron column, that creates a beam of electrons; a sample chamber, where the electron beam interacts with the sample; detectors, that monitor different signals resulting from the beam- sample interaction (secondary and backscattered electrons, and X-rays) (Johnson. 1996) (Vernon. 2000, pp. 40-44); and a viewing system, that constructs an image from the signals.

An electron gun at the top of the column generates the electron beam. In the gun, an electrostatic field directs electrons emitted from a very small region on the surface of an electrode, through a small spot called the crossover. Then, the gun accelerates the electrons down the column toward the sample with energies typically ranging from a few hundred to tens of thousands of electron volts. Several types of electron guns like tungsten, LaB<sub>6</sub> (lanthanum hexaboride) and field emission (Joseph, et.al. 2003)(Bogner, et. al. 2007, pp. 390-401) (Vernon. 2000, pp. 40-44) use different electrode materials and physical principles, but all share the common purpose of generating a directed electron beam having stable and sufficient current, and the smallest possible size. The electrons emerge from the gun as a divergent beam. A series of magnetic lenses and apertures in the column reconverge and focus the beam into a demagnified image of the crossover. Near the bottom of the column, a set of scan coils deflect the beam in a scanning pattern over the sample surface (Joseph, et. al. 2003). The final lens focus the beam into the smallest possible spot on the sample surface. The beam exits from the column into the sample chamber. The chamber incorporates a stage for manipulating the sample, a door or airlock for inserting and removing the sample, and access ports for mounting various signal detectors and other accessories. As the beam electrons penetrate the sample, they give up energy, which is emitted from the sample in a variety of ways. Each emission mode is potentially a signal from which to create an image (Johnson. 1996).

## 2. Material and Methods

### 2.1 Stereo Pair Acquisition

3D reconstruction from stereoscopic images (acquired at varying specimen tilt angles (Kayaalp. 1990, pp. 21-246) is based on the measurement of the disparity, which is the shift (in pixels) of the specimen features from one image to the other (Pouchou. 2002, pp- 135-144). In SEM, images are formed by scanning a focused electron beam rectilinearly over the sample surface and synchronously detecting one of the signals generated by the beam-specimen interaction processes (Marinello. 2008).

Two main kinds of detection principle exist for imaging surfaces: secondary electron emission and back scattered electron emission. Secondary electrons, which produce an emission from less than ~ 10 nm below the surface, are usually preferred since they provide topographies with an optimal signal-to-noise ratio and eventually with very high resolution (Joseph, et.al. 2003) (Marinello. 2008). Viewing an object from two separated viewpoints which subtend an angle  $\theta$  at that object is equivalent to taking two images from a single viewpoint, but rotating the object through the same angle  $\theta$  between the images (Figure 1). It is this procedure that is usually employed in the SEM (Joseph, et.al. 2003). To produce a stereoscopic reconstruction, it is necessary to tilt the sample a few degrees to acquire two images (stereopair), and capture approximately the same region of interest (Joseph, et.al. 2003) (Marinello, 2008) (Roy, et. al. 2012, pp. 4361-4364).

To facilitate the acquisition of the images, it is hardly recommended to use a SEM with a manual or automatic specimen stage (Oliveira, et. al. 1999, pp. 256-263), with movement along five axes (x, y, z, tilt and rotation).

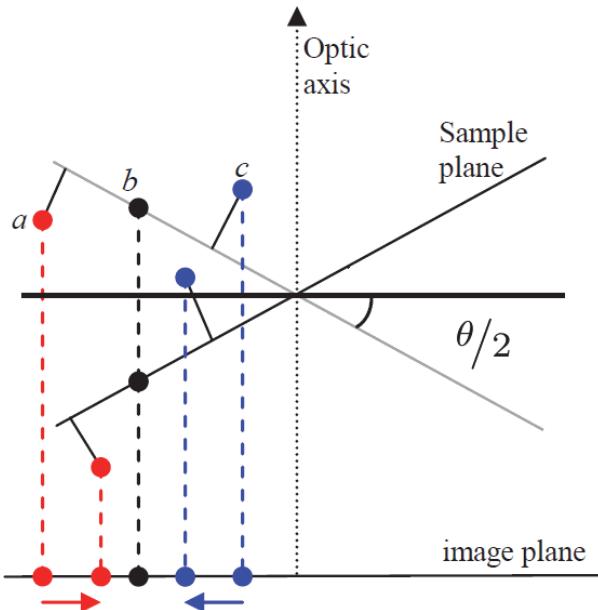


Figure 1. Tilt concept to acquire a stereo pair (Roy, et. al, 2012, pp. 4361-4364).

For simplicity in further calculation, we take the two images with equal tilt, but different sign.

## 2.2 3D Reconstruction using Dynamic Programming

### 2.2.1 Preprocessing

Before the stereo matching it is necessary to assure two main characteristics in the stereo pair to obtain accurately results: the brightness and the alignment.

The global histogram contrast must be similar for both images. A histogram matching was chosen to assure the same brightness and contrast in both images while preserving relative positions of edges and other textural characteristics (Oliveira, et. al. 1999, pp. 256-263). Specifically, the mean and standard deviation of one image is matched regarding the other.

The first step is set one image as the reference image, and the other one as the work image. To equalize the histograms, it is necessary to obtain the difference between each pixel of the work image and its mean. Then the result is divided by the standard deviation of the work image. This step makes the work image to have a mean equal to zero and a standard deviation equal to 1. Finally, the resulting image is multiplied by the standard deviation of the reference image and the mean value of the image is added to the equalized work image. Thus, both images will have approximately the same brightness and contrast.

There is uncertainty during the acquisition process that can produce image misalignments in the stereo pairs. Alignment is needed to ensure that only the parallax displacement components will be measured (Oliveira, et. al. 1999, pp. 256-263). Aiming to correct such misalignments, a geometric transformation that aligns the stereo pair is approximated.

To obtain a reasonable alignment between two images it is necessary to find corresponding points between the stereo pair. With this information is possible to calculate a geometric transformation, and finally use the matrix obtained to align one image with respect to the other.

The correspondence problem consists in finding correct point-to-point correspondences between images (Olof Enqvist. 2009). We have two images of the same 3D scene; each image tilted a few degrees. The objectives to find a set of distinctive features in one image which can be identified as the same features in the other image.

To search image point correspondences, the task can be divided into three main steps. Interest points are selected at distinctive locations in the image.

Next, the neighborhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and at the same time robust to noise, detection displacements and geometric and photometric deformations.

Finally, the descriptor vectors are matched between different images (Bay, et. al. 2008, pp. 346-359). Speeded-Up Robust Features (SURF) algorithm is used to find landmarks in the stereo pair. The landmarks detected in the two images are matched, conforming a set of corresponding points.

To find outliers in the set of points is used the Random Sample Consensus (RANSAC) algorithm, that makes the points considered as outliers can not be used to determine the geometric transformation. RANSAC is capable of interpreting/smoothing data containing a significant percentage of gross errors and is thus ideally suited for applications in automated image analysis where interpretation is based on the data provided by error-prone feature detectors (Fischler & Bolles. 1981, pp. 381-395). With this constraint, the transformation can be more accurate.

We estimate an affine geometric transform, which returns a transformation function, that can be applied to the positions of the corresponding points in one image to align them with the positions in the other image.

## 2.2.2 Stereo Matching and Dynamic Programming

Stereo matching is the process of taking two or more images from different viewpoints simultaneously and estimating a 3D model of the scene by finding matching pixels in the images and converting their 2D positions into 3D depths (Tingbo, et. al. 2012, pp. 908-921) (Szeliski. 2010) (Kayaalp. 1990, pp. 21-246)].

The basic principle involved in the recovery of depth using passive imaging is triangulation. The triangulation needs to be achieved with the help of only the existing ambient illumination. Hence a correspondence need to be established between the features from two images that correspond to some physical feature in space. The major steps involved in the process of stereopsis are preprocessing, establishing correspondence, and recovering depth (Dhond & Aggarwal. 1989, pp. 1489-1510).

The solution may be represented as a disparity field specifying the positional differences of corresponding feature points relative to the image coordinate systems. The distance to the scene may then be computed from the disparity field, and from knowledge on the transformation relating the two image coordinate systems (Olsen. 1990).

A main problem in stereo analysis is to detect disparities over a large range of values. To reduce the search space, it is commonly assumed that the epipolar line geometry is known. This knowledge reduces the search space to a line segment (Olsen. 1990) (Roy, et. al. 2012, pp. 4361-4364).

Dynamic programming solves problems by combining the solutions of subproblems (Cormen, et. al. 2009). The idea behind dynamic programming is given a problem, solve different parts of the problem (subproblems), and then combine the solution of the subproblems to reach an overall solution. Dynamic programming is applied to optimization problems. Such problems can have many possible solutions. Each solution has a value, and the objective is to find a solution with the optimal (minimum or maximum) value.

For every pixel in the left image, the goal is to find corresponding pixel in the right image (a match). Matching single pixels is nearly impossible, therefore every pixel is represented by a small region containing it, called correlation window. The window is centered around the pixel and has a constant size (Mhlmann, et. al. 2002, pp. 79-88).

The choice of the size of the window has some drawbacks. If the window is too small, it is going to find small details, but the result is going to be noisy. On the other hand, if the window is too big, it is not going to find small details, but the result is going to be smooth (Scharstein & Szeliski 2002, pp. 7-42). In general, the robustness of matching is increased with large areas. The dimensions of the correlation window have to be uneven; otherwise there would be an offset of half a pixel between the disparity map and the left image. (Mhlmann, et. al. 2002, pp. 79-88)

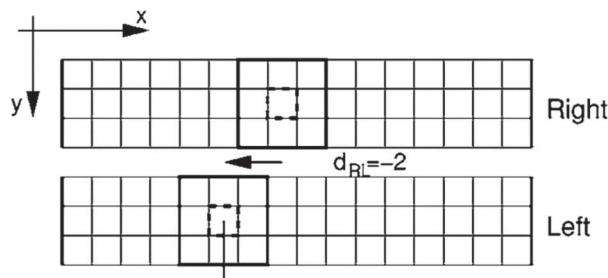


Figure 2. Disparity search using a correlation window [23]

The matching cost is calculated for a left image pixel from its intensity and the suspected correspondence in the

right image. It is used a sum of absolute differences (SAD) to correlate the windows of both images. Pixel-wise cost calculation is generally ambiguous and wrong matches can easily have a lower cost than correct ones, due to noise, etc (Hirschmller. 2008, pp. 328-341). An additional constraint is added that supports smoothness penalizing changes of the neighboring disparities. Finally, the matching cost has a term that is the sum of all pixel matching costs for the disparities, the second term adds a constant penalty for all pixels in the neighborhood of a pixel, for which the disparity changes a little bit, and a third term adds a larger constant penalty, for all larger disparity changes (Roy, et. al. 2012, pp. 4361-4364) (Hirschmller. 2008, pp. 328-341).

The problem of stereo matching can now be formulated as finding the disparity image that minimizes the cost function.

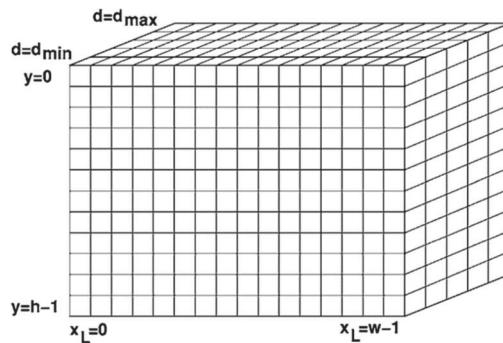


Figure 3. Disparity space volume defined by the dimensions of the left image and the disparity search range (Mhlmann, et. al, 2002, pp. 79-88)

The cost function values are stored in memory in a cuboid (Figure 3) whose dimensions are given by the width and the height of the images and the disparity ( $d$ ) search range from  $d_{\min}$  to  $d_{\max}$ . Every position  $(x, y, d)$  in the volume contains a similarity or correlation measure between the window representing position  $(x, y)$  in the left image and the one representing  $(x + d, y)$  in the right image (Mhlmann, et. al. 2002, pp. 79-88).

The disparity image that corresponds to the left image is determined by selecting for each pixel the disparity that corresponds to the minimum cost.

For subpixel estimation, a quadratic curve is fitted through the next higher and lower disparity, and the position of the minimum is calculated (Hirschmller. 2008, pp. 328-341).

The horizontal disparity map is finally converted into heights according to the acquisition parameters: tilt angle, magnification and pixel size, with simple trigonometric equations. If the tilt angles are different the trigonometric equations carry to a complicated relation between disparity and height (Dhond & Aggarwal, 1989, pp. 1489-1510). To simplify, we use symmetrical stereo pair. Height  $h$  and disparity  $d$  in  $\theta$  microns are related by the following equation (Szeliski, 2010) (Dhond. 1989, pp. 1489-1510) (Roy, et. al. 2012, pp. 4361-4364):  $d = 2 \cdot h \cdot \sin(\theta / 2)$ . Therefore, the height  $h$  (in microns) of a point whose disparity is  $d$  (in pixels) is:

$$h = (d \cdot p) / 2 \sin(\theta / 2) \quad (1)$$

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.

### 3. Results

The 3D height map of each surface using stereo pairs of SEM images were computed with an algorithm developed with the software Wolfram Mathematica §R 10. To evaluate our methodology, a validation test was performed using as reference a microscopic grid. The microscopic grid was imaged at x1000 magnifications with a scanning electron microscope (Jeol JSM-6300F, Jeol Ltd.). The reconstructed surface was then compared with manufacturer specifications (10µm step width and 2µm step height).

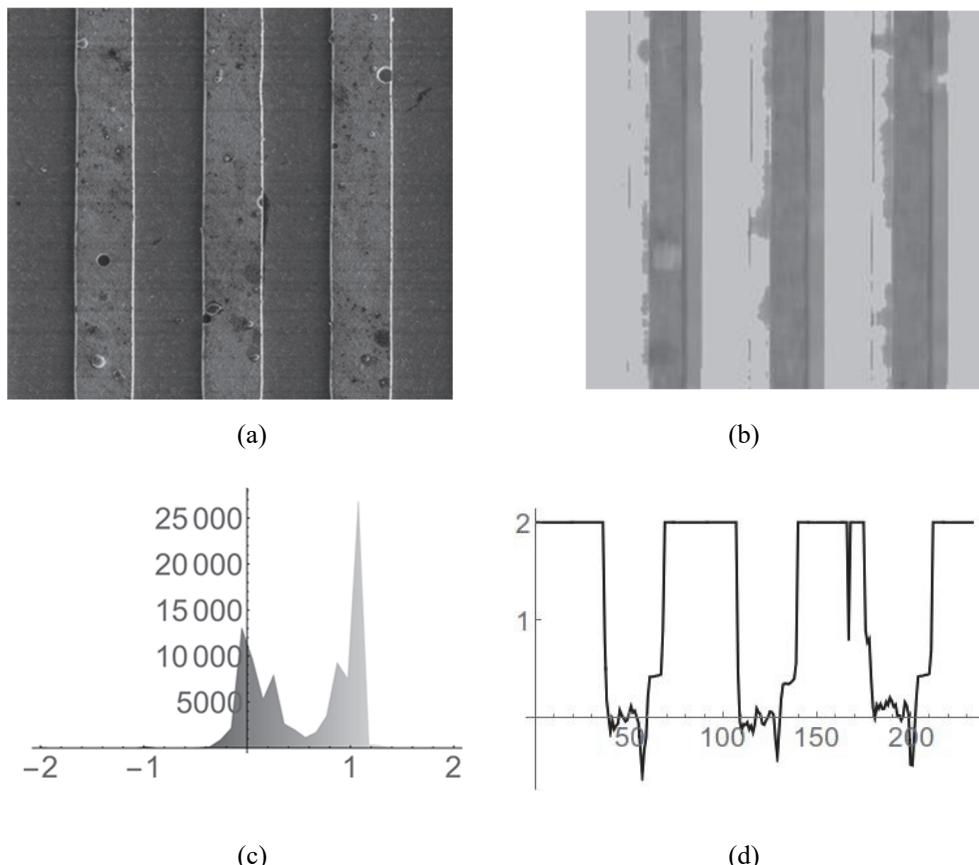


Figure 4. (a) Original image, (b) Gray level disparity map, (c) Histogram of disparities and (d) Profile of the sample

Figure 4 shows the result of the microscopic grid. At the top of the figure, the disparity map shows how the algorithm found a stable disparity map with two heights (top and bottom) as we can expect from a grid. The areas of equal height are discriminated according to the physical characteristics of the grid. Despite the edges are usually difficult to compute, the algorithm has an acceptable performance, and gives valuable information of the surface of interest, without smoothing the borders. At the bottom left the histogram of the disparities values shows two large picks representing the two disparity values of the image. Finally, at the bottom right, an image of all the profiles of the disparity map after the computation of the Equation 1, using the disparity map obtained with the stereo matching algorithm shows the stability of the results in homogeneous areas. It can be seen that the values of the height in the profiles of the disparity map are approximately the same value set by the manufacturer of the grid ( $2\mu\text{m}$  step height). They are two reasons because the values of the bottom of the grid in the profile are noisy. First, because the bottom of the grid is not homogeneous and the image shows all the rows of the disparity map, and second because the different kind of details, forms, and disparity information, that makes the profiles change between them.

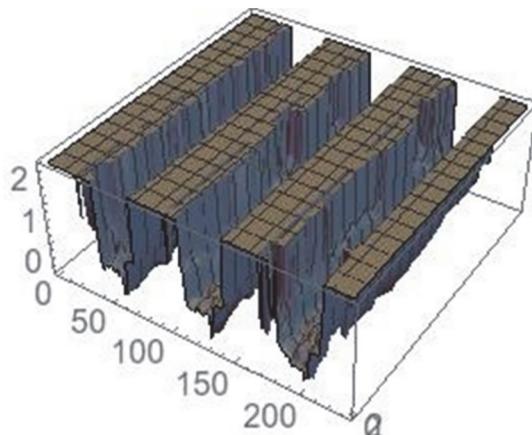


Figure 5. 3D reconstruction of the calibration grid

#### 4. Conclusions

SEM is an excellent tool to obtain information from micro to nano scale. It is robust and has a lot of applications and modules that make it an instrument widely used in research and industry. However, the three-dimensional information that can be obtained from it is limited, needing additional analysis to generate accurate 3D models.

In this work was studied a method using stereo vision and dynamic programming to analyze SEM stereo pairs and generate 3D models from them.

Implementing the 3D reconstruction algorithm was possible to estimate 3D models and surface roughness values using SEM stereo imaging. At the beginning, was used an optical flow algorithm to find the disparity map but the results were not enough accurate for our application. The algorithm uses the basics of stereo vision. Using dynamic programming the algorithm find the disparity map between the stereo pairs, that allows to calculate the 3D values through a trigonometric equation, supported in the geometry of the acquisition. The 3D values are used to plot the 3D model.

The 3D model of the calibration grid gave an accuracy result, showing the values that the manufacturer set. The results in edges and homogeneous areas are satisfactory and the algorithm make an excellent differentiation between the two heights of the sample.

With the 3D reconstruction algorithm and these results on the calibration grid, some applications can be considered to enrich the research. As future work we are looking forward to use the 3D models to measure 2D and 3D surface roughness parameters. We also expect to use the reconstruction methodology to propose a new way to obtain values of wear coefficients of materials. The work with the algorithm is not done yet. We expect to test our methodology using different calibration samples, alternative materials and new geometries.

#### Acknowledgments

This research was supported by the Administrative Department of Science, Technology, and Innovation - Colciencias with the program Jovenes Investigadores e Innovadores 2013 cooperation agreement No. 0737-2013.

#### References

- Bay, Herbert., Ess, Andreas, Tuytelaars, Tinne & Van Gool, Luc. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110, 346-359. <https://doi.org/10.1016/j.cviu.2007.09.014>
- Bogner, A., Jouneau, P. H., Thollet, G., Bassett, D., & Gauthier. C. (2007). *A history of scanning electron microscopy developments: Towards “wet-STEM” imaging.* *Micron*, 38, 390-401. <https://doi.org/10.1016/j.micron.2006.06.008>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*. The MIT Press, Cambridge, Massachusetts.
- Dhillon, S., & Kant, R. (2013). Quantitative roughness characterization and 3D reconstruction of electrode surface using cyclic voltammetry and SEM image. *Applied Surface Science*, 282, 105-114. <https://doi.org/10.1016/j.apsusc.2013.05.071>
- Dhond, Umesh R., & Aggarwal, J. K. (1989). Structure from Stereo – A Review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 1489-1510. <https://doi.org/10.1109/21.44067>

- Enqvist Olof. (2009). *Correspondence Problems in Geometric Vision*. Lund University.
- Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24, 381-395. <https://doi.org/10.1145/358669.358692>.
- Goldstein, J. I., Lyman, Ch. E., Newbury, D. E., Lifshin, E., Echlin, P., Sawyer, L., David, C. J., & Joseph, R. M. (2003). *Scanning Electron Microscopy and X-Ray Microanalysis*. Kluwer Academic/Plenum Publishers, New York.
- Hirschmller Heiko. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 328-341. <https://doi.org/10.1109/TPAMI.2007.1166>
- Jean-Louis, P., Denis, B., Pierre, B., Guy Le B., & Fabien, V. (2002). 3D Reconstruction of Rough Surfaces by SEM Stereo Imaging. *Mikrochim. Acta*, 139, 135-144. <https://doi.org/10.1007/s006040200052>.
- Johnson, R. (1996). *Environmental Scanning Electron Microscopy*. Philips Electron Optics.
- Karsten, M., Dennis, M., Jrgen, H., & Reinhard, M. (2002). Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation. *Int. J. of Comput. Vis.*, 47, 79-88. <https://doi.org/10.1023/A:1014581421794>
- Kayaalp, Ali, Ravishankar, A., Rao & Ramesh Jain. (1990). Scanning Electron Microscope-Based Stereo Analysis. *Machine Vision and Applications*, 3, 231-246. <https://doi.org/10.1007/BF01211849>
- Luis Rogerio de Oliveira Hein & Fernando de Azevedo Silva and Ana Maria Martinez Nazar and Jean Jacques Ammann. (1999). Three-Dimensional Reconstruction of Fracture Surfaces: Area Matching Algorithms for Automatic Parallax Measurements. *Scanning*, 21, 256-263. <https://doi.org/10.1002/sca.4950210405>.
- Mahovic Poljacek, S., Risovic, D., Furic, K. & Gojo, M. (2008). Comparison of fractal and profilometric methods for surface topography characterization. *Applied Surface Science*, 254, 3449-3458. <https://doi.org/10.1016/j.apsusc.2007.11.040>.
- Marinello, F., Bariani, P., Savio, E., Horsewell, A., & De Chiffre, L. (2008). Critical factors in SEM 3D stereo microscopy. *Meas. Sci. Technol.*, 19, 12. <https://doi.org/10.1088/0957-0233/19/6/065705>.
- Olsen, Soren I. (1990). Stereo Correspondence by Surface Reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 309-315. <https://doi.org/10.1109/34.49055>
- Risovi, D., Mahovi Poljacek, S., Furi, K., & Gojo, M. (2008). Inferring fractal dimension of rough/porous surfaces - A comparison of SEM image analysis and electrochemical impedance spectroscopy methods. *Applied Surface Science*, 255, 3063-3070. <https://doi.org/10.1016/j.apsusc.2008.08.106>
- Roy, S., Meunier, J., Marian, A. M., Vidal, F., Brunette, I., & Constantino, S. (2012). Automatic 3D reconstruction of quasi-planar stereo Scanning Electron Microscopy (SEM) images. 34th Annual International Conference of the IEEE EMBS, 4361-4364. <https://doi.org/10.1109/EMBC.2012.6346932>.
- Scharstein, D., & Richard, S. (2002). A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *Int. J. of Comput. Vis.*, 47, 7-42. <https://doi.org/10.1109/SMBV.2001.988771>
- Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Draft. Springer. <https://doi.org/10.1007/978-1-84882-935-0>.
- Tingbo, H., Baojun, Q., Tao, W., Xin, X., & Hangen, He. (2012). Stereo matching using weighted dynamic programming on a single-direction four-connected tree. *Computer vision and Image Understanding*, 116, 908-921. <https://doi.org/10.1016/j.cviu.2012.04.003>
- Vernon-Parry, K. D. (2000). Scanning Electron Microscopy: an introduction. *III-Vs Review*, 13, 40-44. [https://doi.org/10.1016/S0961-1290\(00\)80006-X](https://doi.org/10.1016/S0961-1290(00)80006-X).

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).