Model-Based 2.5-D Deconvolution for Extended Depth of Field in Brightfield Microscopy

François Aguet, *Student Member, IEEE*, Dimitri Van De Ville, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

***Abstract*—Due to the limited depth of field of brightfield microscopes, it is usually impossible to image thick specimens entirely in focus. By optically sectioning the specimen, the in-focus information at the specimen's surface can be acquired over a range of images. Commonly based on a high-pass criterion, extended-depth-of-field methods aim at combining the in-focus information from these images into a single image of the texture on the specimen's surface. The topography provided by such methods is usually limited to a map of selected in-focus pixel positions and is inherently discretized along the axial direction, which limits its use for quantitative evaluation. In this paper, we propose a method that jointly estimates the texture and topography of a specimen from a series of brightfield optical sections; it is based on an image formation model that is described by the convolution of a thick specimen model with the microscope's point spread function. The problem is stated as a least-squares minimization where the texture and topography are updated alternately. This method also acts as a deconvolution when the in-focus PSF has a blurring effect, or when the true in-focus position falls in between two optical sections. Comparisons to state-of-the-art algorithms and experimental results demonstrate the potential of the proposed approach.**

***Index Terms*—Biomedical image processing, deconvolution, inverse problems, optical transfer functions.**

I. Introduction

THE limited depth of field of conventional brightfield microscopes is a significant shortcoming when imaging specimens whose thickness and surface profile extend beyond the system's focal range. It is impossible to image such samples entirely in focus with a single acquisition; only those portions that lie within the depth of field appear in focus and sharp, whereas the remaining regions are blurred by the system's point spread function (PSF). A common way around this limitation consists in acquiring a series of optical sections of the sample by gradually moving it through the focal plane. This results in a "z-stack" of images that collectively contains all available in-focus information of the specimen. Such z-stacks can be difficult to interpret, and the success of an automated analysis essentially depends on its ability to correctly identify in-focus information. Many methods have been proposed for extending the depth of field of microscopes, either by optical means [1] or through image processing [2]–[4], and sometimes via a combination of both [5]. In practice, it is often achieved via the fusion of images by means of extended-depth-of-field (EDF) algorithms [6]. In this work, we propose a new model-based alternative to these approaches which combines the recovery of both the texture and the topography of the specimen in a single, global optimization process.

Manuscript received November 23, 2007; revised March 11, 2008. This work was supported in part by the Swiss National Science Foundation under Grant 200020-109415 and in part by the Hasler Foundation under Grant 2033. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Peyman Milanfar.

The authors are with the Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland (e-mail: francois.aguet@epfl.ch; dimitri.vandeville@epfl.ch; michael.unser@epfl.ch).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2008.924393

A. Review of Previous Work

1) Optical Techniques: The first type of solution relies on modified microscope optics along with nonstandard acquisition schemes. McLachlan described a system for illuminating the focal plane with a sheet of light that permits integration by axially scanning over the desired range of the specimen, and thus, avoids collecting out-of-focus information [7]. Similarly, Sheppard *et al.* later proposed confocal microscopy as a means of scanning and integrating the sample along the axial direction [1]. An acquisition scheme using structured illumination was presented by Burke *et al.* [8]. A further approach consists in using pupil masks [9]; wave-front coding to increase the depth-of-field has been demonstrated in [10], [11]. It is also worth noting that promising results have been achieved via annular illumination in the case of two-photon fluorescence microscopy [12].

2) Computational Techniques: None of the above methods are practical or possible to implement for conventional brightfield microscopy, and, thus, image processing was exploited early on to propose alternate solutions. Regrouped under the denomination of EDF algorithms, this second type of solution facilitates the visualization and interpretation of z-stack acquisitions by combining the in-focus information from multiple images into a single fusion image that depicts the specimen entirely in focus. Due to the thickness and staining of samples prepared for brightfield microscopy, it is generally assumed that the information that appears in focus lies on the surface of the specimen (we shall refer to this information as the specimen's texture). Given a z-stack acquisition, the goal of EDF algorithms thus lies in determining the in-focus position z that corresponds to the surface of the specimen for every point (x, y) in the image plane. Almost all existing EDF methods are variations of the following scheme:

1) slice-by-slice application of a high-pass criterion;
2) energy measurement in a local neighborhood around each pixel (x, y) in every slice z;
3) construction of an in-focus map by selection of the slice index z with maximal energy at every position (x, y);
4) generation of the EDF image based on the in-focus index map.

1057-7149/$25.00 © 2008 IEEE

A notable byproduct of the above is the topographical information obtained in the form of the in-focus index map.

One of the earliest descriptions of EDF algorithms can be traced back to Pieper and Korpel; these authors discussed the use of pointwise criteria based on intensity, as well as nondirectional finite difference filtering to discriminate for in-focus regions [2]. Subsequently, Sugimoto and Ichioka suggested using the variance within fixed windows of the acquired z-stack as a possible sharpness criterion [13]. They also showed that the texture and topography can be used to generate a stereoscopic representation of the sample. Itoh *et al.* adapted this approach to account for possible intensity variations in reflected light microscopy by introducing a cylindrical lens into the system and evaluating the local variance according to the resulting astigmatic aberrations [14]. Other proposals include high-pass filtering [15] and gradient-based approaches, where the presence of sharp edges is used to detect in-focus areas [16].

To date, the most successful approaches rely on some form of multiresolution analysis; e.g., steerable pyramids [17]–[19] and wavelet decompositions [3], [21]–[24]. By applying a wavelet transform to every slice, one automatically performs high-pass filtering at different resolutions. This approach avoids the choice of a fixed-size filter. The selection of the in-focus slice is performed in the wavelet domain too; i.e., at every resolution level. For a state-of-the-art wavelet approach, we refer to [6]; this work also includes a comparative evaluation of the primary EDF methods.

B. Towards a New Model-Based Approach

The primary goal of all methods outlined above is to yield an accurate rendition of the sample's texture. Although the map of selected in-focus slice positions for every pixel yields some topographical information, this distance map is noisy and coarsely discretized, and is ill-suited for accurate 3-D reconstruction and texture mapping. Some approaches (see, e.g., [20]) use local consistency checks and smoothing to improve the appearance of the topography, but they do not guarantee an accurate interpretation of the specimen's surface profile, nor do they suppress the inherent "staircase" effects.

In this paper, we propose a new algorithm for EDF that uses a parametric model of image formation to recover the specimen's texture and topography through an optimization process. Specifically, we assume that image formation is the result of a 3-D convolution between the PSF of the system and the sample, where the latter is modeled as a texture mapped onto a thin surface, which is described by a topography map [25]. We then formulate the EDF reconstruction as a least-squares estimation problem, and propose a solution by alternate optimization of the texture and the topography. Compared to the previously discussed EDF techniques, the topography is not limited to discretized values anymore, since it can change in a continuous way to maximally match the measurements. A further advantage is that the texture estimation process is capable of acting as a deconvolution in cases where a residual blur remains at the in-focus position, or when the true in-focus position falls in between two slices. Compared to classical deconvolution, the texture estimation is much better conditioned since the whole z-stack contributes to its estimation—hence, this process can be interpreted as a 2.5-D deconvolution operation.

C. Organization of the Paper

In the next section, we introduce our image formation model. After this, in Section III, we formulate the least-squares estimation problem, along with its iterative, gradient-based solution. Subsequently, we present a simple, computationally efficient Gaussian PSF model and compare it to accurate optical models (Section IV). The proposed algorithm is then demonstrated in Section V, where simulations and experimental results are shown. Finally, there are a number of practical issues raised by our technique that are investigated in Section VI, where we also discuss the influence of the PSF model's accuracy on the estimation results.

II. Image Formation in Brightfield Microscopy

Brightfield microscopes can be configured to use either transmitted or reflected light (called diascopic or episcopic configurations, respectively), depending on the specimen's optical properties. The transmission mode is appropriate for samples that are sufficiently transparent, whereas the reflection mode makes it possible to image opaque specimens. Consequently, we propose an image formation model that applies to both modalities. For an opaque specimen imaged in reflection, the object can be modeled as a 3-D surface, which leaves image formation unchanged. It turns out that this thin surface model is a valid approximation for the diascopic configuration as well, under the condition that the specimen is sufficiently thick, such that its surface alone appears in focus. It should be noted that in the latter case, this is a somewhat idealized representation that ignores potential out-of-focus contributions from beneath the surface of the specimen (note, furthermore, that these contributions vary depending the local thickness and density of the specimen). However, it is a necessary simplification to make our approach feasible, and our experiments indicate that it does not ensue in a loss of precision in the estimation results (see Section V). For this approximation to apply, transparent specimens should be significantly thicker than the system's depth of field, which is defined as

$$d = \frac{\lambda n_i}{\mathrm{NA}^2} \tag{1}$$

where λ is the illumination wavelength, n_i is the refractive index of the immersion medium, and NA designates the numerical aperture of the objective.

The theoretical developments below and in the following sections are stated for grayscale or single-channel data. The extension of these results as well as other issues pertaining to multichannel data are presented in Section III-E.

As mentioned, we express the sample $o(x, y, z)$ as a 3-D surface, described by a topography $p(x, y)$ onto which a texture $f(x, y)$ is mapped

$$o(x, y, z) = f(x, y)\delta\left(z - p(x, y)\right) \tag{2}$$

where the Dirac distribution represents the surface. The image formation and acquisition process is modeled as a convolution

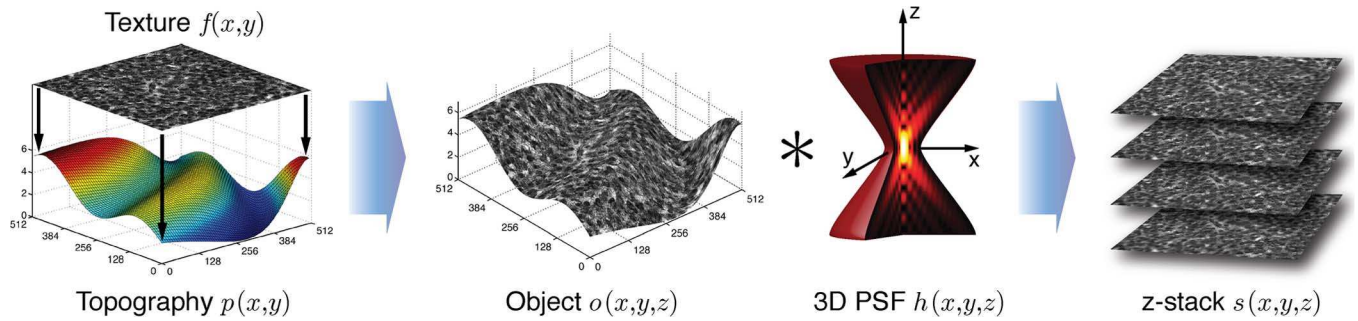


Fig. 1. Image formation: Illustration of the forward model. A texture mapped surface constitutes the object, which is convolved with the 3-D PSF of the system and sampled to obtain a z-stack acquisition. Phantom data for use in simulations was generated using this model.

between the object and the microscope's 3-D PSF $h(x,y,z)$, yielding the volume of acquisitions $\tilde{s}(x,y,z)$

$$\tilde{s}(x,y,z) = \int_{\mathbb{R}^2} f(u,v)h\left(x-u, y-v, z-p(u,v)\right)\mathrm{d}u\mathrm{d}v. \tag{3}$$

In practice, this is followed by sampling to generate discrete data (see Fig. 1).

III. Joint Texture and Topography Estimation

We propose to recover the topography and texture based on the quadratic cost function

$$J(f,p) = \int_{\mathbb{R}^3} \left(s(x,y,z) - \tilde{s}(x,y,z)\right)^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{4}$$

where s is the measured image stack. Given this formulation, it is easy to see that an estimate of the topography will depend on the texture, and vice versa. Therefore, we resort to an iterative two-phase optimization method that alternately updates the texture and the topography, where the texture and the topography at the τth iteration are denoted as $f^{(\tau)}$ and $p^{(\tau)}$, respectively.

Note that this solution can be interpreted in the context of maximum likelihood. Under the hypothesis of an additive white Gaussian noise component on the image formation model described by (3), the minimization of (4) is equivalent to maximizing the corresponding likelihood function. In that setting, the estimation is performed using the expectation-maximization algorithm (EM) [26], where the texture represents the quantity to be estimated, and the topography corresponds to the hidden state. Specifically, the E-step performs the update of the texture given the current estimate of the topography (i.e., the hidden state), whereas the M-step updates the topography. This noise model is a valid assumption in the context of brightfield microscopy, where Gaussian-distributed read-out noise is the primary source. Nonetheless, in state-of-the-art cameras, read-out noise can usually be deemed negligible when compared to signal intensity.

In the following sections, we present the texture and topography estimation phases of the algorithm, and discuss some implementation issues. Two methods for texture estimation are described: a general one based on a nonseparable PSF model, and a computationally more efficient one based on a simplified, separable PSF model.

A. Texture Estimation

1) Exact Model: Given an initial or current estimation $p^{(\tau)}$ of the topography, we estimate the texture by minimizing the cost function (4) with respect to f

$$f^{(\tau+1)} = \arg\min_f J\left(f, p^{(\tau)}\right). \tag{5}$$

This minimization cannot be obtained in closed form, and is achieved by performing a steepest descent on J with respect to the texture $f(x,y)$. The partial derivatives, easily obtained by applying the chain rule, are given by

$$\begin{aligned}\frac{\partial J}{\partial f(x,y)} &= -2\int_{\mathbb{R}^3} e(x,y,z)h\left(x-u, y-v, z-p^{(\tau)}(u,v)\right) \\ &\quad \times \mathrm{d}x\mathrm{d}y\mathrm{d}z\Big|_{u=x,v=y} \\ &= -2(e * h^T)\left(x, y, p^{(\tau)}(x,y)\right)\end{aligned} \tag{6}$$

where $\cdot^T$ stands for the spatial transpose; i.e., $h^T(x,y,z) = h(-x,-y,-z)$, and where $e(x,y,z) = s(x,y,z) - \tilde{s}(x,y,z)$. The gradient descent update step then consists of

$$f_{i+1}^{(\tau+1)}(x,y) = f_i^{(\tau+1)}(x,y) - \alpha\frac{\partial J}{\partial f(x,y)} \tag{7}$$

where $f_{i=0}^{(\tau+1)}(x,y)$ is initialized with $f^{(\tau)}(x,y)$. The factor $\alpha > 0$ controls the strength of the update; its optimal value is obtained by performing a line search, which leads to

$$\alpha_{\mathrm{opt}} = -\frac{\int_{\mathbb{R}^3} e(x,y,z)\omega(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z}{\int_{\mathbb{R}^3} \omega(x,y,z)^2\mathrm{d}x\mathrm{d}y\mathrm{d}z} \tag{8}$$

where

$$\omega(x,y,z) = \int_{\mathbb{R}^2} \frac{\partial J}{\partial f(u,v)}h\left(x-u, y-v, z-p(u,v)\right)\mathrm{d}u\mathrm{d}v. \tag{9}$$

Starting from $f^{(\tau)}$, several subiterations (index i) defined by (7) are performed to finally obtain a new estimate of the texture $f^{(\tau+1)}$.

2) Simplified, Separable Model: We now consider the above developments in conjunction with the simplest PSF model satisfactory for our image formation model, and show how this leads to an efficient and practical algorithm. The model

$$h(x,y,z) = \delta(x)\delta(y)h(z) \tag{10}$$

can be related to the hypothesis of in-focus information being correlated with intensity extrema along z [2]. Applying this, the forward model of (3) becomes

$$\begin{aligned}\tilde{s}(x,y,z) &= \int_{\mathbb{R}^2} f(u,v)\delta(x-u)\delta(y-v)h\left(z-p(u,v)\right)\mathrm{d}u\mathrm{d}v \\ &= f(x,y)h\left(z-p(x,y)\right)\end{aligned} \tag{11}$$

and the cost function accordingly simplifies to

$$J(f,p) = \int_{\mathbb{R}^3} \left(s(\mathbf{x}) - f(x,y)h\left(z-p(x,y)\right)\right)^2 \mathrm{d}\mathbf{x}, \tag{12}$$

where $\mathbf{x} = (x,y,z)$. The main point is that the global cost function can now be decomposed as

$$J(f,p) = \int_{\mathbb{R}^2} J(f,p)(x,y)\mathrm{d}x\mathrm{d}y \tag{13}$$

where

$$J(f,p)(x,y) = \int_{\mathbb{R}} \left(s(x,y,z) - f(x,y)h\left(z-p(x,y)\right)\right)^2 \mathrm{d}z \tag{14}$$

and, consequently, the minimization can be performed in a decoupled fashion around the current point (x,y). Specifically, the minimization of $J(f,p)(x,y)$ with respect to $f(x,y)$ yields

$$f(x,y) = \frac{\int_{\mathbb{R}} s(x,y,z)h\left(z-p(x,y)\right)\mathrm{d}z}{\int_{\mathbb{R}} h\left(z-p(x,y)\right)^2 \mathrm{d}z}. \tag{15}$$

This expression is essentially a weighted interpolation among the slices of the z-stack around the position indicated by the topography. The above development can also be considered in the context of nearest-neighbor deconvolution, where analogous assumptions concerning the decoupling of the image formation model are made [27]. The key advantage of this simplified algorithm is the stability it lends to the overall estimation when used in conjunction with the topography estimation algorithm discussed below. This aspect, together with the use of the gradient-based estimation algorithm to perform deconvolution of the texture, is detailed at the end of this section and in the discussion (Section VI).

B. Topography Estimation

In a similar fashion, the topography can be updated using the latest texture estimate $f^{(\tau+1)}$. To this end, we minimize the cost function with respect to the topography

$$p^{(\tau+1)} = \arg\min_p J\left(f^{(\tau+1)}, p\right). \tag{16}$$

The partial derivatives are also obtained using the chain rule as

$$\begin{aligned}\frac{\partial J}{\partial p(x,y)} &= 2f^{(\tau+1)}(u,v) \\ &\quad \times \int_{\mathbb{R}^3} e(x,y,z)h_z\left(x-u, y-v, z-p(u,v)\right) \\ &\quad \times \mathrm{d}x\mathrm{d}y\mathrm{d}z\Big|_{u=x,v=y} \\ &= 2f^{(\tau+1)}(x,y)\left(e * h_z^T\right)\left(x,y,p(x,y)\right)\end{aligned} \tag{17}$$

where h_z stands for the partial derivative of the PSF with respect to z. Then, the gradient descent update step for the topography is

$$p_{i+1}^{(\tau+1)}(x,y) = p_i^{(\tau+1)}(x,y) - \alpha\frac{\partial J}{\partial p(x,y)}. \tag{18}$$

C. Coarse-to-Fine Optimization

The direct dependence of the topography update upon the current texture estimate $f^{(\tau+1)}$ can perturb the algorithm in its initial stages when strong global adjustments are being made, and make it converge towards an erroneous local optimum. To deal with this potential instability, we propose to perform the topography estimation in a coarse-to-fine framework, which has the effect of imposing regularity and robustness to the topography update, as the following development will show (see also Fig. 3).

In our implementation, we have chosen to represent the topography at resolution m in a shift-invariant basis generated by symmetric B-splines of order n dilated by $m \in \mathbb{N}^+$

$$p_m(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} c_m[\mathbf{k}]\beta^n\left(\frac{\mathbf{x}}{m} - \mathbf{k}\right) \tag{19}$$

with $\mathbf{x} = (x,y)$ [28], and where the 2-D basis functions are formed from the tensor product of their 1-D counterparts. The minimization problem (16) then amounts to finding the "optimal" B-spline coefficients $c_m[\mathbf{k}]$. The corresponding partial derivatives of the criterion are

$$\begin{aligned}\frac{\partial J}{\partial c_m[k,l]} &= 2\int_{\mathbb{R}^2} \beta^n\left(\frac{u}{m}-k, \frac{v}{m}-l\right) f(u,v) \\ &\quad \times \Bigg(\int_{\mathbb{R}^3} e(x,y,z)h_z\left(x-u, y-v, z-p_m(u,v)\right) \\ &\quad \times \mathrm{d}x\mathrm{d}y\mathrm{d}z\Bigg)\mathrm{d}u\mathrm{d}v \\ &= \left((\beta_m^n)^T * \frac{\partial J}{\partial p}\right)(mk, ml) \\ &= \left(\beta_m^n * \frac{\partial J}{\partial p}\right)(mk, ml)\end{aligned} \tag{20}$$

where $\beta_m^n(x) = \beta^n(\frac{x}{m})$. Thus, the gradient on the coefficients is essentially a smoothened and downsampled version of $\partial J/\partial p$, which can be implemented according to the block diagram shown in Fig. 2(a), where the smoothing filter

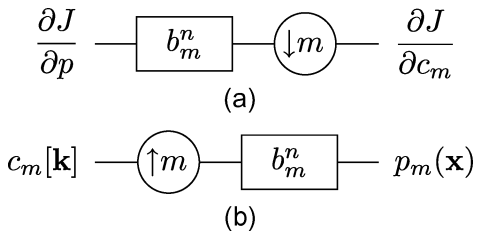


Fig. 2. Downscaling and upscaling steps in the coarse-to-fine topography estimation algorithm. (a) B-spline gradient coarsening. (b) Fine-scale interpolation of the B-spline topography model.

$b_m^n[\mathbf{k}] = \beta^n(\mathbf{k}/m)$ is the sampled version of the basis function in (20). Accordingly, the coefficient update at scale m becomes

$$c_m^{(i+1)}[k,l] = c_m^{(i)}[k,l] - \alpha\frac{\partial J}{\partial c_m}. \tag{21}$$

The fine scale sampling of the topography (required at each iteration) is finally obtained by upsampling c_m and postfiltering the result, as summarized in Fig. 2(b).

Due to the dependence of h on p in (4), the line search in (20) cannot be achieved in closed form. Here, it is desirable to apply an iterative minimization scheme that does not depend on a parameter regulating the update step. Unfortunately, $J(f, p-\alpha J_p)$ is not necessarily convex, and, thus, a bounded optimization approach (i.e., via a quadratic upper bound) cannot be deployed. Instead, we resort to inverse parabolic interpolation, which converges reasonably fast and can by easily controlled via Brent's algorithm [29]. In practice, $J(f, p-\alpha J_p)$ is smooth and has a unique minimum, and, thus, this scheme proves to be very efficient, requiring only a minimal number of iterations.

D. Implementation

The joint estimation of the texture and topography from a volume of acquisitions is achieved by alternating iterations of the methods outlined above. Starting from an initialization of the texture, we iterate between estimating the topography at scale m and refining the texture. Upon convergence at that scale, the algorithm switches to a finer scale ($m' = m/2$ in the dyadic case) and replicates this process. This is repeated until convergence at scale $m = 1$ is reached (see Fig. 3). In order to make this procedure as computationally efficient as possible, the texture estimation is performed using the simplified model at all scales except $m = 1$. At that stage, the higher precision and especially the deconvolution capabilities of the exact, gradient-based method are applied to generate the final fusion image. For the description of the topography, we used cubic B-splines ($n = 3$).

E. Refinements for Processing Color Data

We propose to process multichannel data sets via a straightforward extension of the algorithm above. While it would be possible to formulate our method to simultaneously operate on multiple channels, we believe that doing so during the estimation of the topography does not provide sufficient advantages when weighted against the resulting increase in computational cost, and, to a lesser degree, the influence of further degrees of freedom in the PSF model on the estimation result. Consequently, we chose to perform the estimation on grayscale data, appropriately converted from the color input. Dyes binding to a particular structure such as cell membranes or nuclei are frequently employed in biology; as a result, large specimens tend to exhibit a strongly dominant color after preparation. Taking this bias into account by performing the grayscale conversion using appropriate weights (obtained, for example, using principal component analysis) can lead to markedly improved results, as shown in [20].

The only modification to the single-channel algorithm then consists of extending the last texture estimation step (after the topography has converged at scale $m = 1$) to operate on all channels in parallel, i.e., by applying each step of (7) individually to each channel.

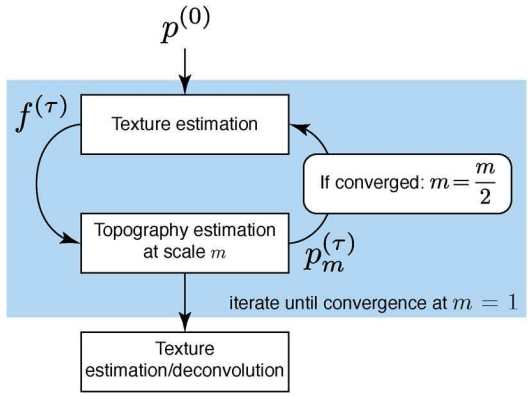


Fig. 3. Schematic representation of the optimization algorithm. Starting from an estimate of the topography $p_m^{(\tau)}$ at scale m, the method alternates between texture and topography estimation until convergence at that scale, after which the process is repeated at successively finer scales.

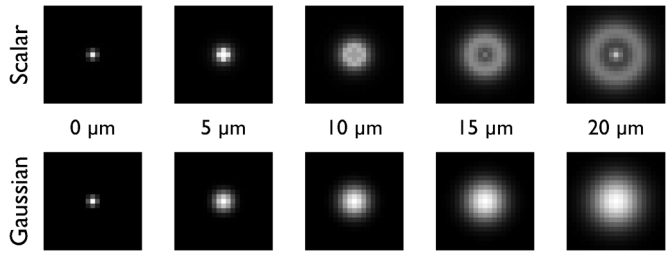


Fig. 4. Comparison between a scalar PSF model [30] and its Gaussian approximation for a 10×, 0.3 NA objective for a CCD with 6.7×6.7 μm^2 pixels. Different focal settings are shown. For a given amount of defocus, the same dynamic range is used for illustration purposes. The corresponding parameters of the Gaussian model (22) are $\eta_0 = 0.670$ and $\eta_1 = 0.734$. A 5-μm increase in defocus corresponds to a unitary increase in z.

IV. Theoretical PSF Model

The gradient-based algorithms presented above are not bound to any particular PSF model. In principle, it is therefore possible to use a highly accurate theoretical model that is potentially axially shift-variant and takes into account spherical aberrations, such as the scalar model proposed by Gibson and Lanni [30]. However, given that the problem is reasonably well posed, and for the sake of decreasing computational complexity, we argue for the use of a Gaussian approximation of the microscope's PSF. Since PSF models are usually formulated for a single wavelength, they should be integrated over the spectrum of the light source used. In modern brightfield microscopes, the light source is well balanced over the visible spectrum, which

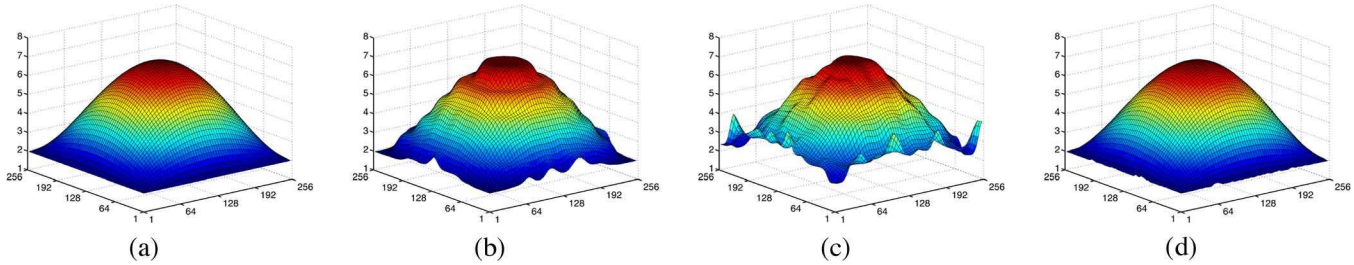


Fig. 5. Simulation results for topography estimation using the variance-based, CWT-based, and model-based algorithms. (a) Ground truth; (b) variance method; (c) CWT-EDF; (d) model based.

TABLE I
QUANTITATIVE PERFORMANCE COMPARISON FOR TOPOGRAPHY AND TEXTURE ESTIMATION FOR VARIOUS EDF METHODS

Method	Texture SNR	Topography SNR
Variance	32.07 dB	26.94 dB
CWT-EDF	32.41 dB	35.89 dB
Model-based	54.22 dB	53.36 dB

means that the PSF needs to be uniformly integrated over the interval of visible wavelengths. The result of this integration is essentially a smoothing effect; in fact, the xy-sections of a white-light PSF for a low magnification objective closely resemble Gaussians.

At each iteration of the topography estimation, the PSF needs to be sampled over a given support for each point of the topography. It is therefore highly advantageous to use a Gaussian PSF model that is computationally much less complex to evaluate than an optically accurate PSF that involves numerical integration. As illustrated in Fig. 4, a good estimate of the PSF can be obtained by fitting a model of the form

$$h(x,y,z) = \frac{1}{2\pi\left(\eta_0 + \eta_1|z|\right)^2} \exp\left(-\frac{(x^2+y^2)}{2\left(\eta_0+\eta_1|z|\right)^2}\right) \tag{22}$$

to a theoretical PSF evaluated according to the optical properties of the microscope used, or to an experimentally measured PSF.

A key property of brightfield microscopes consists in Köhler illumination, which is designed to uniformly illuminate the sample independently of the position of the focus [31]. To account for this, the PSF needs to be normalized such that the intensity within each xy-section is equal to unity. This is not necessarily the case when the Gaussian PSF model is sampled over the integers. In particular, when $\sigma = \eta_0 + \eta_1|z|$ tends towards zero, $\sum_{i,j\in Z^2} h(i,j,k_0)$ goes to infinity, which is physically nonplausible. In order to avoid numerical blow up when $\sigma < 0.8$, we further normalize h by

$$\frac{1}{2\pi\sigma^2}\left(1+2e^{\frac{-1}{2\sigma^2}}\left(1+e^{\frac{-3}{2\sigma^2}}\right)\right)^2 \tag{23}$$

which is the sum over the essential support of the Gaussian.

V. RESULTS

We now present estimation results obtained from several data sets, first in simulation and then on acquisitions of biological specimens. Where appropriate, we compare our results to two reference methods for validation: one based on a local variance criterion and the other on a wavelet transform. As mentioned earlier, the former constitutes the classical approach to EDF. It has a low computational complexity and generally yields results of reasonable quality, which explains its ongoing popularity. Specifically, we compute the variance in a 5×5 window, and filter the resulting topography using a Gaussian kernel with $\sigma = 2$ to enforce local smoothness.

Among state-of-the-art wavelet-based methods, we chose the complex wavelet-based EDF algorithm of [20], co-developed by two of us (from this point on, we will refer to it as the CWT-EDF method), which was shown to outperform competing methods in many instances. This choice is further justified by it being the only recent method that aims at providing a coherent estimate of the topography by imposing local smoothness and consistency constraints.

A. Performance Comparison in Simulation

In order to have a ground truth at our disposal for the comparison and qualitative assessment of the results produced by the above cited methods, we performed a first series of experiments using a simulated set of acquisitions. A z-stack of images was generated from a given topography and texture using the object and image formation model, i.e., by applying the nonstationary convolution of (9) (see Fig. 1).

In Fig. 5, we compare the topography estimation results obtained on a simulated stack containing eight images of a dome-shaped object onto which a 256×256 texture was mapped. The latter was derived from a brightfield image of a histological slide containing murine liver tissue. Our implementation of the CWT-EDF method permits to control the smoothness of the estimated topography; here, we empirically optimize this step to maximize the quality of the topography. As expected, the model-based algorithm correctly recovers the topography, while the two reference methods only approximately approach the correct shape. The quality of the estimation results, in terms of the signal-to-noise ratio $\mathrm{SNR} = 10\log_{10}(||\mathrm{signal}||^2/||\mathrm{noise}||^2)$, is documented in Table I. The continuous-domain formulation of the object model in our approach not only leads to a better estimation of the topography, but also deconvolves the texture at locations between slices, which accounts for the gain in estimation quality of both signals. Care was taken to generate phantom data containing in-focus information for all (x,y)-positions; this explains the relatively good quality of the texture estimation results for all methods. We initialized our method

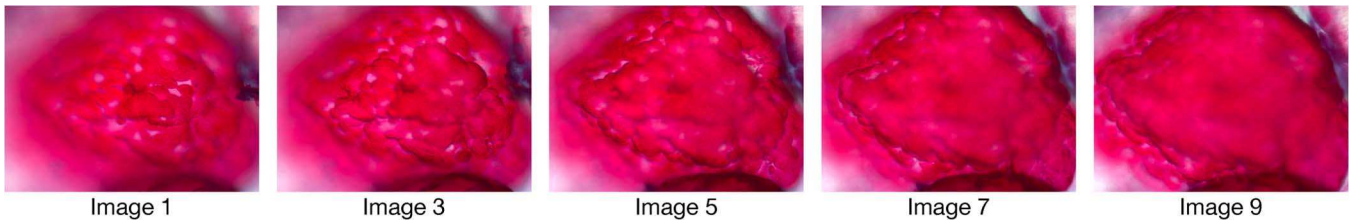
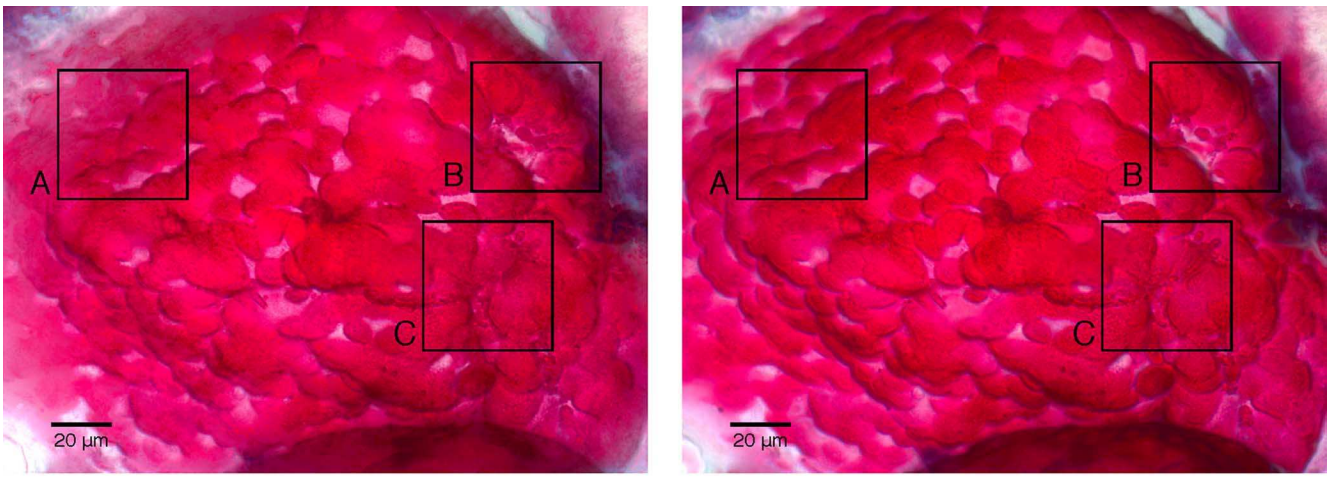


Fig. 6. Individual image slices from the Peyer's patches stack. Distinct in-focus areas are clearly visible.



(a) (b)

Fig. 7. Comparison of texture estimation against a wavelet-based method. The specimen is a sample of Peyer's patches in the mouse intestine. (a) Texture obtained using the complex wavelet-based algorithm described in [20], with maximal quality settings. (b) Texture resulting from our joint estimation algorithm. The coarse-to-fine estimation approach ensures local smoothness, which prevents mis-selection artifacts that are typical of approaches that rely on a high-frequency criterion.

with a flat topography, and performed iterations at the scales $m = \{8, 4, 2, 1\}$. In general (also for the results on experimental data), the topography estimation typically requires around five subiterations, whereas the texture requires only 2–3 steps of the gradient update to reach convergence.

B. Results on Experimental Data

We now demonstrate the potential of our method on two sets of experimental acquisitions, one for a transmission setup and the other for a configuration using reflected light. The first specimen consists of a section of mouse intestine containing Peyer's patches, stained using a membrane-labeling dye. The sample was imaged using a Zeiss Plan–Neofluar 40×, 0.75 NA objective in air immersion, with a Jenoptik ProgRes CCD camera with $4.3 \times 4.3\ \mu\text{m}^2$ pixels. 16 optical sections in steps of 3 μm were acquired. The acquisitions have a resolution of 1996×1450 pixels, several of which are shown in Fig. 6.

The topography resulting from the joint estimation, shown in Fig. 9, clearly reveals the heap-like structure of the sample. This result was obtained by initializing the algorithm with the variance method, and by performing iterations at the scales $m = \{16, 8, 4, 2, 1\}$. In Fig. 7, we compare the texture obtained with CWT-EDF and our method. At a first glance the results look similar, but the inspection of some details (see highlights of Fig. 7 in Fig. 8) reveals a marked improvement with our method. Notably, there is a complete absence of color artifacts typical of wavelet-based methods. Also, the deconvolving action of the texture estimation restitutes several in-focus regions missing from the acquisitions. This is most noticeable around

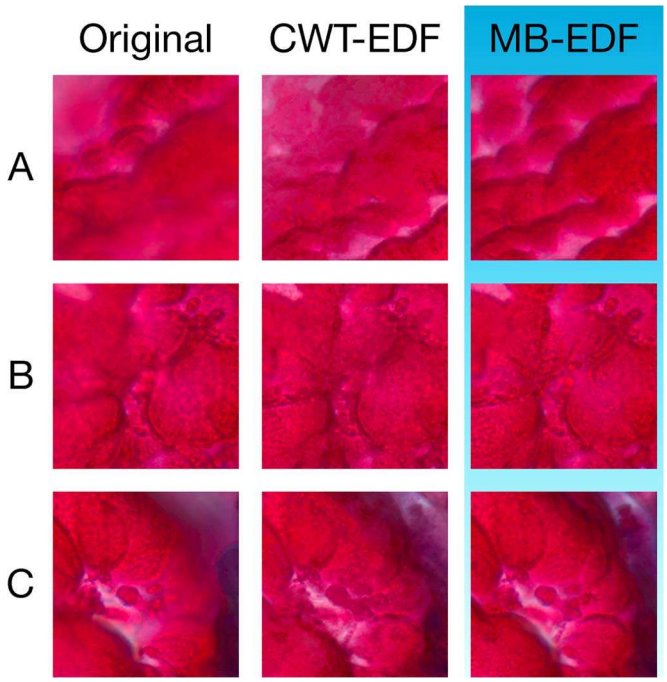


Fig. 8. Details of the texture estimation comparison from Fig. 7.

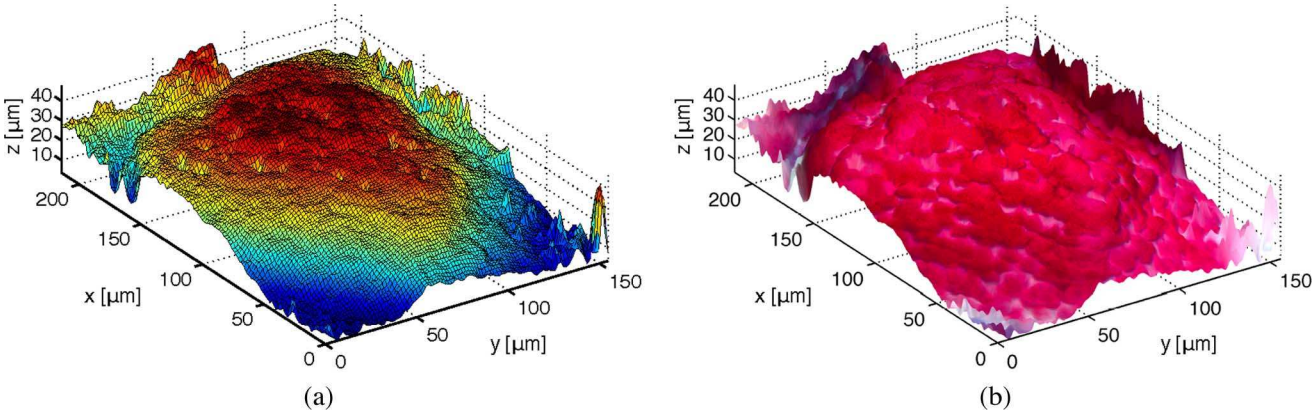


Fig. 9. Texture mapping of the estimated texture from the Peyer's patches stack onto the corresponding estimated topography.

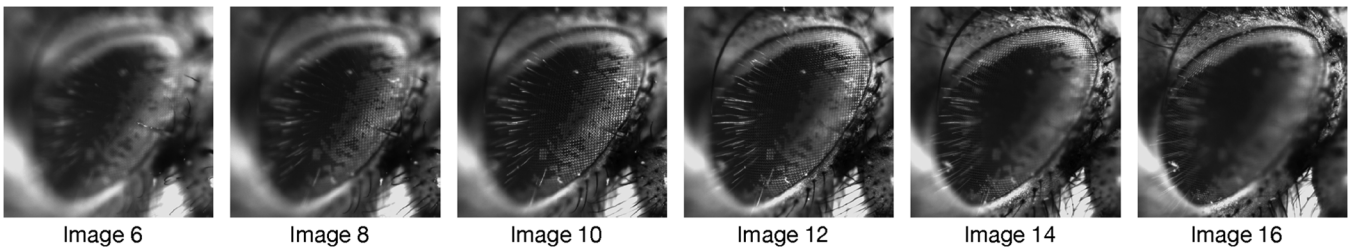


Fig. 10. Individual images from the fly eye stack. Distinct in-focus areas are clearly visible.

the cell boundaries, which appear much sharper (compare also with [24]).

Some artifacts are present in the topography, especially around the borders where structures are visible but out of focus in the acquisitions. These artifacts are direct effects of the convolution between the topography gradient and the basis functions during the coarse-scale iterations. As the algorithm moves to finer scales, a lack of in-focus information prevents further adjustments in the concerned regions.

Our second test specimen consisted of a common house fly imaged in reflection using a Zeiss Achroplan $4\times$, 0.1 NA objective in air immersion, with a QImaging Retiga CCD camera with $6.45 \times 6.45\ \mu\text{m}^2$ pixels. The acquisitions were cropped to a 1000×1000 pixel region around the fly's eye. The z-stack consists of 32 images with a z-step of 20 μm.

The topography resulting from the joint estimation, shown in Fig. 11, reveals the homogeneous structure of the eye. The peaks present in the topography correspond to small hairs, which are also visible in the texture (Fig. 12). This result was obtained by initializing the algorithm with the variance method, and by performing iterations at the scales $m = \{32, 16, 8, 4, 2, 1\}$.

As these results attest to, recovering an accurate representation of the topography not only facilitates the analysis of thick specimens, but also leads to marked improvements in the estimated texture. Furthermore, the experiments confirm that the 2.5-D deconvolution leads to a sharp estimate of the texture when in-focus information is missing from the acquired data, or when it falls between acquisitions.

VI. Discussion

There are two key factors that can strongly impact on the performance of our method. We begin this section by examining

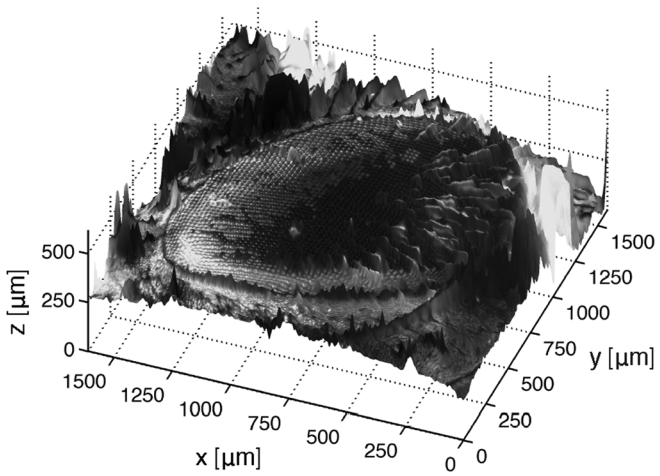


Fig. 11. Texture mapping of the estimated texture from the fly stack onto the corresponding estimated topography.

the influence of the parameters controlling our Gaussian PSF model on the quality of the results, and subsequently discuss the computational performance of our approach in comparison to competing methods, in connection to which we briefly comment on possible initializations of the algorithm.

A. Sensitivity to the PSF Model Parameters

When dealing with experimental acquisitions, it frequently occurs that the PSF can only be approximately determined or measured. However, since the optimization of the texture and topography is generally well posed (this is largely dependent upon the number of acquisitions in the z-stack), the reconstruction algorithm can be expected to be relatively insensitive to the

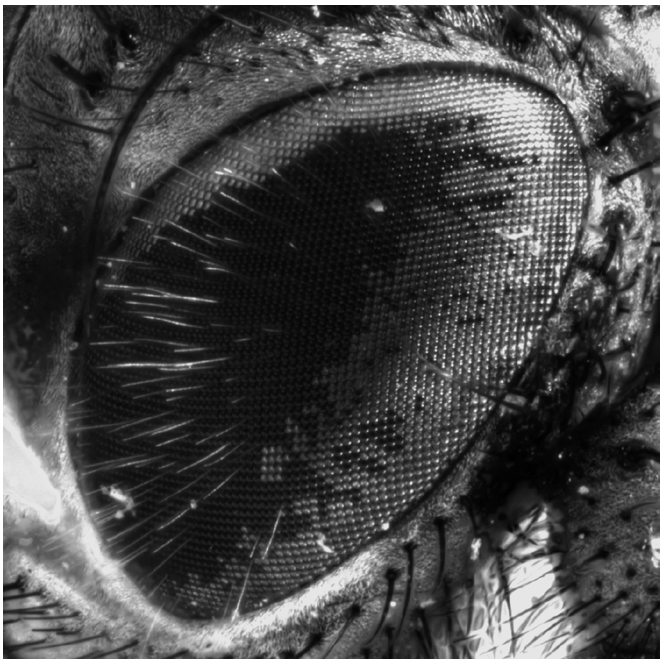


Fig. 12. Fusion image estimated from a z-stack of a fly's eye.

parameterization of the PSF. This is indeed the case; to a certain extent, only the in-focus PSF $h(x, y, 0)$ needs to be precise. Otherwise, due to our assumptions on the specimen, the deconvolution of the fusion image is either too weak, or results in out-of-focus information being falsely reallocated to the texture. The Gaussian PSF approximation holds very well at the in-focus position and, consequently, does not significantly affect the results.

In theory, an erroneous PSF can have an effect on the gain of the texture, i.e., lead to a scaling in intensity. This can, in turn, affect the topography estimation, and is especially problematic when operating on color data, where it can lead to false color artifacts. In our algorithm, however, due to the use of the separable texture estimation step throughout the joint estimation, this effect is avoided altogether. Indeed, when considering the texture estimation as an interpolation, the possible values for the texture are constrained to the dynamic range of the input volume.

Simulations using data generated from the model showed that an error up to 30% on the η_1 parameter in the PSF model (22) leads to the recovery of the topography without any significant loss in accuracy.

B. Computational Aspects

Although, due to its sophistication, our approach incurs a higher computational cost than wavelet-based EDF methods (which require a forward and inverse transform in combination with what are essentially point-wise operations), this cost is largely compensated for by the increased quality of the results. Nevertheless, the cost can by minimized by employing the separable texture estimation step whenever deconvolution of the texture is not necessary (note that it cannot be diminished for the topography estimation). In practice, it therefore makes sense to reduce the total number of iterations by computing an initial estimate of the topography via the variance-based method. Given such an initialization, many coarse-scale iterations can be avoided. Using this scenario, our algorithm requires 64 s on a $512 \times 512 \times 10$ stack, whereas the CWT-EDF algorithm requires 16 s. Further means of acceleration could be a topic for future research.

VII. Conclusion

We have presented a new approach to extended depth of field that constitutes a significant improvement over current methods in terms of the quality of the results. By stating the problem as a joint estimation of the sample's texture and topography, we are able to successfully recover a continuous and characteristic representation of the topography, devoid of the discretization artifacts present in existing methods. Additionally, the joint estimation model acts as a 2.5-D deconvolution operation for the resulting fusion image, which yields sharper textures, especially when the true in-focus information lies between acquisition planes. We demonstrated the validity of the approach on simulated data sets that were generated with a known topography, and obtained promising results on experimental acquisitions; improvements due to the continuous topography and texture deconvolution are clearly visible.

Since the recovered topography is continuous, it would be feasible to reconstruct a stereoscopic representation to further enhance the 3-D visualization of the specimen, as was suggested in [13]. While we presented our method in the context of bright-field microscopy where it is likely to find its most immediate application, it is clear that, due to the relatively low sensitivity to accuracy of the PSF, the method is applicable to a wider range of optical systems with a limited depth of field. For example, a 3-D vision system combining shallow depth of field optics with the proposed method could be envisaged. In microscopy still, the method raises the possibility of performing profilometry, where it should be interesting to see with what accuracy the surface profile of an object can be quantitatively determined.

Foremost, our results illustrate the potential of a model-based approach in the context of EDF. Our solution has no pretense of optimality but should rather be viewed as a proof of concept. There likely is room for exploration and improvement, i.e., by applying a different regularization criterion, or by devising other algorithms that optimize the proposed cost function. A Java implementation (ImageJ plugin) of this work is available at http://bigwww.epfl.ch/demo/edf/.

Acknowledgment

The authors would like to thank Dr. N. Garin for providing the Peyer's patches acquisitions, and Dr. P. Thévenaz for the fly dataset.

References

[1] C. J. R. Sheppard, D. K. Hamilton, and I. J. Cox, "Optical microscopy with extended depth of field," *Proc. Roy. Soc. Lond. A*, vol. 387, pp. 171–186, 1983.

[2] R. J. Pieper and A. Korpel, "Image processing for extended depth of field," *Appl. Opt.*, vol. 22, no. 10, pp. 1449–1453, 1983.

[3] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[4] V. S. Petrović and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Trans. Image Process.*, vol. 13, no. 2, pp. 228–237, Feb. 2004.

[5] S. C. Tucker, W. T. Cathey, and E. R. Dowski, "Extended depth of field and aberration control for inexpensive digital microscope systems," *Opt. Expr.*, vol. 4, no. 11, pp. 467–474, 1999.
[6] A. G. Valdecasas, D. Marshall, J. M. Becerra, and J. J. Terrero, "On the extended depth of focus algorithms for bright field microscopy," *Micron*, vol. 32, pp. 559–569, 2001.
[7] D. McLachlan, "Extreme focal depth in microscopy," *Appl. Opt.*, vol. 3, no. 9, pp. 1009–1013, 1964.
[8] J. F. Burke, G. Indebetouw, G. Nomarski, and G. W. Stroke, "White-light three dimensional microscopy using multiple-image storing and decoding," *Nature*, vol. 231, pp. 303–306, 1971.
[9] G. Indebetouw and H. Bai, "Imaging with Fresnel zone pupil masks: Extended depth of field," *Appl. Opt.*, vol. 23, no. 23, pp. 4299–4302, 1984.
[10] E. R. Dowski and W. T. Cathey, "Extended depth of field through wave-front coding," *Appl. Opt.*, vol. 34, no. 11, pp. 1859–1866, 1995.
[11] A. Castro and J. Ojeda-Castañeda, "Asymmetric phase masks for extended depth of field," *Appl. Opt.*, vol. 43, no. 17, pp. 3474–3479, 2004.
[12] E. J. Botcherby, R. Juškaitis, and T. Wilson, "Scanning two photon fluorescence microscopy with extended depth of field," *Opt. Commun.*, vol. 268, pp. 253–260, 2006.
[13] S. A. Sugimoto and Y. Ichioka, "Digital composition of images with increased depth of focus considering depth information," *Appl. Opt.*, vol. 24, no. 14, pp. 2076–2080, 1985.
[14] K. Itoh, A. Hayashi, and Y. Ichioka, "Digitized optical microscopy with extended depth of field," *Appl. Opt.*, vol. 28, no. 15, pp. 3487–3493, 1989.
[15] K. Kaneda, S. Ishida, A. Ishida, and E. Nakamae, "Image processing and synthesis for extended depth of field of optical microscopes," *Vis. Comput.*, vol. 8, no. 5–6, pp. 351–360, 1992.
[16] M. Antunes, M. Trachtenberg, G. Thomas, and T. Shoa, "All-in-focus imaging using a series of images on different focal planes," *Lecture Notes Comput. Sci.*, vol. 3656, pp. 174–181, 2005.
[17] E. H. Adelson, "Depth-of-focus imaging process," U.S. Patent 4 661 986, 1987.
[18] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *Proc. 4th Int. Conf. Computer Vision*, 1993, pp. 173–182, IEEE.
[19] Z. Liu, K. Tsukada, K. Hanasaki, Y. K. Ho, and Y. P. Dai, "Image fusion by using steerable pyramid," *Pattern Recognit. Lett.*, vol. 22, pp. 929–939, 2001.
[20] B. Forster, D. Van De Ville, J. Berent, D. Sage, and M. Unser, "Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images," *Microsc. Res. Tech.*, vol. 65, pp. 33–42, 2004.
[21] A. P. Bradley and P. C. Bamford, "A one-pass extended depth of field algorithm based on the over-complete discrete wavelet transform," in *Proc. Image and Vision Computing*, 2004, pp. 279–284.
[22] S. Li, J. T. Kwok, I. W. Tsang, and Y. Wang, "Fusing images with different focuses using support vector machines," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1555–1561, Dec. 2004.
[23] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Image fusion using a new framework for complex wavelet transforms," in *Proc. Int. Conf. Image Processing*, 2005, pp. 1338–1341.
[24] L. Tessens, A. Ledda, A. Pižurica, and W. Philips, "Extending the depth of field in microscopy through curvelet-based frequency-adaptive image fusion," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2007, pp. I861–I864.
[25] F. Aguet, D. Van De Ville, and M. Unser, "Joint texture and topography estimation for extended depth of field in brightfield microscopy," in *Proc. 3rd IEEE Int. Symp. Biomedical Imaging: From Nano to Macro—Proceedings*, 2006, pp. 778–781.
[26] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.
[27] D. A. Agard, Y. Hiraoka, P. Shaw, and J. W. Sedat, "Fluorescence microscopy in three dimensions," *Methods Cell Biol.*, vol. 30, pp. 353–377, 1989.
[28] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I—Theory," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.
[29] R. P. Brent, *Algorithms for Minimization Without Derivatives*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[30] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Amer. A*, vol. 8, no. 10, pp. 1601–1613, 1991.
[31] A. Köhler, "Ein neues Beleuchtungsverfahren für mikrophotographische Zwecke," *Zeitschrift für wissenschaftliche Mikroskopie*, vol. 10, pp. 433–440, 1893.

François Aguet (S'06) received the M.S. degree in computer science in 2004 from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He is currently pursuing the Ph.D. degree at the Biomedical Imaging Group, EPFL.

His research interests include inverse problems in optics, feature detection, and biomedical image processing.

Dimitri Van De Ville (M'02) received the M.S. degree in engineering and computer sciences from Ghent University, Ghent, Belgium, in 1998, and the Ph.D. degree from the Medical Signal and Image Processing Group, Ghent University.

He obtained a grant as Research Assistant with the Fund for Scientific Research Flanders Belgium (FWO). In 2002, he joined Prof. M. Unser's Biomedical Imaging Group at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, where he continues his research today. In December 2005, he became responsible for the Signal Processing Antenna at the University Hospital of Geneva, Geneva, Switzerland, as part of the Centre d'Imagerie Biomédicale (CIBM). His research interests include wavelets, statistical analysis, multidimensional splines, and applications in biomedical imaging, such as functional magnetic resonance imaging, spectroscopy, electro-encephalography, and microscopy.

Dr. Van De Ville serves as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING (since February 2006) and was previously an Associate Editor for IEEE SIGNAL PROCESSING LETTERS (2004–2006). Since 2003, he has also been an Editor and Webmaster of *The Wavelet Digest*, a noncommercial electronic newsletter for the wavelet community with more than 22 000 subscribers. He organized the "Wavelets and Applications" semester of EPFL's Bernoulli Center, together with M. Vetterli and M. Unser, and WavE2006, its associated international conference (July 2006). He was also General Co-Chair of the Wavelets XII conference in San Diego, CA, held in August 2007.

Michael Unser (M'89–SM'94–F'99) received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 1981 and 1984, respectively.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. He is now Professor and Director of the Biomedical Imaging Group, EPFL. His main research topics are biomedical image processing, splines, and wavelets. He is the author of over 120 published journal papers in these areas.

Dr. Unser has been actively involved with the IEEE TRANSACTIONS ON MEDICAL IMAGING, holding the positions of Associate Editor (1999–2002, 2006–present), member of steering committee, and Associate Editor-in-Chief (2003–2005). He has acted as an Associate Editor or member of the editorial board for eight more international journals, including the *IEEE Signal Processing Magazine*, the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1995), and the IEEE SIGNAL PROCESSING LETTERS (1994–1998). He organized the first IEEE International Symposium on Biomedical Imaging (ISBI 2002). He currently chairs the technical committee of the IEEE-SP Society on Bio Imaging and Signal Processing (BISP), and well as the ISBI steering committee. He is the recipient of three Best Paper Awards from the IEEE Signal Processing Society.