Efficient stripe artefact removal by a variational method: application to light-sheet microscopy, FIB-SEM and remote sensing images

NIKLAS ROTTMAYER[1,*], CLAUDIA REDENBACH[1] AND FLORIAN O. FAHRBACH[2]

[1] *Mathematics Department, RPTU Kaiserslautern-Landau, Gottlieb-Daimler-Straße, 67663 Kaiserslautern, Germany*
[2] *Leica Microsystems CMS GmbH, Am Friedensplatz 3, 68165 Mannheim, Germany*

Abstract: Light-sheet fluorescence microscopy (LSFM) is used to capture volume images of biological specimens. It offers high contrast deep inside densely fluorescence labelled samples, fast acquisition speed and minimal harmful effects on the sample. However, the resulting images often show strong stripe artifacts originating from light-matter interactions. We propose a robust variational method suitable for removing stripes which outperforms existing methods and offers flexibility through two adjustable parameters. This tool is widely applicable to improve visual quality as well as facilitate downstream processing and analysis of images acquired on systems that do not provide hardware-based destriping methods. An evaluation of methods is performed on LSFM, FIB-SEM and remote sensing data, supplemented by synthetic LSFM images. The latter is obtained by simulating the imaging process on virtual samples.

1. Introduction

The observation by microscopy plays a fundamental role in understanding multicellular life and biological processes. Since the beginning of this century, light-sheet (fluorescence) microscopy (LSFM) has gained popularity in 3D-imaging of biological specimens. It provides high resolution and fast acquisition speeds [1] while minimizing harmful exposure to light [2]. However, the acquired images are often accompanied by strong stripe artifacts caused by light absorption and scattering, see Figure 1 for reference. The corruptions can occur in large parts of the imaged volume and obscure underlying structures, thus impairing their appearance and complicating further data analysis. There is a large quantity of articles concerned with stripe artifacts in LSFM [3–9]. In a recent paper, Ricci and colleagues [10] present a detailed summary of previous research on removing stripes from LSFM images. They discuss hardware solutions to prevent stripe formation, algorithms to remove stripes through post-processing and hybrid methods which combine ideas from both worlds.

At first glance, the prevention of artifacts is clearly the preferential choice. However, this requires modifications of the imaging setup which are costly and often require entirely new microscopes. They may also limit the acquisition speed, the light-efficiency or be incompatible with certain advantageous image acquisition modes such as line-confocal acquisition [11–13]. In comparison, algorithmic post-processing is cheap and requires only computational power. Furthermore, corrupted images may already exist and specimens cannot be re-imaged, e.g., due to degradation and aging. For these reasons it is desirable to have powerful post-processing tools available for stripe removal.

Stripe artifacts are not unique to LSFM. They are also common in various other imaging techniques such as atomic force microscopy (AFM) [14], focused ion beam scanning electron microscopy (FIB-SEM) [15] or remote sensing [16, 17], see Figure 2 for examples. Although having different causes, stripes are very similar in shape showing pronounced elongated structures in a common direction with mostly similar thickness of only few pixels. Their shape deviates strongly from underlying structures and presents a feature that can be used for detection and

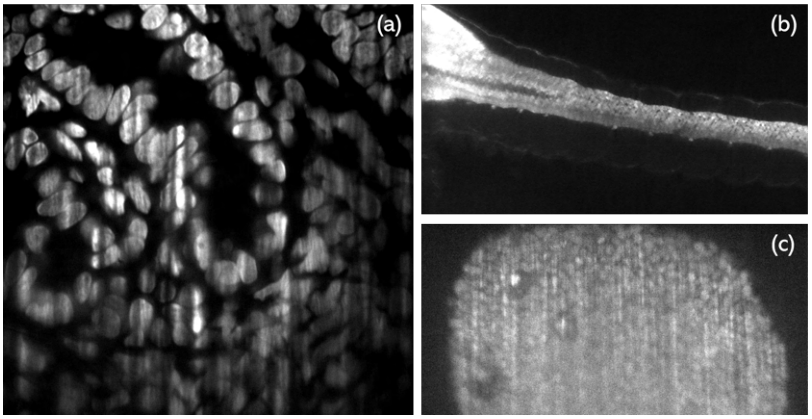

Fig. 1. LSFM images corrupted by stripe artifacts. Cultivated cell organoid of mouse intestine cells (a), tail of Zebrafish larva (b) and cluster of HeLa cells (c).

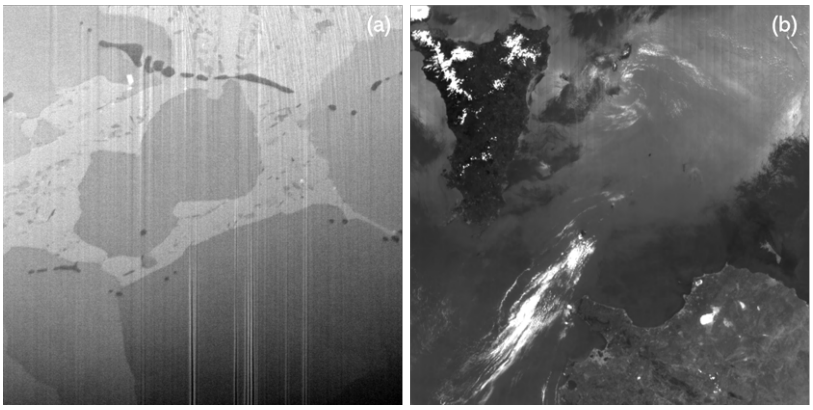

Fig. 2. Stripe artifacts in a FIB-SEM image slice of a tin bronze cast alloy (a) and Terra MODIS data (b).

removal. In LSFM on the other hand, the appearance of stripes is much more diverse. While stripes are still aligned along a common direction, their length, thickness and intensity can vary drastically depending on the imaged structures. For example, a single stripe can be split into smaller separated segments of lengths equal to that of underlying structures, see Figure 1. Additionally, stripes display a much larger range of thickness and depth such that individual stripes can even appear in multiple slices of 3D volume images. This drastically complicates the correct detection and removal of such artifacts, especially since stripe removal methods heavily rely on the distinctive shape differences to image structures.

In this paper, we discuss the two major categories of stripe removal methods in research, Fourier filtering and variational methods. We explain the functionality of Fourier filtering methods on the example of the multi-directional stripe remover (MDSR) [6] and present a modification to increase its performance. For the variational methods we propose a model which is based on the work of Liu and colleagues [16] and performs exceptionally well on all data tested. Furthermore, we suggest possible adaptations to make it applicable to a broad range of stripe removal settings. Afterwards, we evaluate and compare the performance of the methods using real and synthetic image data. The latter is obtained through physically correct simulation of light transport through randomly generated samples and provides a stripe-free reference image. The

simulation is achieved using the Python package biobeam [18] and has to our knowledge not been used before to validate stripe removal. The availability of a ground truth which is unavailable in experimentally acquired images allows for a more precise assessment of the effectiveness of stripe removal methods as common quality metrics such as the peak signal-to-noise ratio (PSNR) or the multi-scale structural similarity index measure (MS-SSIM) can be applied. For the evaluation of real and synthetic data we also rely on visual inspection and the curtaining metric proposed by Roldán [19] which measures the amount of corruptions by stripes and does not require reference images.

2. Methods

The majority of methods proposed for stripe removal fall under one of two categories which we refer to as *Fourier filtering* and *variational methods*. In the following, we concentrate only on these categories, as they generalize well to several imaging methods and variations in the appearance of image structures and stripes. Additional approaches are average filtering [20, 21], histogram matching [22, 23], spline interpolation [24] and only recently neural networks [25, 26]. However, these are usually tailored for a specific appearance of images and stripes such that they are harder to transfer to other scenarios.

2.1. Fourier Filtering

The Fourier transform yields a mathematical decomposition of the image content into constituent frequencies. In image processing, the breakdown into a frequency spectrum is used to detect and modify periodic structures of certain frequencies or entire bands of frequencies. We assume that stripe artifacts point into the same direction, have small widths and appear repeatedly. Applying the Fourier transform, most information about stripes becomes encoded in frequency coefficients around a small band orthogonal to the stripe direction. On the other hand, natural structures tend to differ strongly from stripe artifacts and live on larger scales than the widths of stripes. Thus, image structure information is mostly dominated by low frequencies which are concentrated in central coefficients. This general behavior is visualized in Figure 3.

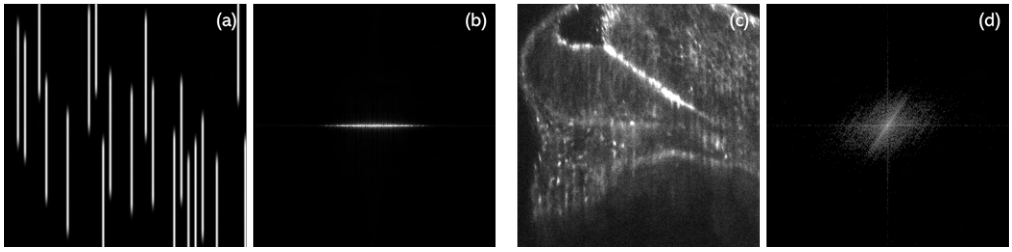

Fig. 3. Image and spectrum of thin stripes (a-b) and an LSFM image (c-d).

By appropriate damping or removal of selected coefficients attributed to stripe artifacts, corruptions can be significantly reduced. However, these coefficients may also contain some information about structures. Therefore, it is essential to select coefficients carefully to minimally destroy structural information. This can be achieved, e.g., by using masked filters [27, 28] or a decision-based algorithm [14]. Nonetheless, using a prior decomposition by structure and scale through wavelets [29] or the non-subsampled contourlet transform (NSCT) [30] reduces the number of modified image structures more effectively. Since the general idea of Fourier filtering approaches remains the same across all methods, we only discuss the multi-directional stripe remover (MDSR) [6] based on the NSCT. However, we use a slightly modified filtering step to improve quality. The MDSR depicts an improvement over wavelet Fourier filtering (WFF) which

was proposed by Münch et al. [29] and is commonly referenced and one of the earliest successful adoptions of Fourier filtering for removing stripes. The workflow of MDSR is visualized in Figure 4 and consists of the following steps:

1. **Image Decomposition:** Initially, stripe and image information are partially separated from each other by application of the NSCT. Using a pyramidal filter bank [30], the image is decomposed into subimages depending on direction and scale until a selected depth $n_{\text{dec}} \in \mathbb{N}$ is reached. The number of decomposition directions is denoted by $n_{\text{dir}} \in \mathbb{N}$.

2. **Fourier Filtering:** Subsequently, coefficients in the Fourier domain are modified and damped through filtering. This step is performed on subimages dependent on their decomposition directions $\theta_i \in [0, \pi], i = 1, \ldots, n_{\text{dir}}$ such that directions close to stripe direction θ_0 are strongly suppressed. Contrary to the proposition of Liang and colleagues [6], we apply damping only to subimages of directions with $|\theta_i - \theta_0| \leq \pi/4$. This reduces artifacting as shown in Figure 5. Considering vertical stripes, i.e. $\theta_0 = \pi/2$, damping is performed via element-wise multiplication with

$$f_i(x, y) = \begin{cases} 1 - \exp\left(-\frac{(y - \frac{n_y}{2})^2}{2\sigma_i^2}\right) & \text{if } |\theta_i - \theta_0| \leq \pi/4, \\ 1 & \text{otherwise,} \end{cases}$$

$$\sigma_i = \sigma \exp\left(-0.5 * \frac{(\theta_0 - \theta_i)^2}{\sigma_a^2}\right),$$

with Gaussian standard deviation σ and vertical image size n_y. The σ_i are adjustable damping parameters which depend on the deviation of directions θ_i to the stripe direction θ_0. The parameter $\sigma_a > 0$ describes how quickly damping is reduced when moving away from the stripe direction.

3. **Backtransformation:** Finally, the modified Fourier coefficients are transformed back into individual subimages. These are then recombined using NSCT reconstruction to yield a destriping result.

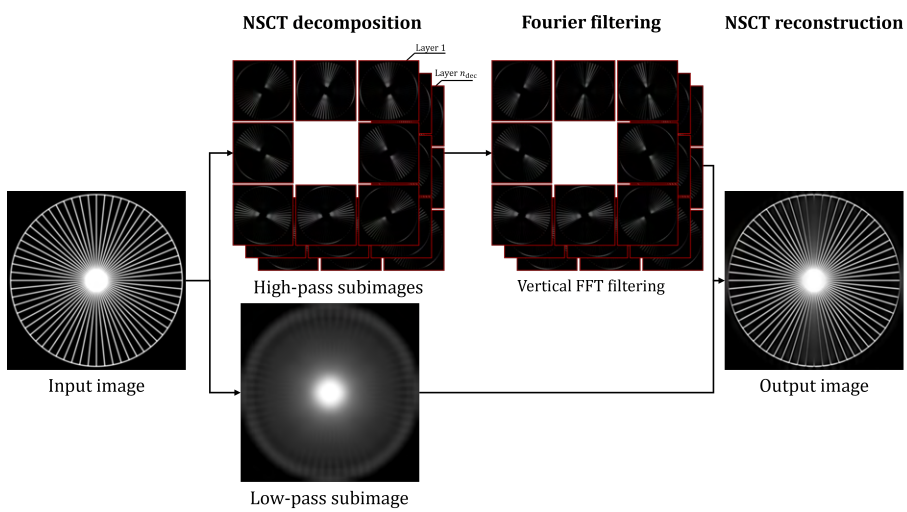

Fig. 4. MDSR workflow for removing vertical stripes.

The MDSR algorithm depends on multiple parameters: the number of directions n_{dir}, the decomposition depth n_{dec}, a Gaussian damping parameter σ, a directional fall-off parameter σ_a and the selected filter banks. However, most of them do not require fine-tuning once chosen adequately. For example, $n_{\text{dir}} = 8$ yields a sufficient directional decomposition and n_{dec} must only be large enough such that stripes are captured by the NSCT. Any further increase in both parameters has a negligible effect on the outcome while significantly increasing the computational complexity. σ is the main parameter to be adjusted and strongly influences the outcome as it depicts the strength of damping. When stripe artifacts are not ideally vertical but slightly oblique, adjusting σ_a may also improve performance.

Fourier filtering approaches such as MDSR rely on the condensation of stripe information in the Fourier domain and separability from image information. These prerequisites are given when stripes and structures have different shapes and appearances. However, stripes in LSFM images cover a wide range of lengths, widths and intensities. This results in stripe information to be less concentrated around the horizontal coefficient band and more mixed up with image information. Therefore, stripe removal becomes more difficult and can lead to the introduction of image artifacts through modification of image information.

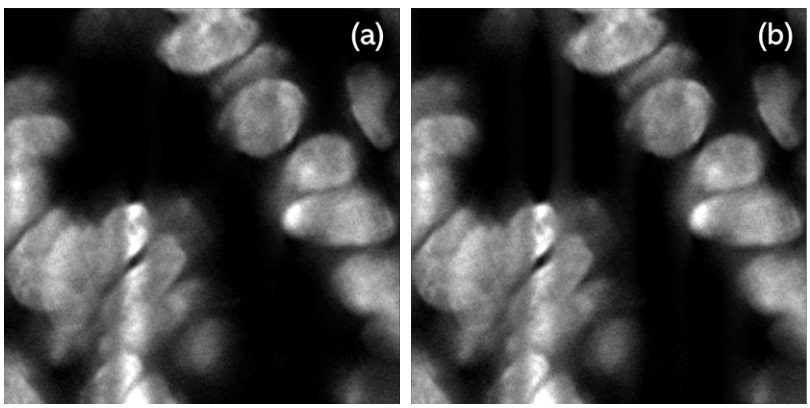

Fig. 5. Visual comparison of restricting Fourier filtering to subimages of directions $\theta_i \in [\frac{\pi}{4}, \frac{3\pi}{4}]$ (a) to unrestricted filtering [6] (b). Restricting filtering reduces artifacts introduced in dark areas while maintaining a similar level of stripe removal.

2.2. Variational Methods

Convex functions possess special properties that make them ideal for optimization problems. Under mild assertions, convex functions have unique minimizers which can be approximated using well-studied optimization algorithms, see [31]. Furthermore, convexity is an additive property of functions that allows for capturing multiple aspects of the stripe removal problem in a single objective function. Building on this concept, variational methods use a problem-specific weighted function consisting of convex building blocks, most commonly norms. The individual blocks penalize unwanted features of the clean image u and stripes s through positive contribution to the function. A corresponding minimizer should then reasonably fulfill the desired properties, i.e., optimizing the objective function yields a solution to the problem. This approach is well known in image processing and commonly used for denoising, segmentation and active contours, see, e.g, [32]. To give further insight into the construction of a suitable objective function, we motivate our model in the following. It has been originally developed by adjusting the functional proposed by Fitschen et al. [15] for curtaining noise in FIB-SEM images to the LSFM setting. The result shows resemblance to the proposition by Liu et al. [16] but extends applicability to

3D images and includes an additional regularization term which improves performance and consistency.

2.2.1. Construction

Consider the corrupted image $u_0 \in [0,1]^N$, the clean image $u \in [0,1]^N$ and the stripe image $s \in [-1,1]^N$ where $N = n_x \times n_y \times n_z$. We assume that $u_0 = u + s$ holds true, i.e., the corrupted image can be expressed as sum of a clean image and stripes. Let ∇_x, ∇_y and ∇_z denote the directional difference operators for the three coordinate directions and assume that stripes point in y-direction.

The objective function is constructed as superposition of individual convex building blocks which encode certain properties of the clean image and stripes. These are generally grouped into a data term $D(u)$ and a noise term $N(s)$ such that the general optimization problem reads

$$\operatorname*{argmin}_{u+s=u_0} D(u) + N(s). \tag{1}$$

The *total variation*

$$\|\nabla u\|_{2,1} = \sum_{i,j,k} \sqrt{(\nabla_x u)^2_{i,j,k} + (\nabla_y u)^2_{i,j,k} + (\nabla_z u)^2_{i,j,k}} \tag{2}$$

is a frequently used data term in image processing, see [16, 33]. It expresses that natural images are piece-wise smooth, i.e., they contain only few strong edges, which is reflected by overall small directional differences in all directions. In contrast, stripe artifacts are usually thin and elongated. This corresponds to large values of the difference operators $\nabla_x s$ and $\nabla_z s$ orthogonal to the stripe direction and small values of the operator $\nabla_y s$ in direction parallel to it. This behaviour is visualized in Figure 6 and incorporated by the function

$$\left\|\nabla_y s\right\|_1 = \sum_{i,j,k} |(\nabla_y s)_{i,j,k}|. \tag{3}$$

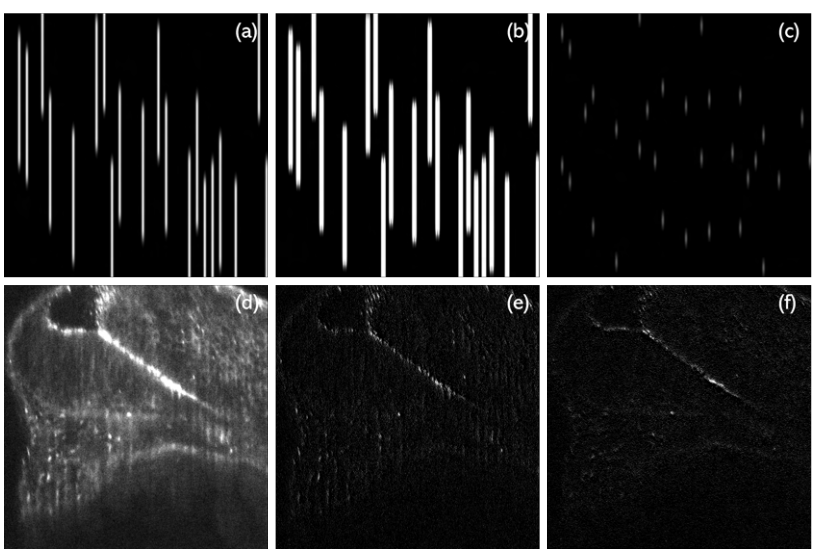

Fig. 6. Images and directional differences ∇_x, ∇_y of thin stripes (a-c) and LSFM image (d-f).

Combined with the total variation (2), these capture the essence of stripe removal in a single data and noise term. However, for increased performance regularization is added through

$$\|s\|_1 = \sum_{i,j,k} |s_{i,j,k}|, \tag{4}$$

$$\iota_{[0,1]^N}(u) = \begin{cases} 0 & \text{if } u \in [0,1]^N, \\ \infty & \text{if } u \notin [0,1]^N. \end{cases} \tag{5}$$

Function (4) is ℓ_1-regularization and promotes sparsity in the stripe image. It reflects the generally small area affected by artifacts and avoids unnecessary detection of image structures as stripes. On the contrary, (5) is included to guarantee the outcome to live in the same value range as the input. Specifically, values remain non-negative which reflects the natural interpretation of the signal. Hence, further post-processing such as adjusting contrast or the intensity profile becomes obsolete. With (2)-(5) the resulting optimization problem reads

$$\underset{u+s=u_0}{\operatorname{argmin}}\ \mu_1 \|\nabla u\|_{2,1} + \left\|\nabla_y s\right\|_1 + \mu_2 \|s\|_1 + \iota_{[0,1]^N}(u), \tag{6}$$

with weighting parameters $\mu_1, \mu_2 > 0$. Due to scaling invariances $\lambda\iota_{[0,1]^N} = \iota_{[0,1]^N}$ and $\operatorname{argmin} f(x) = \operatorname{argmin} \lambda f(x)$ for $\lambda > 0$, only two weights μ_1, μ_2 are required.

2.2.2. Minimization

Convex optimization problems such as (6) can be solved via the primal-dual gradient hybrid method with extrapolation of the dual variable (PDHGMp), see [31]. In principle, the iterative algorithm relies on alternating between minimizing the primal and maximizing the dual problem. Under mild assertions on the objective function, the sequence generated by PDHGMp converges to a solution. More details on the PDHGMp and alternatives can be found in [31, 34].

The weighting parameters μ_1, μ_2 in (6) provide adjustability to control different aspects of the stripe removal, enabling the production of a desired result, see Appendix B. The following rules of thumb can be established:

(i) μ_1 controls the desired smoothness of the outcome. It strongly regulates the amount of stripe removal to be applied. Larger values will result in stronger removal of stripe artifacts. However, vertical and stripe-like image structures may become affected which results in a smoothed image.

(ii) μ_2 acts like a counterpart to μ_1 and controls how strictly attention is paid to the stripe-like appearance of affected structures. Increasing its value generally tends towards a lower reduction of artifacts and image structures. However, in combination with μ_1 it enforces that for the most part only structures with concise stripe properties, e.g., thin, long and vertically aligned, are removed. Structures which deviate from this become less likely to be affected.

(iii) The ratio of μ_1 and μ_2 is crucial and has a strong influence on the general outcome. Scaling both parameters by an equal factor changes the amount of stripe removal while keeping the effect on image structures mostly unchanged.

In addition to offering adjustability through two parameters, variational methods are generally less susceptible to stripes deviating from being thin and long which we will see in Section 3. Furthermore, they can easily be adapted towards specific settings. For example, considering 2D image data we replace the total variation (2) by its lower-dimensional counterpart

$$\left\|\nabla^{(2D)} u\right\|_{2,1} = \sum_{i,j} \sqrt{(\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2}. \tag{7}$$

Alternatively, it is common for LSFM images to have a lower spatial resolution along the displacement direction z of the light-sheet which can be incorporated by using $\nabla_z^* u = \rho_z \nabla_z u$ with $0 \leq \rho_z \leq 1$. This emphasizes that information along the z-direction is less reliable and coherent. The case $\rho_z = 1$ corresponds to (6) and $\rho_z = 0$ to the 2D case. We can also consider oblique stripe artifacts as described in [16] by replacing $\nabla_y s$ with a suitable directional difference operator $\nabla_\theta s$ in (3). Lastly, a simultaneous multi-directional stripe removal is possible using a sum of penalization terms $\left\|\nabla_{\theta_i} s\right\|_1$ of different stripe directions θ_i.

2.2.3. State of the art

Besides our proposition, there exist a variety of different objective functions which are usually tailored for specific cases. An early and more general proposition comes from Fehrenbach and colleagues [35] designed to remove repeating artifacts. The model relies on the representation of artifacts as

$$s = \sum_{i=1}^{m} \lambda_i * \psi_i, \quad m \in \mathbb{N}$$

with elementary noise patterns ψ_i and λ_i containing position information. This approach makes it applicable to not only stripes but any kind of noise created by repeating specific patterns. In a follow-up paper by Escande et al. [36], multiplicative noise of the same form is considered. This represents stripe formation in LSFM more closely. However, the authors have noticed a significant change in contrast as side effect of their method which can strongly alter the image appearance. Furthermore, the selection of reasonable elementary patterns ψ_i is not straightforward since large variations in stripe sizes are observed in LSFM. Other propositions include [17] utilizing a low-rank assumption of s and [15] for the removal of curtaining artifacts including stripes. While these are prime examples of variational methods, their assumptions deviate strongly from observations in LSFM.

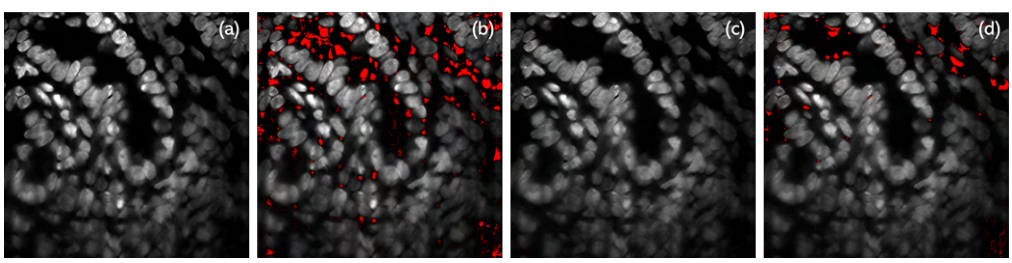

Fig. 7. Stripe removal results using different variations of (6) with $\mu_1 = \frac{1}{3}$ and $\mu_2 = \frac{1}{300}$. 2D with $\iota_{[0,1]^N}$ (a), 2D without $\iota_{[0,1]^N}$(b), 3D with $\iota_{[0,1]^N}$ (c) and 3D without $\iota_{[0,1]^N}$ (d). Areas with values less than 0 are colored in red and were clipped to 0 for comparability.

Another proposition for stripe removal stems from Liu and colleagues [16] who proposed a functional for oblique stripe artifacts which corresponds to (6) for vertical stripes but without the indicator function $\iota_{[0,1]^N}(u)$ to restrict values to the range $[0, 1]$. Additionally, they only consider the 2D-total variation of (2). In Figure 7 we display stripe removal results to visually compare the influence of processing in 2D and 3D and the difference of restricting optimization to the value range $[0, 1]$ or not. At first glance, results are almost identical in terms of their visual appearance. However, at closer inspection some noticeable differences can be spotted. For example, when moving from 2D to 3D, we see a subtle increase in brightness in the background area left to the center. Inspection of the 3D image stack reveals that structural components can be found in the neighbouring slices which explains this phenomenon. Therefore, when interested in

a single image slice it is sufficient to use the 2D variation. However, for 3D the incorporation of depth information ensures reasonably smooth intensity profiles and prevents sudden jumps. This becomes beneficial when further processing, such as image segmentation, is performed on the 3D image. On the other hand, the indicator function does not cause drastically different results if we clip the intensity profile to the same range after processing. However, as the affected areas are not negligible, we argue that removing the need for further steps is beneficial and prevents the introduction of errors.

3. Results

In this section, we evaluate and compare the results of stripe removal methods. Therefore, we consider both, synthetic and real image data and evaluate performance on selected 2D slices.

3.1. Performance Measures

Performance of the destriping methods is evaluated by visual inspection and assisted by three quality metrics: the peak signal-to-noise ratio (PSNR), the multi-scale structural similarity index measure (MS-SSIM) [37], and the 'curtaining' metric proposed by Roldan [19]. PSNR is one of the most common quality metrics used in image processing. For any image $u \in [0,1]^{n_x \times n_y}$ and reference u^*

$$\mathrm{PSNR}(u, u^*) = -10 \log \underbrace{\left(\frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} [u(i,j) - u^*(i,j)]^2 \right)}_{= \text{Mean squared error}} \tag{8}$$

captures the power of corruptions in comparison to the reference as scalar value in $[0, \infty]$ where larger values correspond to higher quality. MS-SSIM on the other hand is an extension of the popular structural similarity index measure

$$\mathrm{SSIM}(u, u^*) = l(u, u^*) \cdot c(u, u^*) \cdot \tilde{s}(u, u^*) \in [0, 1] \tag{9}$$

based on measurements for luminance l, contrast c and structure $\tilde{s}$. It attains values in $[0, 1]$ with larger values corresponding to stronger similarity. To obtain the multi-scale version, measurements of different image scales are combined, see [37].

Lastly, the curtaining metric from [19] measures the presence of stripes. It is based on the same ideas as Fourier filtering methods, see Figure 3 (a), and measures how strongly Fourier coefficients are condensed around a central horizontal line. The range of values is $[0, 1]$ with larger values corresponding to better quality, i.e., less stripes. A detailed description can be found in [19]. Since this metric is based on similar ideas as Fourier filter methods, it is biased toward such and comparably sensitive to deviations from ideal stripes. However, in contrast to the prior measures, the curtaining metric does not require a reference image and is tailor-made for characterizing corruptions by stripes.

3.2. Description of employed stripe removal methods

For our comparison we consider the following methods:

- The Fourier based multi-directional stripe remover (MDSR) proposed by Liang et al. [6] was described in Section 2.1. We consider a modification restricting filtering to subimages of direction close to the stripe direction. It uses the NSCT decomposition followed by filtering of the Fourier domain of selected subimages. As mentioned before, the number of directions $n_{\mathrm{dir}} = 8$ and depth of decomposition $n_{\mathrm{dec}} = 5$ can be fixed for all following results. Processing was performed using the 'Nonsubsampled Contourlet Toolbox' [38] in Matlab. In particular, the function 'nsctdec' with default options for generating filter banks

was used. Therefore, only the standard deviation σ of the Gaussian damping function and the damping fall-off parameter σ_a remain as free parameters.

- The variational stationary noise remover (VSNR) proposed by Fehrenbach et al. [35] was briefly mentioned in Section 2.2. This variational approach follows the idea of decomposing the image of stripe artifacts into elementary stripe patterns ψ_i convolved with location images λ_i. For our application, we chose the real part of three differently sized Gabor filters, see Figure 8, such that short, medium and long stripes can be detected simultaneously. The corresponding objective function reads

$$\underset{\substack{u+s=u_0 \\ \|\lambda\|_\infty \leq 1}}{\operatorname{argmin}} \|\nabla u\|_{1,\epsilon} + \alpha_1 \|\lambda_1\|_1 + \alpha_2 \|\lambda_2\|_1 + \alpha_3 \|\lambda_3\|_1 \quad \text{s.t} \quad s = \sum_{i=1}^{3} \lambda_i * \psi_i, \tag{10}$$

with free weighting parameters $\alpha_1, \alpha_2, \alpha_3, \epsilon > 0$ and

$$\|q\|_{1,\epsilon} = \sum_{i,j} \phi_\epsilon \left(\sqrt{q_{i,j,1}^2 + q_{i,j,2}^2} \right)$$

$$\phi_\epsilon(x) = \begin{cases} \frac{x^2}{2\epsilon} & |x| \leq \epsilon, \\ |x| - \frac{\epsilon}{2} & \text{otherwise.} \end{cases}$$

We use the Matlab implementation [39] provided by the author Pierre Weiss. For a sufficient approximation of the optimum, we chose to perform 25000 steps of the PDHGMp optimization algorithm.

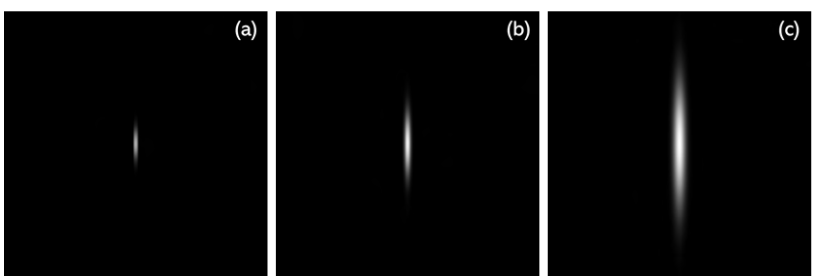

Fig. 8. Elementary stripe patterns $\psi_i, i = 1, 2, 3$ at 4x magnification.

- The variational method (6) has been introduced in Section 2.2 and is an extension of the method proposed by Liu et al. [16]. We will refer to it as "general stripe remover" (GSR) and use the variation

$$\text{(2D)} \quad \underset{u+s=u_0}{\operatorname{argmin}} \mu_1 \left\| \nabla^{(2D)} u \right\|_{2,1} + \left\| \nabla_y s \right\|_1 + \mu_2 \|s\|_1 + \iota_{[0,1]^N}(u),$$

for 2D images. Free parameters are μ_1 and μ_2. An approximation of the optimum is attained using 25000 iteration steps of the PDHGMp.

Our main objective for comparing methods was to achieve outcomes which are visually appealing and optimal in terms of the metrics at the same time. However, the performance measures consider only some aspects of the image quality, e.g., the curtaining metric of a strongly smoothed image would suggest optimal image quality even if the actual image content was destroyed. Therefore, visual inspection is required to confirm numerical results. Initial parameters of the methods were obtained from a coarse grid-search and visual and numerical assessment.

Afterwards, parameters were manually fine-tuned towards a reasonably optimal solution, i.e., any significant perturbation in parameters yielded either visual degradation, significant reductions of the quality metric values or both. For MDSR and VSNR the results were not clipped towards values in $[0, 1]$ when calculating the metrics to reflect the true outcome of the algorithms. However, for proper visualization clipping was applied. From each 3D image used in the following part, an individual slice was selected with prominent stripe artifacts and structures.

3.3. Synthetic Data

Synthetic images are obtained by modeling the imaging setup in LSFM and simulating light transport [7] using the python package *biobeam* [18]. The simulation allows for the retrieval of reference images without stripes by neglecting interactions of light and matter. However, this removes natural light attenuation which is usually visible as an intensity decrease along the illumination direction, see Figures 9 (a-b) and 10 (a-b). Hence, the reference images cannot be interpreted as ideal destriping results and a comparison should be done with caution. For details on the simulation and generation of synthetic images, see Appendix A.

The results shown in Figures 9 and 10 display visually significant improvements in image quality across all stripe removal methods. GSR produces the strongest reduction of stripes including the trailing artifacts behind the spherical body in Figure 9 (e) and inside of individual cells in both examples. VSNR removes most thin stripe artifacts with wide stripes remaining visible in Figure 9 (d) and more artifacts remaining inside of cells. Furthermore, faint smearing artifacts are introduced in cavities in Figure 10 (d). MDSR yields the lowest reduction of stripes in both cases with more visible remainders and stronger smearing artifacts, see Figures 9 (c) and 10 (c). PSNR and MS-SSIM show no significant improvement or deterioration for all results. The degree of corruption is reduced significantly for all methods on the cell cluster image but only for GSR on the dispersed cells.

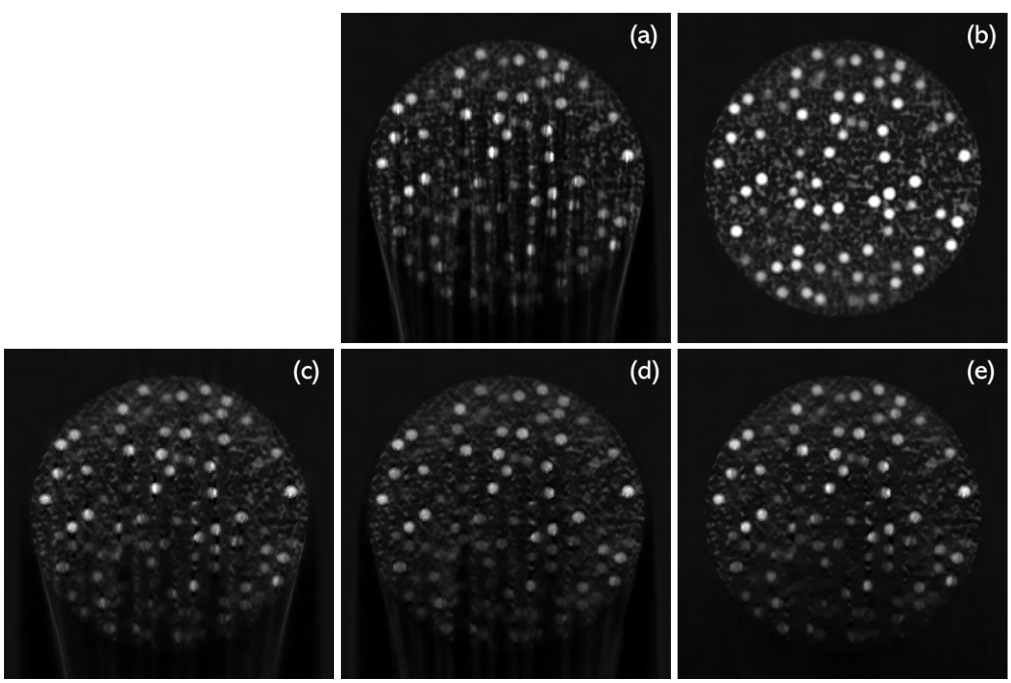

Fig. 9. Synthetic image data and destriping results of a cell cluster. Image (a), reference (b), MDSR, $\sigma = 12$, $\sigma_a = 0.3$ (c), VSNR, $\alpha = (3, 5, 10)$ (d), GSR, $\mu = (\frac{1}{3}, \frac{1}{300})$ (e).

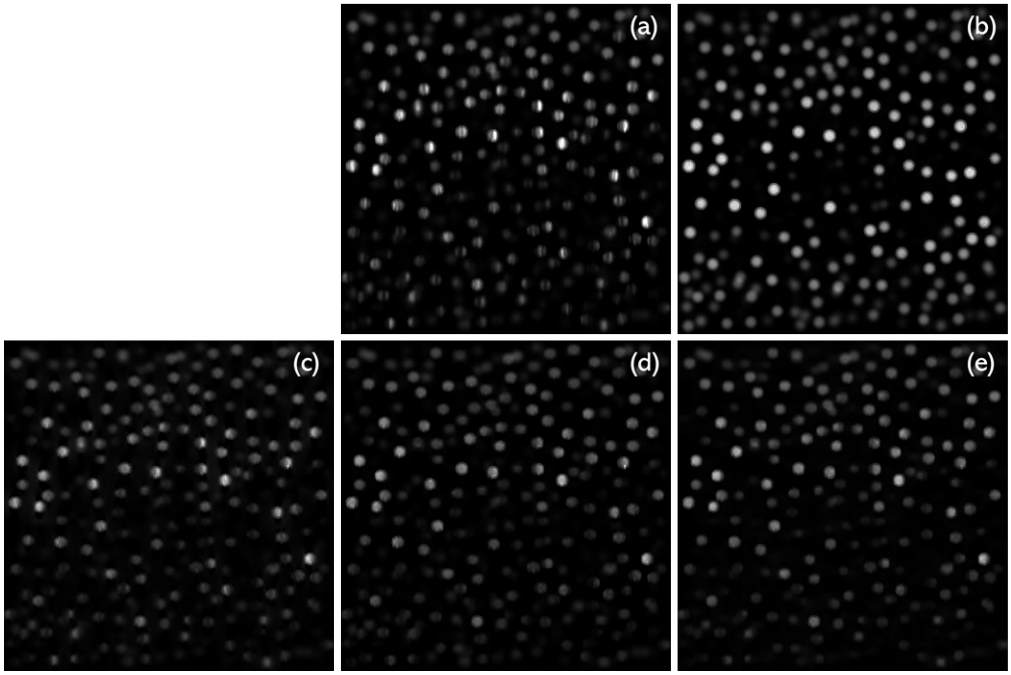

Fig. 10. Synthetic image data and destriping results of dispersed cells. Image (a), reference (b), MDSR, $\sigma = 24$, $\sigma_a = 0.3$ (c), VSNR, $\alpha = (3, 5, 10)$ (d), GSR, $\mu = (\frac{1}{2}, \frac{1}{60})$ (e).

	Metric	**Input**	**Reference**	**MDSR**	**VSNR**	**GSR(2D)**
Cell Cluster	Curtaining	0.78	0.90	0.99	0.90	0.97
	PSNR	24.80	∞	24.64	24.00	24.58
	MS-SSIM	0.86	1	0.85	0.84	0.87
Dispersed Cells	Curtaining	0.91	0.96	0.91	0.91	1
	PSNR	28.22	∞	27.25	27.13	27.20
	MS-SSIM	0.88	1	0.84	0.84	0.84

Table 1. Quality metrics of the images in Figures 9 and 10. Larger is better. The best result is highlighted in green.

3.4. LSFM images of biological samples

As real image data we consider two real LSFM images with different appearances. The image used in Figure 11 shows an organoid of cultivated mouse intenstine cells where indiviudal cells are clearly visible and separated. On the other hand, Figure 12 displays a cluster of HeLa cells which has a more homogeneous appearance.

Similar observations to the synthetic image data can be made. GSR displays the strongest reduction in stripes including wide and slightly oblique artifacts. However, for the mouse intestine cuboid some faint smearing artifacts were introduced in structural cavities, see Figure 11 (d), and few wide stripes remain visible in Figure 12 (d). VSNR removes all visible stripes in the HeLa-cluster image but is unable to remove wide stripes for the mouse intestine cuboid, see Figures 12 (b) and 11 (b). In the latter, thin but oblique artifacts are reduced but remain visible and very faint smearing artifacts were introduced similar to the observation for GSR. The results by MDSR are comparable to VSNR on the mouse intestine cuboid. Some smearing artifacts are introduced but there are even more remainders of thin oblique stripes than for VSNR. Nonetheless, there are barely any remaining artifacts in the HeLa-cluster image with very similar appearance to the result by GSR.

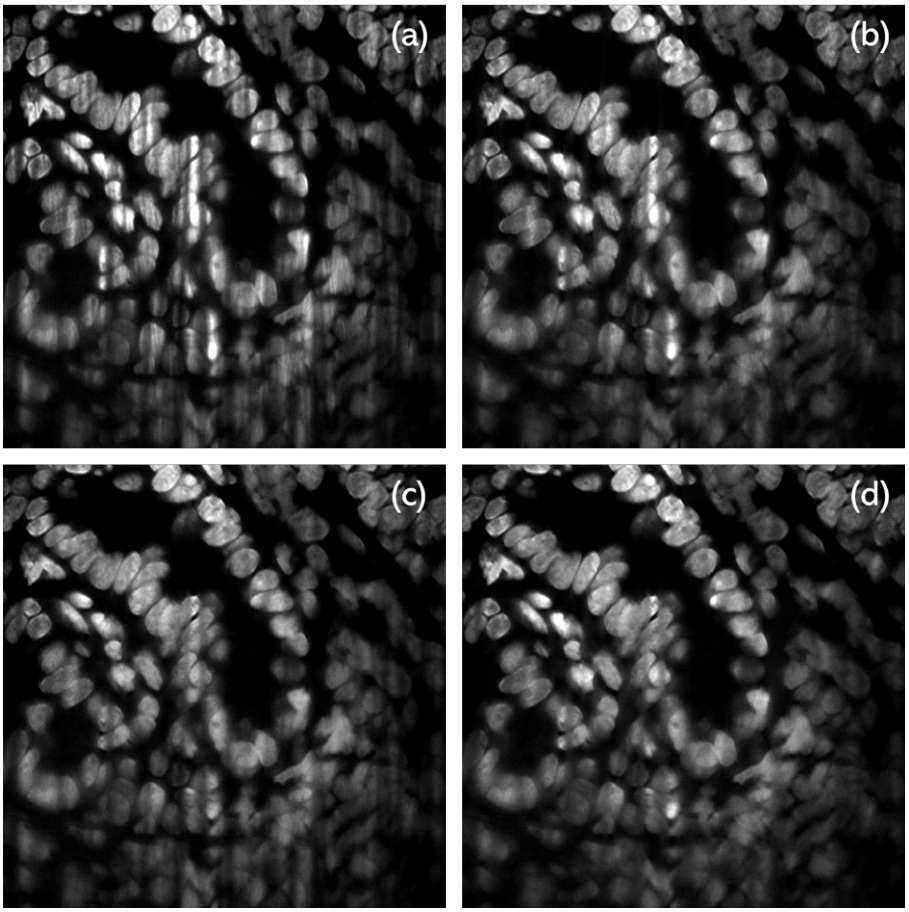

Fig. 11. Mouse intestine cuboid and results. Image (a), MDSR, $\sigma = 12$, $\sigma_a = 0.3$ (b), VSNR, $\alpha = (3, 5, 10)$ (c), GSR(2D), $\mu = (\frac{1}{3}, \frac{1}{300})$ (d). Corresponding curtaining values are 0.58 (a), 0.84 (b), 0.78 (c) and 0.95 (d).

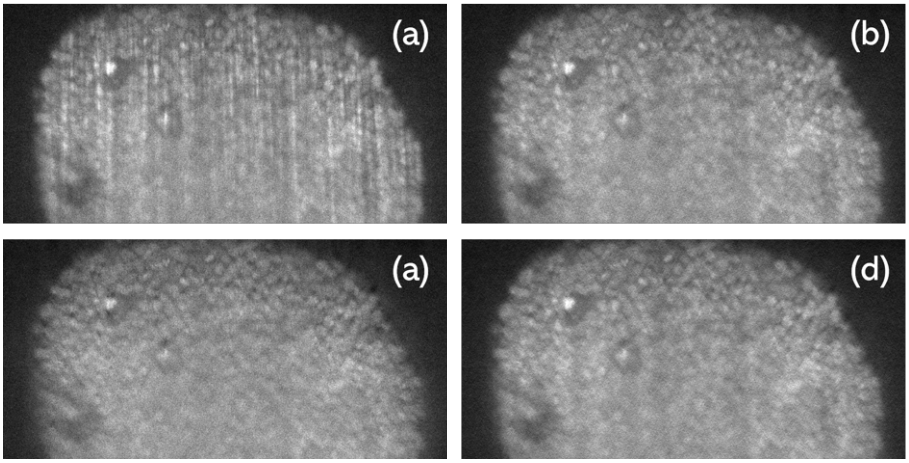

Fig. 12. HeLa-cluster and results. Image (a), MDSR, $\sigma = 5$, $\sigma_a = 0.3$ (b), VSNR, $\alpha = (10, 5, 1)$ (c), GSR(2D), $\mu = (\frac{1}{3}, \frac{1}{300})$ (d). Corresponding curtaining values are 0.24 (a), 0.90 (b), 0.99 (c) and 0.96 (d).

3.5. Application for other imaging methods

We also consider data by other imaging methods where stripe artifacts are commonly encountered. Our examples include a FIB-SEM image of a tin bronze cast alloy and a Terra MODIS satellite image. Both show much more uniformly thin and long stripes compared to the stripes observed in LSFM.

GSR removes stripes entirely from the Terra MODIS image and leaves only few faint remainders of stripes in the FIB-SEM example. Similar results can be observed for VSNR but remainders in the FIB-SEM image are slightly stronger. On the other hand, MDSR introduces visible artifacts in high contrast areas in the Terra MODIS image despite removing all stripes. Additionally, it leaves visible remainders of wide stripes for the FIB-SEM image where multiple stripes occurred close together.

3.6. Discussion

The results shown in Section 3 reveal that general stripe remover (GSR) outperforms existing methods such as multi-directional stripe remover (MDSR) and Variational Stationary Noise remover (VSNR). It produces consistently good results in terms of stripe removal quality and retention of image structures while offering adjustability through the weighting parameters μ_1 and μ_2. On the other hand, MDSR and VSNR also show promising results but cannot keep up with GSR on both synthetic and real image data.

The synthetic data which offer stripe-free reference images showed that only GSR is able to remove stripes of larger widths and slightly oblique artifacts such as the trails in Figure 9. This is caused by the choice of functional (6) where stripes are mainly penalized based on directional differences. Hence, even oblique stripe artifacts are affected despite of a slightly larger penalization. On the contrary, VSNR relies heavily on pre-selected patterns such that the removal of oblique artifacts might require an oblique pattern to work well. Similarly, MDSR assumes that stripe information condenses closely into a horizontal coefficient band in the Fourier domain. Deviations in stripe direction will yield a significantly lower reduction of stripe information. With similar arguments we can explain that GSR performs best on the dispersed cells image in Figure 10 where MDSR and VSNR introduce smearing artifacts and yield a lower reduction of stripes.

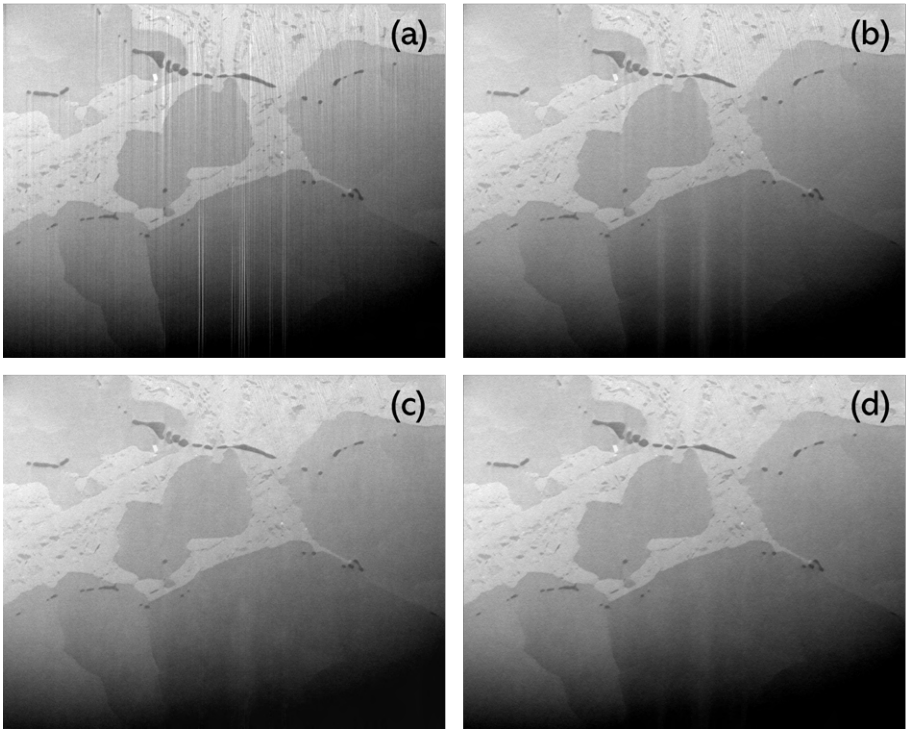

Fig. 13. FIB-SEM example and results. Image (a), MDSR, $\sigma = 12$, $\sigma_a = 0.3$ (b), VSNR, $\alpha = (2, 10, 2)$ (c), GSR(2D), $\mu = (\frac{2}{5}, \frac{1}{300})$ (d). Corresponding curtaining values are 0.18 (a), 0.54 (b), 0.95 (c) and 1 (d).

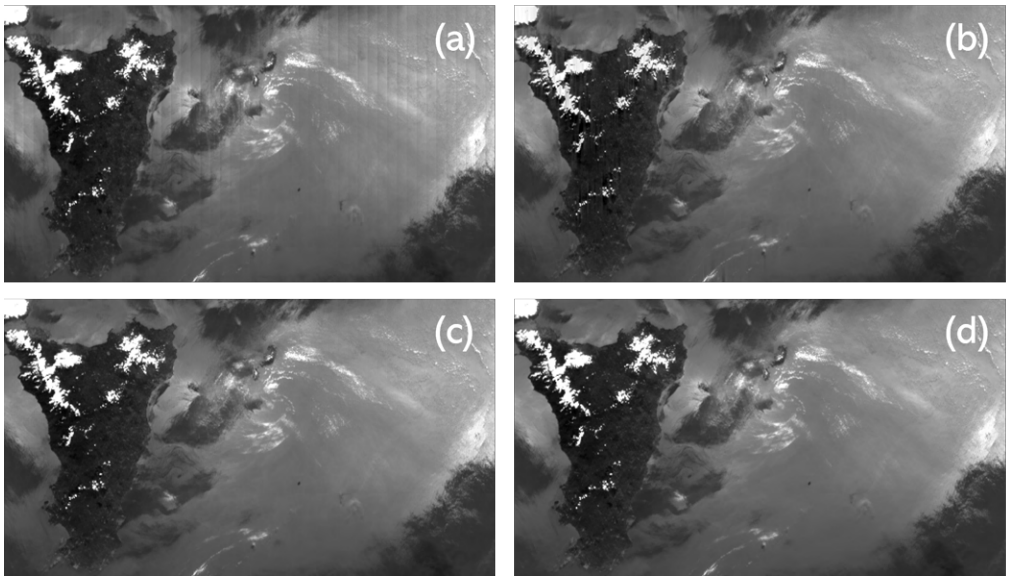

Fig. 14. Terra MODIS example and results. Image (a), MDSR, $\sigma = 10$, $\sigma_a = 0.3$ (b), VSNR, $\alpha = (3, 5, 5)$ (c), GSR(2D), $\mu = (\frac{7}{30}, \frac{1}{300})$ (d). Corresponding curtaining values are 1 (a), 1 (b), 1 (c) and 1 (d).

Including the metrics in Table 1 into our assessment confirms observations for the most part. The high curtaining values produced by MDSR can be attributed to bias as both rely on the same theory and properties of Fourier space. Furthermore, the lack of improvement of MDSR and VSNR on the dispersed cells image is caused by the introduction of smearing artifacts as previously mentioned. Note, that PSNR and MS-SSIM do not improve across all methods. There are multiple reasons that explain this phenomenon. First, the removal of stripe artifacts results in a noticeable decrease in contrast, specifically in areas which were strongly corrupted. Furthermore, the reference image has an overall higher brightness and more details because light attenuation and interactions with matter are neglected during generation. Hence, a full restoration of the original structures is neither anticipated nor possible. Despite of these problems, the fact that PSNR and MS-SSIM have not declined suggests that the structural improvements by removing stripes has compensated any side effect of the removal process.

For the real image data including LSFM and other applications we observe similar results with GSR producing consistently good results. Specifically, the amount of stripe reduction on the mouse intestine cuboid shown in Figure 11 is impressive as almost all artifacts are removed with minimal remainders. This showcases the strengths of GSR as the image is corrupted by a broad range of different stripe artifacts, including various widths, lengths, intensities and deviations in orientations. However, we observe that there are some problems with introducing smearing artifacts in stripe-like cavities of the structure. This can be avoided by adjusting μ_1 and μ_2 at the cost of lower stripe reduction.

VSNR develops its full potential on the HeLa-cluster image. The result is close to perfect showing no visible artifacts while retaining good contrast and image details. This is achieved since the image reflects an ideal scenario of application where stripe patterns fit perfectly to the stripe artifacts, structures have low contrast and a uniformly colored background. The latter hides smoothing artifacts that may be introduced in the process. The results by GSR on this image are on par with slightly better contrast and detail preservation but few large and smooth stripes remain visible. This is most likely caused by a lower smoothing of artifacts and structures in comparison to VSNR.

MDSR performs generally worse on the real image data. For the mouse intestine cuboid even slightly oblique but thin stripes remain visible. The FIB-SEM example shows noticeably more artifacts than other methods. Finally, on the Terra MODIS example a large number of artifacts has been introduced, e.g., in the upper left landmass. This is surprising as stripe artifacts are periodic and very thin which should be the ideal setting for such a method. Nonetheless, the high contrast in certain areas appears to produce unforeseen problems. On the other side, MDSR achieves a close to perfect result on the HeLa-cluster image. Using the same arguments as before, this is due to the low contrast and uniform background which hides imperfections in the artifact removal. All observations combined show that compared to its variational competitors MDSR is more sensitive to deviations from the model assumption of stripe artifacts being thin and long. Any deviation of length, width or direction yields a lower concentration of stripe information around the horizontal coefficient band. This hampers successful filtering.

For evaluating the quality of stripe removal the curtaining metric proved to be a valuable addition to visual inspection. While it is biased towards Fourier filtering methods, it successfully conveys the corruption of stripe artifacts and improvements made by destriping. Furthermore, it is a good indicator for measuring stripe removal that does not require a reference image such as PSNR or (MS)-SSIM. However, an evaluation of quality still depends mostly on visual inspection and cannot be directly deduced from the metric itself. This becomes clear when comparing the value of curtaining between the corrupted images. We argue that the mouse intestine cuboid displays the most severe corruptions by stripes in terms of visual impairment, yet curtaining values suggest otherwise.

The use of synthetically generated LSFM images for evaluation of stripe removal was, to the

best of our knowledge, not considered before in research. However, it provides a broad range of different structures and produces stripe-free reference images which allow us to study the behaviour of stripe removal in much more detail. More precisely, we are able to explore the full capabilities of stripe removal methods as we can produce "extreme" cases and determine limitations. For example, the dispersed cells image in Figure 10 is a challenging setting as stripes appear only as short segments inside of the small cells. It demonstrates limitations such as inconsistent shape reconstruction, insufficient stripe removal or the introduction of smearing artifacts. Furthermore, synthetic image data allows for the use of quality metrics such as PSNR and MS-SSIM which provide more information to asses the stripe removal performance with higher confidence.

4. Conclusion

In this paper, we focused on the application of different stripe removal methods on LSFM images and a subsequent evaluation of results. We explained Fourier filtering and variational methods using selected methods. These incorporated well-founded extensions of existing propositions for increased quality of stripe removal. Afterwards, we used synthetically generated LSFM data and real images to evaluate and compare the performance of stripe removal by using visual inspection. The observations are confirmed using quality metrics. This revealed that our proposed objective function (6), later referred to as GSR, produces consistent results with better stripe removal than prior propositions.

For Fourier filtering, we used a modified version of MDSR. The method applies a structural decomposition by shape and scale called NSCT. Afterwards, components with directions close to stripe direction are reduced by applying appropriate filtering in the Fourier domain where image and stripe information are more locally separated. A transformation back towards the image domain yields the result. On the other hand, for variational methods we discussed the construction of the objective function of GSR in detail. Individual terms are formulated by encoding properties of stripes and structures. These are then combined into a weighted functional and optimized by an minimization algorithm.

During our evaluation on real and synthetic data we found out that the GSR introduced in this paper produces consistently good results across all examples. It offers the best removal of stripe artifacts for all variations of observed stripes. This includes stripes of varying lengths, widths, intensities and slightly deviated directions. Other methods such as the previously published MDSR and VSNR produce a lower reduction of artifacts and show inconsistencies when faced by different image appearances and structures.

The usage of synthetic LSFM data which offer a stripe-free reference image is a novel supplement of real data for comparing quality of stripe removal. It provides valuable insights into the capabilities and limitations of stripe removal methods. Furthermore, it allows for the use of standard image quality metrics such as PSNR and MS-SSIM which further assist in a precise assessment.

The curtaining metric proved to be a valuable addition to visual inspection as indicator for stripe corruption. Despite its shortcomings curtaining values reaffirm our visual inspection and successfully convey a reduction in stripe artifacts in almost all cases. Its major advantages are that values are calculated directly on the image without a need of reference and that it measures corruptions by stripes. However, visual inspection remains indispensable in evaluating results.

References

1. J. Huisken, J. Swoger, F. D. Bene, *et al.*, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," Science **305**, 1007–1009 (2004).
2. M. Jemielita, M. J. Taormina, A. DeLaurier, *et al.*, "Comparing phototoxicity during the development of a zebrafish craniofacial bone using confocal and light sheet fluorescence microscopy techniques," J. Biophotonics **6**, 920–928 (2013).

3. F. O. Fahrbach, P. Simon, and A. Rohrbach, "Microscopy with self-reconstructing beams," Nat. Photonics **4**, 780–785 (2010).
4. J. Huisken and D. Y. R. Stainier, "Even fluorescence excitation by multidirectional selective plane illumination microscopy (mspim)," Opt. Lett. **32**, 2608–2610 (2007).
5. P. Keller and E. Stelzer, "Digital scanned laser light sheet fluorescence microscopy," Cold Spring Harb. protocols **2010**, pdb.top78 (2010).
6. X. Liang, Y. Zang, D. Dong, *et al.*, "Stripe artifact elimination based on nonsubsampled contourlet transform for light sheet fluorescence microscopy," J. Biomed. Opt. **21**, 1–9 (2016).
7. A. Rohrbach, "Artifacts resulting from imaging in scattering media: a theoretical prediction," Opt. Lett. **34**, 3041–3043 (2009).
8. S. M. Salili, M. Harrington, and D. J. Durian, "Note: Eliminating stripe artifacts in light-sheet fluorescence imaging," Rev. Sci. Instruments **89**, 036107 (2018).
9. N. Scherf and J. Huisken, "The smart and gentle microscope," Nat. Biotechnol. **33**, 815–818 (2015).
10. P. Ricci, V. Gavryusev, C. Müllenbroich, *et al.*, "Removing striping artifacts in light-sheet fluorescence microscopy: a review," Prog. Biophys. Mol. Biol. **168**, 52–65 (2022).
11. F. O. Fahrbach and A. Rohrbach, "Propagation stability of self-reconstructing bessel beams enables contrast-enhanced imaging in thick media," Nat. Commun. **3**, 632 (2012).
12. F. O. Fahrbach, V. Gurchenkov, K. Alessandri, *et al.*, "Light-sheet microscopy in thick media using scanned bessel beams and two-photon fluorescence excitation," Opt. Express **21**, 13824–13839 (2013).
13. F. O. Fahrbach, V. Gurchenkov, K. Alessandri, *et al.*, "Self-reconstructing sectioned bessel beams offer submicron optical sectioning for large fields of view in light-sheet microscopy," Opt. Express **21**, 11425–11440 (2013).
14. S.-W. W. Chen and J.-L. Pellequer, "DeStripe: frequency-based algorithm for removing stripe noises from AFM images," BMC Struct. Biol. **11**, 7 (2011).
15. J. H. Fitschen, J. Ma, and S. Schuff, "Removal of curtaining effects by a variational model with directional forward differences," Comput. Vis. Image Underst. **155**, 24–32 (2017).
16. X. Liu, X. Lu, H. Shen, *et al.*, "Oblique stripe removal in remote sensing images via oriented variation," CoRR **abs/1809.02043** (2018).
17. Y. Chang, L. Yan, T. Wu, and S. Zhong, "Remote sensing image stripe noise removal: From image decomposition perspective," IEEE Trans. on Geosci. Remote. Sens. **54**, 7018–7031 (2016).
18. M. Weigert, K. Subramanian, S. T. Bundschuh, *et al.*, "Biobeam—multiplexed wave-optical simulations of light-sheet microscopy," PLoS Comput. Biol. **14**, 1–11 (2018).
19. D. Roldán, C. Redenbach, K. Schladitz, *et al.*, "Image quality evaluation for FIB-SEM images," J. Microsc. **293**, 98–117 (2024).
20. A. N. M. Ashrafuzzaman, S. Y. Lee, and M. K. Hasan, "A self-adaptive approach for the detection and correction of stripes in the sinogram: Suppression of ring artifacts in CT imaging," EURASIP J. on Adv. Signal Process. **2011**, 183547 (2010).
21. M. Boin and A. Haibel, "Compensation of ring artefacts in synchrotron tomographic images," Opt. Express **14**, 12071–12075 (2006).
22. B. Cao, Y. Du, D. Xu, *et al.*, "An improved histogram matching algorithm for the removal of striping noise in optical remote sensing imagery," Optik **126**, 4723–4730 (2015).
23. P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Stripe noise reduction in MODIS data by combining histogram matching with facet filter," IEEE Trans. on Geosci. Remote. Sens. **45**, 1844–1856 (2007).
24. F. Tsai, S.-Q. Lin, J.-Y. Rau, *et al.*, "Destriping hyperion imagery using spline interpolation," in *Proc. 26th Asian Conf. Remote Sensing, November, Hanoi, Vietnam,* (2005).
25. Y. Liu, K. Weiss, N. Navab, *et al.*, "Destripe: A self2self spatio-spectral graph neural network with unfolded hessian for stripe artifact removal in light-sheet microscopy," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022,* L. Wang, Q. Dou, P. T. Fletcher, *et al.*, eds. (Springer Nature Switzerland, Cham, 2022), pp. 99–108.
26. Z. Wei, X. Wu, W. Tong, *et al.*, "Elimination of stripe artifacts in light sheet fluorescence microscopy using an attention-based residual neural network," Biomed. Opt. Express **13**, 1292–1311 (2022).
27. D. Wilken, P. Feldens, T. Wunderlich, and C. Heinrich, "Application of 2d Fourier filtering for elimination of stripe noise in side-scan sonar mosaics," Geo-Marine Lett. **32**, 337–347 (2012).
28. S. Liu, L. Sun, and F. Xiao, "A staged filtering approach to kill curtain noise in FIB-SEM images," "" **2016**, 1–5 (2016).
29. B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet — Fourier filtering," Opt. Express **17**, 8567–8591 (2009).
30. A. Da Cunha, J. Zhou, and M. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," IEEE Trans. on Image Process. **15**, 3089–3101 (2006).
31. M. Burger, A. Sawatzky, and G. Steidl, "First order algorithms in variational image processing," in *Splitting Methods in Communication, Imaging, Science, and Engineering,* R. Glowinski, S. J. Osher, and W. Yin, eds. (Springer International Publishing, Cham, 2016), pp. 345–407.
32. L. A. Vese and C. Le Guyader, *Variational methods in image processing* (CRC Press Boca Raton, FL, 2016).
33. J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition–modeling, algorithms, and

parameter selection," Int. J. Comput. Vis. **67**, 111–136 (2006).
34. A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," J. Math. Imaging Vis. **40**, 120–145 (2011).
35. J. Fehrenbach, P. Weiss, and C. Lorenzo, "Variational algorithms to remove stationary noise: Applications to microscopy imaging," IEEE Trans. on Image Process. **21**, 4420–4430 (2012).
36. P. Escande, P. Weiss, and W. Zhang, "A variational model for multiplicative structured noise removal," J. Math. Imaging Vis. **57**, 43–55 (2017).
37. Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003,* vol. 2 (2003), pp. 1398–1402.
38. Z. Cunha, "Nonsubsampled contourlet toolbox," MATLAB Central File Exchange. Retrieved June 7, 2023.
39. P. Weiss, "VSNR: Variational stationary noise remover. examples of applications," Accessed 19.01.2024, https://www.math.univ-toulouse.fr/~weiss/Codes/VSNR/VNSR_VariationalStationaryNoiseRemover.html.
40. P. Y. Liu, L. K. Chin, W. Ser, *et al.*, "Cell refractive index for cell biology and disease diagnosis: past, present and future," Lab Chip **16**, 634–644 (2016).

A. LSFM Simulation

The creation of synthetic LSFM data such as shown in Figure 15 requires the simulation of light propagation through media to capture scattering and absorption effects correctly. The python package *biobeam* [18] provides the necessary functions for calculating light propagation through a volume of media. The geometric properties of the volume are specified through a distribution of refractive indices (rid) and a fluorescence distribution (fld). The rid is denoted by $\mathbf{n} \in \mathbb{C}^{n_x \times n_y \times n_z}$ and contains complex valued refractive indices in each pixel. The real part is the classical refractive index and describes scattering behaviour while the imaginary part is the absorption coefficient which describes effects such as light attenuation. For generating reference images without light interaction we set $\mathbf{n} \equiv n_0$ to be constant. A reasonable choice is $n_0 = 1.33$ which is approximately the refractive index of water which is the main constituent of biological cells.

The fld is denoted by $f \in \mathbb{R}_{\geq 0}^{n_x \times n_y \times n_z}$. It contains multiplicative non-negative factors that signify the proportion of re-emitted light by fluorescence when subjected to illumination. The fld is a simplified representation of reality. Typically, biological samples are labeled for LSFM using a fluorescent marker, which are specialized proteins that emit light when illuminated at certain wavelengths. The intensity of the emitted fluorescence is linked to the concentration of marker protein within the sample.

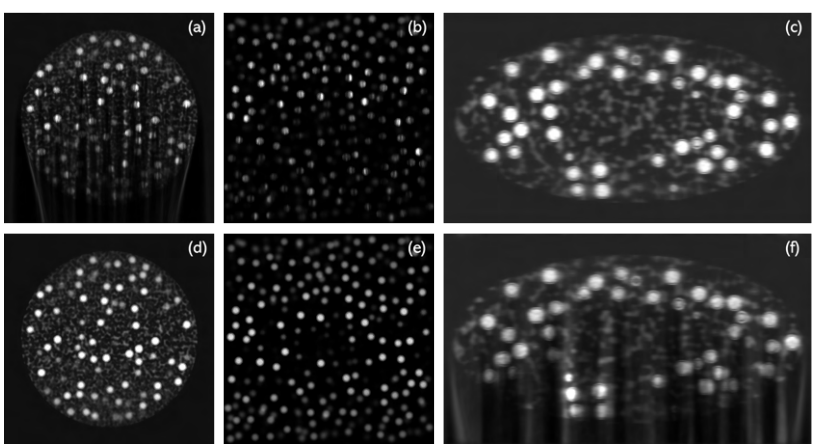

Fig. 15. Simulated images (a-c) and references (d-f) for different simulation examples. Cell cluster (a,d), dispersed cells (b,e) and embryo (c,f).

Besides geometrical properties of a synthetic sample, we have to choose parameters of

the optical setup. We chose the wavelength of illumination light $\lambda = 500$nm, the numerical apertures of illumination and detection optics $\mathrm{NA}_{\mathrm{ill}} = 0.1$, $\mathrm{NA}_{\mathrm{det}} = 0.4$ and the voxel size $v = [0.4, 0.4, 0.4]\mu$m. The corresponding values were fixed for all synthetic image data in this paper. Furthermore, we chose the illumination direction along the y-axis and the detection direction along the z-axis.

With a given setup of geometry and imaging, simulations of synthetic LSFM images can be obtained. The following steps model the real imaging process:

(i) **Illumination:** From $\mathrm{NA}_{\mathrm{ill}}$ and λ a complex-valued x-z cross-section of a cylindrical light sheet is calculated. Its height $\bar{z}$ corresponds to the illuminated slice in the sample and can be chosen arbitrarily. With wave-optical propagation as described in [7], the cross-section is propagated through the sample. This yields the illuminated volume I_{ill}.

(ii) **Fluorescence:** Afterwards, the amount of light re-emitted through fluorescence is obtained via

$$I_{\mathrm{fl}} = I_{\mathrm{ill}} \odot f \tag{11}$$

with element-wise multiplication $\odot$. It represents the 'fluorescence response' to the illumination.

(iii) **Acquisition:** From $\mathrm{NA}_{\mathrm{det}}$, a detection point-spread function (PSF) h_{det} is calculated. It is invariant under rotations around the detection axis. A 2D image of the illuminated slice is obtained by convolution with this PSF and extraction of the corresponding slice

$$I_{\mathrm{det}} = (I_{\mathrm{fl}} * h_{\mathrm{det}})_{z=\bar{z}}. \tag{12}$$

(iv) Repeating steps (i)-(iii) for light sheet heights $\bar{z} \in \{1, \ldots, n_z\}$, slices are combined into a 3D volume image.

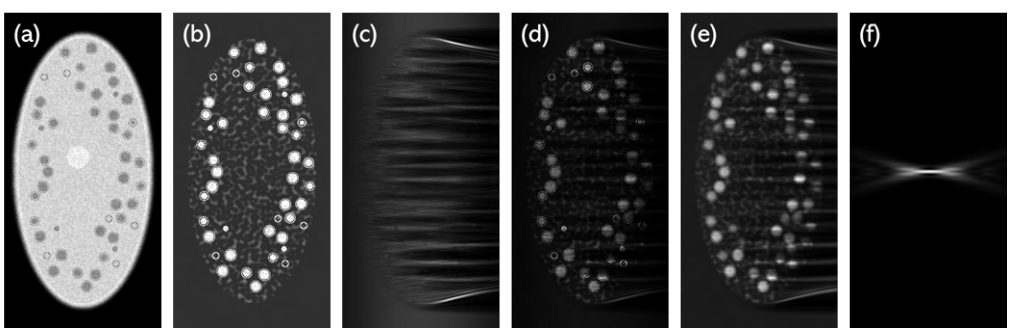

Fig. 16. Visualization of the simulation procedure. |rid| (a), fld (b), I_{ill} (c), I_{fl} (d), I_{det} (e), h_{det} (f). (a)-(e) are y-x-cross-sections of the illuminated slice extracted from the volume. (f) is a z-x-cross-section of h_{det} with 4x magnification compared to (a)-(e).

The essential parts of the simulation procedure are visualized in Figure 16. This process is a simplified model of reality. For example, we use idealized versions of the optical systems. Therefore, the cylindrical light sheet is of constant thickness along the x-direction and the detection PSF remains constant for all pixels. In real systems, these approximations are reasonable for the area of interest but do not hold true further away from it.

The synthetic data shown in Figure 15 depicts some examples of simulated LSFM images with an appearance similar to real data. All samples were created by using non-overlapping spheres as main foreground. In addition, small variations inside of the spheres were created by adding fine

scale 3D Perlin noise. For some structural background a coarse scale 3D Perlin noise was added, except for the dispersed cells image. Values of the rid for different structures were chosen based on reference values [40]. On the other hand, values for the fld were chosen to produce visually appealing images which are in accordance with real LSFM images.

B. Parameter Selection for GSR

In this section of the appendix, we want to give insight into the parameters of the GSR algorithm. More precisely, we want to give advice on how parameters can be selected and tuned for an image. Therefore, we discuss the behaviour of varying μ_1, μ_2 and propose two sets of parameters for producing an initial result depending on the problem at hand. Fine-tuning towards a more desirable result may still be necessary and can then be done according to our following observations.

Figure 17 displays the effects of varying μ_1 and μ_2 for a selection of parameters with reasonable outcomes. We notice that the rules of thumb established in Subsection 2.2 can be observed here. More precisely, an increase in μ_1 corresponds to a stronger reduction in stripes which can lead to smoothing and smearing artifacts as most dominantly shown in (h). On the other hand, increasing μ_2 reduces the amount of stripe reduction and counteracts these kinds of artifacts. However, scaling of both μ_1 and μ_2 by an equal factor, e.g., compare (b) and (i), produces different results. In our experience, scaling with a factor > 1 increases stripe removal with minimal effect on introducing artifacts. Nonetheless, this is only valid in a reasonable range of parameters which we propose to be approximately $\mu_1 \in [0.15, 0.5]$ and $\mu_2 \in [0.0033, 0.017]$. Finding 'optimal' results requires some initial guess followed by manual fine-tuning. In all considered cases choosing either $(\mu_1, \mu_2) = (1/3, 1/300)$ or $(\mu_1, \mu_2) = (1/2, 1/30)$ produced sufficiently good results. The larger parameters usually perform better when stripe artifacts appear in a more difficult setting, e.g., when they are short, wide or slightly oblique. In general, $(\mu_1, \mu_2) = (1/3, 1/300)$ yielded superior results when stripes were thin, long or had smoothed out edges. In our case, fine-tuning was rarely necessary and can be performed according to the previous observations and the rules of thumb in Subsection 2.2. If more stripe removal is wanted, scaling both parameters by an equal factor is a good start. Alternatively, increasing only μ_1 can achieve the same effect. If image structures are too strongly affected, choosing a larger μ_2 typically resolves the issue.

Funding

Federal Ministry of Education and Research (BMBF), Project: Synthetic Data for Machine Learning Segmentation of Highly Porous Structures from FIB-SEM Nano-tomographic Data (poSt), Funding number: 01IS21054A

Acknowledgments

We thank Martin Weigert for guidance on installing biobeam and modelling LSFM imaging with it and Michael Engstler for providing the FIB-SEM image data. Furthermore, we want to thank Diego Roldán for providing and explaining his proposed quality metrics and Jannik Reiser for his help with processing results.

Disclosures

The authors declare no conflicts of interest.

Data Availability Statement

The data and methods presented in this paper are available as supplementary material. This includes MATLAB code for all used methods and scripts for processing images. Furthermore,

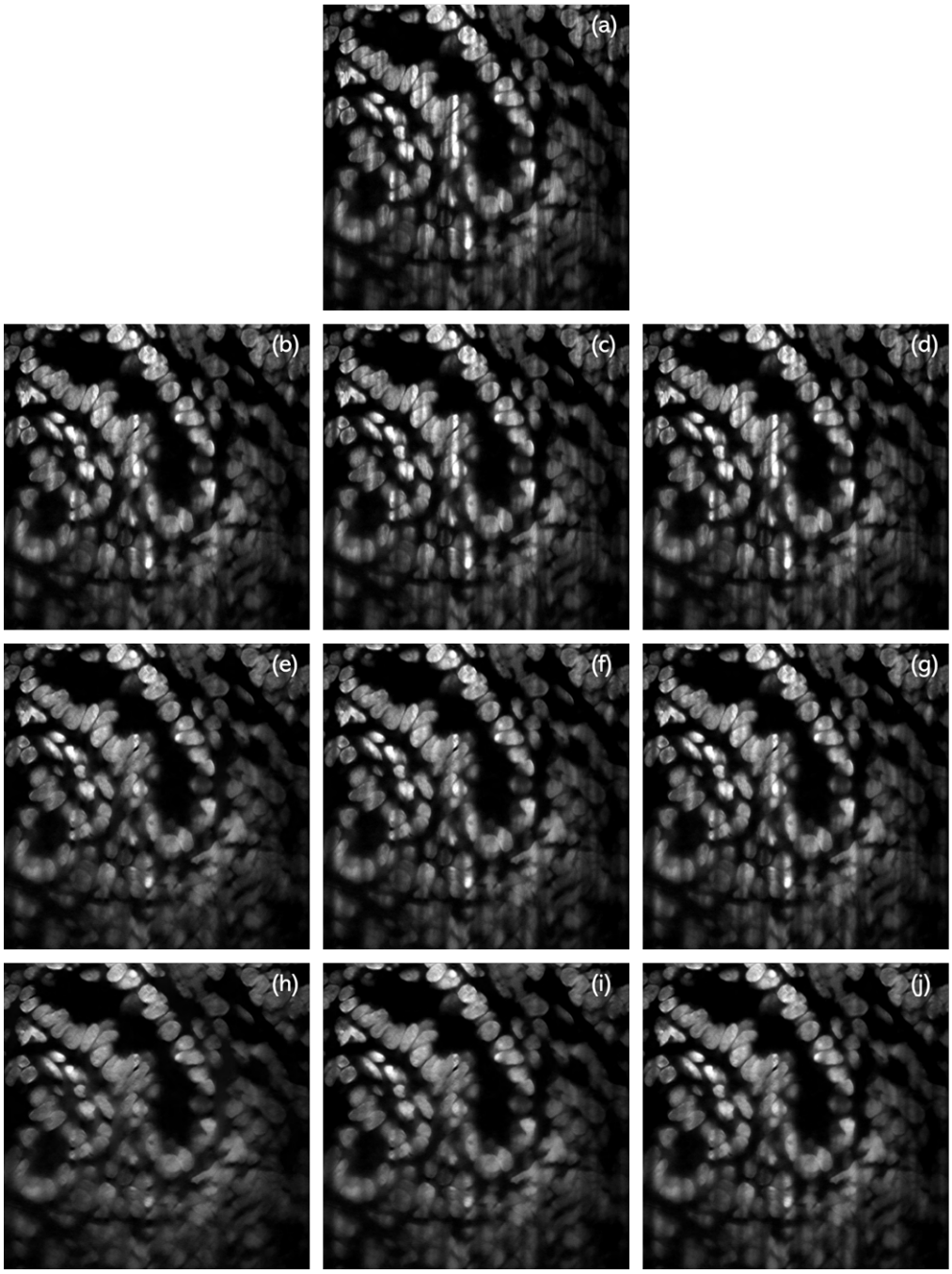

Fig. 17. Results of GSR(2D) on the mouse intestine cuboid with different parameter settings. Image (a), $\mu = (\frac{1}{6}, \frac{1}{300})$ (b), $\mu = (\frac{1}{6}, \frac{1}{100})$ (c), $\mu = (\frac{1}{6}, \frac{1}{60})$ (d), $\mu = (\frac{1}{3}, \frac{1}{300})$ (e), $\mu = (\frac{1}{3}, \frac{1}{100})$ (f), $\mu = (\frac{1}{3}, \frac{1}{60})$ (g), $\mu = (\frac{1}{2}, \frac{1}{300})$ (h), $\mu = (\frac{1}{2}, \frac{1}{100})$ (i), $\mu = (\frac{1}{2}, \frac{1}{60})$ (j).

a Jupyter-notebook for generating synthetic LSFM data and a txt-file for creating the Python environment on Windows 11 are available.