

Surface reconstruction from microscopic images in optical lithography

Virginia Estellers*, Jean-Philippe Thiran*, Maria Gabrani^o

Abstract

We propose a shape-from-shading method to reconstruct surfaces of silicon wafers from images of printed circuits taken with scanning electron microscope. Our method combines the physical model of the optical acquisition system with prior knowledge about the shapes of the patterns in the circuit. The reconstruction of the surface is formulated as an optimization problem with a combined criterion based on the irradiance equation and a shape prior that constrains the shape of the surface to agree with the expected shape of the pattern. To account for the variability of the manufacturing process, the model allows a non-linear elastic deformation between the expected patterns and the reconstructed surface. Our method provides two outputs: a reconstructed surface and a deformation field. The reconstructed surface is derived from the shading observed in the images and the prior knowledge about circuit patterns, which results in a shape-from-shading technique stable and robust to noise. The deformation field produces a mapping between the expected shape and the reconstructed surface, which provides a measure of deviation between the models and the real manufacturing process.

I. INTRODUCTION

As circuit designs become more complex due to technology scaling, engineers in the semiconductor industry require new tools to obtain accurate and useful information about the manufacturing process. In particular, the characterization of surface topography is essential in optical lithography in order to optimize the printing process from layout masks to silicon wafers. Scanning electron microscopy (SEM) is the primary tool currently used in industry to measure geometrical features of semiconductor structures, but it provides only limited information about the third dimension of the surfaces. In fact, several parameters associated to the surface of the wafer are crucial to control the lithographic process (the height of the

V. Estellers and J.P Thiran are with the Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne.
M. Gabrani is with IBM Research, Zurich Laboratory.

step, the slope of the side walls or the curvature of edges) [1], and it has now become necessary for the industry to develop image processing techniques to reconstruct accurate 3D surfaces from 2D SEM images. In this paper, we describe such a technique. It is formulated as a shape-from-shading (SFS) reconstruction that considers prior knowledge about the shape of the printed patterns in order to improve the accuracy and robustness of the method.

In the past, computer vision [2]–[7] has focused on stereoscopy and SFS for the reconstruction of surfaces from SEM images of printed circuits, but the resulting techniques have not been adopted by the industry. The acquisition of pairs of stereoscopic SEM images at the nanometre scale is not only challenging for large wafers, but extremely expensive, and the throughput requirements of manufacturers call for simpler and more economic methods. In particular, SEM images are usually analysed with simple image processing techniques based on edge detection and template matching to define contours and estimates of 1D profiles of the surface [8]–[13]. These techniques coarsely detect edges in the SEM images with simple gradient and thresholding methods, locate a preliminary contour and subsequently refine it to match a physically-meaningful criterion in terms of the underlying surface. To that purpose, they exploit available information about the patterns of the circuit and analyse 1D profiles of the image along the normals to the contours, which are then modified to match a fixed height level of the underlying surface. The resulting profile measurements along the contours are relatively fast to extract, but no real 3D surface is reconstructed. We propose to “fill in the gap” between these two approaches and include the knowledge of the patterns expected in the circuit into a surface reconstruction method based on shape-from-shading. Compared to the stereoscopic method, we do not require SEM image pairs obtained with a dedicated system to reconstruct accurate surfaces. Instead, we exploit the available information about the shapes of the patterns in the circuit to overcome the ill-posed nature of the SFS problem and obtain a robust reconstruction method.

For each pattern in the printed circuit, the nature and accuracy of the available information can be considerably different. While for some technology nodes and designs, reasonable optical proximity correction (OPC) models of the printing process exists, in the first cycles of development the available models are under calibration and only the mask associated to the circuits can be trusted [1]. The reconstruction method that we propose considers both situations and defines different shape prior terms depending on the nature of the available information. Simultaneously, our technique estimates a deformation field between the expected pattern and the reconstructed surface, which can be used a posteriori to quantify unaccounted errors in the OPC modelling or lithographic process.

The contributions of the paper are the following. First, we include a shape prior in the surface

reconstruction method associated to SFS and obtain a method robust to noise and SEM artefacts. Secondly, to account for variability in the lithographic printing process, we estimate a deformation field between the expected shape of the pattern and the reconstructed surface. Third, we use a ridge detector to weight the smoothness penalty commonly used in variational formulations of SFS and improve the reconstruction. Finally, we adopt minimization methods recently introduced to imaging to solve the resulting problem and obtain a fast, efficient and easy-to-code algorithm.

The rest of the paper is organized as follows. Section II presents the proposed technique, reviewing and putting it in context with existing SFS and shape-prior models in imaging. The reconstruction of the surface is then formulated as a minimization problem and solved with efficient optimization techniques in Section III. Section IV presents the experimental results and conclusions are drawn in Section V.

II. SHAPE FROM SHADING WITH SHAPE PRIOR

We formulate the reconstruction of surfaces as an optimization problem with a combined matching criterion based on the irradiance equation and a shape prior that constrains the reconstructed surface to agree with the expected shape of the pattern. To account for the variability of the physical printing process, the model allows a non-linear elastic deformation of the expected patterns. We propose two different objective energies to measure for this deformation, one based on an OPC model of the expected surface and another on the limited information extracted from the mask of the circuit. Our optimization problem has thus two variables or outputs of our reconstruction method: a reconstructed surface and a deformation field. The surface that we recover is derived from the shading observed in the SEM images and the prior knowledge about circuit patterns, while the deformation field produces a mapping between the expected shapes in the circuit and the reconstructed surface.

On the following, we present the different components of our model and review related approaches.

A. *Shape from shading*

In computer vision, shape recovery refers to the reconstruction of a 3D scene from 2D images and covers different techniques, called shape-from-x, depending on the nature of the input information. Shape-from-shading in particular deals with the recovery of shape from a gradual variation of shading in the image.

To understand the problem of SFS it is necessary to study how the images are formed. In our formulation, we attach to the camera a 3D coordinate system (x, y, z) , with the plane (x, y) denoting the image plane and the axis z on the optical axis. If we assume an orthographic projection, the visible part

of the scene defines a height map $z(\mathbf{x})$, where $\mathbf{x} = (x, y)$ ¹ designates a point in the image. The SFS problem is then formulated in terms of the “image irradiance equation”:

$$I_0(\mathbf{x}) = R(\mathbf{n}(\mathbf{x})). \quad (1)$$

Equation (1) relates the grey level $I_0(\mathbf{x})$ measured in the image at point \mathbf{x} with the reflectance function $R(\mathbf{n}(\mathbf{x}))$ of the surface at that point. In physical terms, $R(\mathbf{n})$ gives the value of the light re-emitted by the surface as a function of its orientation, i.e. of the vector \mathbf{n} normal to the surface. This normal can easily be expressed in terms of the gradient of the height map as

$$\mathbf{n} = \frac{(-\nabla_x z, -\nabla_y z, 1)}{\sqrt{|\nabla z|^2 + 1}}. \quad (2)$$

The SFS problem proposed by Horn in [14] consists on the recovery of a height map z , the unknown of the problem, from the shading observed in the image I_0 , the data available from measurements.

Since the introduction of SFS in the 70s [14], most efforts have been directed into the recovery of shapes from images taken under normal lighting conditions and only few investigations [2]–[7], [15], [16] have considered the problem associated to SEM images. In the SEM acquisition system, images are formed scanning the sample with an electron beam and detecting the electrons emitted at the point where the beam strikes the surface. The number of electrons emitted by the surface is related to the angle between the electron beam and the surface normal, resulting in shades in the captured image I_0 proportional to the number of electrons. At high magnifications, the imaging geometry can be modelled as an orthographic projection with a unique light source at infinity. The resulting reflectance function for SEM images is then given by

$$R(\mathbf{n}) = \frac{\nu}{\mathbf{n} \cdot \mathbf{L}}, \quad (3)$$

where ν is the reflectance coefficient of the surface and \mathbf{L} is a unit vector indicating the direction of the light source. In the case of our SEM images of silicon wafers, ν is assumed uniform within an image, $\mathbf{L} \approx (0, 0, 1)$ and both parameters can be directly estimated from the SEM images as explained in Section III-A. Such a reflectance map neglects the effects of mutual illumination, but provides a computationally tractable model of the SEM acquisition system.

In the case of silicon wafers of printed circuits, the features derived from SEM images are generally too sparse for standard SFS techniques. For this reason, in research surface reconstruction from SEM

¹We use boldface for vector fields.

measurements has been formulated as the correspondence problem associated to stereo pairs of images obtained by tilting the microscope stage or deflecting the electron beam [2]–[7]. The combination of stereoscopy and SFS is able to reconstruct accurate surfaces because the two techniques complement each other: SFS works best where the surface is smooth and featureless, while stereo works well where the surface is rough and features are dense.

In the industry, however, the main approach to 3D surface reconstruction from SEM images of circuit patterns is based on 1D measurements and Monte Carlo simulations [17]–[19]. The Monte Carlo method models statistically the interaction of incident electrons with the wafer by repeated detailed simulations, which results in extremely long running times. Consequently, Monte Carlo simulations are only used to obtain a library of SEM signals for simple shapes defined by a few parameters. A measured SEM signal is then compared to the library and a 1D surface profile is reconstructed estimating the parameters that best match the observed measurements from the signals in the library. In this case it is important to obtain a robust estimate of the SEM profile, and several 1D measurements are usually averaged, that is, smoothed to avoid the effects of noise. This technique has proved considerably accurate for simple structures available in the libraries, but lacking for the complicated 2D patterns currently used in circuit designs. Compared to the SFS approach adopted in computer vision, Monte Carlo simulations provide an accurate model of the reflectance map but the reconstruction is limited to 1D profiles and does not consider the 3D nature of the surface.

Our approach combines SFS techniques from computer vision with the prior knowledge available in industry about the expected patterns and OPC models. In the proposed approach, Equation (3) is only a rough approximation to the SEM acquisition system and we expect deviations even in clean images in comparison to the Monte Carlo method. To overcome the limitations of Equation (3) our formulation considers the 2D nature of the image measurements, the expected smoothness of the surface and the prior knowledge about its shape. Similar to the combination of SFS with stereoscopy, our approach integrates the complementary techniques of SFS and image registration to merge the best of both worlds: while SFS works well where the image is featureless and the surface is smooth, the registration of the shape prior is guided by edges in the image, i.e., abrupt changes in the surface topology. Compared to stereoscopy, we do not require special acquisition systems to obtain simultaneous SEM image pairs and only make use of the information available from circuit designs about the printed patterns. It is worth noting, however, that our approach could also be modified to include stereoscopic pairs.

We formulate our SFS reconstruction algorithm in a variational framework, i.e., as a minimization problem, for three reasons. First, variational SFS techniques are generally more robust than other available

approaches [20], [21], and therefore better suited to its application with industrial SEM data. Secondly, the objective minimization energy can easily include a term describing the prior information provided by the mask or the OPC model in the shape reconstruction algorithm. Thirdly, the resulting minimization problem can be solved with efficient optimization techniques that are guaranteed to converge and can benefit from a multi-resolution approach to speed up the reconstruction. We refer the reader to the review papers [20]–[22] and references therein, for a review of the different available SFS approaches.

Variational SFS methods are defined by the choice of three elements: the unknown z used to describe the surface, the objective functional \mathcal{E} that is minimized, and the minimization method. Functional \mathcal{E} is defined by two terms or constraints: the brightness and the smoothness constraint. The brightness term \mathcal{A} ensures that the reconstructed shape produces a similar irradiance as the input image, while the smoothness term \mathcal{G} ensures the reconstruction of a smooth surface and is introduced to overcome the ill-posed nature of the problem. As both constraints can be directly written in terms of the gradient ∇z or normal \mathbf{n} of the surface, the minimization can be solved in terms of ∇z or \mathbf{n} and the height map subsequently recovered by integration [23]–[26]. In this case, the integrability constraint $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$ is included in the formulation [27] to reconstruct physically meaningful surfaces and improve accuracy. The integrability of the surfaces can also be ensured by directly formulating the brightness constraint in terms of the height map z , which results in variational methods of higher order and, consequently, in slower or more complex minimization algorithms [28]–[30]. We adopt such a formulation, but use recent advances in optimization techniques to obtain a fast algorithm.

In our formulation, the brightness term reads

$$\mathcal{A}(I) = \frac{1}{2} \int_{\Omega} (R(\mathbf{d}) - I_0)^2 \text{ with } \mathbf{d} = (d_x, d_y) = \nabla z, \quad (4)$$

which assumes Gaussian noise in the SEM image acquisition process, but other norms than L_2 could be used for more robust methods.

We adapt the common smoothness term $\frac{1}{2} \int_{\Omega} (\Delta z + 2\nabla_x \nabla_y z)^2$ based on image segmentation techniques. In particular, our approach is inspired by the geodesic active contour model [31], and weights the common smoothness term with an inverse ridge detector to only penalize locations where high slopes are unlikely. The resulting penalty is

$$\mathcal{G}(z) = \frac{1}{2} \int_{\Omega} w (\Delta z + 2\nabla_x \nabla_y z)^2 = \frac{1}{2} \int_{\Omega} w |\nabla \mathbf{d}|^2 \quad \text{with } \mathbf{d} = \nabla z, \quad (5)$$

where the weighting function w is designed to satisfy $w \approx 0$ in positions where changes in the shape

of the surface are likely and $w \approx 1$ elsewhere. Weighting the smoothness term is equivalent to the constraint proposed by Zheng and Chellappa in [32], which specifies that the intensity gradients of the reconstructed image and the input image should be close to each other. In the same way that the geodesic active contour model [31] integrates the original snake model of Kass, Witkin and Terzopoulos [33], our weighted smoothness term forces the intensity gradients of the reconstructed image to be close to the input ones. In our experiments we define w as the inverse ridge detector proposed in [34], [35] applied to the SEM image I_0 , but other weight functions are equally possible.

B. Deformation field and similarity to the layout

Humans, and in particular engineers, interpret SEM images of circuits by incorporating high-level knowledge about the expected patterns. Such a knowledge is necessary to disambiguate low-level intensities and artefacts in noisy images and make any automatic evaluation method robust. Computer vision has successfully imitated the integration of prior knowledge done by humans into the image segmentation problem, but this information has not yet been exploited in the context of SFS. Inspired by image segmentation techniques, we incorporate into our surface reconstruction process the prior knowledge about the shapes of the patterns in the circuit.

Shape priors are included in image segmentation techniques, see e.g. [36]–[42] and references therein, to ensure robustness to occlusions and noise. The standard approach in variational methods is to incorporate an additional term in the objective energy that defines the segmentation criteria. This additional term measures the similarity between the evolving shape and a reference one (called shape prior) and penalizes deviations with respect to the expected shape. The resulting method must deal with two important issues: alignment of the shape prior within the image and variability of the shape with respect to the prior. For the question of shape alignment, affine transformations (rotation, translation and scaling) can be taken into account in the model. This is done explicitly in the definition of the energy, but at the price of increasing the complexity of the optimization; or by intrinsic alignment of more complex shape descriptors [37], [42], [43]. The second issue is the question of shape variability from the reference template, which is usually handled through statistical [36], [37], [44], [45] or trained [36], [38], [39], [43], [44], [46] shape models.

Shape alignment techniques could be adapted and included in our SFS model, increasing the complexity of the minimization. For SEM images of silicon wafers, however, it is not necessary because this issue is handled during image acquisition and we can assume correct scaling and rotation for the sake of simplicity. The question of shape variability is more difficult to handle in our case because obtaining an

accurate model of surfaces is extremely expensive at the nanometre scale. To model the shape variations in terms of SFS we require the 3D surfaces, not the available SEM images, and only few options are available. Atomic force microscopy can be used to measure the surfaces for the different patterns, but it is an extremely expensive process specially if several images of each pattern should be acquired. OPC modelling can be used to obtain surface models by simulations, but are partly based on heuristics and only predict failures in the printing process due to certain causes. Due to these limitations, we include a dense deformation field in our surface recovery model instead of a model of shape variability. In our approach we simultaneously recover the surface z and estimate a deformation field \mathbf{u} between the expected shape and the recovered one. This strategy does not require a shape model and offers two additional advantages. First we can incorporate information of the expected shapes in terms of 3D models or 2D masks of the patterns, depending on the accuracy of the available information. Second, the approach directly provides a measure of erosion in the printing process compared to the expected shape (larger deformation fields are associated to larger printing errors) and can be directly used in metrology or OPC modelling to detect failing regions, wrong measurements or models. The problem of shape variability is thus formulated as the estimation of a deformation field between two surfaces or, equivalently, the registration of their height maps. The registration technique that we adopt is based on the non-linear elastic model of [47].

In agreement with our variational formulation of SFS, we consider the registration problem as the minimization of a functional whose unknown is the displacement field \mathbf{u} and refer the reader to [48] for an overview of other registration methods. Denoting by z_0 the template or shape prior and z the reconstructed surface described in terms of the height map, image registration is performed minimizing a functional that combines a measure of similarity between z and z_0 and a penalty on the displacement field \mathbf{u} .

If we can use OPC simulations to provide a prior model of the surface and its height map z_0 , we can directly measure the distance between the surface and the expected OPC shape with the L_2 -norm. The term associated to the shape prior is then

$$\mathcal{B}(z, \mathbf{u}) = \int_{\Omega} (z(\mathbf{x}) - z_0(\mathbf{x} + \mathbf{u}))^2. \quad (6)$$

If a reliable OPC model is not available, we measure the similarity of the reconstructed surface directly with the binary mask of the pattern provided in the circuit design. In that case the minimization of the energy will be slower and the reconstruction less robust to the optical and physical proximity effects modelled in OPC. To compare our surface to the mask, we need first to binarize it in order to have a

meaningful measure. For simplicity, we threshold the surface at a fixed threshold θ and compare it to the mask with the L_2 -norm. The thresholding is written in terms of the Heaviside function $H(z - \theta)$ and the mask described in terms of the signed distance function ϕ to its contour, i.e. $H(\phi)$ corresponds to the binary mask. The similarity to the layout is therefore measured with the following shape prior term

$$\mathcal{B}(z, \mathbf{u}) = \int_{\Omega} (H(z(\mathbf{x}) - \theta) - H(\phi(\mathbf{x} + \mathbf{u})))^2. \quad (7)$$

We choose to describe the mask by means of the distance function to its contour in order to improve the numerical convergence associated with the registration of binary texture-less images.

To regularize the displacement field we can either rely on physical models, such as elastic and viscous fluid models, or rely on image regularization [49]–[51]. In general, physical models are more interesting than image regularizers because they model real materials and follow our intuition. In our case, we consider that the deformation of the mask or OPC model follows a physical model and borrow energies from material sciences to penalize deformation fields \mathbf{u} that are non-elastic.

The concept of elasticity is incorporated in image registration by considering the template and reference images as observations of the same elastic body before and after being subjected to a deformation. Deformations of elastic bodies are then measured in terms of the Green-St. Venant strain tensor

$$\varepsilon = \frac{\nabla \mathbf{u} + \nabla \mathbf{u}^T + \nabla \mathbf{u}^T \nabla \mathbf{u}}{2}. \quad (8)$$

In Equation (8), $\varphi = Id + \mathbf{u}$ is the deformation matrix, and the strain tensor measures thus the deviation between the given deformation and a rigid one. As penalty for our deformation field \mathbf{u} , elastic models use the St. Venant-Kirchoff energy stored by the deformed material, which is given by

$$\mathcal{W}(\varepsilon(u)) = \int_{\Omega} \frac{\lambda}{2} (\text{tr } \varepsilon(u))^2 + \mu \text{tr } \varepsilon(u)^2, \quad (9)$$

where λ and μ are the first and second Lamé parameters characteristic to the material.

The first linear elastic model, proposed by Broit in [52], neglects the quadratic term in the strain tensor and the resulting model is fast to optimize but does not allow large deformations. To overcome this limitation, Christensen et al. [53] proposed a viscous fluid model where objects evolve as fluids in accordance to Navier-Stokes equations. Compared to the elastic models, fluid equations perform regularization on the velocity field and allow larger deformations, but they are computationally more expensive.

It has been observed recently in [47] that we can obtain an elastic model comparable to the viscous

model of [53] by simply considering the quadratic term in the strain tensor and defining Equation (9) as smoothness penalty. The non linearity in this elastic model allows small strains even when the stress is relatively large and, consequently, large deformation fields between the shape obtained from the OPC model and the recovered surface or their respective masks.

III. PROPOSED MINIMIZATION MODEL

The variational formulation of our surface reconstruction method results in the following non-convex minimization problem

$$\min_{T, \mathbf{u}} \alpha \mathcal{A}(\nabla z) + \beta \mathcal{B}(z, \mathbf{u}) + \gamma \mathcal{G}(\nabla z) + \mathcal{W}(\nabla \mathbf{u}). \quad (10)$$

To solve this optimization problem, we make use of recent advances in minimization techniques [54], [55] and apply variable splitting and augmented Lagrangians [56] to obtain an efficient and easy-to-code algorithm.

The minimization problem that we need to solve has terms involving the surface height map z , its spatial gradients $\mathbf{d} = \nabla z$, the deformation field \mathbf{u} and its spatial derivatives $V = \nabla \mathbf{u}^2$. To efficiently solve the problem, we design an iterative algorithm that considers these variables as independent and use Lagrangian multipliers to ensure that the relationships $\mathbf{d} = \nabla z$ and $V = \nabla \mathbf{u}$ hold. In image processing, similar techniques have been recently applied to variational image denoising, segmentation and reconstruction that were previously solved with PDE methods [54], [55], [57].

Minimization techniques based on PDEs derive the Euler-Lagrange equations associated to the objective functional and solve them with iterative time marching algorithms. The resulting PDE evolution is inefficient because its time step is restricted by the Courant-Friedrichs-Lewy stability conditions. These conditions assure convergence and accuracy in the solution, but make explicit PDE evolution extremely slow for models with higher order derivatives, like the one proposed here. In the last years, however, PDE methods have been combined and replaced by more efficient optimization techniques to provide faster and more stable algorithms in imaging. These optimization techniques are usually based on the dual formulation of convex problems [58]–[60] or on variable-splitting and equality constrained optimization; which can be solved by quadratic-penalties [54], Bregman iterations [55], [61] or the equivalent augmented Lagrangian method [62]. The algorithm that we propose falls in this last category. We rewrite the problem

² \mathbf{u} is a 2-dimensional vector field in \mathbb{R}^2 and we denote as $V = \nabla \mathbf{u}$ the 2×2 matrix field obtained of applying the gradient operator to each component of \mathbf{u} .

as a constraint minimization and use augmented Lagrangians to solve it. The resulting Lagrangian is minimized with respect to each variable independently and the multipliers are then updated in a cyclic way. The independent minimizations for each variable are easier to solve because they reduce to convex problems involving only first order derivatives or are decoupled for each pixel. The initial complex minimization of Equation (10) is thus divided in small and easier subtasks that are efficiently solved, and the resulting algorithm is faster, easier to control and more stable.

We also adopt a multi-resolution approach to speed up the surface reconstruction method. Multi-resolution or scale space tracking is commonly used in computer vision to solve large non-convex minimization problems efficiently. An approximate solution is found at a coarse scale and then tracked through scale as the problem is solved at higher resolutions. At a large scale, the problem will not suffer from local minima and a first coarse solution is easily found. This solution is used to initialize the algorithm at smaller scales. As the scale is reduced, local minima appear in the minimization and tracking of the initial solution guarantees that the solution is kept meaningful. In fact, scale space is not certified to find the global minimum of a non-convex minimization problem, only a significant one that appears at large scales. In terms of efficiency, the algorithm is designed to perform most of its iterations at large scale. It then initializes the same minimization algorithm at smaller scales with this coarse solution and requires only a few iterations, which are now more expensive due to the finer resolution of the grid, to converge. Multi-resolution makes our method more fast and robust to local minima.

A. Estimation of the reflectance map

In order to improve the approximation of the reflectance function (3), we alternatively iterate between the refinement of the reflectance map (in terms of the estimated reflectance coefficient ν and light source direction \mathbf{L}) and the refinement of the reconstructed surface.

Given a first estimate of the surface and its normal \mathbf{n} , we can refine the parameters of the reflectance map solving the following least squares minimization problems.

$$\nu = \arg \min_{\nu} \int_{\Omega} (\nu - I_0(\mathbf{x})\mathbf{n}(\mathbf{x}) \cdot \mathbf{L})^2. \quad (11)$$

In Equation (11), \mathbf{L} is assumed known and takes the value of the previous estimate. We can then analytically compute the solution to this least-square problem, which reads $\nu = \int_{\Omega} I_0(\mathbf{x})\mathbf{n}(\mathbf{x}) \cdot \mathbf{L}$. Similarly, for the estimation of \mathbf{L} , we assume ν known, giving it the value of the previous estimate, and

solve the following minimization problem

$$\mathbf{L} = \arg \min_{|\mathbf{L}|=1} \int_{\Omega} \left(\frac{\nu}{I_0(\mathbf{x})} - \mathbf{n}(\mathbf{x}) \cdot \mathbf{L} \right)^2. \quad (12)$$

In this case, the constraint $|\mathbf{L}| = 1$ is not handled explicitly in the minimization by a closed-form solution and we adopt the same augmented Lagrangian technique used for the other minimizations.

With the resulting estimate of the reflectance map, the estimate of the surface can be improved with the proposed SFS algorithm. In our algorithm, the process is iterated by alternating the refinements of surface and reflectance maps.

B. Notation

We discretize the image domain $\Omega \subset \mathbb{R}^2$ with a regular grid of size $n = n_x \times n_y$. In Ω we consider images and height maps as scalar functions with $z(i) \in \mathbb{R}$ and their gradients as vector-valued functions with $\nabla z(i) \in \mathbb{R}^2$. For vector fields $\mathbf{d} = (d_x, d_y)$, we define the gradient operator on each of its components, that is $\nabla \mathbf{d} = (\nabla d_x, \nabla d_y)$.

We use forward differences to compute the discrete gradients and backward differences for the divergence in order to preserve the adjoint relationship $\text{div} = -\nabla^*$ in the discrete setting. In vector notation, we can efficiently compute the spatial derivatives multiplying the discrete functions arranged as a column vector with the sparse finite difference matrices $\nabla_x z = D_x z$, $\nabla_y z = D_y z$.

For numerical purposes we also relax the Heaviside function to a smooth approximation $H_\epsilon(x) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{x}{\epsilon} \right) \right)$, with its corresponding derivative δ_ϵ approximating the Dirac distribution.

C. Numerical minimization

To reconstruct the height map z associated with the surface and estimate the deformation field u , we need to solve the minimization problem of Equation (10). We propose an iterative algorithm to solve this problem based on splitting and constraint minimization techniques. The main idea is to split the original problem into sub-optimization problems which are easy to solve, and combine them together using an augmented Lagrangian.

Let us consider the following constrained minimization problem, which is equivalent to (10):

$$\min_{\substack{\mathbf{d}=\nabla z \\ \mathbf{V}=\nabla \mathbf{u}}} \alpha \mathcal{A}(\mathbf{d}) + \beta \mathcal{B}(z, \mathbf{u}) + \gamma \mathcal{G}(\mathbf{d}) + \mathcal{W}(\mathbf{V}). \quad (13)$$

Next, we reformulate this constrained minimization problem as an unconstrained optimization task. This can be done with an augmented Lagrangian approach, which translates the constraints into pairs of

Lagrangian multiplier and penalty terms. The augmented Lagrangian energy associated to (13) reads

$$\begin{aligned} \mathcal{L}(z, \mathbf{d}, \mathbf{u}, \mathbf{V}, \boldsymbol{\lambda}_z, \boldsymbol{\lambda}_u) &= \mathcal{E}_c(z, \mathbf{d}, \mathbf{u}, \mathbf{V}) + \int_{\Omega} \boldsymbol{\lambda}_z \cdot (\mathbf{d} - \nabla z) + \\ &\frac{r_z}{2} |\mathbf{d} - \nabla z|_2^2 + \boldsymbol{\lambda}_u \cdot (\mathbf{V} - \nabla \mathbf{u}) + \frac{r_u}{2} |\mathbf{V} - \nabla \mathbf{u}|_2^2, \end{aligned} \quad (14)$$

where the vectorial functions $\boldsymbol{\lambda}_z, \boldsymbol{\lambda}_u$ are Lagrange multipliers and r_z, r_u are positive constants.

The constraint minimization problem (13) reduces to finding the saddle-point of the augmented Lagrangian energy \mathcal{L} . The solution to the saddle point problem (14) can be approximated iteratively by the following algorithm: initialize the variables and Lagrange multipliers to zero; at iteration k find an approximate minimizer of $\mathcal{L}(z, \mathbf{d}, \mathbf{u}, \mathbf{V}, \boldsymbol{\lambda}_z^{k-1}, \boldsymbol{\lambda}_u^{k-1})$ with respect to the variables $z, \mathbf{d}, \mathbf{u}, \mathbf{V}$ and update the Lagrange multipliers with the residuals associated to each constraint.

In general, it is difficult to find the exact minimizer of the Lagrangian \mathcal{L} with respect to all the variables $z, \mathbf{d}, \mathbf{u}, \mathbf{V}$. However, experiments show that a good approximation can be found by alternating the minimization of \mathcal{L} with respect to each variable while considering the others fixed. The resulting method is equivalent to the alternating direction method of multipliers [63]. The minimization method is summarized in Algorithm 1.

Algorithm 1 Augmented Lagrangian method to solve the constraint minimization (13)

- 1: Initialize $z, \mathbf{d}, \mathbf{u}, \mathbf{V}, \boldsymbol{\lambda}_z, \boldsymbol{\lambda}_u$
- 2: For each iteration $l = 1, 2, \dots$, find an approximate minimizer of \mathcal{L} with respect to variables $(z, \mathbf{d}, \mathbf{u}, \mathbf{V})$ with fixed Lagrange multipliers $\boldsymbol{\lambda}_z^l, \boldsymbol{\lambda}_u^l$:

$$z^l = \arg \min_z \mathcal{L}(z, \mathbf{d}^{l-1}, \mathbf{u}^{l-1}, \mathbf{V}^{l-1}, \boldsymbol{\lambda}_z^{l-1}, \boldsymbol{\lambda}_u^{l-1}) \quad (15)$$

$$\mathbf{d}^l = \arg \min_{\mathbf{d}} \mathcal{L}(z^l, \mathbf{d}, \mathbf{u}^{l-1}, \mathbf{V}^{l-1}, \boldsymbol{\lambda}_z^{l-1}, \boldsymbol{\lambda}_u^{l-1}) \quad (16)$$

$$\mathbf{u}^l = \arg \min_{\mathbf{u}} \mathcal{L}(z^l, \mathbf{d}^l, \mathbf{u}, \mathbf{V}^{l-1}, \boldsymbol{\lambda}_z^{l-1}, \boldsymbol{\lambda}_u^{l-1}) \quad (17)$$

$$\mathbf{V}^l = \arg \min_{\mathbf{V}} \mathcal{L}(z^l, \mathbf{d}^l, \mathbf{u}^l, \mathbf{V}, \boldsymbol{\lambda}_z^{l-1}, \boldsymbol{\lambda}_u^{l-1}) \quad (18)$$

- 3: Update Lagrange multipliers

$$\boldsymbol{\lambda}_z^l = \boldsymbol{\lambda}_z^{l-1} + r_z(\mathbf{d}^l - \nabla z^l) \quad (19)$$

$$\boldsymbol{\lambda}_u^l = \boldsymbol{\lambda}_u^{l-1} + r_u(\mathbf{V}^l - \nabla \mathbf{u}^l) \quad (20)$$

- 4: Stop the iterative process when $\frac{\|\mathbf{u}^l - \mathbf{u}^{l-1}\|}{\|\mathbf{u}^l\|} < \epsilon$, $\frac{\|z^l - z^{l-1}\|}{\|z^l\|} < \epsilon$.
-

The next step is to determine the solutions of the sub-minimization problems (15),(18), which can be computed efficiently. To simplify notation we omit the super-index in the sub-minimization problems.

D. Minimization w.r.t. z

If we have an OPC model z_0 of the printed surface, the sub-minimization problem with respect to z reads

$$\min_z \int_{\Omega} \frac{\beta}{2} (z - \bar{z}_0)^2 + \boldsymbol{\lambda}_z \cdot (\mathbf{d} - \nabla z) + \frac{r_z}{2} |\mathbf{d} - \nabla z|_2^2, \quad (21)$$

where $\bar{z}_0 = z_0(\mathbf{x} + \mathbf{u})$ does not depend on z . We observe that it is a convex minimization problem and therefore we can compute its optimality conditions by means of its first variation. Euler-Lagrange gives us the following optimality condition

$$\beta z - r_z \Delta z = \boldsymbol{\lambda}_z + \beta \bar{z}_0 - \operatorname{div}(\boldsymbol{\lambda}_z + r_z \mathbf{d}). \quad (22)$$

If we discretize the optimality conditions and consider each scalar function in Ω as a vector of its pixel values, equation (22) is discretized as a system of linear equations $A_z z = b$ with a positive definite matrix. In particular we have

$$\begin{aligned} A_z &= \beta I_n + r_z D_x^T D_x + r_z D_y^T D_y \\ b &= \beta \bar{z}_0 + D_x^T (r_d d_x + \lambda_{z,x}) + D_y^T (r_d d_y + \lambda_{z,y}). \end{aligned} \quad (23)$$

Note that matrix A_z is block circulant and we can use the Fourier transform \mathcal{F} to decompose it as $A_z = \mathcal{F}^T D \mathcal{F}$, with D a diagonal matrix. Consequently, the system (23) can easily be solved in the Fourier domain. In practice we use the FFT transform instead of doing the matrix multiplications with \mathcal{F} and \mathcal{F}^T , which gives us the solution $z = \mathcal{F}^T (D^{-1} \mathcal{F} b)$ with a complexity $\mathcal{O}(n \log n)$.

If we compare the surface to the expected mask in terms of the Heaviside function, we have the following minimization problem

$$\min_z \int_{\Omega} \frac{\beta}{2} (H_{\theta}(z) - h)^2 + \boldsymbol{\lambda}_z \cdot (\mathbf{d} - \nabla z) + \frac{r_z}{2} |\mathbf{d} - \nabla z|_2^2, \quad (24)$$

where $h = H(\phi(\mathbf{x} + \mathbf{u}))$ does not depend on z and we have denote as H_{θ} the composition of the Heaviside function with the translation by θ , i.e., $H_{\theta}(x) = H(x - \theta)$. In this case we do not have a closed-form solution for the minimization and require an iterative algorithm. In practical terms, we derive a fast semi-implicit minimization algorithm, that we initialize with the value of the previous iterate in the augmented Lagrangian Algorithm 1. The resulting minimization with respect to z converges within 2-5 iterations.

The Euler-Lagrange equations associated with Equation (24) are

$$-r_z \Delta z = \beta(H_\theta(z) - h)\delta_\theta(z) + \boldsymbol{\lambda}_z + \beta \bar{z}_0 - \operatorname{div}(\mathbf{v}), \quad (25)$$

where $\mathbf{v} = \boldsymbol{\lambda}_z + r_z \mathbf{d}$ and $\delta(\cdot)$ is the derivative of the Heaviside function, that is, the Dirac distribution. Given a small $\epsilon > 0$, the system is equivalent to

$$\epsilon z - r_z \Delta z = \epsilon z + \beta(H_\theta(z) - h)\delta_\theta(z) + \boldsymbol{\lambda}_z + \beta \bar{z}_0 - \operatorname{div} \mathbf{v} \quad (26)$$

and can be solved with the following fixed-point iterative method (iterations on k)

$$\begin{aligned} \epsilon z^{k+1} - r_z \Delta z^{k+1} = & \quad \epsilon z^k + \beta(H_\theta(z^k) - h)\delta_\theta(z^k) + \\ & \quad \boldsymbol{\lambda}_z + \beta \bar{z}_0 - \operatorname{div} \mathbf{v}. \end{aligned} \quad (27)$$

At each iteration we have a linear system equivalent to (22), that we solve again in the Fourier domain.

E. Minimization w.r.t. \mathbf{d}

If we define $\mathbf{v} = \nabla z - \frac{\boldsymbol{\lambda}_z}{r_z}$, the minimization problem with respect to $\mathbf{d} = (d_x, d_y)$ reads

$$\min_{\mathbf{d}} \int_{\Omega} \alpha (R(\mathbf{d}) - I_0)^2 + \gamma w |\nabla \mathbf{d}|^2 + r_z |\mathbf{d} - \mathbf{v}|_2^2. \quad (28)$$

The ideal reflectance function for SEM is given by Equation (3), but in fact, at the nanometre scale the reflectance map is considerably more complex and has to account for proximity effects. As a result, Equation (3) is only an approximation of the SEM image acquisition system, which can be equally modelled by its first order Taylor expansion in the minimization process. Such an approach has also been shown to improve convergence in [32]. At iteration $k + 1$ we approximate the reflectance map around the previous slope estimate as

$$R(\mathbf{d}) \approx R^k + (d_x - d_x^k) R_x^k + (d_y - d_y^k) R_y^k, \quad (29)$$

where $R^k = R(\mathbf{d}^k)$, $R_x^k = \partial_{d_x} R(\mathbf{d}^k)$ and $R_y^k = \partial_{d_y} R(\mathbf{d}^k)$. Then, at each iteration we have to solve

$$\min_{\mathbf{d}} \int_{\Omega} \alpha \left(R_x^k d_x + R_y^k d_y + \bar{R}^k \right)^2 + \gamma w |\nabla \mathbf{d}|^2 + r_z |\mathbf{d} - \mathbf{v}|^2, \quad (30)$$

with $\bar{R}^k = R(\mathbf{d}^k) - R_x^k d_x^k - R_y^k d_y^k - I_0$.

For simplicity we minimize independently for each component of \mathbf{d} and correct this approximation

through the iterative method. A simultaneous minimization could also be performed efficiently solving a 2×2 linear system of equations for each pixel in the frequency domain, see e.g., [64]. The Euler-Lagrange equations associated with the minimization with respect to d_x and d_y are

$$\begin{aligned} r_z d_x + \alpha R_x^k d_x + \gamma \nabla w \cdot \nabla d_x - \gamma w \Delta d_x &= r_z v_x - \alpha R_x^k (R_y^k d_y + \bar{R}^k) \\ r_z d_y + \alpha R_y^k d_y + \gamma \nabla w \cdot \nabla d_y - \gamma w \Delta d_y &= r_z v_y - \alpha R_y^k (R_x^k d_x + \bar{R}^k). \end{aligned}$$

If we discretize the optimality conditions, we have a linear system of equations $A_x d_x = b_x$ and $A_y d_y = b_y$ with

$$\begin{aligned} A_x &= r_z I_n + \alpha D_{R_x k} + \gamma D_x^T D_w D_x + \gamma D_y^T D_w D_y \\ b_x &= \alpha D_{R_x k} (D_{R_y k} d_y - D_{\bar{R}^k}) + r_z v_x \\ A_y &= r_z I_n + \alpha D_{R_y k} + \gamma D_x^T D_w D_x + \gamma D_y^T D_w D_y \\ b_y &= \alpha D_{R_y k} (D_{R_x k} d_x - D_{\bar{R}^k}) + r_z v_y, \end{aligned}$$

where $D_{R_x k}$, $D_{R_y k}$ and D_w are diagonal matrices, with entries associated with the values of functions R_x^k , R_y^k and w in Ω . Matrices A_x , A_y are then symmetric, sparse and diagonally dominant and the systems can be inverted with an iterative Gauss-Seidel method. As a consequence of the sparsity of the matrices, each iteration of the Gauss-Seidel method is extremely fast and convergence is achieved within a few updates.

F. Minimization w.r.t. \mathbf{u}

If we have an OPC model z_0 of the printed surface, the sub-minimization problem with respect to \mathbf{u} reads

$$\begin{aligned} \min_{\mathbf{u}} \int_{\Omega} \frac{\beta}{2} (z(\mathbf{x}) - z_0(\mathbf{x} + \mathbf{u}))^2 + \boldsymbol{\lambda}_u \cdot (V(\mathbf{x}) - \nabla \mathbf{u}(\mathbf{x})) + \\ \frac{r_u}{2} |V(\mathbf{x}) - \nabla \mathbf{u}(\mathbf{x})|^2. \end{aligned} \quad (31)$$

The associated Euler-Lagrange equation is

$$\begin{aligned} -r_u \Delta u(\mathbf{x}) &= b(u) \\ b(u) &= v + \beta z_0(\mathbf{x} + \mathbf{u}) \nabla z_0(\mathbf{x} + \mathbf{u}), \end{aligned}$$

where we define $v = -\beta z(\mathbf{x}) - \text{div}(\boldsymbol{\lambda}_u(\mathbf{x}) + r_u V(\mathbf{x}))$

If we compare the surface to the expected mask in terms of the Heaviside function, we have the following minimization problem

$$\begin{aligned} \min_{\mathbf{u}} \int_{\Omega} \frac{\beta}{2} (h_{\theta}(z(\mathbf{x})) - H(\phi(\mathbf{x} + \mathbf{u})))^2 + \\ \boldsymbol{\lambda}_u \cdot (V(\mathbf{x}) - \nabla \mathbf{u}(\mathbf{x})) + \frac{r_u}{2} |V(\mathbf{x}) - \nabla \mathbf{u}(\mathbf{x})|^2. \end{aligned} \quad (32)$$

The associated Euler-Lagrange equation is now

$$\begin{aligned} -r_u \Delta u(\mathbf{x}) &= b(u) \\ b(u) &= v + \beta H(\phi(\mathbf{x} + \mathbf{u})) \delta(\phi(\mathbf{x} + \mathbf{u})) \nabla \phi(\mathbf{x} + \mathbf{u}), \end{aligned}$$

where now $v = -\beta h_z(\mathbf{x}) - \text{div}(\boldsymbol{\lambda}_u(\mathbf{x}) + r_u V(\mathbf{x}))$.

In both cases, the optimality conditions reduce to

$$-r_u \Delta \mathbf{u} = \mathbf{b}(\mathbf{u}) \quad (33)$$

and can be solved iteratively with the same scheme. To that purpose, we choose a small $\epsilon > 0$ and initialize the following fixed-point method with the solution to the previous Lagrangian iterate

$$\epsilon \mathbf{u}^{t+1} - r_u \Delta \mathbf{u}^{t+1} = \mathbf{b}(\mathbf{u}^t) + \epsilon \mathbf{u}^t. \quad (34)$$

The previous fixed-point schema is discretized as a system of linear equations equivalent to (23) and can be efficiently solved by means of the FFT transform or a few iterations of Gauss-Seidel. We have experimentally observed that 2-5 iterations of this fixed-point algorithm are enough to achieve convergence.

G. Minimization w.r.t. \mathbf{V}

Defining $\mathbf{W} = \nabla \mathbf{u} - \frac{\boldsymbol{\lambda}_u}{r_u}$, the minimization w.r.t. \mathbf{V} reads

$$\min_{\mathbf{V}} \int_{\Omega} |\mathbf{V} - \mathbf{W}|_F^2 + \mathcal{W}\left(\frac{\mathbf{V} + \mathbf{V}^T + \mathbf{V}^T \mathbf{V}}{2}\right), \quad (35)$$

where $|\cdot|_F$ designates the Frobenius norm. In this case the minimization functional contains no derivatives is decoupled for each pixel and we only need to solve a scalar minimization problem independently for each pixel. To that purpose we use a semi-implicit gradient descent scheme with time step τ . At each pixel we denote $\mathbf{V} = \begin{pmatrix} V_{1,1} & V_{1,2} \\ V_{2,1} & V_{2,2} \end{pmatrix}$ and iteratively update \mathbf{V} pixel-wise with the following rule

$$\begin{aligned}
V_{1,1}^{k+1} &= \frac{V_{1,1}^k + \tau (r_u W_{1,1} - \lambda c_{0,1}^k - \mu c_{0,2}^k - \mu c_{0,4}^k V_{1,2}^k)}{1 + \tau (r_u + \lambda c_{0,1}^k + \mu c_{0,2}^k)} \\
V_{1,2}^{k+1} &= \frac{V_{1,2}^k + \tau (r_u W_{1,2} - \mu c_{0,4}^k (1 + V_{1,1}^k))}{1 + \tau (r_u + \lambda c_{0,1}^k + \mu c_{0,3}^k)} \\
V_{2,1}^{k+1} &= \frac{V_{2,1}^k + \tau (r_u W_{2,1} - \mu c_{0,4}^k (1 + V_{2,2}^k))}{1 + \tau (r_u + \lambda c_{0,1}^k + \mu c_{0,2}^k)} \\
V_{2,2}^{k+1} &= \frac{V_{2,2}^k + \tau (r_u W_{2,2} - \lambda c_{0,1}^k - \mu c_{0,3}^k - \mu c_{0,4}^k V_{2,1}^k)}{1 + \tau (r_u + \lambda c_{0,1}^k + \mu c_{0,3}^k)},
\end{aligned}$$

where

$$\begin{aligned}
c_{0,1} &= V_{1,1} + V_{1,2} + \frac{1}{2} (V_{1,1}^2 + V_{1,2}^2 + V_{2,1}^2 + V_{2,2}^2) \\
c_{0,2} &= 2V_{1,1} + V_{1,1}^2 + V_{2,1}^2 \\
c_{0,3} &= 2V_{2,2} + V_{1,2}^2 + V_{2,2}^2 \\
c_{0,4} &= V_{1,2} + V_{2,1} + V_{1,1}V_{1,2} + V_{2,1}V_{2,2}.
\end{aligned}$$

A similar scheme has been proposed in [47] in the context of image registration, but the relation $\mathbf{V} = \nabla \mathbf{u}$ is here considered a constraint instead of a penalty in the minimization. Experimentally we have found that few iterations are enough to reach the accuracy required for the Lagrangian update rule of Algorithm 1.

IV. EXPERIMENTAL RESULTS

In this section we present the experimental results of our surface reconstruction method. For our experiments we have used large field of view SEM images of circuits corresponding to diffusion layer, covering an area of $1.25 \times 1.25 \mu\text{m}$. We have used a multi-resolution approach to speed up convergence, adapting the shape prior at each resolution with the solution of the previous scale in order to have smaller and smoother deformation fields at each scale. Therefore, at each scale the deformation should be computed concatenating the deformations at previous resolutions correctly scaled. For simplicity, in all the experiments we have set $\alpha = 500$, $\beta = 50$, $\gamma = 100$, $\lambda = \mu = 10^{-4}$, but the value of these parameters could be refined.

The shape prior adopted are based on the mask, that is on Equation (7), as we did not have access to

OPC models. Convergence of the algorithm should be faster with a shape prior of Equation (6) based on OPC models because the sub-minimization tasks with respect to z and \mathbf{u} would be easier. However, the reconstructed surface would only be more accurate to the same extent that the OPC models are. In that sense, it is always better to use limited but reliable mask information, than unreliable OPC models. Figure 1 shows the masks associated to the patterns in the experiments and Figure 3 shows the reconstructed surface for a simple structure, where we have included the reconstructed surface and the estimated deformation field \mathbf{u} . For quick visual inspection, we have superposed the isoline of height $\theta = 0.5z_{max} + 0.5z_{min}$ of the reconstructed surface in the SEM image of Figure 2, where we observe that the surface reconstruction correctly estimates the position of the edges in the image.

Figure 4 presents the results for the reconstruction of more complex patterns of Figure 1(b)-1(d). For these experiments we can only provide qualitative results for our reconstruction method because no ground truth of the reconstructed surfaces is currently available, but we observe that results agree with our intuitive reconstruction from the SEM images.

To measure the contribution of the shape-prior term in the reconstruction, we compare the results to the equivalent variational method ignoring the shape-prior, that is, defining z as the minimizer of $\min_z \alpha \mathcal{A}(\nabla z) + \gamma \mathcal{G}(\nabla z)$. Figures 5(a) present the results of the reconstruction without the shape prior, where SFS alone fails due to noise, shadows and SEM artefacts. For instance, the shadow present under the white band in the SEM image results in a deep and narrow valley in the surface reconstructed in Figure 5(e).

In order to test the robustness of our method to noise, we have artificially added white Gaussian noise to the previous SEM images in Figures 5(b)-5(d). We observe that our SFS method is robust to noise. In particular, the only difference visible in the reconstructed shape in comparison to our reconstruction of Figure 3 and 4, are the gentle and almost negligible waves present in the dark flat areas. This waves are the only residual left in the reconstruction from the noise.

It is interesting to analyse the reconstructions of Figure 6, where the printing process partially fails. In the first pattern, we observe that the height of the walls of the reconstructed surface degrades in the vertical direction, while for the second surface we can see how the walls of the pattern become thinner and almost break. In both cases, our method produces surfaces that agree with the human interpretation of the corresponding SEM images and can detect failures in the lithographic process.

Due to the lack of reliable ground truth data at the nanometre scale, we propose the following experiment to indirectly evaluate the performance of our reconstruction method in the context of optical lithography. In particular we will compare one-dimensional distance measurements associated to different



Fig. 1: Binary masks associated to different patterns.

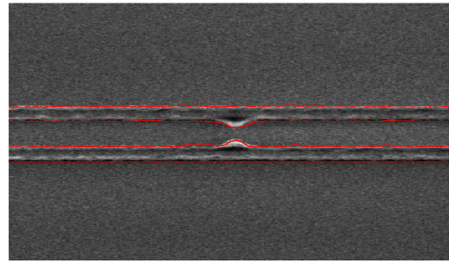


Fig. 2: SEM image with associated mask 1(a). In red we superposed the iso-contours of height θ associated to the reconstructed surface of Figure 3.

patterns obtained by OPC simulations with the corresponding distances measured from our reconstructed surfaces. To that purpose, it is first necessary to define some concepts from lithography. One of the basic tasks of computational lithography is the quantitative characterization of the printing process, from patterns and masks in circuit designs to the shapes carved in silicon wafers. In particular, numerical simulations are performed to model optical and physical properties of the light source and materials and estimate the shape of the wafer's surface with OPC models. For accurate results, these models are iteratively refined by calibration with measurements of the wafer in several cycles of printing-simulation. To that purpose, OPC modellers describe the patterns printed in a wafer in terms of critical distances, that is, one-dimensional measurements of distances associated to the surface of the wafer at a critical position designated by a gauge. Gauges are in fact line segments that mark the specific location for each pattern where the engineers expect that the printing process might fail. In the next experiment we analyse the critical distance associated to the reconstructed surface with its equivalent generated with calibrated OPC models for 24 patterns that printed correctly. With an accurate SFS reconstruction, we expect these distances to coincide with the OPC models; while wrong SFS reconstructions would lead to a mismatch of the pairs of critical distances. Consequently, we use this process as validation of our SFS reconstruction method. For the OPC simulations, we use the compact resist model *CMI* of *Mentor Graphics* and run simulations with *Calibre* to estimate the critical distance associated to each one of the 24 patterns. In both

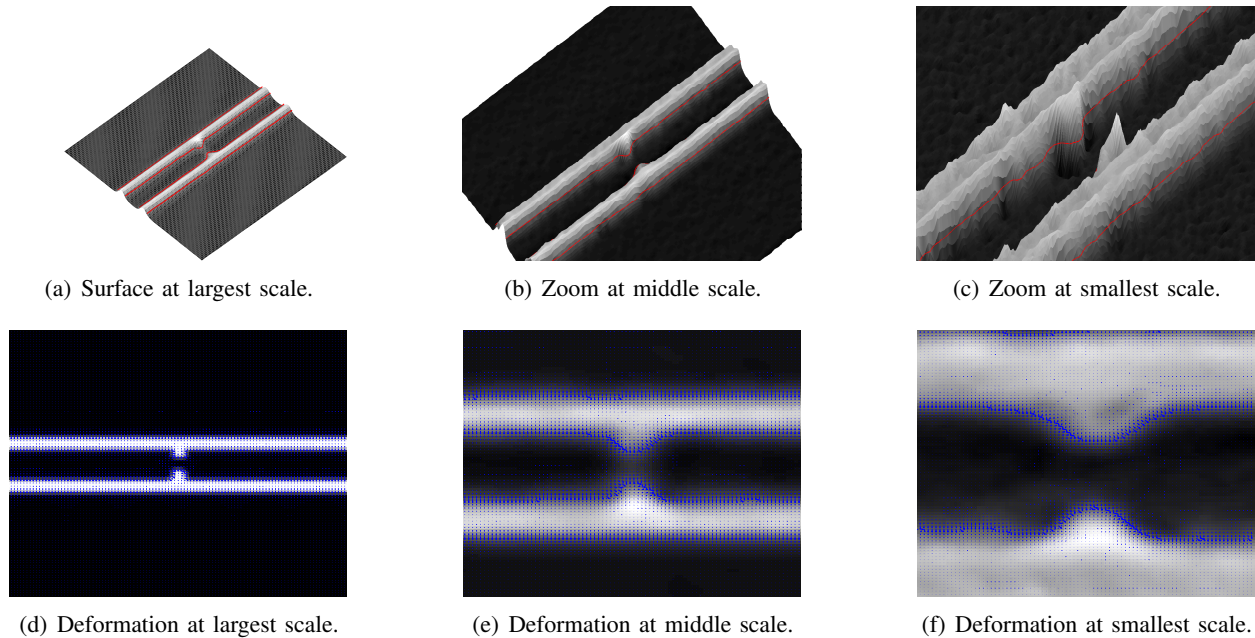


Fig. 3: Reconstructed surface and deformation fields at three different scales from SEM image of Figure 2.

simulations and SFS reconstruction, the critical distance is measured at the location of the gauge where the simulated and reconstructed surfaces attain height 0.3 with respect to a normalized surface. Figure 7 shows the scatter plots between the critical distances in OPC simulations and the proposed method. This Figure provides a quantitative evaluation of the proposed method, but it is only representative to the extent that the OPC model is reliable. We observe that the critical distances are nicely correlated (the difference in scaling is due to the pixel-to-nanometre conversion), except in a few outlier cases. We conclude that the proposed reconstruction method agrees with the calibrated OPC estimates. Currently we cannot assess if the existing differences are due to errors in the OPC models, in our SFS reconstruction method or they simply show effects of the manufacturing process that are not considered in the OPC model. To clarify this point it is necessary to compare the surfaces estimated with our method and OPC models to real and accurate 3D measurements on the wafer, possibly with atomic force microscopy. At present, this information is not available and the question remains open to future investigations.

V. CONCLUSIONS

In this paper we have proposed a surface reconstruction method from SEM images of silicon wafers. Our experiments show that the method is robust to noise and SEM artefacts because it incorporates

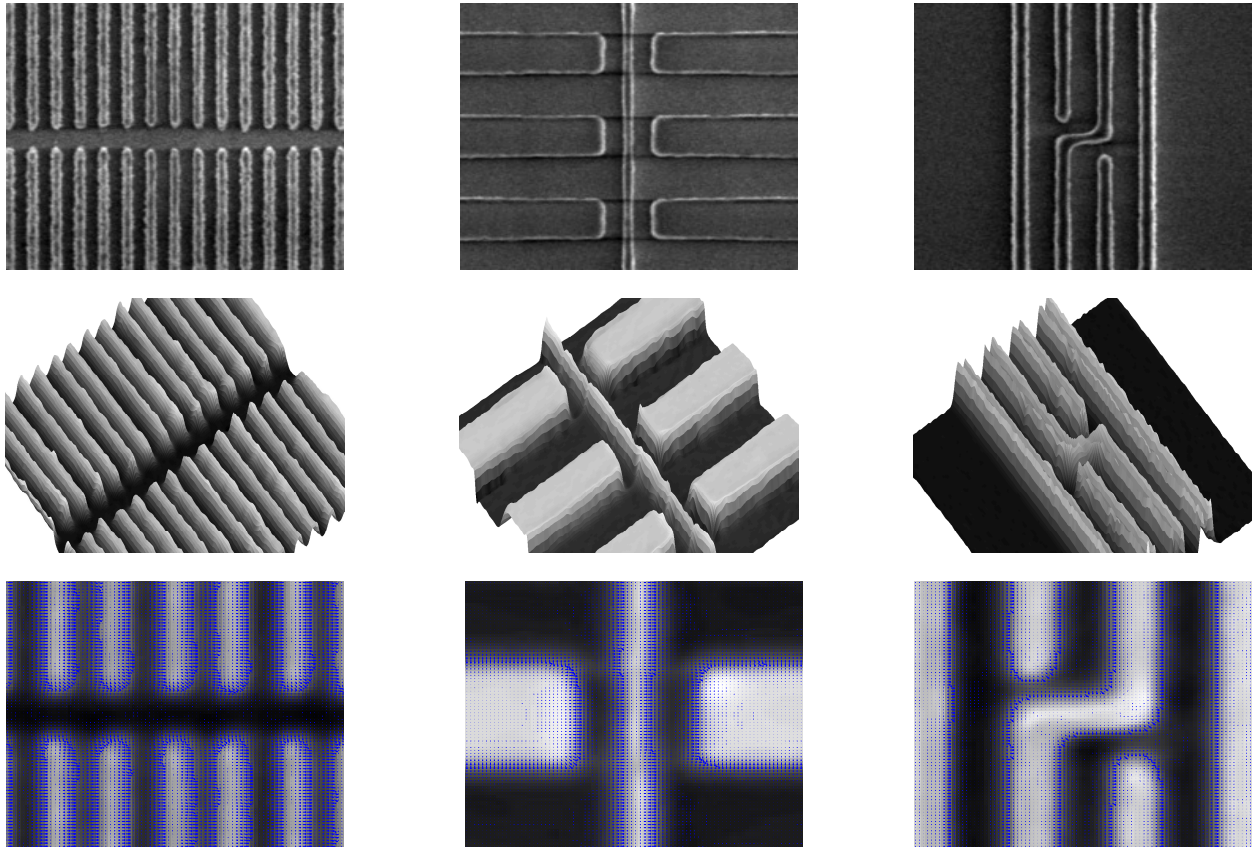


Fig. 4: Results of our surface reconstruction method for different SEM images and patterns of Figure 1(b) (left), 1(c) (middle) and 1(d) (right). First row shows the SEM images, the second column the reconstructed surfaces and the third one a zoom on the estimated deformation fields.

knowledge about the shapes of the patterns in the reconstruction process. In particular, we adopt an intuitive variational formulation that includes models of the SEM acquisition systems, the expected smoothness of the wafer and a term penalizing deviations from the patterns in the circuit. The resulting minimization is solved by variable splitting and augmented Lagrangian techniques, which results in a fast, efficient and easy-to-code algorithm. Our algorithm also estimates a deformation field between the ideal pattern of the circuit and the reconstructed surface, which is included to model variability of printing process and can be directly used to detect failures and critical regions in the wafer. From the point of view of applications, we believe that the proposed technique can have an important impact on the chip manufacturing industry, where an accurate description of the surface topography is essential for the characterization of the lithographic process.

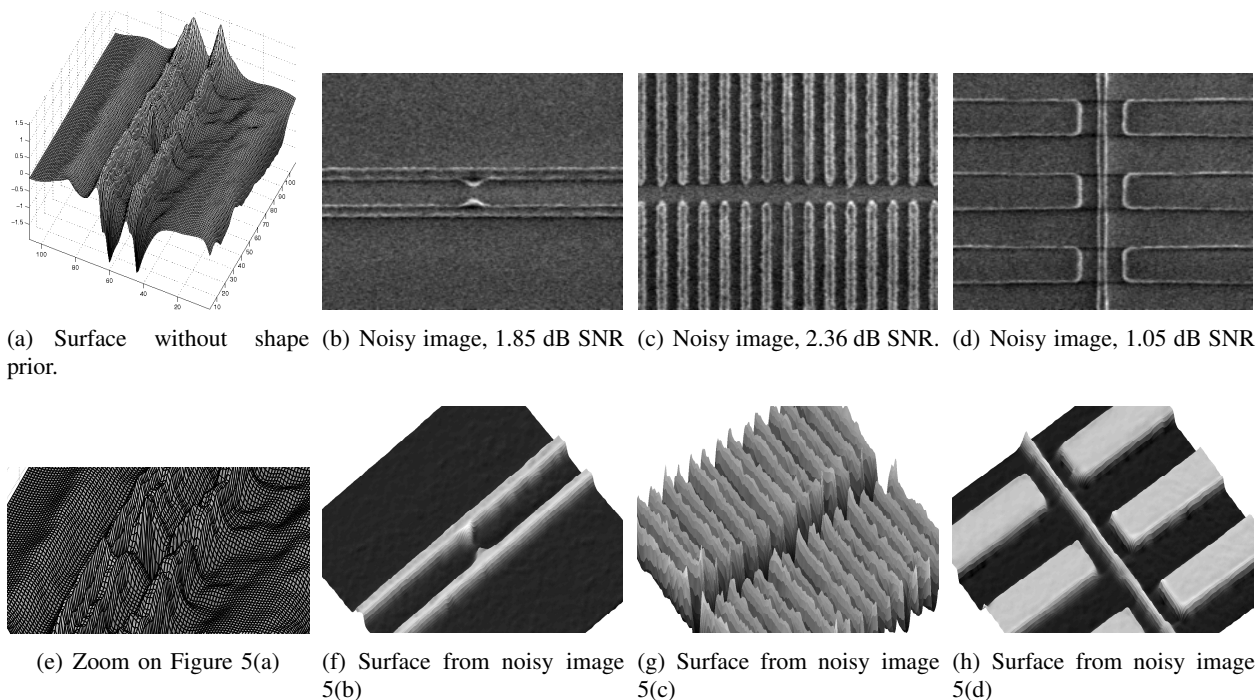


Fig. 5: First column shows the reconstruction results for the SEM image of Figure 2 ignoring the shape prior in the SFS algorithm. The last three columns correspond to reconstruction of surfaces with the proposed method with noisy images.

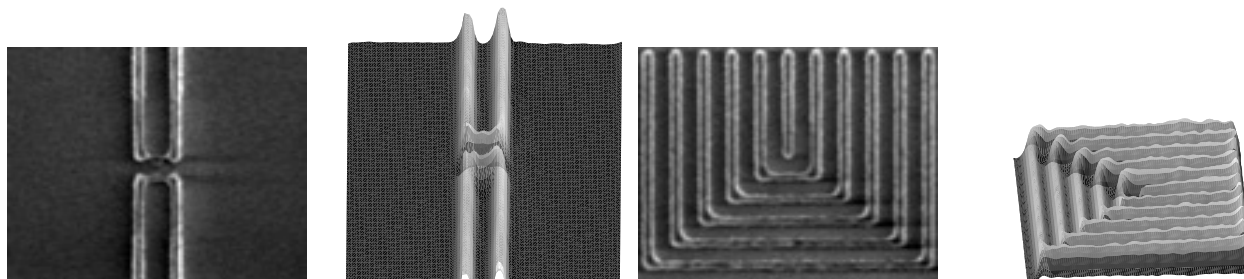


Fig. 6: SEM images and reconstructed surface for patterns of Figure 1(e) and 1(f). The lithographic process resulted in pinching in the center of the left' pattern and thin and fading the outer walls on the right pattern.

ACKNOWLEDGMENT

The authors would like to thanks R. Viswanathan, A. Abdo, B. Dai, D. Shao, J. Oberschmidt and D. Demaris from IBM for technical discussions on lithography.

REFERENCES

- [1] C. Mack, "extract_CLbook.pdf," in *Fundamental principles of optical lithography: the science of microfabrication*, 220, Ed. Wiley, 2007, pp. 121–124.

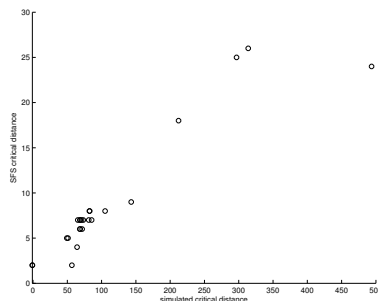


Fig. 7: Scatter plots of the critical distances obtained with our SFS reconstruction method and OPC simulations.

- [2] A. Kayaalp, a. Ravishankar Rao, and R. Jain, "Scanning electron microscope-based stereo analysis," *Machine Vision and Applications*, vol. 3, no. 4, pp. 231–246, Sep. 1990.
- [3] W. Beil and I. C. Carlsen, "Surface reconstruction from stereoscopy and shape from shading in SEM images," *Machine Vision and Applications*, vol. 4, no. 4, pp. 271–285, Sep. 1991.
- [4] T. P. Ellison and C. J. Taylor, "Calculating the surface topography of integrated circuit wafers from SEM images," *Image and Vision Computing*, vol. 9, no. 1, pp. 3–9, Feb. 1991.
- [5] A. Jones and C. Taylor, "Scale space surface recovery using binocular shading and stereo information," *Proceedings of the BMVC*, pp. 77–86, 1995.
- [6] R. Danzl and K. Karner, "Integrating Shape from Shading and Shape from Stereo for Variable Reflectance Surface Reconstruction from SEM Images," in *Proceedings of 26th Workshop of the Austrian Association for Pattern Recognition*, 2002, pp. 281–288.
- [7] E. Reithmeier, T. Vynnyk, and T. Schultheis, "3D-measurement using a scanning electron microscope," *Applied mathematics and computation*, vol. 217, no. 3, pp. 1193–1201, 2011.
- [8] H. Morokuma, A. Sugiyama, and Y. Toyoda, "A new matching engine between design layout and SEM image of semiconductor device," *Proceedings of SPIE*, 2005.
- [9] C. Tabery, H. Morokuma, R. Matsuoka, L. Page, G. E. Bailey, I. Kusnadi, and T. Do, "SEM image contouring for OPC model calibration and verification," *Proceedings of SPIE*, 2007.
- [10] C. Tabery, H. Morokuma, A. Sugiyama, and L. Page, "Evaluation of OPC quality using automated edge placement error measurement with CD-SEM," *Proceedings of SPIE*, 2006.
- [11] H. Shindo, A. Sugiyama, H. Komuro, Y. Hojo, R. Matsuoka, J. L. Sturtevant, T. Do, I. Kusnadi, G. Fenger, P. De Bisschop, and J. Van de Kerckhove, "High-precision contouring from SEM image in 32-nm lithography and beyond," *Proceedings of SPIE*, vol. 7275, pp. 72 751F–72 751F–9, 2009.
- [12] C. Shishido, M. Tanaka, and M. Osaki, "Cd bias reduction in cd-sem of very small line patterns: sidewall shape measurement using model-based library matching method," pp. 76 383I–76 383I, 2010.
- [13] T. Shibahara, T. Minakawa, M. Oikawa, H. Shindo, H. Sugahara, and Y. Hojoyo, "A cd-gap-free contour extraction technique for opc model calibration," vol. 7971, 2011.
- [14] B. Horn, "Shape from shading: A method for obtaining the shape of a smooth opaque object from one view," Ph.D. dissertation, Massachusetts Institute of Technology, 1970.
- [15] A. Jones and C. Taylor, "Robust shape from shading," *Image and Vision Computing*, vol. 12, no. 7, pp. 411–421, Sep. 1994.
- [16] L. B. Rad, H. Feng, J. Ye, R. F. W. Pease, D. G. Seiler, A. C. Diebold, R. McDonald, C. M. Garner, D. Herr, R. P. Khosla, and E. M. Secula, "Computational Scanning Electron Microscopy," in *Conference on Frontiers of Characterization and Metrology*, vol. 931, no. 1. Aip, 2007, pp. 512–517.
- [17] J. S. Villarrubia, A. E. Vladar, B. D. Bunday, and M. Bishop, "Dimensional Metrology of Resist Lines using a SEM Model-Based Library Approach," in *Proceedings of SPIE*, vol. 199, May 2004, pp. 199–209.
- [18] J. Villarrubia, R. Dixon, and A. Vladár, "Proximity-associated errors in contour metrology," *Proceedings of SPIE*, vol. 7638, pp. 76 380S–76 380S, 2010.
- [19] J. S. Villarrubia and Z. J. Ding, "Sensitivity of SEM width measurements to model assumptions," *Proceedings of SPIE*, vol. 7272, no. Cd, pp. 72 720R–72 720R–15, 2009.
- [20] R. Zhang, P. Tsai, J. Cryer, and M. Shah, "Shape-from-shading: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690–706, 1999.
- [21] J.-D. Durou, M. Falcone, and M. Sagona, "Numerical methods for shape-from-shading: A new survey with benchmarks," *Computer Vision and Image Understanding*, vol. 109, no. 1, pp. 22–43, Jan. 2008.
- [22] E. Prados, F. Camilli, and O. Faugeras, "A unifying and rigorous shape from shading method adapted to realistic data and applications," *Journal of Mathematical Imaging and Vision*, vol. 3, 2006.
- [23] K. Ikeuchi and B. K. Horn, "Numerical shape from shading and occluding boundaries," *Artificial Intelligence*, vol. 17, no. 1-3, pp. 141–184, Aug. 1981.

- [24] M. Brooks and B. Horn, "Shape and Source from Shading," in *Shape from Shading*, B. Horn, Ed. MIT Press, 1989, ch. Shape and, pp. 53–68.
- [25] P. Worthington and E. Hancock, "New constraints on data-closeness and needle map consistency for shape-from-shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1250–1267, 1999.
- [26] P. Daniel and J. Durou, "From deterministic to stochastic methods for shape from shading," *Proc. 4th Asian Conf. on Comp. Vis.*, pp. 1–23, 2000.
- [27] R. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 439–451, Jul. 1988.
- [28] B. K. P. Horn, "Height and gradient from shading," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 37–75, Aug. 1990.
- [29] Y. Leclerc and A. Bobick, "The Direct Computation of Height from Shading," in *Computer Vision and Pattern Recognition*, 1991, pp. 552–558.
- [30] R. Szeliski, "Fast shape from shading," *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 129 – 153, 1991.
- [31] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 10, no. 10, pp. 1467–75, Jan. 1997.
- [32] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading," *Computer Vision and Pattern Recognition*, pp. 540–545, 1991.
- [33] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 331, pp. 321–331, 1988.
- [34] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Comput. Soc. Press, 1996, pp. 465–470.
- [35] A. Frangi and W. Niessen, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 1496, no. 3, 1998, pp. 130–137.
- [36] M. Leventon, "Statistical shape influence in geodesic active contours," in *International Conference in Computer Vision and Pattern Recognition*, 2000, pp. 1316–1323.
- [37] D. Cremers and F. Tischhäuser, "Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional," *International journal of computer vision*, vol. 50, no. 3, pp. 295–313, 2002.
- [38] M. Rousson and N. Paragios, "Shape priors for level set representations," in *European Conference on Computer Vision*, 2002, pp. 78–92.
- [39] Y. Chen and H. Tagare, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.
- [40] T. Chan and W. Zhu, "Level Set Based Shape Prior Segmentation," *Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 1164–1170, 2005.
- [41] X. Bresson, P. Vandergheynst, and J.-P. Thiran, "A Variational Model for Object Segmentation Using Boundary Information and Shape Prior Driven by the Mumford-Shah Functional," *International Journal of Computer Vision*, vol. 68, no. 2, pp. 145–162, Mar. 2006.
- [42] A. Foulonneau, P. Charbonnier, and F. Heitz, "Affine-invariant geometric shape priors for region-based active contours." *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 8, pp. 1352–7, Aug. 2006.
- [43] D. Cremers, S. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk," *Pattern Recognition*, pp. 1–28, 2004.
- [44] A. Tsai and A. Y. Jr, "Model-based curve evolution technique for image segmentation," in *IEEE International Conference of Computer Vision and Pattern Recognition*, vol. 00, no. C, 2001, pp. 2–7.
- [45] N. Paragios, "A level set approach for shape-driven segmentation and tracking of the left ventricle." *IEEE transactions on medical imaging*, vol. 22, no. 6, pp. 773–6, Jun. 2003.
- [46] D. Cremers, N. Sochen, and C. Schnörr, "Towards recognition-based variational segmentation using shape priors and dynamic labeling," *Scale Space Methods in Computer Vision*, pp. 388–400, 2003.
- [47] C. L. Guyader and L. Vese, "A combined segmentation and registration framework with a nonlinear elasticity smoother," *Scale Space and Variational Methods in Computer Vision*, vol. 2, pp. 600–611, 2009.
- [48] J. Modersitzki, *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. Oxford university press, 2004.
- [49] B. Fischer and J. Modersitzki, "Fast diffusion registration," *Contemporary Mathematics*, vol. 313, pp. 117–128, 2002.
- [50] —, "Curvature based image registration," *Journal of Mathematical Imaging and Vision*, pp. 81–85, 2003.
- [51] —, "A unified approach to fast image registration and a new curvature based registration technique," *Linear Algebra and its Applications*, vol. 380, pp. 107–124, Mar. 2004.
- [52] C. Broit, "Optimal registration of deformed images." Ph.D. dissertation, University of Pennsylvania, 1981.
- [53] G. Christensen, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.
- [54] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging*, pp. 1–24, 2008.
- [55] T. Goldstein and S. Osher, "The Split Bregman Method for L1-Regularized Problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, p. 323, 2009.
- [56] R. Glowinski, J.-L. Lions, and R. Tremolieres, *Numerical Analysis of Variational Inequalities*. Elsevier, 1981.
- [57] T. Goldstein, X. Bresson, and S. Osher, "Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1–3, pp. 272–293, Nov. 2009.
- [58] A. Chambolle, "An algorithm for Mean Curvature Motion," *Interfaces and Free Boundaries*, vol. 6, pp. 195–218, 2004.

- [59] T. F. Chan, G. H. Golub, and P. Mulet, "A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration," *SIAM Journal on Scientific Computing*, vol. 20, no. 6, p. 1964, 1999.
- [60] M. Zhu, S. J. Wright, and T. F. Chan, "Duality-based algorithms for total-variation-regularized image restoration," *Computational Optimization and Applications*, vol. 47, no. 3, pp. 377–400, Dec. 2010.
- [61] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman Iterative Algorithms for l_1 -Minimization with Applications to Compressed Sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.
- [62] C. Wu and X.-C. Tai, "Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 300–339, 2010.
- [63] R. Glowinski and P. L. Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM, 1989.
- [64] J. Hahn, C. Wu, and X.-C. Tai, "Augmented Lagrangian Method for Generalized TV-Stokes Model," *Journal of Scientific Computing*, vol. 50, no. 2, pp. 235–264, Apr. 2011.