

Bootstrapping de-shadowing and self-calibration for scanning electron microscope photometric stereo

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 Meas. Sci. Technol. 25 105402

(<http://iopscience.iop.org/0957-0233/25/10/105402>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 133.87.177.243

This content was downloaded on 25/12/2015 at 04:47

Please note that [terms and conditions apply](#).

Bootstrapping de-shadowing and self-calibration for scanning electron microscope photometric stereo

Atsushi Miyamoto¹, Deshan Chen² and Shun'ichi Kaneko²

¹ Yokohama Research Laboratory, Hitachi Ltd, Yokohama 244-0817, Japan

² Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan

E-mail: atsushi.miyamoto.ts@hitachi.com, chen@ssc.ssi.ist.hokudai.ac.jp and kaneko@ssi.ist.hokudai.ac.jp

Received 26 March 2014, revised 15 July 2014

Accepted for publication 17 July 2014

Published 15 September 2014

Abstract

In this paper, we present a novel approach that addresses the blind reconstruction problem in scanning electron microscope (SEM) photometric stereo. Using only two observed images that suffer from shadowing effects, our method automatically calibrates the parameter and resolves shadowing errors for estimating an accurate three-dimensional (3D) shape and underlying shadowless images. We introduce a novel shadowing compensation model using image intensities for both cases of presence and absence of shadowing. With this model, the proposed de-shadowing algorithm iteratively compensates for image intensities and modifies the corresponding 3D surface. Besides de-shadowing, we introduce a practically useful self-calibration criterion by enforcing a good reconstruction. We show that incorrect parameters will engender significant distortions of 3D reconstructions in shadowed regions during the de-shadowing procedure. This motivated us to design the self-calibration criterion by utilizing shadowing to pursue the proper parameter that produces the best reconstruction with least distortions. As a result, we develop a bootstrapping approach for simultaneous de-shadowing and self-calibration in SEM photometric stereo. Extensive experiments on real image data demonstrate the effectiveness of our method.

Keywords: SEM photometric stereo, scanning electron microscope, shadowing compensation, self-calibration

(Some figures may appear in colour only in the online journal)

1. Introduction

The scanning electron microscope (SEM) is an essential tool for observing micro-structures. The need for three-dimensional (3D) surface measurements is significantly increasing. For example, in the semiconductor industry, the measurements of heights and 3D geometric structures are becoming even more crucial for precise control of critical dimensions as the sizes of electronic devices continue to shrink.

However, SEM only produces 2D images rather than 3D shapes directly. Attempts have been made to transform SEM into a 3D measuring tool for over 30 years, and scientific research on 3D surface reconstruction in SEM continues to

be intensively studied. Different approaches, which can be broadly classified into two groups, have been utilized. One popular group is generally referred to as the photogrammetry method [1–6], where a pair of stereo images of a specimen at different inclination angles are taken, and the shape or depth of surface features is determined by measuring the deviations of corresponding points on the two images. However, this method is practical only for measuring surface points where recognizable fine structures exist. In general, it cannot be used on smooth surfaces or to reconstruct a continuous surface profile at high magnification. The second approach is the so-called SEM photometric stereo method where multiple secondary or backscattered electron detectors (generally two or

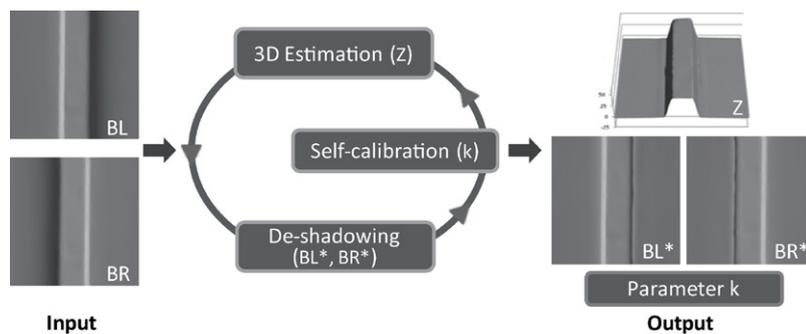


Figure 1. Framework of bootstrapping de-shadowing and self-calibration (BDS). Given observed images (BL and BR) that suffer from shadowing effects, the method iteratively eliminates shadowing errors and corrects the parameter via evaluation of reconstruction distortions. The algorithm provides outputs for the reconstructed surface (Z), underlying shadowless images (BL^* and BR^*) and the parameter (k).

four) are symmetrically positioned about the beam axis, so that the gradient information on the surface being observed is estimated from multiple images that are simultaneously obtained by multiple detectors [7–13]. The 3D surface is consequently computed from the gradient information. Despite extensive studies, SEM photometric stereo has not yet become a practical measurement technology. There exist several hindrances: two key issues are negative influences of shadowing effects and practical restrictions in the calibration process.

Shadowing is one of the most significant difficulties encountered in SEM photometric stereo, which frequently occurs in imaging processes and results in a smaller observed image intensity than the underlying shadowless one. Therefore, the gradient measurements in shadowed regions generally contain large deviations and consequently engender serious distortions in the reconstructed 3D surface. Paluszyński and Słówko [14] introduced several techniques for reconstructing the shape in the shadowed regions that utilize two or three unshadowed detectors under an arrangement of four secondary electron detectors. However, this method requires numerous manual interactions and interventions, such as shadowed regions that must be manually identified from the base images each time, and is restrictive in various actual applications.

Another practical difficulty is calibration, that is the recovery of parameters arising in gradient measurement formulas. Without proper calibrations, the shape cannot be reconstructed to an absolute level. Suganuma [9] utilized an off-line procedure with a sphere pattern as the calibration object. However, such an off-line approach is restrictive and inflexible. This is because a suitable calibration object may not always be available and the parameter may frequently vary due to a number of factors.

We present, in this study, a bootstrapping approach (refer also to figure 1) for joint surface reconstruction, a latent shadowless images estimation, and a self-calibration of the parameter to handle the challenging task of absolute height measurement from only two observed images via SEM photometric stereo in a blind setting, i.e. with no calibration objects or prior knowledge of the parameter. In the de-shadowing procedure, once the parameter is correctly associated, the method performs well in the reconstruction of the 3D surface and in the estimation of latent shadowless images. However, if the parameter is incorrect, the reconstruction will suffer from non-uniform distortions in the shadowed regions. To improve reconstruction,

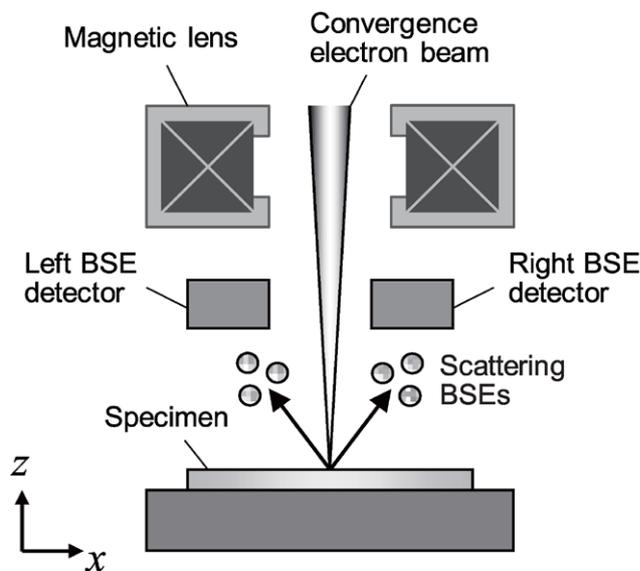


Figure 2. Schematic diagram of measuring system with two BSE detectors.

the proposed method combines de-shadowing and calibration and allows the two tasks to benefit from each other in the pursuit of producing an accurate 3D surface and determining the correct parameter. We investigate the interactions between de-shadowing and calibration and introduce an objective function to evaluate the reconstruction after de-shadowing and to correct the parameter. Our algorithm iteratively performs between de-shadowing and self-calibration to simultaneously eliminate shadowing errors and correct the parameter to obtain a reconstruction to the absolute level.

The main contributions of this study are as follows:

- We mathematically model the shadowing generation process and derive a shadowing compensation model. This model provides an important clue for estimating underlying shadowless images from shadowed ones.
- We present a novel de-shadowing approach utilizing our shadowing compensation model that produces both shadowless images and the 3D shape.
- We introduce a self-calibration criterion based on investigation of the internal relationships between the parameter and shadowing. We propose a solution to resolve the

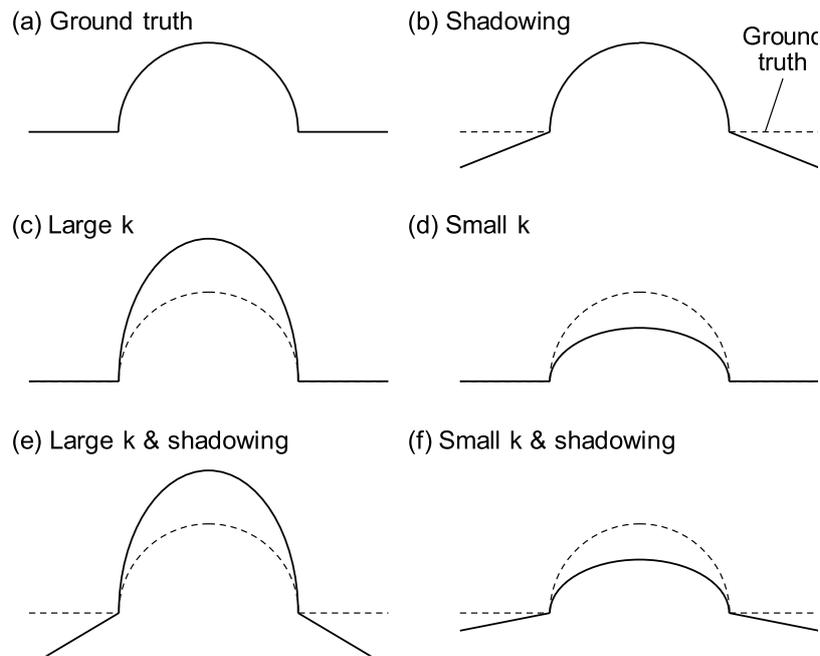


Figure 3. Schematic illustration of the effects of shadowing and parameter k .

challenging blind reconstruction problem which corrects the parameter while handling the shadowing errors.

The remainder of the paper is organized as follows. Section 2 briefly introduces the basis of SEM photometric stereo and its existing practical difficulties. Section 3 proposes the framework of our bootstrapping approach and presents an efficient optimization procedure. Experiments on real image data are carried out in section 4. Finally, we provide some discussion and conclude the paper in section 5.

2. Problem statement

In this section, we briefly introduce the basis of the SEM photometric stereo method and its existing problems.

2.1. The basis of SEM photometric stereo

SEM photometric stereo is substantially related to optic photometric techniques using an optical camera as the measuring device [15–18] in which both methods are used to estimate gradient information for an object being observed from shading cues in the image data (the special case, where the data is a single image, is generally known as shape from shading [19–21]). However, because of different imaging principles, gradient estimation formulas are derived in different ways.

As schematically illustrated in figure 2, in our SEM measurement system, two backscattering electron (BSE) detectors are symmetrically positioned along a direction. It is convenient for us to define this direction as the x -axis. Such a system simultaneously provides two BSE images from the two detectors, denoted as BL and BR, respectively. In general, we use BL and BR to refer to the left and right image, respectively. Under such system settings, Suganuma [19] proposed an experimentally derived model to measure the gradient component in the direction of the x -axis (i.e. the slope $\tan \alpha$):

$$\frac{\partial Z}{\partial x} = k \frac{BL^2 - BR^2}{(BL_n + BR_n)^2}, \quad (1)$$

where k is a parameter and BL_n and BR_n are signal outputs from two detectors on the horizontal specimen surface.

3D shapes are obtained from gradient measurements by either straightforward numerical integration [19] or advanced optimization techniques [22, 23].

2.2. Problems

The accuracy of the reconstructed shape highly depends on the corresponding accuracy of the gradient measurements. It is thus obvious that shadowing and calibration are the two key issues hindering absolute level reconstructions in SEM photometric stereo, as they directly affect image intensity (BL and BR) and the parameter (k) separately arising in (1).

As illustrated in figure 3(b), shadowing errors engender significant distortions in the reconstruction. In fact, shadowing phenomena frequently occur in imaging processes due to the exceptional features of sample topography. As shown in figure 4, a portion of the electrons is screened out because of occlusion, resulting in fewer electrons collected by the BSE detector and a smaller observed image intensity than the underlying shadowless one. Therefore, the gradient measurements generally contain large deviations when applying (1) in a region with shadowing effects. It is thus obvious that shadowing effects will significantly contribute to the final error of a reconstructed 3D surface.

The parameter k is also essential for absolute level reconstruction. It is clear that it can linearly change the non-zero measurements of the gradient in (1) and consequently deform the shape of non-flat regions (refer also to figures 3(c) and (d)). However, calibrating the parameter k is not a trivial task in practice. The conventional off-line calibration procedure is to acquire images of a calibration object with a known shape,

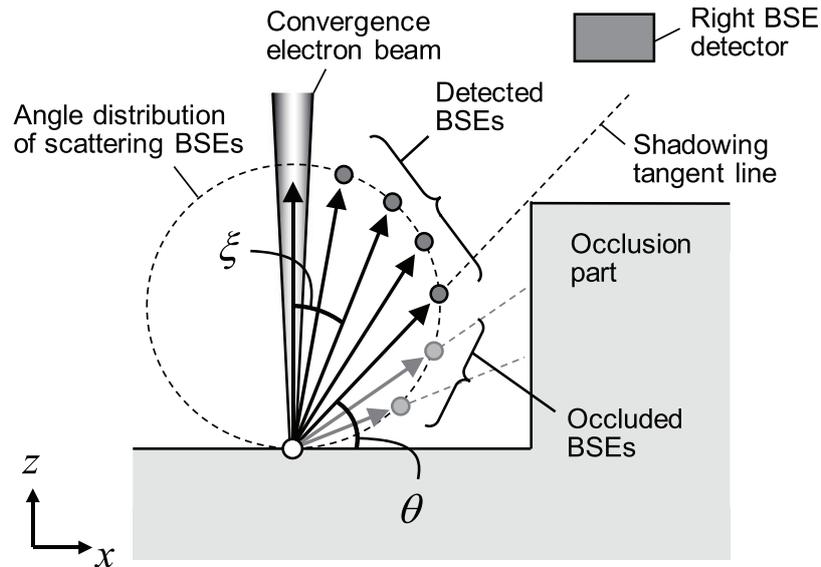


Figure 4. The shadowing effect. BSEs emitted in a direction below the tangent line are generally absorbed by the occlusion part and consequently are not collected by the detector.

such as a sphere, and compute the parameter by minimizing the gradient error between the estimation made from the images and the true ones calculated geometrically. Such off-line approaches generally provide good but constrained results. First, the parameter generally depends on a number of factors, such as the material of the specimen, and the imaging conditions, such as accelerating voltage, emission current, etc [24]. As these factors frequently change in practice, the parameter will vary accordingly. Thus one has to recalibrate the parameter whenever a factor changes, causing significant practical difficulties and restrictions. Second, the material of the calibration objects should be identical to the specimen to be observed, and their shapes must be precisely controlled. So the accessibility to proper calibration objects in practice is also limited.

To achieve an automatic and highly accurate SEM photometric stereo, we solve the joint of the two problems (refer also to figures 3(e) and (f)) by simultaneous de-shadowing and self-calibration. The overview of our proposed method is illustrated in figure 1. On the one hand, we seek to remove shadowing errors in reconstruction (de-shadowing). On the other hand, we automatically adjust parameter k by utilizing shadowing cues in the de-shadowing procedure (self-calibration).

3. Bootstrapping de-shadowing and self-calibration

In this section, we present our bootstrapping de-shadowing and self-calibration framework in a blind situation, i.e. with only shadowed left and right images as input and with no calibration object available, and develop an efficient algorithm to solve the problem.

A natural procedure for SEM photometric stereo would be to first calibrate the parameter k in (1) and then perform a de-shadowing algorithm to obtain an accurate reconstruction. However, we may only have two observed images in practice, which suffer from shadowing effects and for which no calibration object or prior information for the parameter k is accessible. As previously discussed, the shape in non-flat regions may be deformed

with incorrect values of k . In addition, as we will see, proper de-shadowing also requires a correct k . Therefore, reconstruction to the absolute level from only shadowed images is a very challenging task (dubbed as blind reconstruction). To the best of our knowledge, few studies have involved this problem. Here, we address the task of blind reconstruction by exploiting the interactions between de-shadowing and calibration.

We divide the problem into two steps. First, we focus on resolving the shadowing problem under a temporary assumption that the parameter k is known. We introduce a shadowing compensation model, which relates the underlying shadowless image to the observed one. Using the model, we formulate the proposed de-shadowing into a constrained optimization problem via a variational approach. We adopt an iterative scheme to solve the optimization problem, which effectively provides both an accurate 3D shape and compensated shadowless images after convergence. Second, we investigate how improper values of the parameter k affect reconstructed shapes in de-shadowing. Consequently, we design a novel self-calibration criterion by penalizing the distortions in the reconstructed shape caused by an incorrect k . We ultimately address the blind reconstruction problem by fusing de-shadowing and self-calibration.

3.1. De-shadowing

3.1.1. Shadowing compensation model. We model the mechanism of the generation of shadowing effects based on the following assumptions:

- The angle distribution of BSEs [24] is as follows:

$$\frac{d\eta}{d\Omega} = \frac{\eta_0}{\pi} \cos^n \xi,$$

where ξ describes the direction of the electron emission; η_0 and n are coefficients; and η and Ω denote the quantity of BSEs and the solid angle, respectively.

- The image intensity can be modeled as the amount of BSEs collected by the detector.

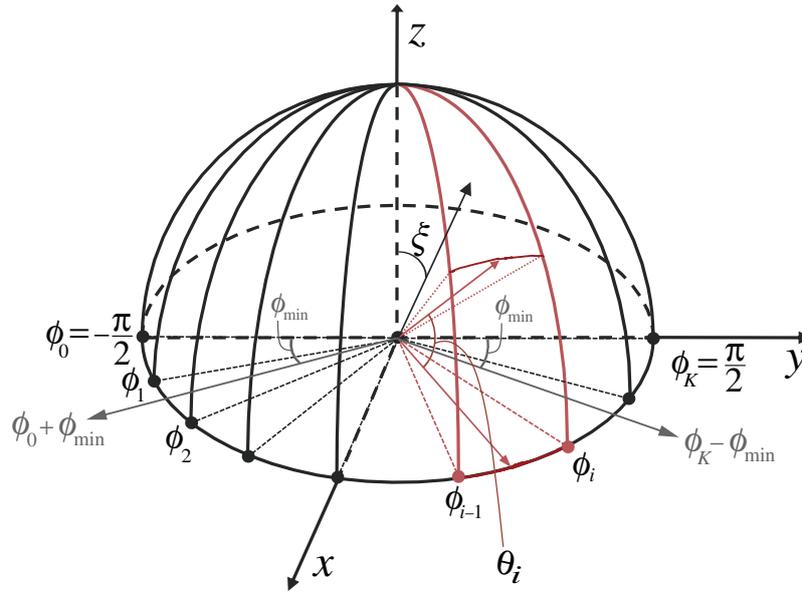


Figure 5. Segmentation of the detectable region of a BSE detector.

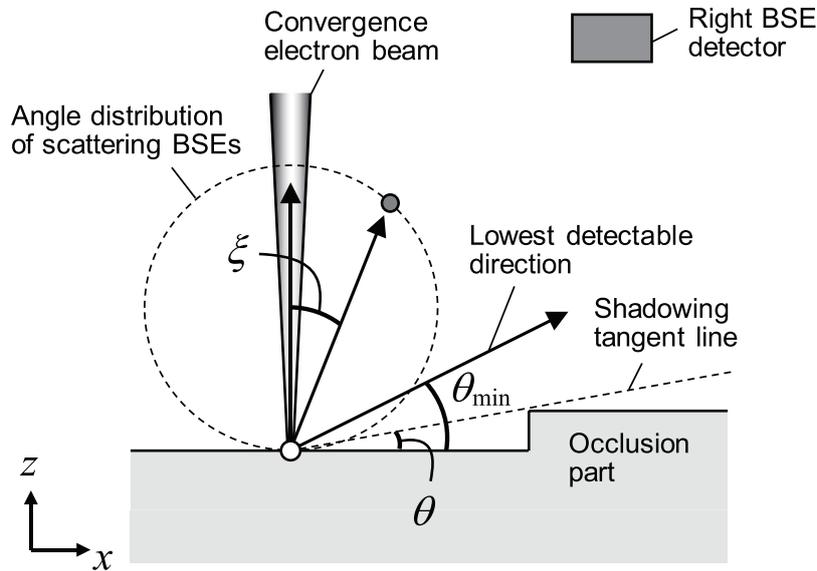


Figure 6. Geometric representations of practical factors θ_{\min} .

We first model the image intensity in the case of the presence of shadowing that in turn is convenient when modeling the shadowless case, where the shadowless case is considered just a special case of shadowing.

BSE intensity in the presence of shadowing. The ideal range of the azimuth angle ϕ for detectors should be $-\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$. Nonetheless, BSEs in the region where ϕ is near $\pm\frac{\pi}{2}$ generally cannot be detected owing to the practical limitations of the detector. Therefore, we introduce the parameter ϕ_{\min} (refer also to figure 5) to model this shortfall; then the range of the azimuth angle in the practical detectable region should be $-\frac{\pi}{2} + \phi_{\min} \leq \phi \leq \frac{\pi}{2} - \phi_{\min}$.

In the shadowing case, as shown in figure 4, the range of the zenith angle ξ is $0 \leq \xi \leq \frac{\pi}{2} - \theta(\phi)$ in general, i.e. the range

is between the z axis ($\xi = 0$) and the shadowing tangent line ($\xi = \frac{\pi}{2} - \theta(\phi)$), where $\theta(\phi)$ is the shadowing angle along the ϕ direction and generally varies for different values of ϕ . However, due to the finite size of practical detectors, BSEs with an emission direction near the horizontal direction may not be collected. We therefore introduce another parameter θ_{\min} to model the lowest detectable direction. In the case of figure 6, when the shadowing tangent line is below the lowest detectable direction, i.e. $\frac{\pi}{2} - \theta(\phi) > \frac{\pi}{2} - \theta_{\min}$, there is in fact no shadowing in the ϕ direction and the range of ξ is only up to the lowest detectable direction. Including both cases for the zenith range, the upper limit of the zenith angle should be $\min\left\{\frac{\pi}{2} - \theta_{\min}, \frac{\pi}{2} - \theta(\phi)\right\}$. Therefore, the detectable region takes the following form:

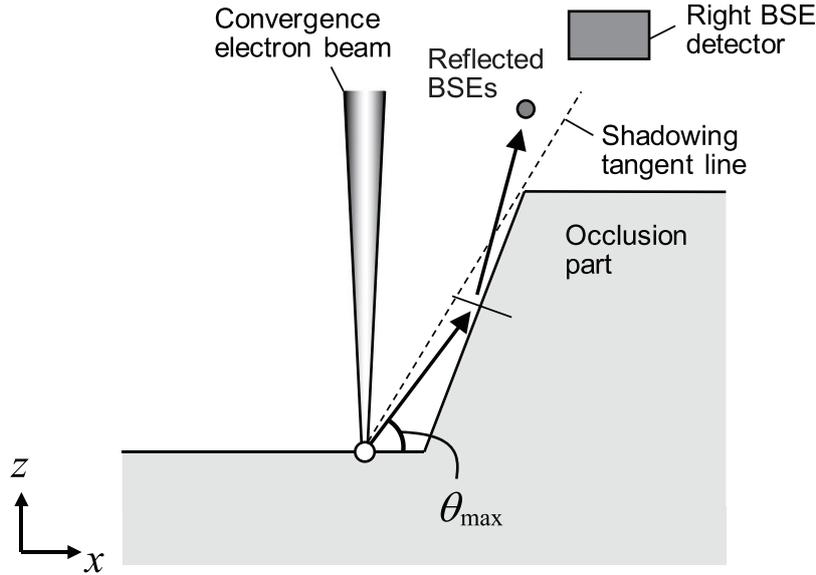


Figure 7. Geometric representations of practical factors θ_{max} .

$$\Omega^S: \begin{cases} -\frac{\pi}{2} + \phi_{min} \leq \phi \leq \frac{\pi}{2} - \phi_{min}, \\ 0 \leq \xi \leq \frac{\pi}{2} - \theta(\phi), (\theta(\phi) = \theta_{min}, \text{ if } \theta(\phi) < \theta_{min}). \end{cases}$$

Here, for convenience, we have transformed the expression of the upper limit of ξ via a thresholding operator.

With the assumption that the image intensity is modeled as the amount of BSEs collected by the detector, the observed BSE intensity is modeled by the following integral:

$$BSE = \int_{\Omega^S} \frac{d\eta}{d\Omega} d\Omega = \int_{\Omega^S} \frac{\eta_0}{\pi} \cos^n \xi d\Omega.$$

However, the problem is that an analytical expression for $\theta(\phi)$ is unlikely to be accessible. This issue can be resolved by dividing the detectable region into small segment regions with azimuth step $\Delta\phi = \pi/K$ (refer to figure 5). It is convenient to approximate ϕ_{min} : $\phi_{min} = S\Delta\phi$, where $S = [\phi_{min}/\Delta\phi]$ and $[\]$ denotes the rounding operator. It is reasonable to assume that the shadowing angle within each small region Ω_i^S remains approximately constant. Thus, the detectable region is approximated by

$$\Omega^S = \bigcup_{i=S+1}^{K-S} \Omega_i^S,$$

where

$$\Omega_i^S: \begin{cases} \phi_{i-1} \leq \phi \leq \phi_i, \\ 0 \leq \xi \leq \frac{\pi}{2} - \theta_i, (\theta_i = \theta_{min}, \text{ if } \theta_i < \theta_{min}). \end{cases}$$

Therefore, the BSE intensity can be derived as

$$\begin{aligned} BSE &= \int_{\Omega^S} \frac{\eta_0}{\pi} \cos^n \xi d\Omega = \sum_{i=S+1}^{K-S} \int_{\Omega_i^S} \frac{\eta_0}{\pi} \cos^n \xi d\Omega \\ &= \frac{\eta_0}{K(n+1)} \cdot \sum_{i=S+1}^{K-S} (1 - \sin^{n+1} \theta_i). \end{aligned} \tag{2}$$

Here we have employed the fact that $d\Omega = \sin \xi d\xi d\phi$ in the calculation of the integral.

BSE Intensity in the absence of shadowing. In fact, the shadowless case is considered as a special shadowing case where all the shadowing angles are zero. The shadowless BSE intensity is derived by setting $\theta_i = 0$ in (2). With the thresholding operator ($\theta_i = \theta_{min}$, if $\theta_i < \theta_{min}$), the BSE intensity in the case of absence of shadowing takes the following form:

$$BSE^* = \frac{\eta_0}{K(n+1)} \cdot (K - 2S)(1 - \sin^{n+1} \theta_{min}). \tag{3}$$

Practical modeling of θ_{max} . We previously gave the ideal assumption on occlusion phenomena, where the BSEs below the shadowing tangent line were totally absorbed. However, as illustrated in figure 7, when the emission direction of the BSEs is nearly parallel to the slope of the shadowing object (a considerably large incident angle), the electrons are likely to be re-scattered rather than absorbed into the shadowing objects. Such a reflected portion of BSEs is consequently detectable, which thus increases the corresponding image intensity and decreases shadowing effects. While this phenomenon is generally difficult to model exactly due to the complicated dependence on the structure of the occlusion feature, we thus approximately model it in a simple yet practical manner by setting another threshold, θ_{max} , for the shadowing angle. When the shadowing angle is larger than θ_{max} , it is set to θ_{max} to eliminate false shadowing effects resulting from the unshadowed portion (reflected portion) of BSEs.

Shadowing compensation model. Combining the image intensity of both cases of presence (2) and absence (3) of shadowing and summarizing this practical modeling of parameters, we obtain the following shadowing compensation model:

$$BSE^* = \frac{1}{R} \cdot BSE, \tag{4}$$

where the detection ratio is

$$R = \frac{\sum_{i=S+1}^{K-S} (1 - \sin^{n+1} \theta_i)}{(K - 2S)(1 - \sin^{n+1} \theta_{\min})}, \quad (5)$$

and the thresholding operator is

$$\theta_i = \begin{cases} \theta_{\min} & \theta_i \leq \theta_{\min}, \\ \theta_{\max} & \theta_i \geq \theta_{\max}, \\ \theta_i & \text{otherwise.} \end{cases}$$

The detection ratio R describes the mechanism of generation of shadowing via shadowing angles that imply the occlusion amount. In fact, shadowing angles are determined by the 3D surface. Therefore each shadowing angle should be a function of the 3D surface, i.e. $\theta_i = \theta_i(Z)$.

One advantage of the shadowing compensation model is that there is no need to treat the regions with shadowing errors separately from those without such errors, because the shadowless case is a special case of shadowing where detection ratio $R = 1$ (shadowing angles should be θ_{\min} after implementation of the thresholding operator). Therefore, in contrast to a literature example [14], this case does not require an image segmentation process to extract shadowing regions, a process that is substantially difficult to implement automatically.

3.1.2. Estimation of 3D surface and shadowless images. With the gradient measurement model described in (1) and the shadowing compensation model described in (4), the de-shadowing problem is formulated as follows: minimize

$$\mathbf{E}_1 = \int \int_{\mathcal{D}} \left[\left(\frac{\partial Z}{\partial x} - k \frac{BL^{*2} - BR^{*2}}{(BL_n + BR_n)^2} \right)^2 + \lambda |\nabla Z| \right] dx dy, \quad (6)$$

$$\text{s.t.} \quad \begin{cases} BL^* = \frac{BL}{R_L}, \\ BR^* = \frac{BR}{R_R}, \end{cases} \quad (7)$$

over some domain of interest \mathcal{D} in the plane, where BL^* and BR^* are the underlying shadowless left and right images, R_L and R_R denote corresponding detection ratios, λ is a parameter and $|\nabla Z| = \sqrt{\left(\frac{\partial Z}{\partial x}\right)^2 + \left(\frac{\partial Z}{\partial y}\right)^2}$.

The formulation can be explained as follows. The first term in (6), named the fidelity term, is to guarantee that the gradient of the reconstructed surface is consistent with the gradient measurement. To eliminate shadowing effects in reconstruction, the gradient should be evaluated from the shadowless images (BL^* and BR^*). As constraints, shadowless images are related to the corresponding observed images (BL and BR) through our shadowing compensation model. The second term is total variation (TV), a regularizer used to penalize the roughness of the solution and consequently make our method robust to noise influence. A TV model preserves edges well while removing noise effects and has proven to be an effective method applicable to a range of problems in image processing and computer vision, such as reconstruction, denoising, and deblurring [25]. λ is a positive scalar factor that weighs the

relative contributions of these two terms. Note that λ depends upon the noise level of the gradient measurements; the noisier the measurements, the larger should the value of λ be.

Solving such a constrained optimization problem is not trivial. As previously mentioned, shadowing angles arising in the detection ratio described in (5) is a function of surface Z , i.e. $\theta_i = \theta_i(Z)$. Therefore, substituting (7) into (6) results in \mathbf{E}_1 as a functional of Z . However, the dependence of shadowing angles on Z is very complicated. It is futile to present $\theta_i(Z)$ in an analytic form. Therefore, it is difficult to directly minimize it. We adopt the alternating minimization scheme that performs commutatively between the following two steps:

(a) **Surface reconstruction**

Fixing BL^* and BR^* in (6), we recover Z by solving the resulting unconstrained optimization problem (refer also to the Appendix).

(b) **Shadowing compensation**

Given surface Z , shadowing angles are numerically calculated. BL^* and BR^* are updated with (5).

The algorithm procedure is summarized in algorithm 1.

Algorithm 1. De-shadowing

Input: k , BL and BR

Output: Z , BL^* , and BR^*

Initialize: $BL^* = BL$, $BR^* = BR$;

While no convergence **do**

1: Fix BL^* and BR^* , update the surface Z from $Z = \arg \min_Z \mathbf{E}_1$
The implementation is presented in the appendix.

2: Fix Z , update BL^* , BR^* via (7) with (5)

end

return Z , BL^* , and BR^*

3.2. Self-calibration

As previously discussed, the off-line calibration technique is practically inflexible and restrictive due to its prerequisite requirement of calibration objects and awkwardness in handling frequent variations. In this section, we explore an alternate self-calibration strategy.

One immediate idea may be to design a calibration criterion by modeling the physical dependencies of the parameter k . However, such dependencies on specimen material and imaging conditions, for example, are physically complicated in general. Therefore, it is difficult to model these physical processes to obtain a self-calibration criterion.

In fact, the need for good reconstructions is generally a strong motivation for the pursuit of a proper parameter. We will introduce a self-calibration criterion to choose the proper k by enforcing a good reconstruction with minimum distortions. Such an approach may not be physically tenable, but it turns out to be practical.

We introduce our self-calibration method based on the following assumptions:

- The surface to be measured is piece-wise smooth and has a low total variation in general.

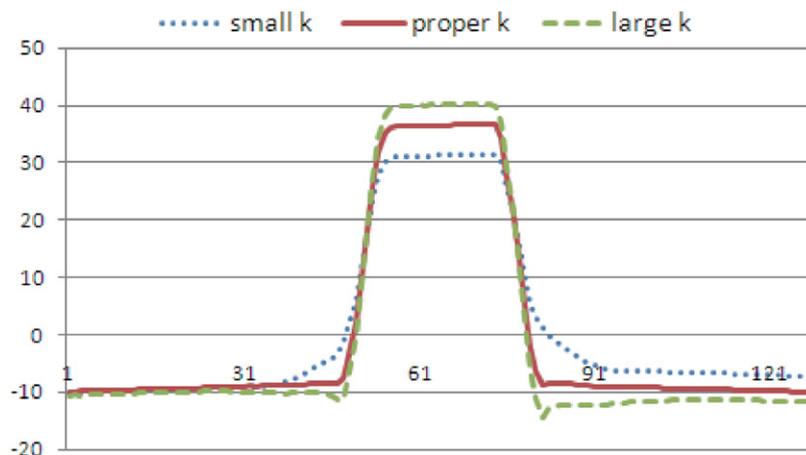


Figure 8. Under-compensation and over-compensation effects caused by improper k in de-shadowing. Section profiles of a line pattern are for demonstration purposes.

- Shadowing effects exist, i.e. there are shadowing errors somewhere in at least one of the images (BL or BR).

The first assumption, in fact, has already been used in the de-shadowing procedure, where, from the statistical point of view, the low total variation assumption is equivalent to the Laplacian prior. The second assumption may appear somewhat strange, because, in the previous section, we focused our efforts on removing shadowing errors. In this section, we will make use of shadowing information as an important cue for self-calibration. However, the assumption of the existence of shadowing effects is reasonable in general, as shadowing phenomena normally occur in actual measurements.

We will first investigate how improper values of k engender reconstruction distortions in de-shadowing which turns out to be a key constraint for our self-calibration. With such a constraint, we will design our self-calibration criterion and introduce the blind reconstruction method.

3.2.1. Constraints between k and reconstruction in de-shadowing. Figure 8 shows how incorrect values of k affect the reconstruction results in the de-shadowing procedure. The de-shadowing results of a line pattern with three different values of k are presented. As shown, in the case of small k , the shadowing effects remain in the reconstructed shape, named under-compensation; in the case of large k , an overshooting phenomenon occurs, called over-compensation.

The main cause of these distortions is that an improper k can incorrectly reflect shadowing in the shadowing compensation stage, and can be explained as follows. As previously discussed, k may deform the shape of non-flat regions. Thus, once k is incorrect, the estimation of detection ratio R in the de-shadowing process will generally be inaccurate. This is because the corresponding shadowing angles that are calculated from the deformed 3D surface are incorrect. In short, when k is smaller than the true value, the shadowing angles will also be calculated smaller and consequently engender under-compensation effects. Conversely, large values of k

will result in over-compensation effects. Note that incorrect k may also influence the boundary between estimated shadowing and shadowless regions in the de-shadowing process, i.e. the under-compensation phenomenon provides insufficient compensation in shadowing regions and may even mistreat some shadowing regions as shadowless regions, and vice versa; and the over-compensation phenomenon may mistakenly treat some shadowless regions as shadowing regions.

In summary, the constraint between parameter k and reconstruction is that proper reconstruction in de-shadowing requires a correct k . In short, if k is not correctly associated (either large or small), the reconstruction after de-shadowing will contain significant distortions due to improper compensations of shadowing.

3.2.2. Self-calibration criterion. Inspired by the previous constraint that an incorrect parameter causes reconstruction distortions in de-shadowing, the main idea of our self-calibration is to automatically adjust parameter k by enforcing the best reconstruction with least distortions in the de-shadowing procedure. However, one cannot know anything about distortions without any information on the shape. The most straightforward idea for evaluating distortions would be measuring the difference with the true shape, such as by the use of root mean square error, but this is obviously too strong an assumption and impossible for us to utilize, as the true shape is the goal we are pursuing. To overcome such a dilemma, we again employ the prior shape, i.e. the low total variation assumption that we have posited.

We thus introduce the following evaluation function as the reconstruction distortion metric:

$$E_2 = \frac{\int \int_{\mathcal{D}} \omega(\min\{R_L, R_R\}) |\nabla Z| dx dy}{\int \int_{\mathcal{D}} \omega(\min\{R_L, R_R\}) dx dy}, \tag{8}$$

where the weight function is

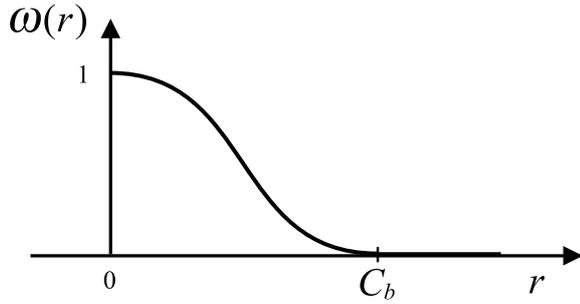


Figure 9. Weight function.

$$\omega(r) = \begin{cases} \left[1 - \left(\frac{r}{C_b}\right)^2\right]^2 & r \in [0, C_b), \\ 0 & r \in [C_b, 1]. \end{cases} \quad (9)$$

Here C_b is a positive constant.

The basic idea of this model is that if the value of k is incorrect, the reconstruction distortions after de-shadowing will give rise to a large deviation from the low total variation assumption. We thus seek the proper k that minimizes \mathbf{E}_2 , which implies that the obtained reconstruction is the best in the sense of minimum distortion. Here we adopt a weighted TV model rather than the pure TV model given in (6). This is due to the fact that reconstruction deviations caused by incorrect k do not obey a uniform distribution over the domain \mathcal{D} . As shown in figure 8, the distortions mainly occur in the shadowed regions, while in the shadowless regions the shape is seldom influenced. Therefore, if we use the pure TV model, it has a strong regularization effect on shadowless regions and the estimation of k may easily fail. Moreover, we also observe that in the shadowing regions the stronger the shadowing, the larger the deviation is. Based on these observations, our weight function (refer also to figure 9) is designed as a continuous monotone decreasing function, where C_b is a threshold that controls the shape of the weight function. Thus when $\min\{R_L, R_R\}$ is small, implying that at least one of the detectors suffers considerably from strong shadowing at this position, the weight will be large enough to let this point contribute significantly to \mathbf{E}_2 and vice versa.

3.2.3. Blind reconstruction with self-calibration. With the de-shadowing method and the self-calibration criterion, we ultimately address the challenging blind reconstruction problem via a bootstrapping approach that performs iteratively between de-shadowing and self-calibration:

(a) **De-shadowing**

Given k , de-shadowing procedure reconstructs surface Z , and its corresponding shadowless images.

(b) **Self-calibration**

With reconstructed surface Z , the self-calibration procedure updates k by decreasing the energy of \mathbf{E}_2 .

Since k is implicitly hidden in \mathbf{E}_2 , it is difficult to analytically compute the gradient to apply a gradient descent method.

Algorithm 2. Bootstrapping de-shadowing and self-calibration

Input: BL, BR
Output: Z, BL^*, BR^*, k
Initialize: k ;
Set: step length Δ , coefficient m ($0 < m < 1$);
 $Z = \arg \min_{Z, BL, BR} \mathbf{E}_1(Z|k)$;
for $t = 1$ **to** T **do**
 $k^+ = k + \Delta$;
 $Z^+ = \arg \min_{Z, BL, BR} \mathbf{E}_1(Z|k^+)$;
 if $\mathbf{E}_2(Z^+) < \mathbf{E}_2(Z)$ **then**
 $k = k^+$;
 end
 if $\mathbf{E}_2(Z^+) > \mathbf{E}_2(Z)$ **then**
 $k^- = k - \Delta$;
 $Z^- = \arg \min_{Z, BL, BR} \mathbf{E}_1(Z|k^-)$;
 end
 if $\mathbf{E}_2(Z^-) < \mathbf{E}_2(Z)$ **then**
 $k = k^-$;
 end
 if $(\mathbf{E}_2(Z^+) \geq \mathbf{E}_2(Z)) \&\& (\mathbf{E}_2(Z^-) \geq \mathbf{E}_2(Z))$ **then**
 $\Delta = m\Delta$;
 end
end
return Z, k, BL^* and BR^* ;

We thus adopt a numerical method to search descent directions. Algorithm 2 describes bootstrapping de-shadowing and a self-calibration algorithm.

4. Experimental results

In this section, we present several experiments on real image data to demonstrate the robustness and effectiveness of the proposed method. Two different patterns with strong shadowing effects are used to evaluate the proposed reconstruction method, including a standard particle with a sphere pattern and a semiconductor device with a line pattern. The SEM used in our experiment performs raster scanning along the direction located 45° from the x -axis. (The x -axis is defined as the direction along which the detectors are positioned.) Therefore, the x' -axis and the y' -axis consistent with the image array are defined for convenience in displaying the results. For computational efficiency, we embedded denominator $(BL_n + BR_n)^2$ into k in the Suganuma's gradient evaluation equation (1). We set $C_b = 0.8$ for the weight function in (9). The iteration number in Algorithm 2 is set to $T = 100$.

To give a vivid view of how our proposed Bootstrapping De-shadowing and Self-Calibration (BDS) method automatically corrects the parameter k and eliminates shadowing errors, we compare the proposed method with the following methods: (1) the traditional Suganuma method [9] applying the Depth From Gradient method (DFG) using (6) with $BL^* = BL$ and $BR^* = BR$; and (2) the de-shadowing method: the Shadowing Compensation Model (SCM) method via Algorithm 2.

Reconstruction results. Figure 10(a) and (b) show the original left and right images and corresponding estimated shadowless images. The sphere part/line part plays the role of the shadowing object for the surrounding parts. Figures 11 and 13

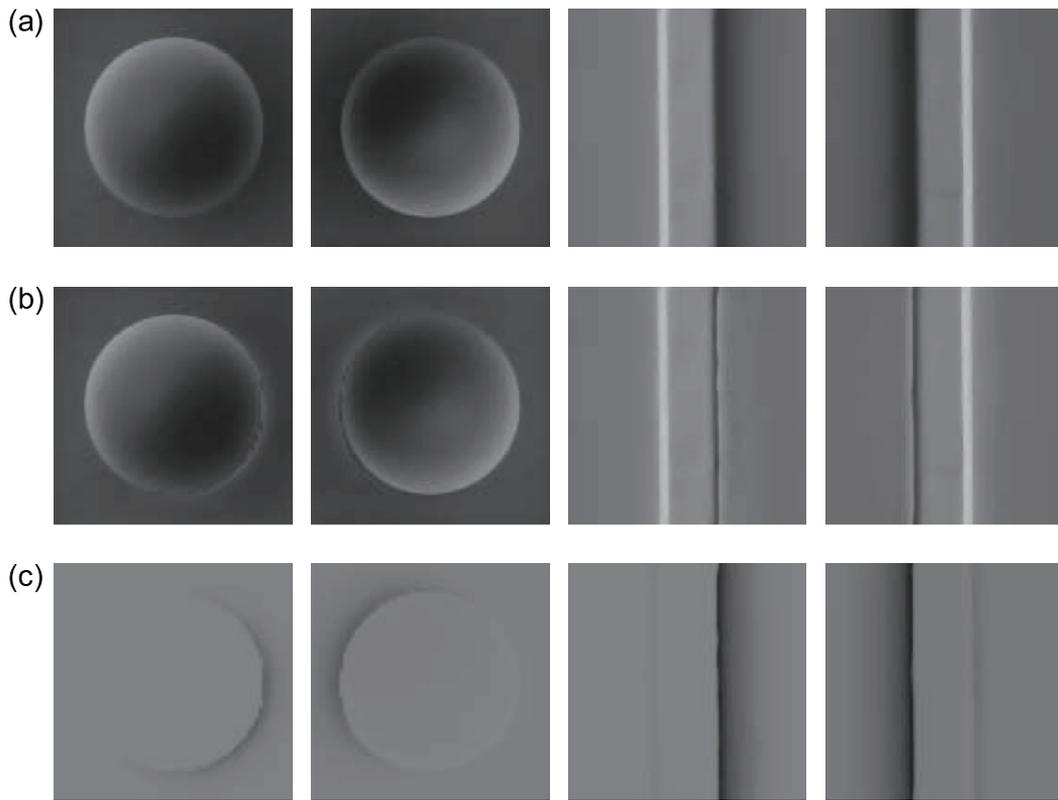


Figure 10. Original and estimated shadowless images: (a) original left and right images of a sphere pattern and line pattern; (b) corresponding estimated shadowless images; (c) difference images of original and shadowless images.

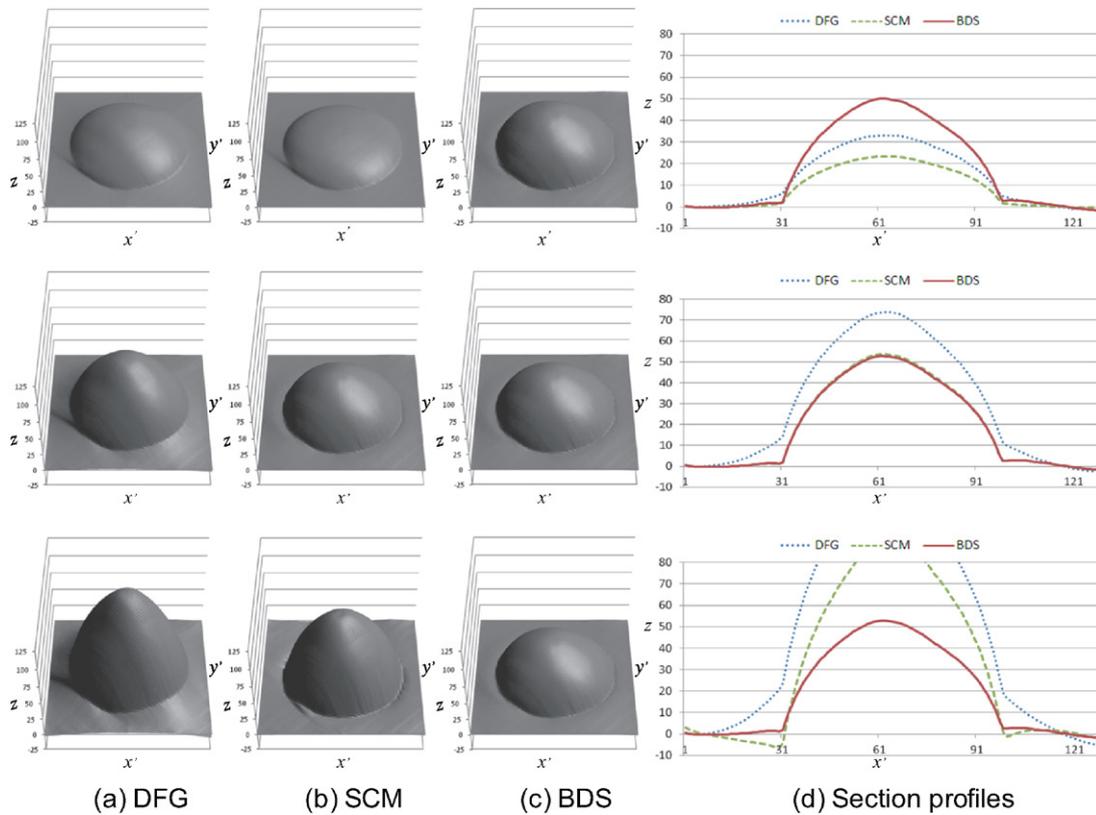


Figure 11. Comparison of 3D surface of the sphere pattern with initial value of k : top row $k = 0.00009$, middle row $k = 0.00019$ and bottom row $k = 0.00030$.

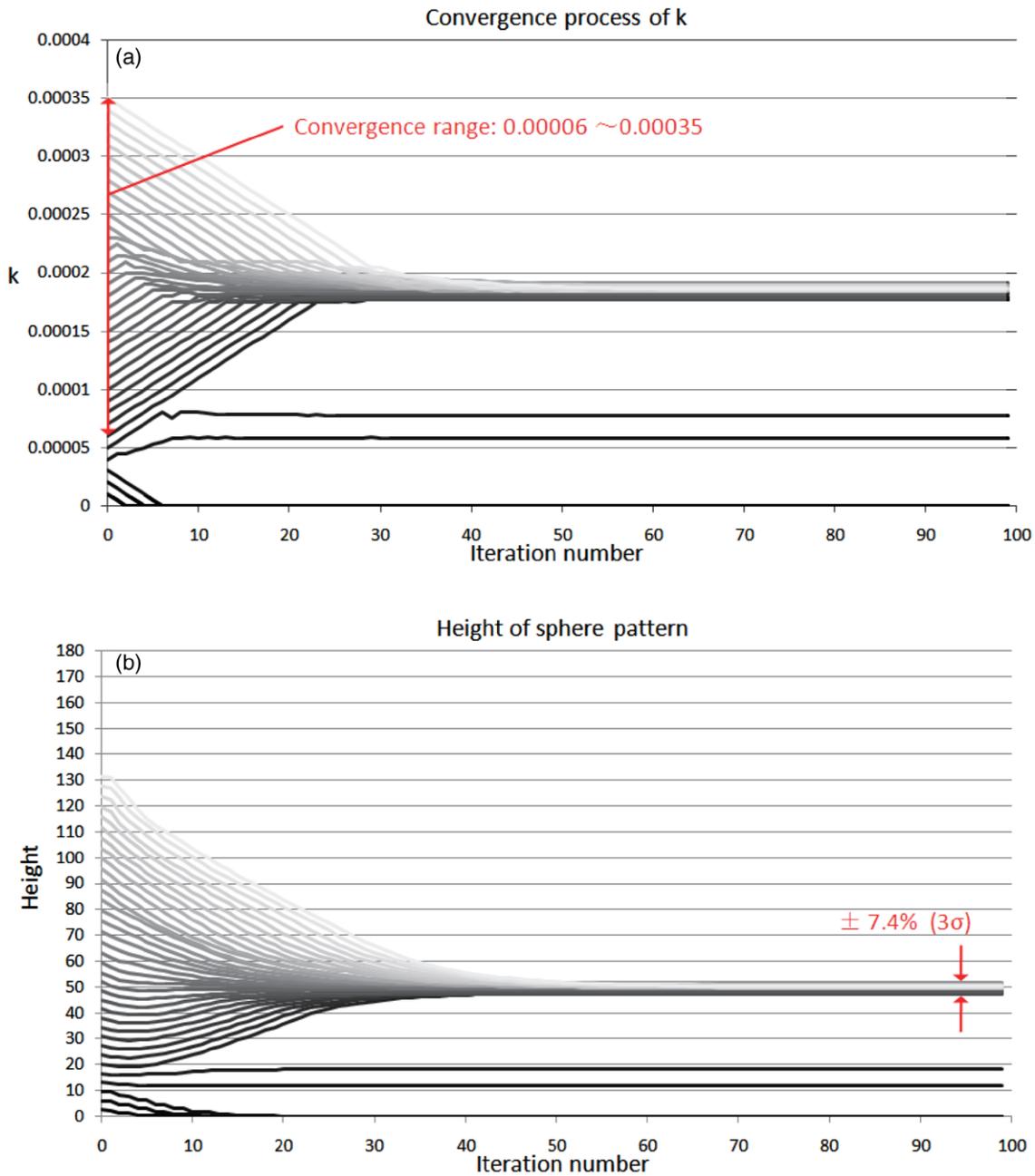


Figure 12. Convergence process of the sphere pattern with various initial values of parameter k : (a) convergence processes of k and (b) convergence processes of height.

show the reconstruction results of the sphere pattern and line pattern, respectively. As we can see, the DFG method always suffers from the so-called rounding-slope effects in the actually flat region, which is in fact caused by shadowing effects, and the non-shadowing sphere part/line part shows stretch effects if k is incorrect. The SCM method, given proper k , eliminates shadowing effects well to achieve an accurate reconstruction, while, if k is too small or too large, the actual flat parts suffer from either under-compensation effects or over-compensation effects and non-shadowing parts show stretch effects. The proposed method provides stable and accurate reconstruction results for both flat parts and sphere/line parts in all three cases, as our method is able to evaluate and correct the parameter k when de-shadowing.

As shown in figure 10(a), the original images show that the brightness in the shadowing regions is less intense than that in the flat regions without shadowing, whereas, in the shadowless images (figure 10(b)), the brightness of the flat regions around the spherical particle/line part in both the left and right compensated image demonstrates a certain uniformity, which in turn implies good compensation of the image intensities. Figure 10(c) shows the subtract images of the original and shadowless images, where the darkness indicates the strength of the difference in image intensity. The subtract images reflect the distributions of shadowing effects. As observed in figure 10(c), the regions near to the spherical particle/line part are displayed darker than the regions far away from them. This visual observation is in accordance with the actual occurrence

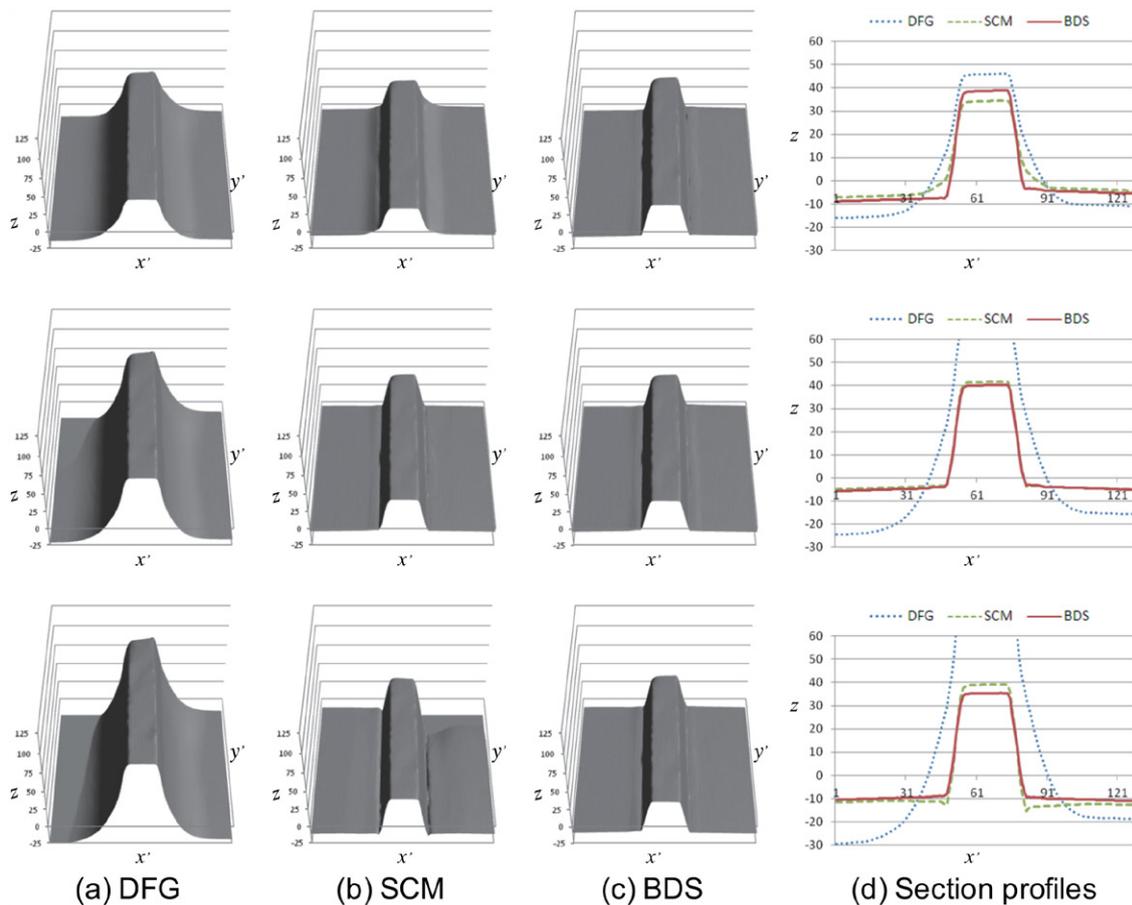


Figure 13. Comparison of 3D surface of the line pattern with initial value of k : top row $k = 0.00018$, middle row $k = 0.00025$ and bottom row $k = 0.00029$.

of the shadowing phenomena in both cases, which further implies the effective compensation of image intensities.

Convergence analysis. To analyze the convergence property of our method, we conducted extensive experiments with the initial value of k varying from 0.0001–0.00035 with step lengths of 0.00001. The convergence processes of k for both patterns are shown in figures 12(a) and 14(a), respectively. As observed, in both cases, the convergence range is considerably large. In the case of the line pattern, the convergence range is [0.0001, 0.0003]. In the case of the sphere pattern, the convergence range of k is not smaller than [0.00006, 0.000035]; initial values larger than 0.00035 are not verified and some of them may converge properly as well. Although there is a requirement for a proper initialization of the parameter k , the proposed method converges well with a very wide range of initialization values, which indicates its practicability.

Quantitative analysis. For a quantitative assessment of the metric properties of our reconstruction system, we evaluate the reconstruction errors of the height of each pattern, i.e. the length from the bottom to the top of the pattern. The sphere pattern is a standard particle and the accuracy of the shape can be guaranteed. The true height of the sphere pattern is thus the radius of the sphere that is 50 (in pixels). According to the design data, the height of the line pattern is approximately 44 (in pixels) and is assumed as the true height. Taking all the measurement results

that are appropriately converged (refer to figures 12(a) and 14(a)), we model the derivation of the measured heights from the true value as 3σ (three times the standard deviation) that reaches a maximum of only 7.4% and 4.0% for sphere and line patterns, respectively (figures 12(b) and 14(b)). Such results indicate that the proposed method can achieve absolute level reconstruction with a high accuracy, once appropriately converged.

5. Conclusion and future work

De-shadowing and automatic calibration is important for SEM photometric stereo to be practically applicable. In this study, we have presented a bootstrapping de-shadowing and self-calibration method for SEM photometric stereo to handle the challenging blind reconstruction problem from only two observed images that suffer from shadowing effects. We solve the de-shadowing problem by modeling the shadowing generation process. We have also developed a self-calibration method through making use of shadowing cues. These two problems are jointly solved by a bootstrapping approach. The proposed method can achieve absolute level reconstruction with a high accuracy and converges well with a very wide range of initialization values of parameter k . The experimental results demonstrate the significance of the proposed method.

For future work, the method may be extended to handle the more challenging scenario of spatially piecewise varying

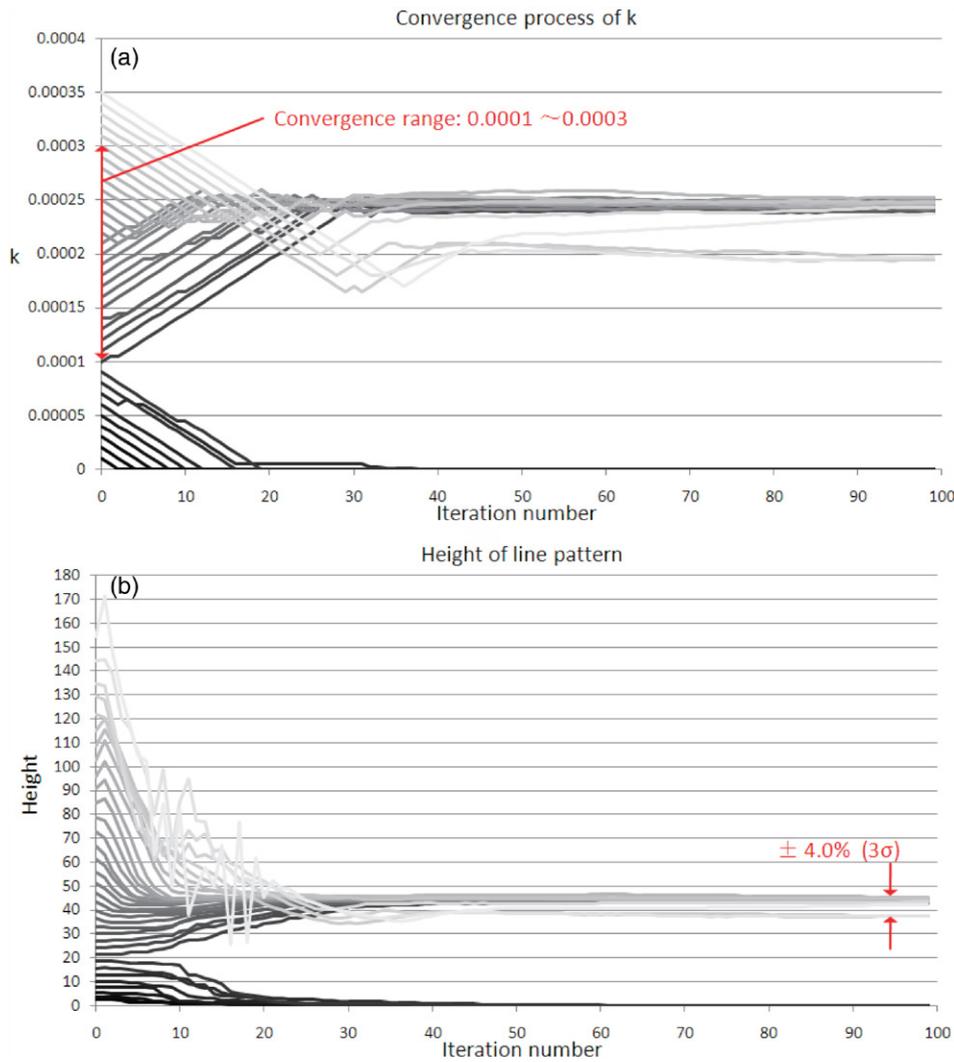


Figure 14. Convergence processes of the line pattern with various initial values of parameter k : (a) convergence processes of k and (b) convergence processes of height.

parameters for compound materials. Moreover, using dictionary learning based sparseness prior rather than TV prior [26] is also interesting and worthy of investigation.

Appendix A.

In this appendix, the numerical implementation of surface reconstruction (minimizing (6)) is presented.

In the surface reconstruction procedure, we have assumed BL^* and BR^* are known. Therefore, the optimization problem is reformulated as the following functional minimization,

$$Z^* = \arg \min_Z \int \int_D \left[\left(\frac{\partial Z}{\partial x} - p(x,y) \right)^2 + \lambda |\nabla Z| \right] dx dy, \tag{A.1}$$

where

$$p(x,y) = k \frac{BL_n^2 - BR_n^2}{(BL_n + BR_n)^2}.$$

This optimization is also known as a variational problem in the calculus of variations. A fundamental result of the calculus

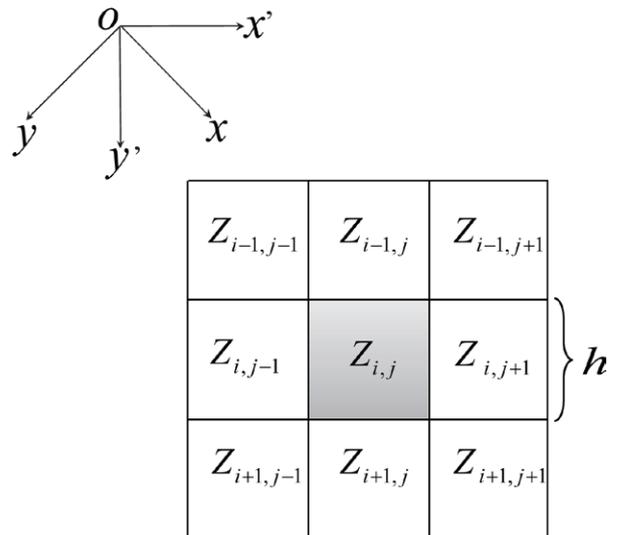


Figure A1. Image coordinates.

of variations is that the extremum of the functional must satisfy the associated Euler–Lagrange equation. Note that the

local minimum and the global minimum are both examples of the extrema. Thus the Euler–Lagrange equation is a necessary condition rather than a sufficient condition. However, the convexity of (A.1) guarantees that any local minimum must be a global minimum. Therefore, addressing (A.1) is equivalent to solving the associated Euler–Lagrange equation

$$2\left(\frac{\partial^2 Z}{\partial x^2} - \frac{\partial p}{\partial x}\right) + \lambda \operatorname{div}\left(\frac{\nabla Z}{|\nabla Z|}\right) = 0, \quad (\text{A.2})$$

where div is the divergence operator.

For numerical implementation, we adopt the finite difference method to discretize (A.2) according to the coordinate system $x - y$. Note that the $x - y$ coordinate system is rotated 45° from the commonly used $x' - y'$ coordinate system using our particular imaging setting (refer also to figure A1). We regard the $x' - y'$ coordinate system to be generally more suitable for a finite difference operator. Therefore, we transform the Euler–Lagrange equation (A.2) into the representation in the $x' - y'$ coordinate system. As the second term $\operatorname{div}(\nabla Z/|\nabla Z|)$ is rotationally invariant, its form will remain the same in the $x' - y'$ coordinate system. Therefore, substitute $\frac{\partial^2 Z}{\partial x^2}$ and $\frac{\partial p}{\partial x}$

$$\frac{\partial^2 Z}{\partial x^2} = \frac{1}{2}\left(\frac{\partial^2 Z}{\partial x'^2} + 2\frac{\partial^2 Z}{\partial x' \partial y'} + \frac{\partial^2 Z}{\partial y'^2}\right),$$

and

$$\frac{\partial p}{\partial x} = \frac{1}{\sqrt{2}}\left(\frac{\partial p}{\partial x'} + \frac{\partial p}{\partial y'}\right),$$

into (A.2), and the Euler–Lagrange equation in the $x' - y'$ coordinate system takes the following form:

$$\begin{aligned} &\left(\frac{\partial^2 Z}{\partial x'^2} + 2\frac{\partial^2 Z}{\partial x' \partial y'} + \frac{\partial^2 Z}{\partial y'^2}\right) + \lambda \operatorname{div}\left(\frac{\nabla Z}{|\nabla Z|}\right) \\ &= \sqrt{2}\left(\frac{\partial p}{\partial x'} + \frac{\partial p}{\partial y'}\right). \end{aligned} \quad (\text{A.3})$$

Moreover, we adopt the commonly used Neumann boundary conditions (also known as natural boundary conditions).

Let h denote the spatial grid size. Utilizing the central difference for the first term in the left-hand side, the right-hand side leads to

$$\begin{aligned} \frac{\partial^2 Z_{i,j}}{\partial x'^2} &= \frac{Z_{i,j+1} - 2Z_{i,j} + Z_{i,j-1}}{h^2} \\ \frac{\partial^2 Z_{i,j}}{\partial x' \partial y'} &= \frac{Z_{i+1,j+1} + Z_{i-1,j-1} - Z_{i+1,j-1} + Z_{i-1,j+1}}{4h^2} \\ \frac{\partial^2 Z_{i,j}}{\partial y'^2} &= \frac{Z_{i+1,j} - 2Z_{i,j} + Z_{i-1,j}}{h^2} \\ \frac{\partial p_{i,j}}{\partial x'} &= \frac{p_{i,j+1} - p_{i,j-1}}{2h} \\ \frac{\partial p_{i,j}}{\partial y'} &= \frac{p_{i+1,j} - p_{i-1,j}}{2h} \end{aligned}$$

The discretization of the second term in the left-hand side is a bit more elaborate. To obtain a high accuracy, we take

$$\begin{aligned} \operatorname{div}\left(\frac{\nabla Z_{i,j}}{|\nabla Z_{i,j}|}\right) &= \frac{1}{h^2} [g_{i+\frac{1}{2},j} \cdot (Z_{i+1,j} - Z_{i,j}) - g_{i-\frac{1}{2},j} \cdot (Z_{i,j} - Z_{i-1,j}) \\ &+ g_{i,j+\frac{1}{2}} \cdot (Z_{i,j+1} - Z_{i,j}) - g_{i,j-\frac{1}{2}} \cdot (Z_{i,j} - Z_{i,j-1})], \end{aligned}$$

where the function

$$g(|\nabla Z|) = \frac{1}{|\nabla Z|} \approx \frac{1}{\sqrt{\left(\frac{\partial Z}{\partial x'}\right)^2 + \left(\frac{\partial Z}{\partial y'}\right)^2 + \epsilon^2}}.$$

Here the small positive parameter ϵ is for stabilization. One can imagine that in the case of a piecewise constant result, g can be unbounded. In our experiments, we set $\epsilon = 0.0001$. In the discretization scheme, $g_{i+\frac{1}{2},j}$, $g_{i-\frac{1}{2},j}$, $g_{i,j+\frac{1}{2}}$ and $g_{i,j-\frac{1}{2}}$ are then evaluated with the following discretization scheme.

$$\begin{aligned} |\nabla Z_{i+\frac{1}{2},j}| &= \sqrt{\left(\frac{Z_{i+1,j} - Z_{i,j}}{h}\right)^2 + \left(\frac{Z_{i,j+1} - Z_{i,j-1} + Z_{i+1,j+1} - Z_{i+1,j-1}}{4h}\right)^2} \\ |\nabla Z_{i-\frac{1}{2},j}| &= \sqrt{\left(\frac{Z_{i,j} - Z_{i-1,j}}{h}\right)^2 + \left(\frac{Z_{i-1,j+1} - Z_{i-1,j-1} + Z_{i,j+1} - Z_{i,j-1}}{4h}\right)^2} \\ |\nabla Z_{i,j+\frac{1}{2}}| &= \sqrt{\left(\frac{Z_{i,j+1} - Z_{i,j}}{h}\right)^2 + \left(\frac{Z_{i+1,j} - Z_{i-1,j} + Z_{i+1,j+1} - Z_{i-1,j+1}}{4h}\right)^2} \\ |\nabla Z_{i,j-\frac{1}{2}}| &= \sqrt{\left(\frac{Z_{i,j} - Z_{i,j-1}}{h}\right)^2 + \left(\frac{Z_{i+1,j-1} - Z_{i-1,j-1} + Z_{i+1,j} - Z_{i-1,j}}{4h}\right)^2} \end{aligned}$$

The finite difference operators result in a large nonlinear system. The nonlinearity is due to g . Let \mathbf{Z} denote the vector obtained from the discretization of $Z(x, y)$. Treat $g(g_{i+\frac{1}{2},j}$, $g_{i-\frac{1}{2},j}$, $g_{i,j+\frac{1}{2}}$ and $g_{i,j-\frac{1}{2}}$) as coefficients, so that the nonlinear system takes the following quasi-linear matrix form:

$$A(\mathbf{Z})\mathbf{Z} = \mathbf{b}, \quad (\text{A.4})$$

where $A(\mathbf{Z})$ is large but sparse.

We adopt a so-called fixed point method to solve the quasi-linear system (A.4). The key idea is to address the nonlinear problem by solving a sequence of linear systems with the nonlinear part arising in $A(\mathbf{Z})$ being fixed. The method is summarized as follows:

- Initialize \mathbf{Z} with a guess of $\mathbf{Z}^{(0)}$.
- At the m^{th} iteration, fix $A(\mathbf{Z})$ with $A(\mathbf{Z}^{(m-1)})$.
- Obtain $\mathbf{Z}^{(m)}$ by solving the resulting linear system

$$A(\mathbf{Z}^{(m-1)})\mathbf{Z}^{(m)} = \mathbf{b},$$

that can be solved effectively through methods such as the Jacobi method, the Gauss–Seidel method, successive over-relaxation, etc.

(d) Iterate between (b) and (c) until convergence.

Note that the linear system need not be solved exactly in each iteration so as to yield fast convergence [27].

References

- [1] Oshima T, Kimoto S and Suganuma T 1970 *Photogrammetric Eng.* **36** 874–9
- [2] Koenig G, Nickel W, Storl J, Meyer D and Stange J 1987 *Scanning* **9** 185–93
- [3] Kholodilov O V, Grigoryev A Y and Myshkin N K 1987 *Scanning* **9** 156–61
- [4] Marschner T and Stief C 2003 *Proc. SPIE* **5038** 892–900
- [5] Morokuma H, Miyamoto A, Tanaka M, Kazui M and Takane A 2004 *Proc. SPIE* **5375** 727–34
- [6] Setoguchi K, Morokuma H, Miyamoto A and Tanaka M 2005 *Proc. SPIE* **5752** 1353–61
- [7] Lebedzik J 1979 *Scanning* **2** 230–7
- [8] Carlsen I C 1985 *Scanning* **7** 169–77
- [9] Suganuma T 1985 *J. Electron Microsc.* **34** 328–37
- [10] Reimer L, Böngerler R and Desai V 1987 *Scanning Microsc.* **1** 963–73
- [11] Czepkowski T and Słowko W 1996 *Scanning* **18** 433–46
- [12] Paluszyński J and Słowko W 2005 *Vacuum* **78** 533–7
- [13] Drzazga W, Paluszyński J and Słowko W 2006 *Meas. Sci. Technol.* **17** 28–31
- [14] Paluszyński J and Słowko W 2006 *Scanning* **224** 93–6
- [15] Wöhler C and Hafezi K 2005 *Pattern Recognition* **38** 965–83
- [16] Woodham R J 1980 *Opt. Eng.* **19** 139–44
- [17] Hernández C, Vogiatzis G and Cipolla R 2008 Shadows in three-source photometric stereo *Eur. Conf. on Computer Vision (Lecture Notes in Computer Science vol 5303)* (Berlin: Springer) pp 290–303
- [18] Barsky S and Petrou M 2003 *IEEE Trans. Pattern Anal. Mach. Intell.* **25** 1239–52
- [19] Ikeuchi K and Horn B K P 1981 *Artif. Intell.* **17** 141–84
- [20] Horn B K P and Brooks M J 1986 *Comput. Vision Graph. Image Process.* **33** 174–208
- [21] Horn B K P 1990 *Int. J. Comput. Vision* **5** 37–75
- [22] Agrawal A, Raskar R and Chellappa R 2006 What is the range of surface reconstructions from a gradient field *Eur. Conf. on Computer Vision (Lecture Notes in Computer Science vol 3951)* (Berlin: Springer) pp 578–91
- [23] Sakaue K, Amano A and Yokoya N 1999 *IEICE Trans. Inform. Syst.* **82** 534–47
- [24] Reimer L 2008 *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis* (Berlin: Springer)
- [25] Fornasier M 2010 *Theoretical Foundations and Numerical Methods for Sparse Recovery* (Berlin: De Gruyter)
- [26] Wright J et al 2010 Sparse representation for computer vision and pattern recognition *Computer Vision and Pattern Recognition* (Champaign, IL: University of Illinois Press)
- [27] Vogel C R and Oman M E 1996 *SIAM J. Sci. Comput.* **17** 227–38