Temporal and Volumetric Denoising via Quantile Sparse Image Prior

Franziska Schirrmacher[1,*], Thomas Köhler[1,4,*], Jürgen Endres[1,2], Tobias Lindenberger[1], Lennart Husvogt[1], James G. Fujimoto[3], Joachim Hornegger[1], Arnd Dörfler[2], Philip Hoelter[2], and Andreas K. Maier[1]

[1] *Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*
[2] *Department of Neuroradiology, Universitätsklinikum Erlangen, Germany*
[3] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, USA*
[4] *e.solutions GmbH, Erlangen, Germany*

* These authors contributed equally to this work.

Abstract

This paper introduces an universal and structure-preserving regularization term, called quantile sparse image (QuaSI) prior. The prior is suitable for denoising images from various medical imaging modalities. We demonstrate its effectiveness on volumetric optical coherence tomography (OCT) and computed tomography (CT) data, which show different noise and image characteristics. OCT offers high-resolution scans of the human retina but is inherently impaired by speckle noise. CT on the other hand has a lower resolution and shows high-frequency noise. For the purpose of denoising, we propose a variational framework based on the QuaSI prior and a Huber data fidelity model that can handle 3-D and 3-D+t data. Efficient optimization is facilitated through the use of an alternating direction method of multipliers (ADMM) scheme and the linearization of the quantile filter. Experiments on multiple datasets emphasize the excellent performance of the proposed method.

Keywords: spatio-temporal denoising, variational approach, QuaSI prior, ADMM

1. Introduction

The reliable reduction of image noise poses a constantly recurring problem in todays imaging systems. In healthcare, noise may limit the reliability of medical image data for subsequent clinical workflows. For instance, in radiology using computed tomography (CT) or related morphological imaging modalities, noise affects the analysis of anatomical structures and thus impedes diagnostic applications. In optical coherence tomography (OCT) for retinal imaging as another example use case, noise limits the measurement of structural features in the human eye, e.g. retinal layer properties. Apart from diagnostic applications, noise reduction is also a major theme for different interventional imaging modalities like fluoroscopically guided procedures. Low dose radiation exposure for patient safety leads to noisy and low-contrast fluroscopic sequences (Amiot et al., 2016).

To mitigate these limitations, *denoising* can be either implemented by means of customized hardware or via postprocessing of captured image data. While hardware-based denoising often leads to increased system complexities, image-based postprocessing facilitates denoising in a cost-effective way using computational methods. Despite the great progress in developing general denoising schemes for natural images, adopting them for medical data poses several challenges. First and foremost, there is a narrow ridge between achieving sufficient noise reduction and unwanted distortions of meaningful medical structures. Moreover, noise distributions in medical data often deviate from the commonly employed models for natural images like additive, white Gaussian noise (AWGN). For example, noise can follow multiplicative models or structured patterns related to acquisition parameters like in CT. General denoising algorithms have been mainly developed for 2-D data, e.g. color photographs, but denoising in medical imaging also needs to handle time-resolved and/or volumetric data. These requirements desire enhanced and robust denoising methods to be applicable within medical workflows.
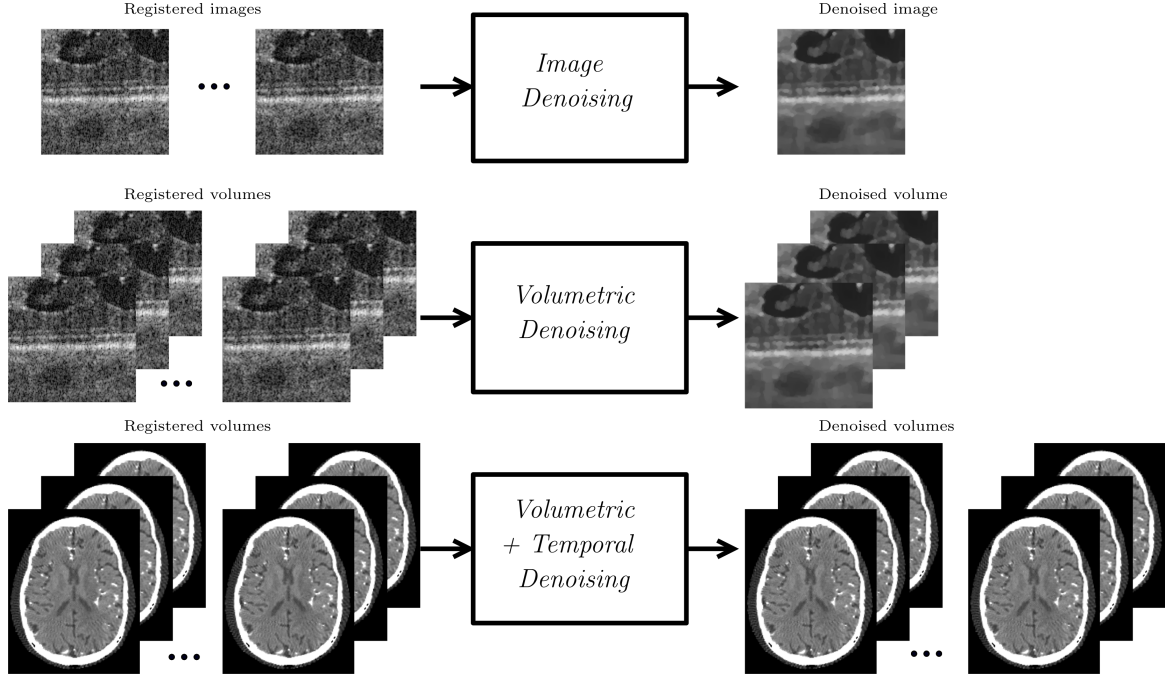
Figure 1: We propose three modi of our spatio-temporal denoising algorithm. In the first modus (top), hereinafter called *image denoising*, single images or a sequence of registered images are processed. The second modus (middle) processes volumes as well as a sequence of registered volumes and is called *volumetric denoising*. The third modus (bottom) processes volumes as well as a sequence of registered volumes, outputs a sequence of volumes, and is called *volumetric + temporal denoising*.

In this paper, we propose denoising for medical image data within a variational framework. As the key contribution, we introduce the class of *quantile sparse image* (QuaSI) priors to model the appearance of noise-free medical data. Specifically, we propose a median filter based regularizer that is based on the QuaSI prior using the 0.5 quantile. This follows the idea that noise-free data should be a fixed point of the median filter and we show that this model facilitates structure-preserving denoising. To approach the resulting non-linear and non-convex optimization problem, we present an alternating direction method of multipliers (ADMM) scheme. Our algorithm can handle *spatio-temporal* denoising by processing either single images or sequences of consecutive images. Furthermore, it enables denoising of volumetric data. Thus, it can be adjusted to the clinical needs within a target application.

This paper is an extension of our prior work in Schirrmacher et al. (2017) and makes the following additional contributions:

- The algorithm as well as the QuaSI prior are extended to process volumetric medical data.

- An investigation of the convergence and parameter sensitivity of our algorithm is conducted.

- An extension of our algorithm is presented to process volumetric data in C-arm CT imaging.

The remainder of this paper is organized as follows. In Section 2, we review related work on spatial and temporal denoising. Section 3 comprises the objective function of the energy minimization problem. In Section 4 the QuaSI prior is introduced. The numerical optimization of our denoising framework is derived in Section 5. In Section 6, an experimental evaluation of our method on publicly available benchmark data, clinical OCT scans as well as CT data is reported. Finally, section 7 contains our conclusion.

2. Related Work

The image-based denoising techniques can be divided into two groups.

2.1. Spatial Denoising Methods

Spatial or *single-image* denoising has been extensively studied in the image processing community and various approaches emerged over the past
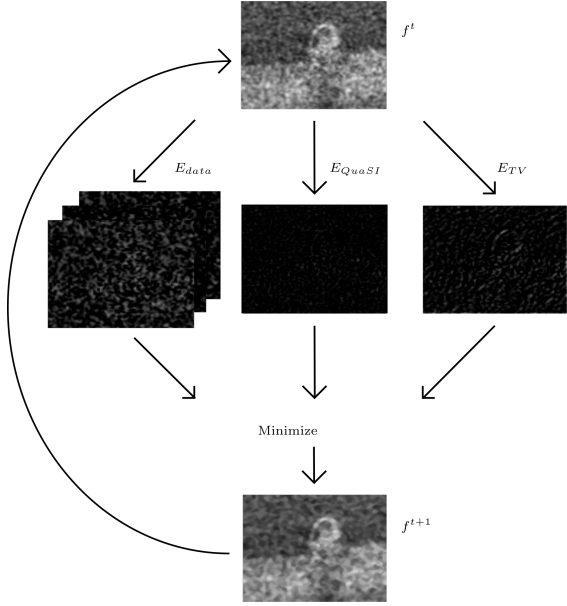
Figure 2: Method overview: The proposed spatio-temporal denoising algorithm is based on an energy minimization formulation with three terms.

decades. Local image filters perform smoothing of noisy images possibly in an adaptive way to preserve image structures (Tomasi and Manduchi, 1998). Non-local filtering also exploits the statistics of similar and repeating patches within images. One representative from this class is the successful BM3D method by Dabov et al. (2007). However, these methods have been mainly designed for natural images under simplified assumptions like additive white Gaussian noise, which is inappropriate to describe speckle noise that is multiplicative in nature. Learning-based denoising, e. g. based on multilayer neural networks (Burger et al., 2012), hold the potential to handle speckle noise by learning noise distributions from training data. However, large-scale training data required for such methods is barely available for OCT.

Some spatial filters that have been adopted for OCT denoising are the hybrid median filter, Lee filter, Wiener filter, or wavelet thresholding as investigated by Ozcan et al. (2007). Global denoising methods for OCT have been introduced by Salinas and Fernandez (2007) using non-linear diffusion and later by Duan et al. (2016) using second-order total generalized variation. Wong et al. (2010) have proposed structure-adaptive Bayesian estimation to handle speckle noise. One interesting approach has been proposed by Fang et al. (2012), where dictionary learning based on B-scans with high signal-to-noise ratio (SNR) is used to denoise low SNR B-sans.

Single-image denoising offers great flexibility in clinical applications of OCT as few assumptions on the scanning protocol are made. However, the noise reduction is limited as such methods can utilize single B-scans only.

2.2. Temporal Denoising Methods

Temporal or *multi-image* denoising methods consider coherence of consecutive images to improve noise reduction over single-image denoising. Such methods have been widely investigated for OCT and exploit sets of B-scans that are acquired sequentially from the same location or nearby positions. A popular approach in commercial systems is to register multiple of these B-scans and to average the registered scans to cancel out random noise. Averaging is computationally efficient but requires many repetitive acquisitions to effectively reduce speckle noise. Mayer et al. (2012) enhance simple averaging based on wavelet decompositions of B-scans to estimate local image structures and noise. Denoising is conducted in the wavelet domain by weighted averaging of wavelet coefficients according to the local image structure. Cheng et al. (2014) formulate OCT denoising from multiple scans as a low-rank matrix completion problem. Thapa et al. (2015) follow a similar notion and exploit the low-rank property on a patch-based level of multiple B-scans using weighted nuclear norm minimization. L. Bian and Dai (2015) have proposed inter-frame and intra-frame priors for denoising using convex optimization. BM4D is an extension of the popular BM3D method to process volumetric data (Maggioni et al., 2013).

All of these multi-image methods have in common that they require multiple input scans. This increases the overall acquisition time and therefore might lead to a higher patient discomfort. Also, they perform denoising on a B-scan level but ignore coherence of nearby B-scans within volumetric OCT data. If denoising of entire volumes is desired, simple consecutive processing of individual B-scans can lead to suboptimal results. In this paper, we mitigate both limitations by proposing a unified approach to handle denoising on a B-scan or volumetric level based on single or multiple scans.

3. Background

This section presents the variational framework for denoising volumetric data. Figure 1 illustrates three modi of this framework, namely image denoising, volumetric denoising, and volumetric+temporal denoising. The pipelines differ in the number of outputs and are therefore divided into multiple-input single-output (MISO) denoising and multiple-input multiple-output (MIMO) denoising.

Throughout this paper, we use the following nomenclature. We denote a volume as a vector $\mathbf{g} \in \mathbb{R}^{N_z N_{xy}}$ composed of N_z images $\mathbf{g}_z$, $z = 1, \ldots, N_z$ of size $N_{xy} = N_x N_y$ pixels. For the sake of convenience, 2-D images of size $N_x \times N_y$ are reshaped to vector notation using a row-wise scanning. A sequence of volumes is denoted as vector $\mathbf{G} \in \mathbb{R}^{N_t N_z N_{xy}}$, where N_t is the number of volumes in the sequence. The input to the proposed framework is a sequence of T volumes, where $1 \leq T \leq N_t$. For volumetric as well as volumetric+temporal denoising, we employ Z consecutive images per volume ($1 < Z \leq N_z$), while image denoising is based on a single image in each volume ($Z = 1$).

3.1. Noise Model

In this paper, we consider several denoising applications with two different underlying noise models. In an *additive* noise model, a noise-free volume $\mathbf{f} = (\mathbf{f}_1, \ldots, \mathbf{f}_Z)^\top$ is related to a noisy volume $\mathbf{g} = (\mathbf{g}_1, \ldots, \mathbf{g}_Z)^\top$ according to:

$$\mathbf{g} = \mathbf{f} + \mathbf{n}, \tag{1}$$

where $\mathbf{n} = (\mathbf{n}_1, \ldots, \mathbf{n}_Z)^\top$ denotes an additive noise term. Common instances of this model are AWGN with stationary distribution of $\mathbf{n}$ or Poisson noise, where the variance of $\mathbf{n}$ depends on the measured image data.

In a *multiplicative* noise model, each captured volume $\mathbf{g}$ is related to a respective noise-free volume $\mathbf{f}$ according to:

$$\mathbf{g} = \mathbf{f} \odot \mathbf{n}, \tag{2}$$

where $\odot$ is the Hadamard (element-wise) product. We can turn the multiplicative model in (2) to the additive one in (1) by transforming it to a logarithmic measurement domain. One common instance of this model is speckle noise that appears in OCT imaging (Wong et al., 2010; Duan et al., 2016).

3.2. Energy Minimization Formulation

Given a sequence of T volumes $\mathbf{g}^{(t)}$ with $t = 1, \ldots, T$ that are either captured from the same position or from nearby positions and registered to each other, we propose MIMO and MISO denoising.

In MISO denoising, we aim at estimating one noise-free volume $\hat{\mathbf{f}}$. We formulate denoising as the minimization of the objective function:

$$\hat{\mathbf{f}} = \underset{\mathbf{f}}{\operatorname{argmin}} \sum_{t=1}^{T} \rho\big(\mathbf{f} - \mathbf{g}^{(t)}\big) + \lambda R_{\mathrm{QuaSI}}(\mathbf{f}) + \mu \|\nabla \mathbf{f}\|_1. \tag{3}$$

The first term in (3) denotes the data fidelity of $\mathbf{f}$ w.r.t. the input volumes $\mathbf{g}^{(t)}$. The second term is the proposed quantile sparse image (QuaSI) prior weighted by $\lambda \geq 0$. The third term denotes anisotropic total variation (TV) weighted by $\mu \geq 0$, which regularizes the spatial gradient $\nabla \mathbf{f} = (\nabla_x \mathbf{f}, \nabla_y \mathbf{f}, \nabla_z \mathbf{f})^\top$. It is worth noting that the general denoising framework in (3) can handle both noise reduction for entire volumes in 3-D as well as for individual images in 2-D by constraining the domain of both regularization terms.

MIMO denoising follows a similar approach but aims at estimating a sequence of volumes $\hat{\mathbf{F}}$. We formulate MIMO denoising as the minimization of the objective function:

$$\begin{aligned} \hat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \rho\big(\mathbf{F} - \mathbf{G}\big) + \lambda R_{\mathrm{QuaSI}}(\mathbf{F}) + \mu \|\nabla \mathbf{F}\|_1 \\ + \omega \|\nabla_t \mathbf{F}\|_1, \end{aligned} \tag{4}$$

where $\nabla_t \mathbf{F}$ denotes the gradient of $\mathbf{F}$ in temporal direction and the associated TV regularization is weighted by $\omega \geq 0$.

In (3) and (4), the data fidelity terms use the loss function $\rho : \mathbb{R}^N \to \mathbb{R}_0^+$ to formulate the image formation. In general, the image formation needs to consider a mixture of noise, potential misalignments between the input volumes, or motion artifacts. Following prior work on mixed noise models in image restoration (Köhler et al., 2016), we propose to use the Huber loss (Ochs et al., 2015):

$$\rho(\mathbf{l}) = \sum_{i=1}^{N} \phi(l_i), \tag{5}$$

where:

$$\phi(l) = \begin{cases} \frac{1}{2} l^2 & \text{if } l \leq \epsilon \\ \epsilon \left(|l| - \frac{1}{2}\epsilon\right) & \text{otherwise,} \end{cases} \tag{6}$$

and $\epsilon > 0$ denotes the threshold of the Huber loss. This leads to an outlier-insensitive model while the underlying data fidelity is a convex term.

(a) Noisy B-scan $\mathbf{f}_{noisy}$

(b) Gold standard B-scan $\mathbf{f}_{gold}$

(c) $\mathbf{r}_{noisy}$

(d) $\mathbf{r}_{gold}$

(e) Histogram of $\mathbf{r}_{noisy}$
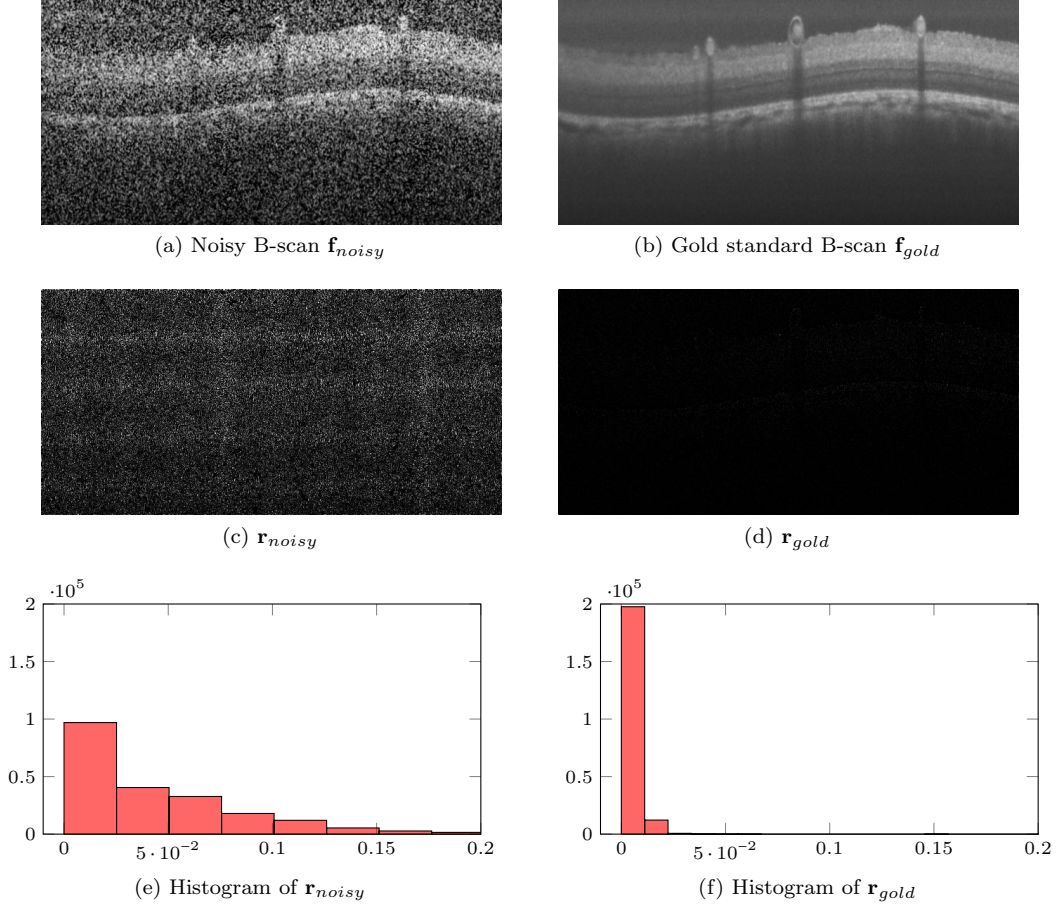
(f) Histogram of $\mathbf{r}_{gold}$

Figure 3: Analysis of our proposed QuaSI prior using median filtering $Q(\cdot)$ to model the appearance of OCT B-scans. (a) and (b) depict a noisy B-scan along with the respective gold standard taken from the pig eye dataset Mayer et al. (2012). (c) and (d) show the residual $\mathbf{r} = \mathbf{f} - Q(\mathbf{f})$ of the QuaSI regularization term, where brighter pixels express higher residuals (contrast enhanced for visualization). (e) and (f) depict the corresponding histograms of the both residuals, where the histogram for the gold standard is sparse. Our QuaSI prior exploits the sparsity of $\mathbf{r} = \mathbf{f} - Q(\mathbf{f})$ for regularization in our variational denoising framework.

4. Quantile Sparse Image (QuaSI) Prior

A robust and efficient regularization term is of importance to achieve results with a high signal-to-noise ratio (SNR). The better the regularization term is able to model natural or medical images, the better the result of the optimization. Structure preservation is a sensitive issue when dealing with medical data. The images might contain small morphological structures that need to be preserved for the purpose of diagnosis. In order to tackle the challenges referred to above, the so called quantile sparse image (QuaSI) prior is introduced.

4.1. Definition of the Prior

The QuaSI prior is based on quantile filtering, where the quantile filter is denoted as $\tilde{\mathbf{f}} = Q(\mathbf{f})$. The p-quantile with $p \in [0,1]$ is determined within a local neighborhood $\mathcal{N}(i)$. The local neighborhood consists of d^3 voxel, where d denotes the width of the cubic filter kernel. For the i-th voxel in $\mathbf{f}$ we filter according to $\tilde{f}_i = \text{quantile}_{\mathcal{N}(i)}(f_i, p)$. Inspired by the regularization by denoising priors by Romano et al. (2016), the denoised volume is a fixed point under the quantile filter. In this way:

$$R_{\text{QuaSI}}(\mathbf{f}) = \|\mathbf{f} - Q(\mathbf{f})\|_1. \qquad (7)$$

Specifically, regularization according to (7) enforces sparsity of the residual $\mathbf{f} - Q(\mathbf{f})$. This offers a gen-
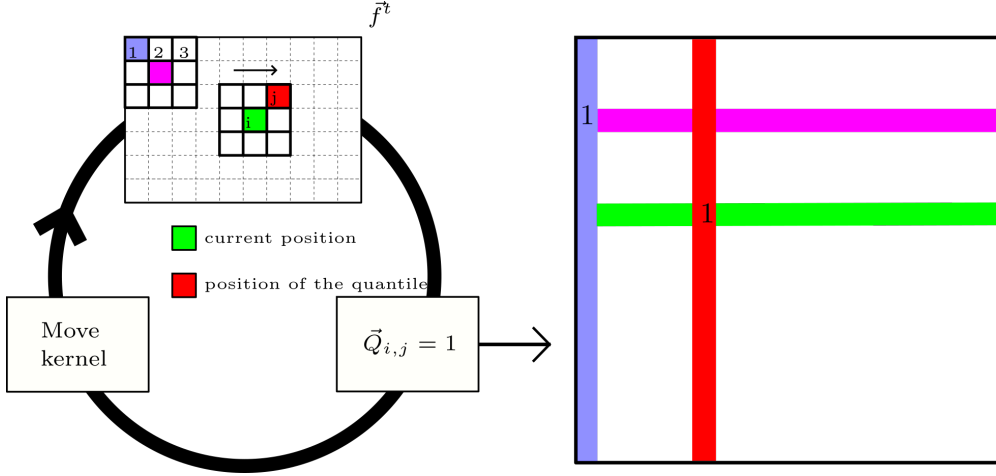
Figure 4: Construction of the binary matrix to approximate the quantile filter $Q(\mathbf{f}) = \mathbf{Q}\mathbf{f}^t$.

eral model for regularization and – depending on the application – various types of statistics can be chosen for $Q(\mathbf{f})$. In this paper, we propose the median filter, where $\tilde{f}_i = \text{median}_{\mathcal{N}(i)}(f_i)$. This follows the rationale that median filtering facilitates structure-preserving denoising under non-Gaussian noise. Further applications including erosion and dilation are not covered in this paper. In the literature (Rohkohl, 2011), quantiles are used to obtain a reference image to estimate non-periodic motion. Those examples are suitable applications that the QuaSI prior can handle.

To validate the QuaSI prior using median filter regularization for denoising, we study its behavior under real measurement noise. For this purpose, we use the publicly available pig eye dataset by Mayer et al. (2012), which provides a gold standard OCT B-scan obtained from the average of 455 registered noisy OCT B-scans. We compare a noisy OCT B-scan $\mathbf{f}_{noisy}$ with the gold standard $\mathbf{f}_{gold}$ in Fig. 3a and Fig. 3b. The residuals $\mathbf{r} = \mathbf{f} - Q(\mathbf{f})$ of the QuaSI regularization term are illustrated in Fig. 3c for the noisy B-scan and in Fig. 3d for the gold standard. Compared to the gold standard, the noisy B-scan yields a less sparse signal as shown in the histograms of both residuals in Fig. 3e and Fig. 3f. Notice that the QuaSI regularization does not penalize image discontinuities. The histogram using the noisy B-scan contains less zero elements, while the histogram for the gold standard is sparse. Our proposed QuaSI prior exploits these observations for structure-preserving regularization in our variational denoising framework.

4.2. Linearization

In order to deal with the non-linearity of the quantile operator $Q(\mathbf{f})$ the linearization $Q(\mathbf{f}) = \mathbf{Q}\mathbf{f}$, similar to the work of Pan et al. (2016), is performed. The binary matrix $\mathbf{Q}$ is assembled element-wise according to:

$$Q_{ij} = \begin{cases} 1 & \text{if } j = q, \\ 0 & \text{otherwise}, \end{cases} \qquad (8)$$

where $q = \arg\text{quantile}_{r \in \mathcal{N}(i)} f_r$. This operation filters the i-th pixel according to the p-quantile in its local neighborhood $\mathcal{N}(i)$. For $\mathbf{f}' = \mathbf{f}$ the linearization fullfills $Q(\mathbf{f}') = \mathbf{Q}\mathbf{f}'$, while otherwise $\mathbf{Q}$ serves as an approximation of the quantile filter.

Figure 4 illustrates the construction of the binary matrix $\mathbf{Q}$ in 2-D. Each pixel is replaced by the quantile within its local neighborhood. The position of the quantile is stored in the binary matrix. In this example, the quantile is at position j. Thus, the i-th row of the matrix contains a one in the j-th column and zeros otherwise. The multiplication $\mathbf{Q}\mathbf{f}$ yields the quantile filtered result.

5. Deploying QuaSI for Denoising

In this section, we show how the proposed QuaSI prior can be deployed for volumetric and temporal denoising. We derive two numerical optimization algorithms for denoising based on a MISO and a MIMO mode.

6

5.1. Multiple-Input Single-Output (MISO) Mode

MISO denoising in our framework is based on the energy minimization formulation in (3). In order to handle the non-smooth L_1 norm terms, we adopt ADMM optimization (Goldstein and Osher, 2009). To this end, (3) is reformulated to the constrained optimization problem:

$$\hat{\mathbf{f}} = \underset{\mathbf{f}}{\operatorname{argmin}} \sum_{t=1}^{T} \rho\left(\mathbf{f} - \mathbf{g}^{(t)}\right) + \lambda \|\mathbf{u}\|_1 + \mu \|\mathbf{v}\|_1 \quad (9)$$
$$\text{such that} \quad \mathbf{u} = \mathbf{f} - Q(\mathbf{f}), \quad \mathbf{v} = \nabla \mathbf{f},$$

where $\mathbf{u}$ and $\mathbf{v}$ are auxiliary variables. Then, an unconstrained optimization problem is obtained from (9) using quadratic penalty functions according to:

$$\hat{\mathbf{f}} = \underset{\mathbf{f}}{\operatorname{argmin}} \sum_{t=1}^{T} \rho\left(\mathbf{f} - \mathbf{g}^{(t)}\right) + \mu \|\mathbf{v}\|_1 + \lambda \|\mathbf{u}\|_1$$
$$+ \frac{\alpha}{2} \|\mathbf{u} - \mathbf{f} + Q(\mathbf{f})\|_2^2 + \frac{\beta}{2} \|\mathbf{v} - \nabla \mathbf{f}\|_2^2. \quad (10)$$

The Lagrangian multipliers $\alpha > 0$ and $\beta > 0$ enforce the constraints $\mathbf{u} = f - Q(\mathbf{f})$ and $\mathbf{v} = \nabla \mathbf{f}$. If $\alpha, \beta \to \infty$, we end up at the original problem (3). In order to strictly enforce the constraint, the Bregman variables $\mathbf{b}_u$ and $\mathbf{b}_v$ are introduced. Then, we minimize the augmented Lagrangian:

$$\mathcal{L}_{\text{AL}}(\mathbf{f}, \mathbf{u}, \mathbf{v}, \mathbf{b}_u, \mathbf{b}_v) = \sum_{t=1}^{T} \rho\left(\mathbf{f} - \mathbf{g}^{(t)}\right)$$
$$+ \frac{\alpha}{2} \|\mathbf{u} - \mathbf{f} + Q(\mathbf{f}) - \mathbf{b}_u\|_2^2 + \lambda \|\mathbf{u}\|_1 \quad (11)$$
$$+ \frac{\beta}{2} \|\mathbf{v} - \nabla \mathbf{f} - \mathbf{b}_v\|_2^2 + \mu \|\mathbf{v}\|_1.$$

We iteratively optimize (11) by alternating minimization w.r.t. the individual parameters. Hence, three subproblems emerge, where the L_1-Norm is decoupled from the L_2-Norm.

The minimization of the augmented Lagrangian (11) w.r.t. $\mathbf{f}$ can be solved in a least square sense. Therefore, the binary matrix $\mathbf{Q}$ is constructed using the result $\mathbf{f}^k$ from the previous iteration, where k denotes the iteration index. In order to cope with the Huber loss, iteratively re-weighted least squares (IRLS) is applied. Solving the resulting least squares

problem leads to the linear system:

$$\mathbf{A}\mathbf{f}^{k+1} = \mathbf{b} \quad (12)$$
$$\mathbf{A} = \sum_{t=1}^{T} \mathbf{W}^{(t)} + \beta \nabla^\top \nabla + \alpha \mathbf{M}^\top \mathbf{M} \quad (13)$$
$$\mathbf{b} = \sum_{t=1}^{T} \mathbf{W}^{(t)} \mathbf{g}^{(t)} \quad (14)$$
$$+ \beta \nabla^\top (\mathbf{v} - \mathbf{b}_v) + \alpha \mathbf{M}^\top (\mathbf{u} - \mathbf{b}_u),$$

where $\mathbf{M} = \mathbf{I} - \mathbf{Q}$ with the identity matrix $\mathbf{I}$. In (12) - (14), $\mathbf{W}^{(t)}$ are diagonal weight matrices constructed from $\mathbf{f}^k$. Using the intermediate result $\mathbf{f}^k$, we can compute the weights for IRLS according to:

$$W_{ii}^{(t)} = \frac{\phi'\left(f_i^k - g_i^{(t)}\right)}{\left|f_i^k - g_i^{(t)}\right|}, \quad (15)$$

where $\phi'(l)$ is the derivative of the Huber loss. The threshold of the Huber loss is set to $\epsilon = 1.345\sigma$ to achieve a 95-percent efficiency of the estimator under Gaussian noise (Ochs et al., 2015). We use the median absolute deviation (MAD) rule to obtain a consistent estimate of the standard deviation according to $\sigma = 1.4826 \cdot \text{MAD}(f_i^k - g_i^{(t)})$ (Rousseeuw and Leroy, 1987). To solve the linear system (12), conjugate gradient (CG) iterations are used.

The minimization of the augmented Lagrangian (11) w.r.t. the auxiliary variables can be done by exploiting the separability of the problem. Given the estimate for the intermediate result $\mathbf{f}^{k+1}$, this leads to the element-wise updates:

$$u_i^{k+1} = \text{shrink}([\mathbf{f}^{k+1} - \mathbf{Q}\mathbf{f}^{k+1} + \mathbf{b}_u^k]_i, \lambda/\alpha), \quad (16)$$
$$v_i^{k+1} = \text{shrink}([\nabla \mathbf{f}^{k+1} + \mathbf{b}_v^k]_i, \mu/\beta)., \quad (17)$$

where $\text{shrink}(z, \gamma) = \text{sign}(z) \max(z - \gamma, 0)$ denotes the shrinkage operator (Goldstein and Osher, 2009).

Given an estimate for the intermediate result $\mathbf{f}^{k+1}$ as well as the auxiliary variables $\mathbf{u}^{k+1}$ and $\mathbf{v}^{k+1}$, the Bregman variables are updated according to:

$$\mathbf{b}_u^{k+1} = \mathbf{b}_u^k + (\mathbf{f}^{k+1} - \mathbf{Q}\mathbf{f}^{k+1} - \mathbf{u}^{k+1}), \quad (18)$$
$$\mathbf{b}_v^{k+1} = \mathbf{b}_v^k + (\nabla \mathbf{f}^{k+1} - \mathbf{v}^{k+1}). \quad (19)$$

Algorithm 1 summarizes the proposed ADMM based iteration scheme. Overall, we use two nested optimization loops to solve (9). We use the mean of the input images as an initial guess $\mathbf{f}^1$ as well as $\mathbf{u}^1 = \mathbf{v}^1 = \mathbf{0}$, $\mathbf{b}_u^1 = \mathbf{b}_v^1 = \mathbf{0}$. The weight matrices for IRLS are updated at every iteration.

Algorithm 1 MISO denoising with QuaSI prior

Set $\mathbf{u}^1 = \mathbf{v}^1 = \mathbf{b}_u^1 = \mathbf{b}_v^1 = \mathbf{0}$, $\mathbf{f}^1 = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}^{(t)}$
for $k = 1,\ldots,K_{\text{outer}}$ **do**
    Assemble $\mathbf{Q}$ from $\mathbf{f}^k$ according to (8)
    for $i = 1,\ldots,K_{\text{inner}}$ **do**
        Update weights $\mathbf{W}^{(t)}$ using (15)
        Update $\mathbf{f}^{k+1}$ using CG for (12)
        Update $\mathbf{u}^{k+1}$ and $\mathbf{v}^{k+1}$ using (16) - (17)
        Update $\mathbf{b}_u^{k+1}$ and $\mathbf{b}_v^{k+1}$ using (18) - (19)
    end for
end for

Algorithm 2 MIMO denoising with QuaSI prior

Set $\mathbf{F}^1 = \mathbf{G}$, $\mathbf{U}^1 = \mathbf{V}^1 = \mathbf{D}^1 = \mathbf{B}_U^1 = \mathbf{B}_V^1 = \mathbf{B}_D^1 = \mathbf{0}$
for $k = 1,\ldots,K_{\text{outer}}$ **do**
    Assemble $\mathbf{Q}$ from $\mathbf{F}^k$ according to (8)
    for $i = 1,\ldots,K_{\text{inner}}$ **do**
        Update weights $\mathbf{W}^{(t)}$ using (15)
        Update $\mathbf{F}^{k+1}$ using CG for (21)
        Update $\mathbf{U}^{k+1}, \mathbf{V}^{k+1}, \mathbf{D}^{k+1}$ using (24) - (26)
        Update $\mathbf{B}_U^{k+1}, \mathbf{B}_V^{k+1}, \mathbf{B}_D^{k+1}$ using (27) - (29)
    end for
end for

The linearization $\mathbf{Q}$ of the quantile filter is updated every K_{inner} iterations, assuming the position of the quantile does not change within the next K_{inner} iterations. This assumption speeds up the algorithm, as the construction of the matrix is time-consuming. Note that K_{inner} should not be chosen too large in order to avoid a bad approximation of the quantile filter. A proper evaluation of the convergence of the algorithm is presented in Sect. 6.1.5.

5.2. Multiple-Input Multiple-Output (MIMO) Mode

MIMO denoising follows a similar optimization approach and is based on the energy minimization formulation in (4). To this end, the augmented Lagrangian is given by:

$$
\begin{aligned}
\mathcal{L}_{\text{AL}}(\mathbf{F},\mathbf{U},\mathbf{V},\mathbf{D},\mathbf{B}_U,\mathbf{B}_V,\mathbf{B}_D) &= \rho(\mathbf{F} - \mathbf{G}) \\
&+ \frac{\alpha}{2}\|\mathbf{U} - \mathbf{F} + Q(\mathbf{F}) - \mathbf{B}_u\|_2^2 + \lambda\|\mathbf{U}\|_1 \\
&+ \frac{\beta}{2}\|\mathbf{V} - \nabla_{x,y,z}\mathbf{F} - \mathbf{B}_V\|_2^2 + \mu\|\mathbf{V}\|_1 \\
&+ \frac{\gamma}{2}\|\mathbf{D} - \nabla_t\mathbf{F} - \mathbf{B}_D\|_2^2 + \omega\|\mathbf{D}\|_1,
\end{aligned}
\tag{20}
$$

where $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{D}$ denote auxiliary variables with the respective Bregman variables $\mathbf{B}_U$, $\mathbf{B}_V$, and $\mathbf{B}_D$

to enforce the constraints of spatial TV, QuaSI, and temporal TV regularization, respectively.

Following MISO denoising as presented in Section 5.1, we linearize the non-linear quantile operator $Q(\mathbf{F}) = (Q(\mathbf{f}_1),\ldots,Q(\mathbf{f}_T))^\top$ using (8). Then, we have $Q(\mathbf{F}) = \mathbf{QF}$, where $\mathbf{Q} = (\mathbf{Q}_1,\ldots,\mathbf{Q}_T)^\top$ and for each volume $\mathbf{f}_t$ in the sequence $\mathbf{F}$ we have $Q(\mathbf{f}_t) = \mathbf{Q}_t\mathbf{f}_t$. Based on this linearization, we solve (20) with an alternating scheme by minimizing w.r.t. the individual parameters. The minimization w.r.t. $\mathbf{F}$ leads to the linear system:

$$\mathbf{AF}^{k+1} = \mathbf{b} \tag{21}$$

$$\mathbf{A} = \mathbf{W} + \beta\nabla_{x,y,z}^\top\nabla_{x,y,z} + \gamma\nabla_t^\top\nabla_t + \alpha\mathbf{M}^\top\mathbf{M} \tag{22}$$

$$
\begin{aligned}
\mathbf{b} &= 2\mathbf{WG} + \beta\nabla_{x,y,z}^\top(\mathbf{D}_{x,y,z} - \mathbf{B}_{x,y,z}) \\
&+ \gamma\nabla_t^\top(\mathbf{D}_t - \mathbf{B}_t) + \alpha\mathbf{M}^\top(\mathbf{U} - \mathbf{B}_u),
\end{aligned}
\tag{23}
$$

where $\mathbf{W}$ is a diagonal weight matrix associated with $\mathbf{F}^k$ and constructed from the Huber loss according to (15). We then solve (23) using CG iterations.

The auxiliary variables $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{D}$ are updated element-wise according to:

$$U_i^{k+1} = \text{shrink}([\mathbf{F}^{k+1} - \mathbf{QF}^{k+1} + \mathbf{B}_U^k]_i, \lambda/\alpha) \tag{24}$$

$$V_i^{k+1} = \text{shrink}([\nabla_{x,y,z}\mathbf{F}^{k+1} + \mathbf{B}_V^k]_i, \mu/\beta) \tag{25}$$

$$D_i^{k+1} = \text{shrink}([\nabla_t\mathbf{F}^{k+1} + \mathbf{B}_D^k]_i, \omega/\gamma). \tag{26}$$

Given the intermediate sequence $\mathbf{F}^{k+1}$ along with the auxiliary variables $\mathbf{U}^{k+1}$, $\mathbf{V}^{k+1}$, and $\mathbf{D}^{k+1}$, the Bregman variables are updated according to:

$$\mathbf{B}_U^{k+1} = \mathbf{B}_U^k + (\mathbf{F}^{k+1} - \mathbf{QF}^{k+1} - \mathbf{U}^{k+1}) \tag{27}$$

$$\mathbf{B}_V^{k+1} = \mathbf{B}_V^k + (\nabla_{x,y,z}\mathbf{F}^{k+1} - \mathbf{V}^{k+1}) \tag{28}$$

$$\mathbf{B}_D^{k+1} = \mathbf{B}_D^k + (\nabla_t\mathbf{F}^{k+1} - \mathbf{D}^{k+1}). \tag{29}$$

An illustration of the proposed optimization scheme is given in Algorithm 2.

6. Applications and Evaluation

In order to show the applicability of the proposed framework for image, volumetric and volumetric+temporal denoising, we evaluate our framework in different diagnostic and interventional imaging workflows namely OCT as well as C-arm CT. Specifically, we benchmark our method on different datasets including comparisons to the state-of-the-art in the respective fields.
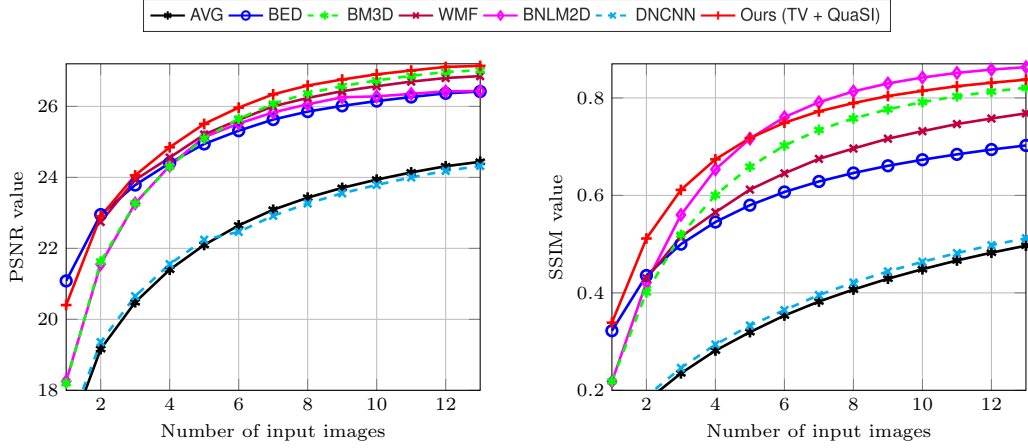
Figure 5: Quantification of noise reduction in terms of mean PSNR and SSIM for different denoising methods on the pig eye dataset for different numbers of input images. The points on the curves denote the average PSNR, and SSIM respectively, over the entire pig eye dataset using the number of input images denoted on the x-axis.

(a) Noisy input image

(b) AVG

(c) BED (Wong et al., 2010)

(d) BM3D (Dabov et al., 2007)

(e) WMF (Mayer et al., 2012)

(f) BNLM2D (Coupe et al., 2009)

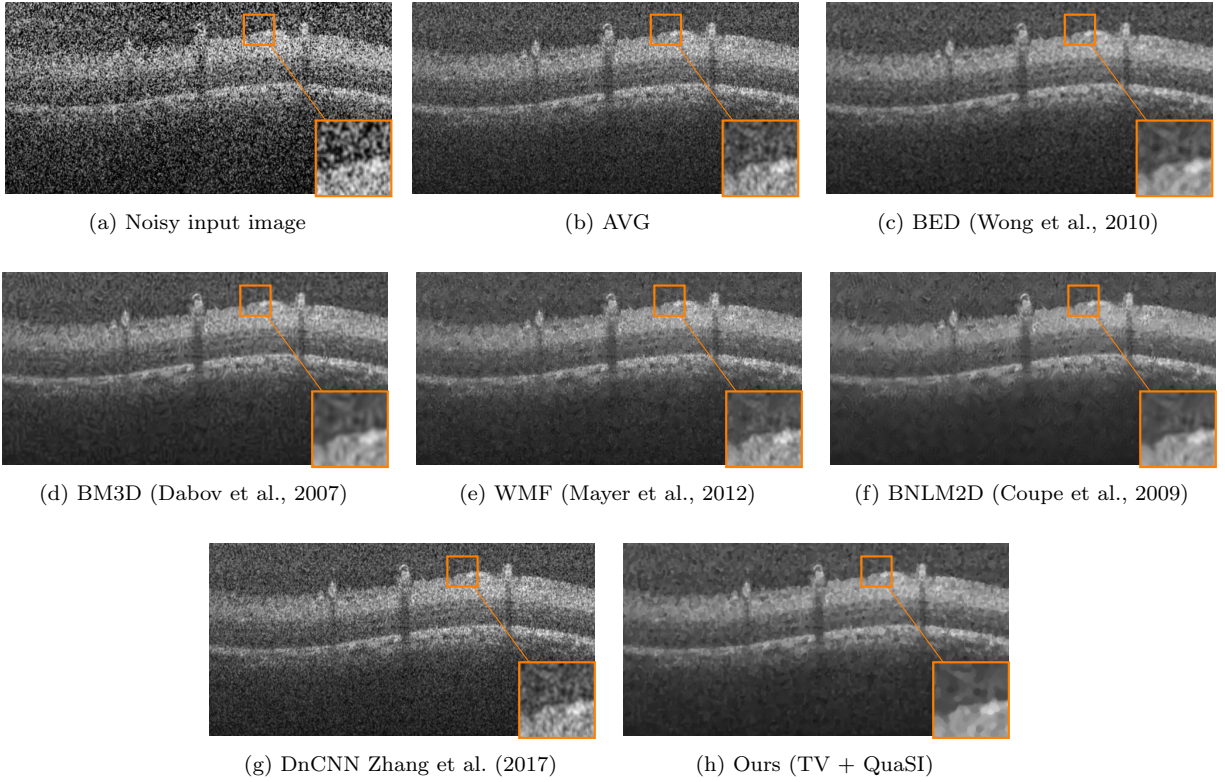(g) DnCNN Zhang et al. (2017)

(h) Ours (TV + QuaSI)

Figure 6: Denoising on position 9 from the pig eye dataset using 5 B-scans. (a) Noisy image, (b) – (h) AVG, BED (Wong et al., 2010), BM3D (Dabov et al., 2007), WMF (Mayer et al., 2012), BNLM2D (Coupe et al., 2009), DnCNN (Zhang et al., 2017) and the proposed method.

6.1. Optical Coherence Tomography Denoising

345   Throughout all experiments on the OCT data, we adopted our framework to image and volumetric denoising. For denoising on a B-scan level, the parameters were set to $\mu = 0.075 \cdot T$, $\lambda = 5.0 \cdot T$, $\alpha = 100.0 \cdot T$, $\beta = 1.5 \cdot T$, $K_{\text{outer}} = 20$ and $K_{\text{inner}} = 2$

for T B-scans and 3×3 median filtering to setup the QuaSI prior. In order to find appropriate standard parameter for the proposed method, we proceeded as follows. The parameter search was conducted on the pig eye dataset, using a clinical relevant image section of eye position 11 and 12 with 5 noisy B-scans each. First, the parameter of the proposed algorithm with pure TV regularization were set using a grid search approach for μ and β. To quantify the image quality, peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were evaluated in addition to a qualitative investigation. Second, the parameter of the proposed algorithm with QuaSI + TV regularization were set, using the optimal TV weights from the previous investigation.

For volumetric denoising based on $Z = 6$ adjacent B-scans, the parameters were set to $\mu = 0.0007 \cdot T$, $\lambda = 1.0 \cdot T$, $\alpha = 120.0 \cdot T$, $\beta = 0.05 \cdot T$, $K_{\text{outer}} = 20$ and $K_{\text{inner}} = 2$ for T volumes and $3 \times 3 \times 3$ median filtering. The proposed algorithm for volumetric denoising was evaluated on clinical data only. The selection of standard parameters was performed in the same way as for denoising on a B-scan level. Using $Z = 6$ adjacent B-scans in $T = 5$ volumes from only 1 patient, the TV weights followed by the QuaSI weights were set.

6.1.1. Datasets

To evaluate the performance of the proposed denoising algorithm, we conducted experiments on two different OCT datasets. This comprises ex-vivo benchmark data and real clinical data.

For an evaluation of denoising on B-scan level, we used the publicly available pig eye dataset provided by Mayer et al. (2012). The dataset comprises 455 B-scans corresponding to 35 eye positions with 13 scans per position and was captured ex-vivo with a Spectralis HRA & OCT. The published B-scans were registered to each other to compensate for geometric shifts. We apply denoising to sets of T registered B-scans with $T \in [1, 13]$ to demonstrate the influence of different numbers of input B-scans on the denoising result. The pig eye dataset provides a gold standard B-scan that was obtained by averaging all 455 registered scans. The quality of the denoising algorithm was evaluated by assessing the fidelity of a denoised B-scan w.r.t. the gold standard using the peak-signal-to-noise ratio (PSNR) as well as the structural similarity index (SSIM).

In order to evaluate and compare B-scans with volumetric denoising, we use clinical data. A prototype ultrahigh-speed swept-source OCT system with 1050 nm wavelength and a sampling rate of 400,000 A-scans per second (W.Choi et al., 2013) was used to acquire volumetric data of 14 human subjects. Proliferative and non-proliferative diabetic retinopathy, early age-related macular degeneration and one healthy subject were imaged on two volumes per subject, where each B-scan was acquired five times in immediate succession. We use 500 A-scans by 500 B-scans for a field size of 3×3 mm.

For denoising on a B-scan level, the central B-scan of each volume is used, while volumetric denoising is performed on adjacent B-scans including the central one. As the clinical data does not provide a gold standard, we follow prior work by Fang et al. (2012); Ozcan et al. (2007); Wong et al. (2010) and measure the noise reduction using the mean-to-standard-deviation ratio (MSR) and the contrast-to-noise ratio (CNR) according to:

$$\text{MSR} = \frac{\mu_f}{\sigma_f} \tag{30}$$

$$\text{CNR} = \frac{|\mu_f - \mu_b|}{\frac{1}{2}\sqrt{(\sigma_f^2 + \sigma_b^2)}}, \tag{31}$$

where μ_f and μ_b as well as σ_f and σ_b are the means and standard deviations of the intensities in a foreground and a background region, respectively. The regions to determine MSR and CNR were manually selected for the central B-scan, see Fig. 8a.

6.1.2. Comparison to the State-of-the-Art

We compared our method against seven competing denoising approaches. As representatives of general-purpose methods, we evaluated BM3D (Dabov et al., 2007) as well as a deep denoising CNN (DnCNN) (Zhang et al., 2017), which are state-of-the-art in the field of natural image denoising. We also evaluated non-local means-based speckle noise filtering (BNLM2D) that has been originally proposed for ultrasound image denoising (Coupe et al., 2009). In terms of spatial filters customized for OCT, we used Bayesian estimation denoising (BED) (Wong et al., 2010). In the field of temporal methods using multiple registered B-scans, we evaluate simple averaging (AVG) as a baseline as well as wavelet multi-frame denoising (WMF) (Mayer et al., 2012). To ensure fair comparisons between spatial and temporal methods, we provide the average of all B-scans as input for single-image denoising (BM3D, BNLM2D, DnCNN, and BED). In contrast, AVG and WMF are pure temporal approaches that process multiple registered B-scans. Notice that all of
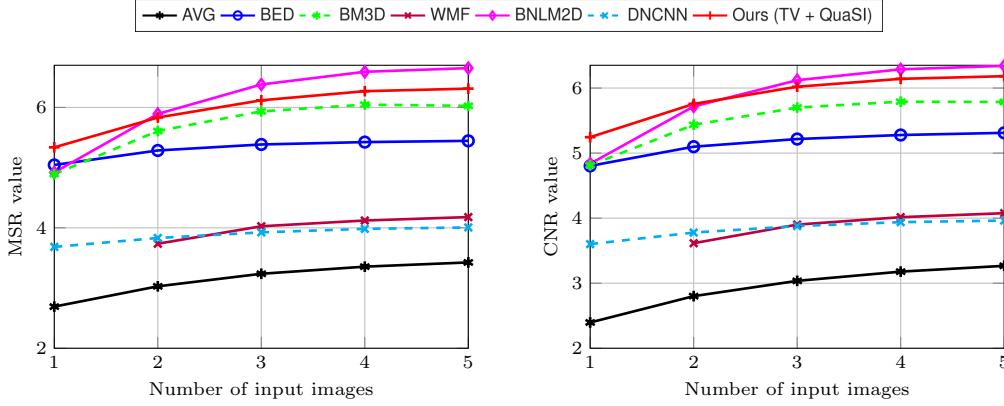
Figure 7: Quantification of noise reduction in terms of mean MSR and CNR measures for denoising on a B-scan level on our clinical dataset for different numbers of input images. The plots illustrate the mean MSR and CNR of the whole clinical dataset and the 5 foreground regions. Each point on the curves denotes the mean MSR and CNR using the number of input images specified on the x-axis as input to state-of-the-art denoising methods and the proposed algorithm with the QuaSI prior.

these methods can only operate on individual 2-D B-scans to denoise volumetric data and are therefore compared to our proposed method on a B-scan level. The parameters of the competing methods were set according to suggestions of the authors and adapted to the OCT data.

First, we conducted experiments for denoising on B-scan level on the pig eye dataset. Figure 5 depicts the mean PSRN and SSIM of the competing denoising methods w.r.t. the gold standard for different numbers of input B-scans. We observed quantitatively that our proposed method consistently outperforms the competing BM3D, BED, and WMF denoising methods regardless of the number of input frames. Moreover, using only $T = 2$ input B-scans, our spatio-temporal method achieved comparable results to averaging $T = 5$ B-scans. The proposed method performs better than BNLM2D for $T < 5$ input B-scans. This reveals that our method is more economic regarding the number of required input scans. This property is essential for clinical applications, where acquiring more scans might lead to unacceptable long acquisition times. Figure 6 depicts qualitative results for $T = 5$ B-scans. Here, the proposed algorithm using the QuaSI prior achieved superior performance in terms of noise reduction, while anatomical structures like retinal layers are preserved. Comparable results are achieved by BNLM2D, but the latter suffers from small streak-like artifacts. DnCNN achieved comparable results to simple averaging both regarding quantitative measures and qualitative assessment.

Second, denoising on a B-scan level was stud-

ied on our clinical datasets using the non-reference MSR and CNR measures for a quantitative evaluation. Figure 7 depicts the averaged MSR and CNR measures for different numbers of input images. Overall, we observed that BNLM2D and our proposed method achieved the best noise reduction expressed by both measures. Figure 8 compares the denoising performance on one example dataset. We found that AVG, WMF, and BED facilitate structure-preserving denoising but were prone to noise breakthroughs in homogeneous areas, which lowers their MSR and CNR. In contrast, BM3D achieved superior noise reduction but suffered from streak artifacts. Similar observations were made in related work on OCT denoising (Fang et al., 2012) and can be explained by the assumption of additive white Gaussian noise used for BM3D. The proposed method achieved a decent tradeoff between noise reduction and structure preservation.

6.1.3. Impact of the QuaSI Prior

We used the pig eye dataset as well as clinical data to evaluate the performance of our spatio-temporal denoising algorithm with and without the QuaSI prior. Figure 9 illustrates the impact of the QuaSI prior on the denoising result for the pig eye data compared to simple averaging and pure TV regularization. In terms of noise reduction, the proposed variational framework outperformed simple averaging. Especially in the enlarged region, a noticeable difference between averaging and the proposed denoising algorithm is shown. In homogeneous areas, the algorithm considerably suppressed speckle noise,

(a) Noisy image (MSR: 2.68, CNR: 2.47)

(b) AVG (MSR: 3.17, CNR: 3.17)

(c) BM3D Dabov et al. (2007) (MSR: 4.61, CNR: 4.85)

(d) BED Wong et al. (2010) (MSR: 4.67, CNR: 4.85)

(e) WMF Mayer et al. (2012) (MSR: 3.67, CNR: 3.55)

(f) DnCNN (MSR: 3.54, CNR: 3.70)

(g) BNLM2D (MSR: 5.04, CNR: 5.32)
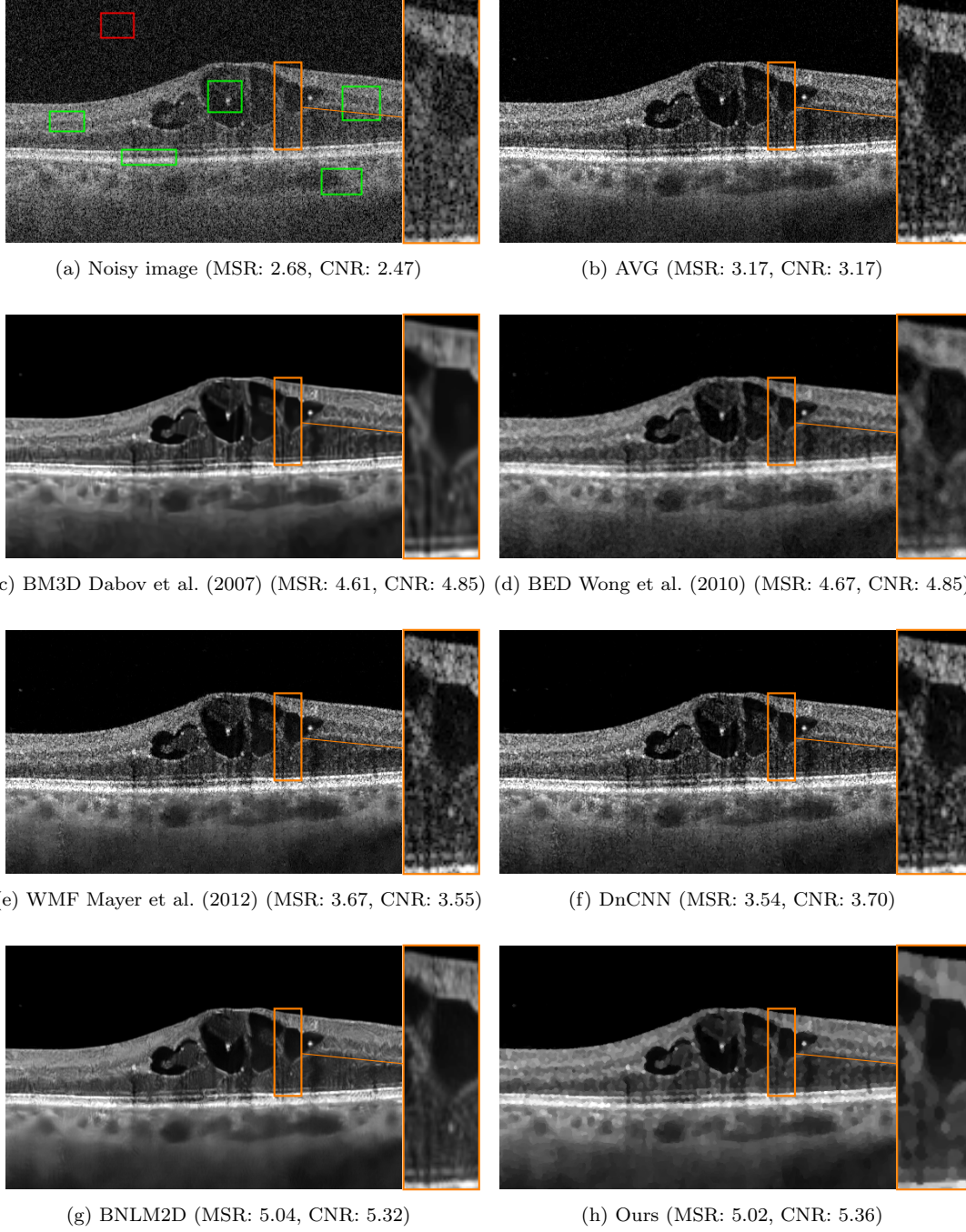
(h) Ours (MSR: 5.02, CNR: 5.36)

Figure 8: Visual comparison of denoising results using our clinical dataset with the central B-scan of $T = 5$ volumes from a 46 years old male patient with diabetic retinopathy. (a) Noisy image with manually selected background (red) and foreground regions (green) to determine MSR and CNR. (b) – (h) AVG, BM3D (Dabov et al., 2007), BED (Wong et al., 2010), WMF (Mayer et al., 2012), DnCNN (Zhang et al., 2017), BNLM2D (Coupe et al., 2009), and the proposed method.

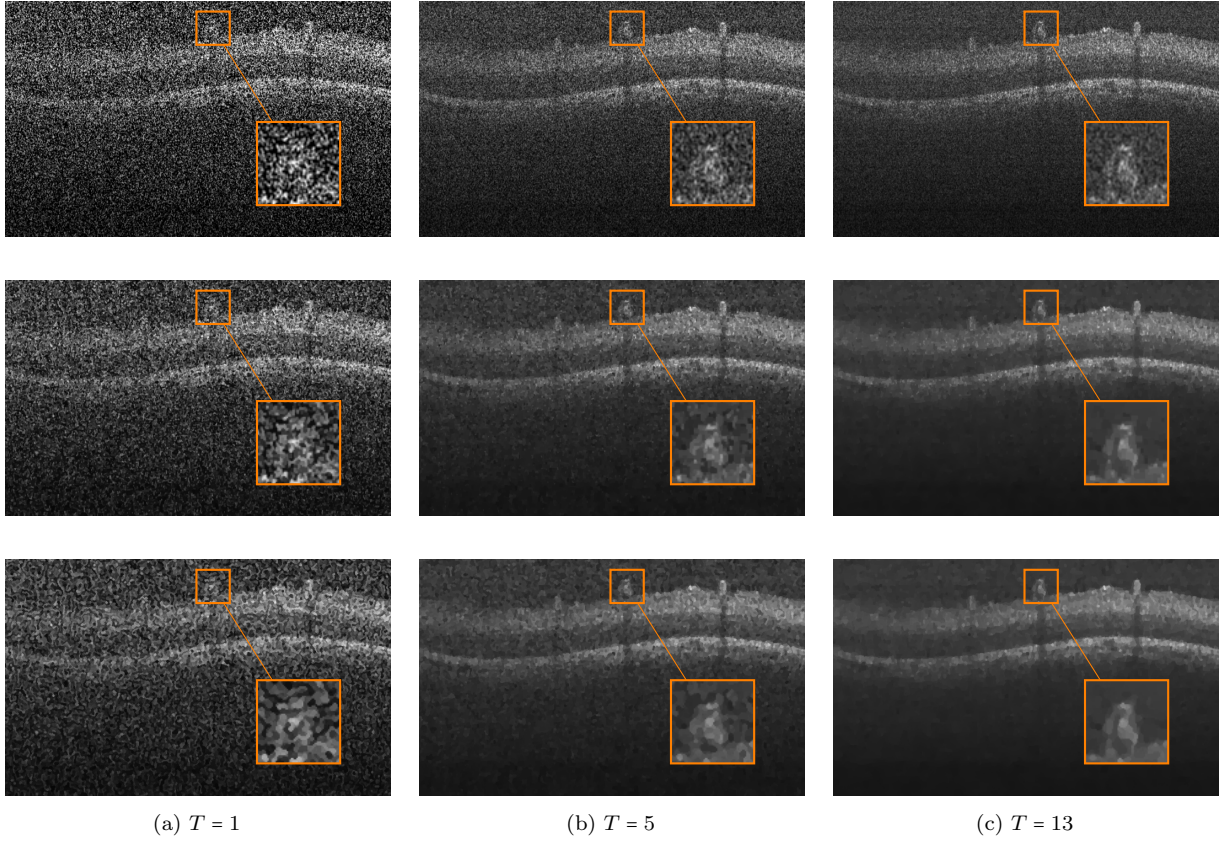(a) $T = 1$　　(b) $T = 5$　　(c) $T = 13$

Figure 9: This comparison aims at demonstrating the improvement of the proposed spatio-temporal denoising with TV + QuaSI regularization (third row) compared to simple averaging of registered B-scans (top row) and the proposed spatio-temporal denoising with TV regularization only (second row) for different numbers of input images. For the comparison, dataset 27 from the pig eye dataset was used to evaluate the proposed algorithm with and without the QuaSI prior using the standard parameter.

while preserving important structures. The noise reduction was superior when using a combination of the QuaSI prior and the TV prior for regularization as shown for the retinal structures in the enlarged region. In addition, the QuaSI prior contributed to structure-preservation and avoided staircasing artifacts that typically appear in TV denoising.

Figure 10 illustrates the impact of the QuaSI prior using PSNR and SSIM (for the pig eye data) as well as MSR and CNR (for clinical data) for different numbers of input scans. Here, our denoising framework with QuaSI prior outperformed TV denoising in terms of all measures.

6.1.4. B-scan vs. Volumetric Denoising

So far, we evaluated denoising of volumetric OCT data by simply processing individual B-scans. In order to evaluate the impact of true volumetric denois-

ing to simple B-scan wise denoising in our proposed framework, we used our clinical dataset. Volumetric denoising processes 6 consecutive B-scans including the central one. That way, CNR and MSR measures from the previous experiments can be used for comparison. Table 1 shows the mean MSR and CNR using $T = 1$ and $T = 5$ registered input volumes. The proposed method is compared to BM4D (Maggioni et al., 2013) using $T = 1$ volume and the average of $T = 5$ volumues as an input. Here, we found that our volumetric denoising achieved better results in terms of noise reduction for $T = 1$ input volume, as adjacent B-scans affect denoising positively. For $T = 5$ input volumes, we found that our B-scan denoising achieved slightly better results in terms of noise reduction. However, as opposed to noise reduction, volumetric denoising achieved superior performance in structure preservation by exploit-
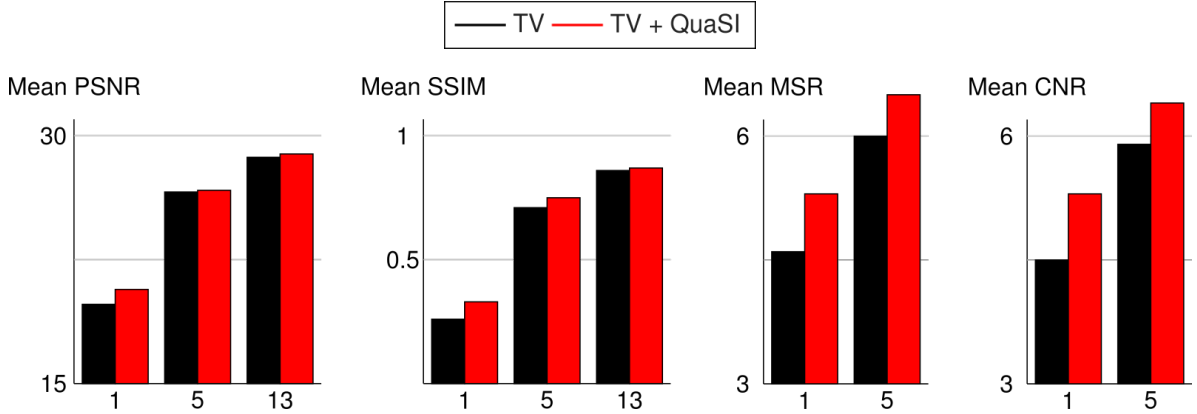
Figure 10: Mean PSNR, SSIM, MSR and CNR measures to quantify noise reduction with and without the QuaSI prior for 1, 5 and 13 input images. The two bar graphs on the left hand side illustrate the average PSNR and SSIM over the entire pig eye dataset using the proposed algorithm with and without QuaSI prior and the standard parameters. The average MSR and CNR over the entire clinical dataset is shown in the two bar graphs on the right hand side.

	$T = 1$ volume			$T = 5$ volumes		
	BM4D (Maggioni et al., 2013)	B-scan denoising	Volumetric denoising	BM4D (Maggioni et al., 2013)	B-scan denoising	Volumetric denoising
MSR	5.16	5.35	5.77	5.38	6.50	6.31
CNR	5.00	5.27	5.60	5.23	6.38	6.18

Table 1: Mean MSR and CNR measures for 1 and 5 registered input volumes on the clinical data. For B-scan denoising, the central B-scan is used and for volumetric denoising 6 adjacent B-scans including the central one are used. The B-scan-wise average of $T = 5$ input volumes served as input to BM4D (Maggioni et al., 2013).

ing coherence between adjacent B-scans. This is depicted in Fig. 11, where the retinal layers in the magnified region can be better distinguished.

6.1.5. Convergence and Parameter Sensitivity

The convergence of the proposed algorithm is shown experimentally on a B-scan level. By our definition, the algorithm converges if a stationary point of the objective function (3) is reached. The value of the objective, hereinafter referred to as energy, is computed after every update of the intermediate image $\mathbf{f}^{k+1}$. In addition, PSNR and SSIM of the intermediate image are computed. Based on the optimal parameter setting $\mu = 0.075 \cdot T$, $\lambda = 5.0 \cdot T$, $\alpha = 100.0 \cdot T$, $\beta = 1.5 \cdot T$, $K_{\text{outer}} = 30$, $K_{\text{inner}} = 10$ and $K_{\text{cg}} = 3$ for B-scan denoising, we denoise the pig eye dataset 9 with $T = 8$ B-scans.

Figure 12 shows the impact of K_{outer}, K_{inner}, and K_{cg} on the convergence using three different parameter settings, where $K_{\text{outer}} \cdot K_{\text{inner}} = 300$ for a fair comparison. The approximation of the QuaSI prior is updated every K_{inner} iterations. We found that increasing numbers of inner iterations ($K_{\text{inner}} = 10$) or CG iterations ($K_{\text{cg}} = 30$) impair the convergence properties of the algorithm as shown by the peaks of the energy and the PSNR. This is mainly caused by the rare update of the linearization $\mathbf{Q}$. If the linearization is updated every iteration ($K_{\text{inner}} = 1$), the convergence is improved as no approximation is necessary but the computational complexity is increased. The optimal setting ($K_{\text{outer}} = 30$, $K_{\text{inner}} = 10$, $K_{\text{cg}} = 3$) provides an excellent tradeoff between stable convergence and low computational complexity.

Figure 13 shows the influence of the QuaSI regularization weight λ to the convergence of our algorithm. We found that with decreasing λ, the PSNR and SSIM measures increase slower due to the low impact of the QuaSI prior. For the optimal setting $\lambda = 5.0$, we observed a fast convergence of our iteration scheme. Notice that further increasing λ does not affect the convergence, which underlines effectiveness of the proposed QuaSI prior and the robustness of our iteration scheme.
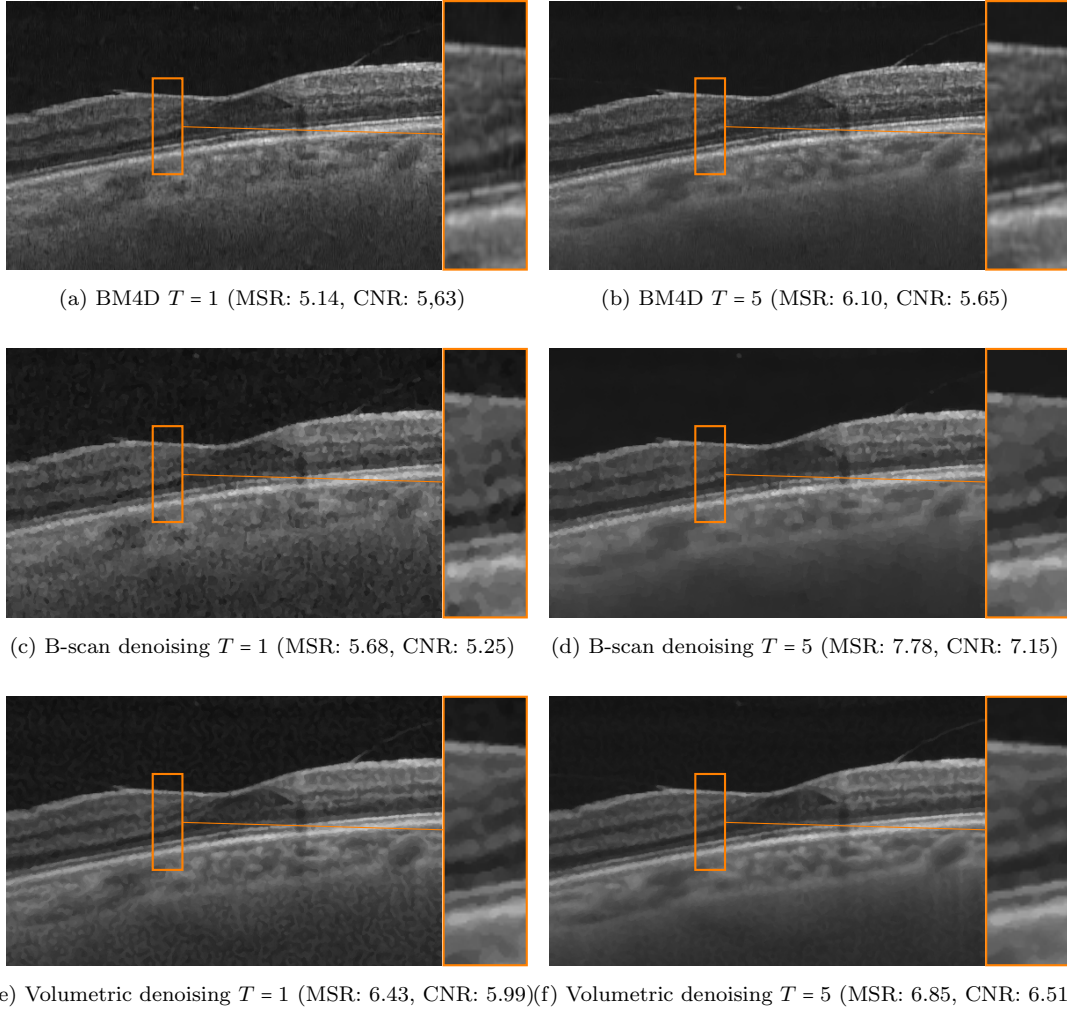
14

(a) BM4D $T = 1$ (MSR: 5.14, CNR: 5,63)

(b) BM4D $T = 5$ (MSR: 6.10, CNR: 5.65)

(c) B-scan denoising $T = 1$ (MSR: 5.68, CNR: 5.25)

(d) B-scan denoising $T = 5$ (MSR: 7.78, CNR: 7.15)

(e) Volumetric denoising $T = 1$ (MSR: 6.43, CNR: 5.99)(f) Volumetric denoising $T = 5$ (MSR: 6.85, CNR: 6.51)

Figure 11: Denoising on the clincial dataset using $T = 5$ registered volumes from a 67 years old male patient with non-proliferative diabetic retinopathy. The left column illustrates the results of the proposed method on a B-scan level with $Z = 1$ scan (c) and on a volumetric level (e) as well as BM4D (a) with $Z = 6$ consecutive scans using $T = 1$ input volume. The right column illustrates the results of the proposed method on a B-scan level with $Z = 1$ scan (d) and on a volumetric level (f) as well as BM4D (b) with $Z = 6$ consecutive scans using $T = 5$ registered input volumes.

Figure 14 depicts the influence of the Lagrangian multiplier α, which enforces the constraint $\mathbf{u} = \mathbf{f} - \mathbf{Q(f)}$ in our ADMM optimization. For $\alpha \to \infty$, the augmented Lagrangian (11) results in the objective function (3). Hence, decreasing α impairs the convergence as shown by the peaks in the PSNR and SSIM measures over the iterations. Choosing α too large resulted in slower convergence compared to the proposed parameter setting $\alpha = 100$.

In order to show the interplay of the QuaSI regularization weight λ and the corresponding Lagrangian multiplier α used for ADMM, Fig. 15 de- picts the influence of different configurations to B-scan denoising using fixed TV parameters ($\mu = 0.075$, $\beta = 1.5$). We evaluated the denoising performance in terms of the PSNR and SSIM measures for a clinical relevant region showing retinal layers. Overall, we observed that increasing λ and thus the impact of QuaSI consistently improved denoising, whereas the sensitivity against α is lower over several orders of magnitudes. Notice that our QuaSI prior was insensitive against oversmoothing as shown by the convergence of PSNR and SSIM for large λ.
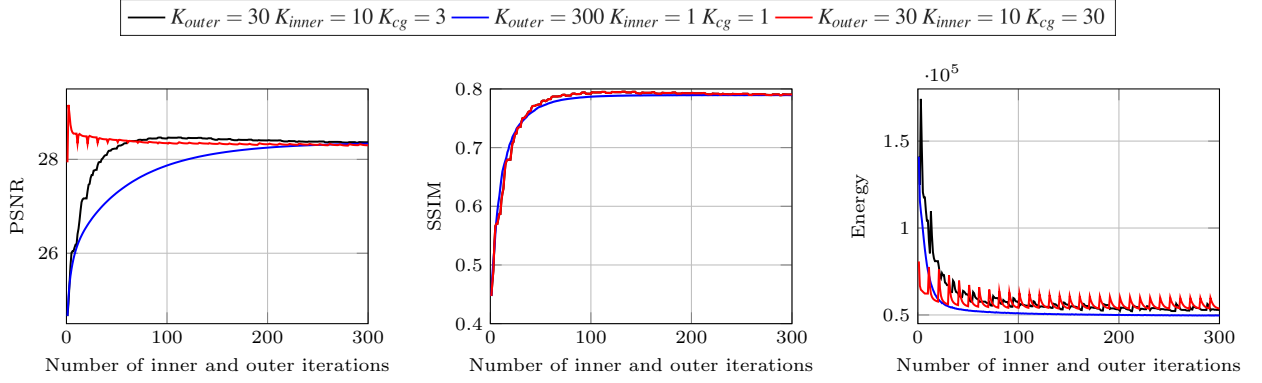
Figure 12: Convergence analysis for our proposed optimization scheme in OCT B-scan denoising using different combinations of iteration numbers K_{outer}, K_{inner} and K_{cg}. For each combination, we depict the value of the energy function optimized by ADMM along the with PSNR of the intermediate denoised images over the iterations.
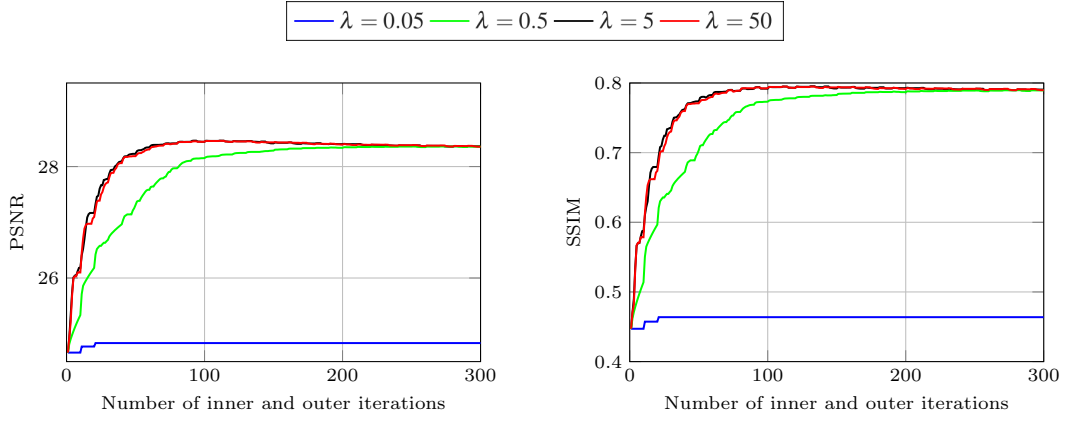

Figure 13: Convergence analysis for our proposed optimization scheme in OCT B-scan denoising using different QuaSI regularization weights λ. For each parameter setting, we depict the influence of λ using the PSNR and SSIM of the intermediate denoised image over the iterations.
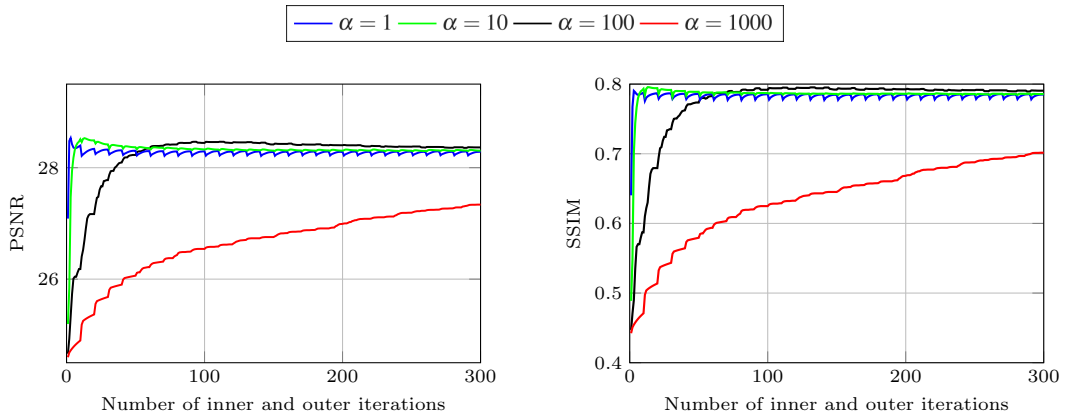

Figure 14: Convergence analysis for our proposed algorithm in OCT B-scan denoising using different Lagrangian multiplier α for ADMM optimization. For each parameter setting, we depict the influence of α using the PSNR and SSIM of the intermediate denoised image over the iterations.
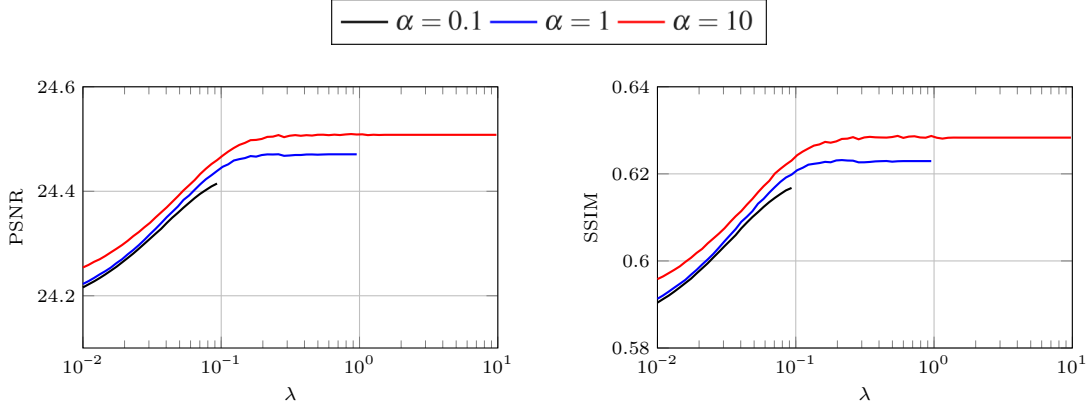
Figure 15: Parameter sensitivity analysis for the interplay of the QuaSI regularization weight λ and the Lagrangian multiplier α used for ADMM to B-scan denoising. The PSNR and SSIM measures were evaluated for a clinical relevant region of position 11 from the pig eye dataset. Each measure was determined for different QuaSI parameters λ and α while keeping the TV regularization weight $\mu = 0.075$ and the corresponding Lagrangian multiplier $\beta = 1.5$ fixed.

6.2. C-Arm Computed Tomography Denoising

C-arm computed tomography (CT) denotes an imaging modality where an X-ray source and detector are mounted on opposing sides of a C-shaped gantry. That gantry is further able to rotate around a patient lying on a table, thus allowing to acquire CT-like projection images. Using image reconstruction techniques (Zeng, 2010; Strobel et al., 2009), these projection images can finally be transformed into a volumetric representation of the object under consideration.

Clinically, C-arm CT is both used for acquiring single volumetric images as well as for acquiring sequences of volumes, as it is for example used in perfusion imaging for acute stroke diagnosis (Fieselmann and Manhart, 2013). While single volumes just provide static information about the morphology itself, the acquisition of volume sequences typically involves injection of contrast agent during the acquisition, thus making the volume sequences provide additional temporal information.

Similar to conventional CT, photon effects as well as patient movement and angular undersampling usually deteriorate the image quality by introducing both structured and unstructured noise, see Fig. 16 b, Fig. 17 a.

For our experiments, the noise $\mathbf{n}$ in reconstructed CT volumes is modeled as additive noise according to (1), and is further composed of both shot noise $\mathbf{p}$ and structured noise $\mathbf{s}$, i.e.

$$\mathbf{n} = \mathbf{p} + \mathbf{s}. \qquad (32)$$

While shot noise in the acquired projection data results from fluctuations measured by the sensor, various processing steps during the reconstruction process complicate an exact statistical description of the noise in the resulting volumetric data (Fessler, 2014). Structured noise comes in the form of high-frequent streak artifacts, causes by angular undersampling.

6.2.1. Datasets

We applied the proposed denoising algorithm on simulated C-arm CT data as well as on acquired, real patient data.

For our application, this results in two cases: single volumes can be denoised using *volumetric denoising* (for the sake of convenience, we further refer to this method as SISO - single volume input, single volume output), while sequences of volumes are processed using *volumetric + temporal denoising* (MIMO - multiple volume input, multiple volume output), cf. Fig. 1.

order to evaluate the denosing, we particularly investigated simulated data since it provides a known ground truth. The simulated data is based on a digital brain CT phantom (Aichert et al., 2013), which was used in combination with a simulation framework mimicking the acquisition process of a C-arm CT system (Maier et al., 2013). We added Poisson noise and simulated minor patient movement during the generation of the simulated data by rotating the head up to a total of 5° around z-axis between the individual scans. After reconstructing the generated projection data, the individual volumes are co-registered again to assert pixel correspondence

17

	Input	BM4D	QuaSI (SISO)	QuaSI (MISO)
PSNR	32.105	32.485	32.462	34.788
SSIM	0.883	0.914	0.925	0.943

Table 2: PSNR and SSIM for the input data, BM4D Maggioni et al. (2013) and the QuaSI methods.

between the volumes. Due to the slight different positions of the head within individual volumes, the resulting streak artifacts slightly differ between the co-registered individual volumes.

For a numerical comparison of different algorithms, we calculate the peak signal to noise ratio (PSNR) and the structured similarity index measure (SSIM) (Wang et al., 2004) using the digital phantom data as ground truth.

In addition to the simulated data, we also apply the proposed methods to real patient data which was clinically acquired during a perfusion imaging procedure.

6.2.2. Comparison to the State-of-the-Art

Current approaches towards noise reduction in CT imaging are, for example, based on anisotropic filtering or rely on a heuristic detection of streaks and vessel structures (Maier and Fahrig, 2015; Maier et al., 2011; Manhart et al., 2014).

We compared the results from the proposed methods to the results from BM4D (Maggioni et al., 2013), which processes volumetric data and is an extension to the well-known BM3D (Dabov et al., 2007). We set the parameters of our method to $\alpha = 0.1$, $\lambda = 0.0005$, $\beta = 0.1$, $\mu = 0.005$, $\gamma = 90$ and $\omega = 0.8$. These parameters have been optimized by investigating grid search on a small patch of the phantom data. The median filter regularization is computed on a $3 \times 3 \times 3$ kernel.

The algorithms are applied to and evaluated on a subset of the brain volume consisting of 30 consecutive slices. The slices, see Fig. 16 for synthetic and Fig. 17 for the real data, show the complete head and contain all structures of interest such as bones, white matter, gray matter and (contrast-enhanced) vessels. The results from the evaluation of the realistic brain phantom show that the proposed denoising algorithm outperforms BM4D with regards to PSNR and SSIM, see Table 2. Vessel structures are well-preserved within both volumes and boundaries between gray and white matter are

(a) Ground truth

(b) Noisy input

(c) SISO with QuaSI

(d) SISO without QuaSI
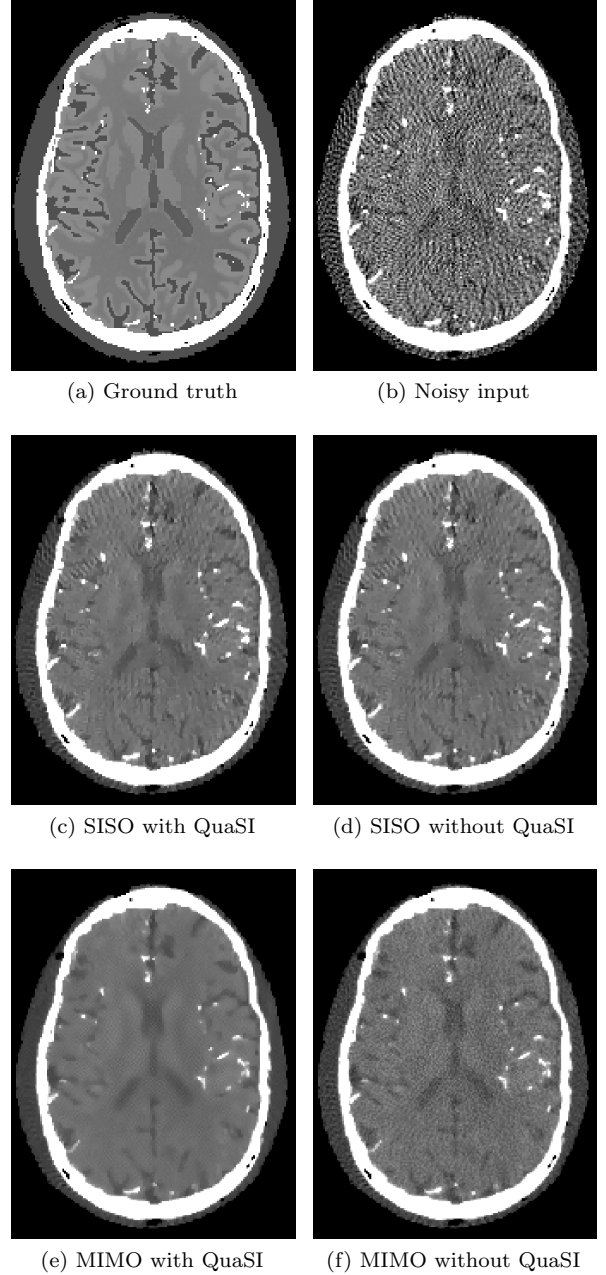
(e) MIMO with QuaSI

(f) MIMO without QuaSI

Figure 16: Denoising on simulated C-arm CT data. (a) and (b) denote the ground truth data and the noisy input to the algorithm, respectively. (c) and (d) denote the denoised result with and without the QuaSI prior when using only a single volume (SISO) of the sequence. (e) and (f) denote the denoised result with and without the QuaSI prior, when using 1 volume (MISO). Note that for MISO, the input to the algorithm is not just the single volume as shown in the figure, but consists of a sequence of volumetric data.

18

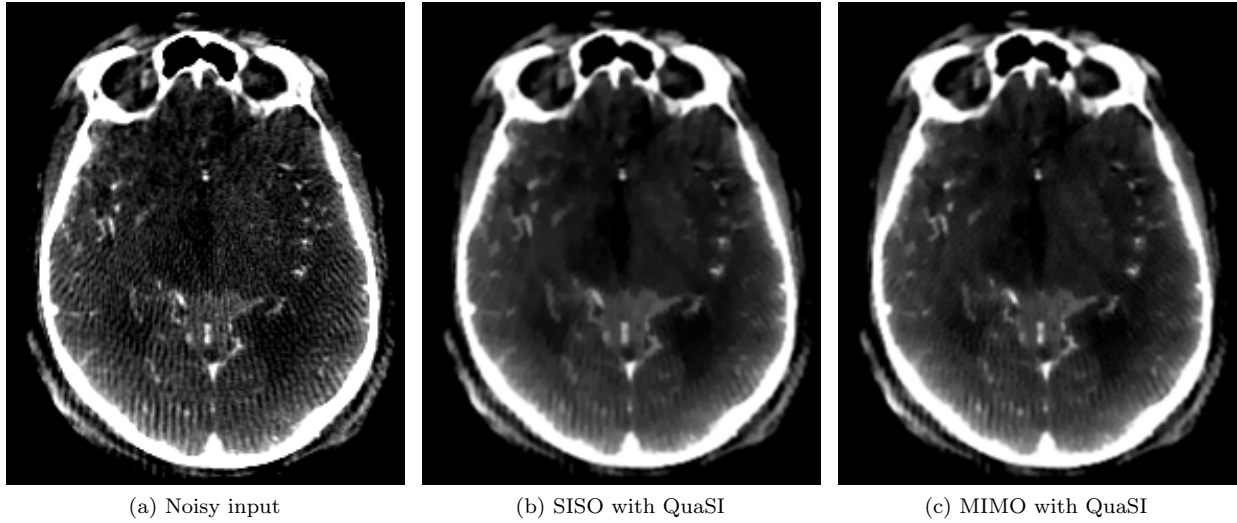| (a) Noisy input | (b) SISO with QuaSI | (c) MIMO with QuaSI |

Figure 17: Denoising on real clinical C-arm CT data. (a) denotes the noisy input, (b) denotes the denoised result when using only a single volume (SISO) of the sequence. (c) denotes the denoised result when using a volume sequence (MIMO). Note that for MIMO, the input to the algorithm is not just the single volume as shown in the figure, but consists of a sequence of volumetric data.

perceivable. Further, a qualitative comparison between processed data with and without the use of the QuaSI prior (Fig. 16 c,d and e,f) shows that the QuaSI prior is able to further lower the amount of noise in the volumetric image data.

7. Conclusion

In this paper, we have presented the quantile sparse image (QuaSI) prior and a corresponding spatio-temporal denoising algorithm suitable for volumetric OCT or CT data. For OCT denoising, we proposed two pipelines to either process B-scans or volumetric OCT data. The numerical optimization is derived using a linearization of the quantile filter and an alternating direction method of multipliers scheme for efficient minimization. We can show that a combination of QuaSI and Total Variation regularization outperforms state-of-the-art methods in terms of quantitative measures. Interestingly, our method can be applied to both CT and OCT data through minor modifications of the denoising pipeline. This suggests that it may be worthwhile to evaluate the potential of the QuaSI prior for inverse problems of other imaging modalities in future work.

References

Aichert A, Manhart MT, Navalpakkam BK, Grimm R, Hutter J, Maier AK, Hornegger J, Doerfler A. A realistic digital phantom for perfusion c-arm CT based on MRI data. In: Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE. IEEE; 2013. p. 1–2.

Amiot C, Girard C, Chanussot J, Pescatore J, Desvignes M. Spatio-temporal multiscale denoising of fluoroscopic sequence. IEEE Transactions on Medical Imaging 2016;35(6):1565–74.

Burger HC, Schuler CJ, Harmeling S. Image denoising: Can plain neural networks compete with bm3d? In: IEEE Conference on Computer Vision and Pattern Recognition. 2012. p. 2392–9.

Cheng J, Duan L, Wong DWK, Tao D, Akiba M, Liu J. Speckle Reduction in Optical Coherence Tomography by Image Registration and Matrix Completion. In: MICCAI 2014, Proceedings, Part I. 2014. p. 162–9.

Coupe P, Hellier P, Kervrann C, Barillot C. Nonlocal means-based speckle filtering for ultrasound images. IEEE Transactions on Image Processing 2009;18(10):2221–9.

Dabov K, Foi A, Kathovnik V, Egiazarian K. Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans Image Process 2007;16(8):2080–95.

Duan J, Lu W, Tench C, Gottlob I, Proudlock F, Samani NN, Bai L. Denoising optical coherence tomography using second order total generalized variation decomposition. Biomed Signal Process Control 2016;24:120–7.

Fang L, Li S, Nie Q, Izatt JA, Toth CA, Farsiu S. Sparsity based denoising of spectral domain optical coherence tomography images. Biomed Opt Express 2012;3(5):927–42.

Fessler JA. Fundamentals of ct reconstruction in 2d and 3d. Comprehensive Biomedical Physics 2014;:263–95.

Fieselmann A, Manhart M. C-arm CT Perfusion Imaging in the Interventional Suite. Current Medical Imaging Reviews 2013;9(2):96–101.

Goldstein T, Osher S. The split bregman method for l1-regularized problems. SIAM J Imaging Sci 2009;2(2):323–43.

19

765 Köhler T, Huang X, Schebesch F, Aichert A, Maier A, Hornegger J. Robust Multiframe Super-Resolution Employing Iteratively Re-Weighted Minimization. IEEE Transactions on Computational Imaging 2016;2(1):42 – 58.

L. Bian J. Suo FC, Dai Q. Multiframe denoising of high-speed
770 optical coherence tomography data using interframe and intraframe priors. Journal of Biomedical Optics 2015;20:20 –.

Maggioni M, Katkovnik V, Egiazarian K, Foi A. Nonlocal transform-domain filter for volumetric data denoising and
775 reconstruction. IEEE Trans Image Process 2013;22(1):119– 33.

Maier A, Fahrig R. GPU Denoising for Computed Tomography; Boca Raton, Florida, USA; volume 1. 1st ed.

Maier A, Hofmann H, Berger M, Fischer P, Schwemmer C,
780 Wu H, Müller K, Hornegger J, Choi J, Riess C, Keil A, Fahrig R. CONRAD - A software framework for cone-beam imaging in radiology. Medical Physics 2013;40(11).

Maier A, Wigström L, Hofmann H, Hornegger J, Zhu L, Strobel N, Fahrig R. Three-dimensional anisotropic adap-
785 tive filtering of projection data for noise reduction in cone beam CT. Medical Physics 2011;38(11):5896–909.

Manhart M, Fahrig R, Hornegger J, Dörfler A, Maier A. Guided Noise Reduction for Spectral CT with Energy-Selective Photon Counting Detectors. In: Noo F, editor.
790 Proceedings of the Third CT Meeting. 2014. p. 91–4.

Mayer MA, Borsdorf A, Wagner M, Mardin CY, Tornow RP. Wavelet denoising of multiframe optical coherence tomography. Biomed Opt Express 2012;3(3):2156–7085.

Ochs P, Dosovitskiy A, Brox T, Pock T. On iteratively
795 reweighted algorithms for nonsmooth nonconvex optimization in computer vision. SIAM J Imaging Sci 2015;8(1):331– 72.

Ozcan A, Bilenca A, Desjardins AE, Bouma BE, Tearney GJ. Speckle reduction in optical coherence tomography images
800 using digital filtering. Journal of the Optical Society of America A 2007;24(7):1901.

Pan J, Sun D, Pfister H, Yang MH. Blind image deblurring using dark channel prior. In: CVPR, Proceedings. 2016. p. 1628–36.

805 Rohkohl C. Motion Estimation and Compensation for Interventional Cardiovascular Image Reconstruction. Ph.D. thesis; Friedrich-Alexander-Universität; 2011.

Romano Y, Elad M, Milanfar P. The Little Engine that Could: Regularization by Denoising (RED). arXiv preprint
810 arXiv:161102862 2016;.

Rousseeuw PJ, Leroy AM. Robust Regression and Outlier Detection. John Wiley & Sons, 1987.

Salinas HM, Fernandez DC. Comparison of PDE-Based Nonlinear Diffusion Approaches for Image Enhancement
815 and Denoising in Optical Coherence Tomography. IEEE Trans Med Imag 2007;26(6):761–71.

Schirrmacher F, Köhler T, Husvogt L, Fujimoto JG, Hornegger J, Maier AK. Quasi: Quantile sparse image prior for spatio-temporal denoising of retinal OCT data. In:
820 MICCAI 2017, Proceedings, Part II. 2017. p. 83–91.

Strobel N, Meissner O, Boese J, Brunner T, Heigl B, Hoheisel M, Lauritsch G, Nagel M, Pfister M, Rührnschopf E, Scholz B, Schreiber B, Spahn M, Zellerhoff M, Klingenbeck-Regn K. 3D Imaging with Flat-Detector C-Arm Systems.
825 In: Reiser M, Becker CR, Nikolaou K, Glazer G, editors. Multislice CT. Springer; 2009. p. 33–51.

Thapa D, Raahemifar K, Lakshminarayanan V. Reduction of speckle noise from optical coherence tomography images using multi-frame weighted nuclear norm minimization
830 method. Journal of Modern Optics 2015;62(21):1856–64.

Tomasi C, Manduchi R. Bilateral Filtering for Gray and Color Images. In: International Conference on Computer Vision (ICCV). Bombay, India: IEEE; 1998. p. 839–46.

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality
835 assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13(4):600–12.

W.Choi , Potsaid B, Jayaraman V, Baumann B, Grulkowski I, Liu JJ, Lu CD, Cable AE, Huang D, Duker JS, Fujimoto JG. Phase-sensitive swept-source optical coherence
840 tomography imaging of the human retina with a vertical cavity surface-emitting laser light source. Opt Lett 2013;38(3):338.

Wong A, Mishra A, Bizheva K, Clausi DA. General bayesian estimation for speckle noise reduction in optical coherence
845 tomography reinal imagery. Opt Express 2010;18(8):8338– 52.

Zeng GL. Medical image reconstruction: a conceptual tutorial. Springer, 2010.

Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond
850 a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Transactions on Image Processing 2017;26(7):3142–55.

20