Multiscale Tikhonov-Total Variation Image Restoration Using Spatially Varying Edge Coherence Exponent Scheme

V. B. Surya Prasath, Dmitry Vorotnikov, Rengarajan Pelapur *Student Member, IEEE*, Shani Jose,
Guna Seetharaman, *Fellow, IEEE*, K. Palaniappan, *Senior Member, IEEE*

Abstract—Edge preserving regularization using partial differential equation (PDE) based schemes are now widely used in image restoration. We propose an adaptive multiscale variable exponent-based anisotropic variational PDE scheme that avoids current limitations such as over smoothing and blockiness artifacts while still retaining and enhancing edge structures across scale. The innovative model automatically balances between L^2 Tikhonov and L^1 total variation (TV) regularization effects using scene content information by adopting a spatially varying edge coherence exponent term constructed from the eigenvalues of the smoothed structure tensor. The multiscale exponent model considered here leads to a novel denoising method which preserves edges and provides selective denoising without generating artifacts for both additive and multiplicative noise models. Mathematical analysis of the proposed method in variable exponent space demonstrates its robustness, unconditional stability of the scheme supporting large (time evolution) step sizes and that the approach theoretically satisfies the maximum-minimum principle which guarantees that artificial edge regions are not created. Extensive experimental results on synthetic and real biomedical images indicate that the proposed Multiscale Tikhonov-Total Variation (MTTV) and Dynamical MTTV (D-MTTV) schemes perform better than sixteen other denoising algorithms in terms of several metrics including signal-to-noise ratio improvement and structure preservation. Promising extensions to handle multiplicative noise models and multichannel imagery are also provided.

I. INTRODUCTION

IMAGE restoration and enhancement to improve image quality under different noise models is a critical requirement across many image processing application domains including defense, space and biomedicine. Regularization and partial differential equations (PDEs) based schemes are very popular for removing noise and directionally smoothing images [1], see [3], [4] for a review. Despite the success enjoyed by these methods, there are problems related to edge and fine structure preservation, staircasing artifacts or over-smoothing of images. Variational regularization approaches use the classical quadratic Tikhonov [5] and total variation (TV) function studied by Rudin et al [6]. The over smoothing nature of the Tikhonov functional results in noise removal but at the expense of edge dislocation. On the other hand the TV regularization functional does a good job in retaining edges

V. B. S. Prasath, R. Pelapur, S. Jose, and K. Palaniappan are with the Department of Computer Science, University of Missouri-Columbia, Columbia, MO, 65211 USA. E-mail: prasaths@missouri.edu. D. Vorotnikov is with the Department of Mathematics, University of Coimbra, Portugal. G. Seetharaman with the Information Directorate, AFRL/RIEA, Rome NY, USA.

while sometimes creating blocky or staircasing artifacts in homogeneous regions. Various remedies have been proposed for mitigating the drawbacks of the functionals including combining both Tikhonov and TV functionals in an adaptive way. Note that the Tikhonov regularization corresponds to $p = 2$ whereas the TV functional corresponds to $p = 1$ using the traditional L^p-norm of the gradient image. That is, the general functional can be written in terms of the gradient p-norm

$$\min_u E(u) = \int_\Omega |\nabla u|^p \, dx, \qquad (1)$$

where the minimizer u is the restored image, and $\Omega \subset \mathbb{R}^2$ is the image domain with $p \geq 1$.

Recently, the pseudo-p-norm with $0 < p < 1$ has been advocated by some researchers for better regularization of images [7]–[10]; see [11] for a review of different regularization terms applicable for image segmentation. On the other hand, generalizations of the exponent to adaptively vary $p \in (1, 2]$ has also been considered. The general energy minimization model can be written as,

$$\min_u E(u) = \int_\Omega \Phi(x, \nabla u) \, dx + \frac{\mu}{2} \int_\Omega f(u, u_0) \, dx \qquad (2)$$

where f denotes the data fidelity function with u_0 being the noisy input image, and depends on the type of noise process contaminating the image. The first term Φ is a gradient-based regularization term. Nonstandard growth functionals like

$$\Phi(x, \nabla u) = |\nabla u|^{p(|\nabla u(x)|)} \qquad (3)$$

where $p(\cdot)$ is a smooth monotonically decreasing function such that $\lim_{s \to 0} p(s) = 2$ and $\lim_{s \to \infty} p(s) = 1$ were first proposed by Blomgren et al., [12]. Bollt et al [13] used smoothed gradients in the exponent function $p(|\nabla G_\sigma \star u_0|)$ and $p(|\nabla G_\sigma \star u|)$. Chen et al [14] utilized the following regularization,

$$\Phi(x, \nabla u) = \begin{cases} \frac{1}{p(x)} |\nabla u|^{p(x)} & \text{if } |\nabla u| < \epsilon, \\ |\nabla u| - \frac{\epsilon p(x) - \epsilon^{p(x)}}{p(x)} & \text{if } |\nabla u| \geq \epsilon, \end{cases} \qquad (4)$$

with $\epsilon > 0$ is fixed, and the exponent is chosen as,

$$p(x) = 1 + \frac{1}{1 + k |\nabla G_\sigma \star u_0(x)|}. \qquad (5)$$

where $k > 0$ is a parameter. This functional is connected to the classical Chambolle and Lions model [15], see also [16].

Recently, Li et al [17] used a variable exponent functional $p = p(x)$ which is studied in the Musielak-Orlicz spaces. Guo et al [18] studied an adaptive exponent based on anisotropic diffusion models, see also [16], [19]–[23] for related adaptive anisotropic diffusion PDE models. Tremendous progress has been made on the variable exponent spaces, corresponding functionals, and $p(x)$-Laplacian diffusion models with applications in various applied mathematical areas, as reviewed in [24].

In this paper our novel contribution is to develop an adaptive variable exponent functional which automatically and continuously balances between edge and corner preserving L^1 total variation and L^2 Tikhonov-smoothing using image-based local feature information that is multiscale and anisotropic. The exponent $p(x)$ is chosen according to the pointwise structure tensor computed from the image which allows the regularization method to denoise the image by modifying the amount of smoothness automatically to fit different image features. The smoothed structure tensor can identify different image features such as flat, edge and corner regions [25]. The variational model which we derive adapts between total variation which is edge preserving to the quadratic Tikhonov case which helps in removing noise effectively. In contrast other gradient only choice of exponents [14], [17], [26] inherit the undesirable blockiness artifacts associated with TV models, whereas our multiscale structure tensor approach provides better image restoration that preserves edge and corner regions more accurately. The adaptive anisotropic models such as Chen et al [26] or Guo et al [18] are adhoc methods and lack a strong theoretical foundation in contrast to the model proposed in this paper.

The proposed adaptive exponent operator can be combined with other regularization functions which makes the proposed approach a global image model for selective smoothing and restoration. Well-posedness results for our adaptive smoothed structure tensor (SST) driven variational - PDE model are studied. Mathematical analysis of the proposed method in variable exponent spaces is described for our method. Experimental results on different noisy images indicate the advantage of the proposed adaptive variable exponent variational - PDE model. Moreover, we extend it to handle multichannel and multiplicative noise corrupted images. Further applications in smoothing based thresholding segmentation on real and biomedical images highlight the applicability of the method. The proposed approach is proven to provide better edge preserving smoothing with effective noise removal as illustrated by variety of examples.

The rest of the paper is organized as follows. Section II introduces the structure tensor driven variational scheme along with detailed mathematical analysis of the variational - PDE scheme. Section III illustrates the implementation details along with extensive noisy image denoising, segmentation and restoration results. Finally, Section IV concludes the paper indicating future directions.

(a) $Corner$ (b) $Steps$

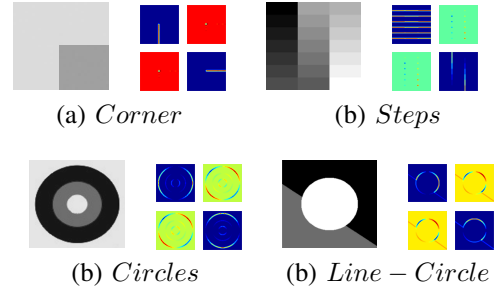(b) $Circles$ (b) $Line - Circle$

Fig. 1. Structure tensor components characterize shape features in a given image. In each sub-figure we show the entries of the tensor (6) as images to the right of various synthetic (noise-free) images including $Corner$, $Steps$, $Circles$, and $Line - Circle$.

II. MULTISCALE TIKHONOV-TOTAL VARIATION (MTTV) BLENDED RESTORATION

A. Adaptive smoothed scale space structure tensor

The $p(x)$ growth regularization functionals studied before typically are gradient based, i.e., $p(x) = p(|\nabla u(x)|)$ or $p(x) = p(|G_\sigma \star \nabla u(x)|)$. Thus, the edge maps given by these can inherit the traditional problems associated with such gradient based functionals such as blocky artifacts [27], edges dislocation [22], corner smoothing [28], etc. To mitigate such problems one can consider the smoothed structure tensor (SST),

$$K_\sigma(u) = G_\sigma \star (\nabla u \nabla u^T) = \begin{pmatrix} G_\sigma \star u_x^2 & G_\sigma \star u_x u_y \\ G_\sigma \star u_y u_x & G_\sigma \star u_y^2 \end{pmatrix} \quad (6)$$

where $G_\sigma(\mathbf{x}) = (\sigma\sqrt{2\pi})^{-1} \exp(-|\mathbf{x}|^2/2\sigma^2)$ is the 2D Gaussian kernel, $\star$ denotes the convolution operator and superscript T is the transpose. The use of the structure tensor in locally adaptive regression steering kernels for adaptive image filtering is noted as being exceedingly robust to noise and perturbations of the data [29]. In Figure 1 we show the entries of the matrix as images for different images with commonly occurring features. Let the eigenvalues of the SST K_σ be (λ_+, λ_-), with eigenvectors (θ_+, θ_-). The eigenvalues (λ_+, λ_-) which are the maximum and minimum respectively, describe average contrast within a neighborhood of size $O(\sigma)$ along the eigen-directions. The eigenvectors (θ_+, θ_-) describe the orientation which maximizes gray value fluctuations and preferred local direction of smoothing respectively. The SST characterizes different image regions based on the range of eigenvalues to indicate flat, edge and corner regions. The SST has been advocated as an unifying choice and represents truly anisotropic features as evidenced in Figure 1. We refer to [30] for more discussion on anisotropic diffusion using SST.

Remark 1. *The SST characterizes orientation energy and is also called the second moment matrix with the pre-smoothing by G_σ in each entry to avoid the ill-posedness of computing derivatives. Thus, this smoothing makes it insensitive to small scale details less than $O(\sigma)$. This σ parameter is referred to as integration scale and it provides a window size over which orientation is analyzed.*

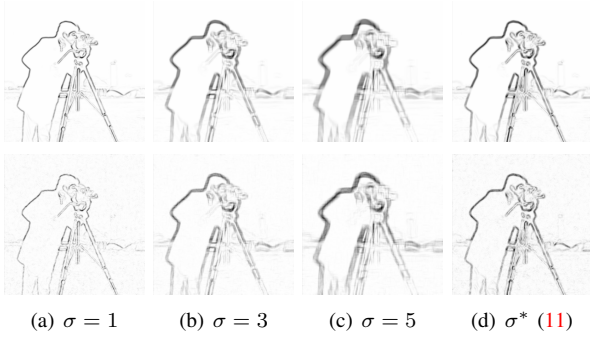**Remark 2.** *An additional Gaussian convolution of the tensor entries with standard deviation ρ can also be incorpo-*

(a) $\sigma = 1$ (b) $\sigma = 3$ (c) $\sigma = 5$ (d) σ^* (11)

Fig. 2. Effect of parameter σ on the eigenvalue λ_+ of the SST matrix (6) for noise-free (top row) and noisy std $\sigma_n = 30$ (bottom row) *Cameraman* image. λ_+ is shown here with single scale based SSTs (6) (a-c), and multiscale version of the SST (11).

rated [31],

$$J_{\rho,\sigma}(u) := G_\sigma \star (\nabla u_\rho \nabla u_\rho^T),$$

where $u_\rho := G_\rho \star u$ is the additional smoothing and ρ can be thought of as noise scale.

B. Growth variational regularization using adaptive exponents

Following the above discussion, we make use of a variable exponent based minimization functional driven by the smoothed structure tensor. The proposed non-standard growth variational functional,

$$\min_u E(u) = \int_\Omega |\nabla u|^{p(x,\sigma)} \, dx + \frac{\mu}{2} \int_\Omega (u - u_0)^2 \, dx \quad (7)$$

where $\mu \geq 0$ is the regularization parameter; this variational PDE operator is local, non-linear, anisotropic and spatially varying. The quadratic data fidelity term in Eqn. (7) is by the additive Gaussian model[1] and can also be modified according to other noise models, see Section III-C2. The variable exponent function is chosen as,

$$p(x,\sigma) = 1 + \exp\left(-\frac{2\,\lambda_+(x,\sigma)\,\lambda_-(x,\sigma)}{k + \lambda_+(x,\sigma) + \lambda_-(x,\sigma)}\right)$$
$$\times \exp\left(-2\left(\frac{\lambda_+(x,\sigma) - \lambda_-(x,\sigma)}{k + \lambda_+(x,\sigma) + \lambda_-(x,\sigma)}\right)^2\right) \quad (8)$$

where parameter $k > 0$ is added for numerical stability.

Here, we use the notation $p(x,\sigma)$ to indicate that the eigenvalues of the SST matrix (6) are computed for the particular scale $\sigma > 0$. The first exponential term is

$$\mathcal{H}(x,\sigma) = \left(\frac{\lambda_+(x,\sigma)\,\lambda_-(x,\sigma)}{k + \lambda_+(x,\sigma) + \lambda_-(x,\sigma)}\right) = \frac{det K_\sigma}{trace K_\sigma} \quad (9)$$

which we note is the harmonic mean of the eigenvalues (and is also referred to as the Förstner corner operator [31] or Noble corner measure [32], [33] that is sometimes used as an alternative to the closely related Harris corner measure [34]) and emphasizes responses to local edge and corner structures

---

[1]Additive noise, $u_0(x) = u(x) + n(x)$. Noise follows a Gaussian distribution with mean zero and variance σ_n^2, i.e., $n(x) \sim \mathcal{N}(0, \sigma_n^2) = (\sigma_n\sqrt{2\pi})^{-1} e^{(-|\mathbf{x}|^2/2\sigma_n^2)}$.

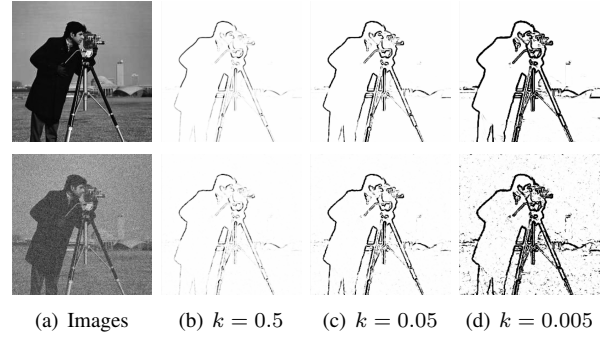(a) Images   (b) $k = 0.5$ (c) $k = 0.05$ (d) $k = 0.005$

Fig. 3. Effect of parameter k on the variable exponent $p(x,\sigma^*)$ in Eqn. (12) for noise-free (top row) and noisy $\sigma_n = 30$ (bottom row) *Cameraman* image.

in the image that contain high spatial frequencies. The second term

$$\mathcal{C}(x,\sigma) = \left(\frac{\lambda_+(x,\sigma) - \lambda_-(x,\sigma)}{k + \lambda_+(x,\sigma) + \lambda_-(x,\sigma)}\right)^2 \quad (10)$$

is known as *coherence* and we normalize the eigenvalues to $[0, 1]$ before computing the coherence based exponent (8). The motivation for introducing the variable exponent $p(x,\sigma) \in (1, 2]$ based on the structure tensor (8) is due to the fact that the amount of regularization on the gradient norm can be automatically adjusted according to different image features. The following image features are adaptively handled by the above SST driven variational scheme:

- If $0 \ll \lambda_- \leq \lambda_+$, then a corner is present as both eigenvalues are large and $p(x,\sigma) \to 1$ which in turn implies that corner points are not smoothed out with $\mathcal{H}$ large and $\mathcal{C} \approx 1$.
- If $0 \approx \lambda_- \ll \lambda_+$, then an edge is present and $p(x,\sigma) \to 1.135 = (1 + 1/e^2)$, thus object boundaries are preserved with $\mathcal{H} \approx 0$ and $\mathcal{C} \approx 1$.
- Finally, if both $\lambda_+, \lambda_- \approx 0$, then it represents homogeneous or noisy regions and $p(x) \to 2$ which implies stronger smoothing and hence the noise is removed from flat regions with $\mathcal{H} \approx 0$ and $\mathcal{C} \approx 0$.

C. Multiscale TTV (MTTV)

Figure 2 shows the effect of pre-smoothing parameter σ used in the SST (6) on λ_+ eigenvalue for the *Cameraman* test image. As can be seen a right balance of σ is important, as the edge map provided by the exponent (8) depends on the eigenvalues computed from the SST (6). Larger σ values tend to blur the edge map and smaller can miss certain edge pixels. We instead use a multi scale strategy for selecting the right pre-smoothing parameter σ which is described in detail next.

Multi-scale structure tensor responses are computed by selecting a range of G_σ where σ is the standard deviation of the Gaussian distribution.

$$\sigma^* = \underset{\sigma}{\mathrm{argmax}}\ |\lambda_+(x,\sigma)| \quad (11)$$

We take the corresponding λ_- value and use it in the exponent,

$$p(x,\sigma^*) = 1 + \exp\left(-2\mathcal{H}(x,\sigma^*)\right) \exp\left(-2\mathcal{C}(x,\sigma^*)\right) \quad (12)$$

and utilize the minimization of the functional (7). Controlling the range of σ in (11) is important for getting the desired

response and it is dependent on the amount of noise. The normalized derivatives as shown in Eqn. (6) assume local maxima over scales [35], [36] which corresponds to locally strongest normalized feature response. Taking the maximum over a range of σ results in responses from multiple scales. Natural images have a variety of edge and corner features of different sizes and intuitively, using a range of scales will help capture responses that a single scale cannot do so.

Figure 3 shows the test image $Cameraman$ used in our experiments and its corresponding edge map shown for the exponent $p(x, \sigma^*)$ defined in Eqn. (12) for different parameter values k with multi scale SST based eigenvalues (11). For medium noise level $10 \leq \sigma_n \leq 30$, the threshold value $k = 0.05$ works fine. As can be seen the flat regions are indicated by the value greater than one and other edge and ramp regions are correctly labeled as unity. Although the adaptive SST driven variable exponent $p(x, \sigma^*)$ in Eqn. (12) is utilized in the simple energy minimization of the form in Eqn. (7), it can be used in other variational and PDE models as well[2]. Note that when the exponent attains the critical value one ($p(x) \to 1$, the TV case) the existence of a solution to the corresponding minimization problem (7) is non-trivial, see [37]. It can be seen that $p(x) \in (1, 2]$ and thus the smoothness is restricted from the TV case, $p = 1$, which can lead to staircasing in the results. Detailed analysis (existence theorems, decay of solutions) of the proposed MTTV method using multiscale exponents in the variational PDE model for blended restoration are given in the Appendix.

Remark 3. *The best scale at each pixel* (11) *can also be determined by taking the maximum response of each eigenvalue separately (See Chap. 13 in [38]):*

$$\lambda_{max+} = arg \max_{\sigma} |\lambda_+(x, \sigma)| \, , \lambda_{max-} = arg \max_{\sigma} |\lambda_-(x, \sigma)| \, .$$

Surprisingly, we observed that using a mixture of scales for the min and max eigenvalues, in Eqn. (8), *instead of the same scale as shown in Eqn.* (11) *leads to visually similar results for edge preserving smoothing.*

Remark 4. *The best scale can be found by maximizing the coherence function directly over scale space instead of for the largest eigenvalue. But this selection of σ^* may be more sensitive to noise.*

III. EXPERIMENTAL RESULTS

A. Discretization and implementation details

To implement the proposed SST feature driven variational exponent Multiscale Tikhonov Total Variation (MTTV) method with $p(x)$ in (12) computed from the initial noisy image u_0 and the dynamic D-MTTV method using $p(x)$ computed from the previous iteration image $u(x, t-1)$ we use the corresponding dynamical PDE (24-26). We use an explicit finite difference approximation scheme (central differences for the spatial variables, forward difference for the time variable) to solve the above PDE. Following [39] we can prove that the

---

[2]In what follows we use the simplified notation $p(x)$ for the exponent with the understanding it represents the multi scale exponent (12).

(a) Noisy (b) TIREG (1), $p = 2$ iteration 10, 40 (c) TVREG (1), $p = 1$ iteration 100, 200

(d) $p = 1.25$ (e) $p = 1.5$ (f) $p = 1.75$ (g) $p(x)$ (h) D-MTTV
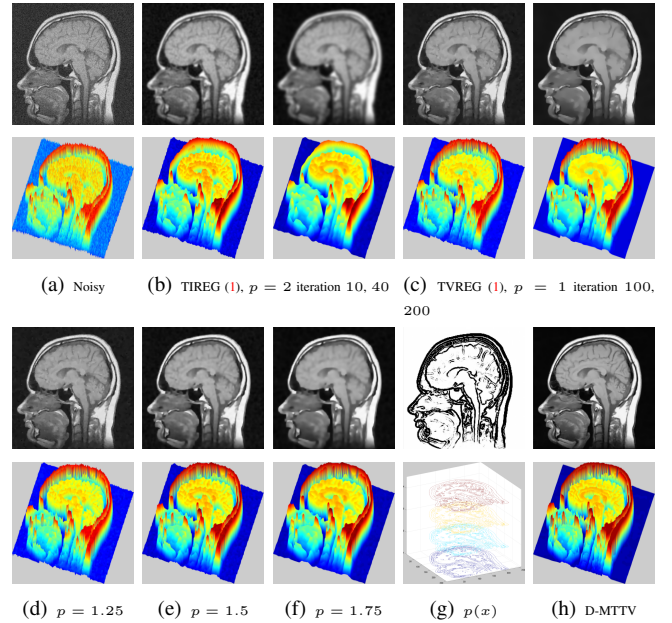
Fig. 4. Comparison results for additive Gaussian noise corrupted synthetic $Shapes$ gray scale image with classical regularization methods and different scalar p-exponent based regularizations. (a) Noisy image (b) TIREG (1) with $p = 2$ (c) TVREG (1) with $p = 1$ (d)-(f) Different scalar p value based results (g) Multiscale SST based exponent $p(x)$ (12) from noisy image for scales $\sigma = \{1, 2, 3, 4\}$ (h) D-MTTV. Top row: images Bottom row: surface.

discrete scheme is unconditionally stable and thus the time step size (taken here $\triangle t = 0.2$) can be chosen big without worrying about instabilities. Moreover, the discrete scheme satisfies the maximum-minimum principle, which ensures that artificial edge regions are not created during the restoration process. For images with Gaussian white noise level $\sigma_n = 10$ to 30 we set $k = 0.05$ and the range[3] of σ in (11) to be the interval $[1, 4]$.

The smoothing with Gaussian kernel in (6) is approximated by a fast box filtering in all our experiments[4]. Our scheme takes on average 0.2 seconds for a gray scale image of size 256×256 using a MATLAB implementation on a 2.3 GHz Intel Core i7, 8GB 1600 MHz DDR3 Mac Laptop. All the schemes listed below for comparison were solved using their corresponding time dependent PDE formulation and explicit finite difference schemes [39]. The termination condition[5] for all the schemes is chosen automatically according to the stopping criteria $\left\| u^{t+1} - u^t \right\|_2 \leq 10^{-4}$.

B. Gaussian noise removal

1) Comparative results: We first compare our scheme with classical schemes such as the regularization methods of Tikhonov (TIREG, $p = 2$ in (1)), total variation of Rudin et al [6] (TVREG, $p = 1$ in (1)) as well as scalar p exponent based regularization. Figure 4 shows a comparison of different regularization results for the $Brain$ gray scale image of size

---

[3]We take maximum on discrete scale levels $\{1, 2, 3, 4\}$ to find σ^*, see Eqn. (11).

[4]See Supplementary for the justification of using Box filter, see also Section Chap. 14 in [38].

[5]Other options are also possible, for e.g., stopping when a maximum PSNR/SSIM value is reached, but will require extensive tuning of related parameters in each scheme and noise-free image for reference.

256×256 which is corrupted by Gaussian noise of strength 20%. Figure 4(b) illustrates the over smoothing effect of Tikhonov regularization, though effective in noise removal excessive smoothing of edges renders the result at higher iteration not useful. In contrast TV regularization obtains a better restoration though small staircasing is visible in low iterations and piecewise constant result is obtained for higher iteration. Intermediate scalar p exponent based regularizations are given in Figure 4(d-f). The proposed variable exponent $p(x)$ map (12) is shown in Figure 4(g) which captures multi scale edges. Figure 4(h) show that proposed scheme D-MTTV provides better restoration without artifacts and edges are well preserved. To show edge preservation properties of different schemes we show the corresponding surface maps where the traditional regularization methods such as TIREG, TVREG and scalar exponent p-regularization have various artifacts whereas the proposed MTTV schemes performs better overall.

We next compare some contemporary denoising filters widely used in the image processing area. Although we do not provide details of the different methods compared due to space limitations[6], the connections between a variety of image denoising methods in the context of nonparametric point estimation and kernel functions is discussed in [29]. Milanfar [29] observes that the most successful modern filtering approaches are nonparametric, adaptive to signal content and iterative (repeated application of a filter or sequence of filters followed by data aggregation). We note that the proposed MTTV method and its variants incorporate all three elements for improved performance.

Figures 5 illustrate the results along with their level lines given to highlight various problems with other schemes. Figure 5(a) is the original synthetic test image, and Figure 5(b) is the noisy image. Figures 5(c)-(t) show the denoising results with the bilateral filtering (BLF) [40], Gauss curvature-driven diffusion (GCDD) [41], Perona–Malik (PM) [1], nonlinear complex diffusion (NCD) [42], adaptive smoothing via contextual, local discontinuities (ASCL) [43], nonlocal means (NLM) [44], linear regression Yaralovsky neighborhood filter (LYNF) [45], robust anisotropic diffusion (RAD) [46], adaptive TV method (ATV) [47], fourth-order PDE (FPDE) [48], TVREG method [6], combined model of TV filter and fourth-order PDE (TV2&4) filter [49], edge-flat-grey (EFG) scale filtering [50], Gauss-TV (GTV) filtering [51], nonlinear diffusion filtering with an additive operator splitting (AOS) scheme [39], and ATVDC model [26] (see Eqn. (13) below), respectively. Figure 5(s-t) are our MTTV and DMTTV results which shows that we obtain improved noise reduction, edge preservation and fewer staircasing artifacts especially for smooth gradients in the horizontal and vertical grayscale ramps in the upper left and lower right of the image.

Finally, we compare various adaptive exponent based variational-PDE schemes from the recent literature. Adaptive variable exponent based scheme of Chen et al [14] (VAREG, Φ is given in (4)), Li et al [17] (IGEDV, $p(x) = 1 + (1 + |G_\sigma \star \nabla u_0(x)|)^{-1}$ in (7)). Further, we compare with three recent works which consider ad-hoc formulations for

[6]See Supplementary materials.

(a) Original (b) Noisy (c) BLF (d) GCDD (e) PM

(f) NCD (g) ASCL (h) NLM (i) LYNF (j) RAD

(k) ATV (l) FPDE (m) TVREG (n) TV2&4 (o) EFG
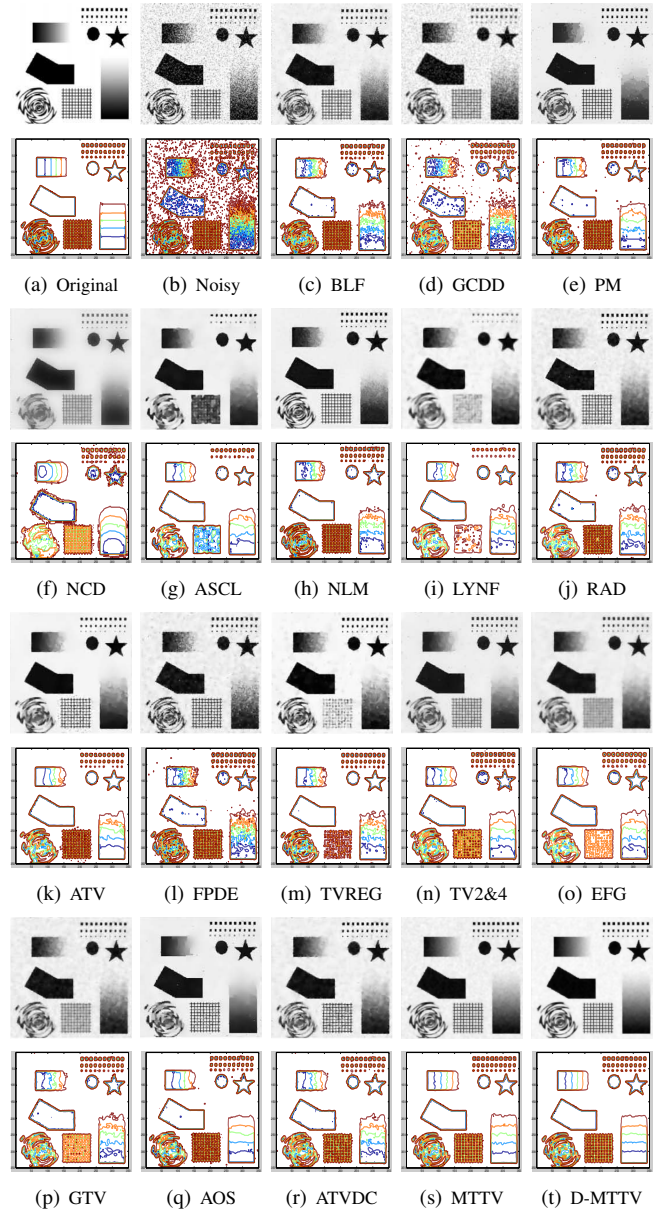
(p) GTV (q) AOS (r) ATVDC (s) MTTV (t) D-MTTV

Fig. 5. Comparison results for additive Gaussian noise corrupted synthetic *Shapes* gray scale image with different denoising methods. (a) Original image (b) Noisy image obtained by adding Gaussian noise $\sigma_n = 30$ (c) [40] (d) [41] (e) [1] (f) [42] (g) [43] (h) [44] (i) [45] (j) [46] (k) [47] (l) [48] (m) [6] (n) [49] (o) [50] (p) [51] (q) [39], (r) [26] (s) MTTV (t) D-MTTV. Better viewed online and zoomed in.

controlling the smoothness[7]:

- Adaptive total variation with difference curvature (ATVDC) model of Chen et al [26]:

$$\min_u E(u) = \int_\Omega |\nabla u|^{p(D)} \, dx \tag{13}$$

where the exponent $p(D) = 2 - \sqrt{\overline{D}}$ with $\overline{D}$ the normalized difference curvature, $D = ||u_{\eta\eta}| - |u_{\xi\xi}||$, with derivatives parallel and perpendicular to the gradient direction.

[7]Note that we have excluded the fidelity term in ATVDC, LAADE, $\alpha(x)$-PM, D-$\alpha(x)$-PM models and our scheme i.e., we take $\mu = 0$ in Eqn. (7). Adding the fidelity term in the Gaussian noise case decreased the final PSNR values and adaptive fidelity can be used to improve the results, see [23].

- A locally adaptive version of the classical Chambolle-Lions model [15] (LAADE) recently studied by Zhou et al [16]:

$$\min_u E(u) = \int_{|\nabla u| < \beta(x)} \frac{1}{2\beta(x)} |\nabla u|^2 \, dx$$
$$+ \int_{|\nabla u| \geq \beta(x)} |\nabla u| \, dx - \frac{\beta(x)}{2} \qquad (14)$$

where the adaptive parameter $\beta(x)$ is obtained by solving a separate energy minimization problem

$$\min_\beta F(\beta) = \frac{1}{2} \int_\Omega |\nabla \beta|^2 \, dx + \frac{\lambda}{2} \int_\Omega \left(\beta - \frac{|\nabla u|}{2} \right)^2 \, dx.$$

- Adaptive Perona - Malik models of Guo et al [18]:

$$\frac{\partial u}{\partial t} = div \left(\frac{\nabla u}{1 + (\nabla u / K)^{\alpha(x)}} \right) \qquad (15)$$

with

$$\alpha(x) = 2 - \left(\frac{2}{1 + k |\nabla G_\sigma \star u_0|^2} \right) \qquad (16)$$

called $\alpha(x)$-PM, and

$$\alpha(x) = 2 - \left(\frac{2}{1 + k |\nabla G_\sigma \star u|^2} \right) \qquad (17)$$

called the dynamic $\alpha(x)$-PM (or D-$\alpha(x)$-PM) methods respectively.

We consider restoring the $Cameraman$ gray scale image of size 256×256 which is corrupted by Gaussian noise of standard deviation $\sigma_n = 30$ for testing different variable exponent models. Figure 6 shows noisy images and their SSIM maps (in $[0, 1]$). Figure 7 show corresponding denoised results and close-up shots which highlight some of the problematic regions in other schemes. Overall, our MTTV and D-MTTV methods outperform other related methods in terms of multi-scale edge preservation and improved noise reduction without introducing staircasing artifacts seen in other variable exponent and adaptive models. For example, the mouth region of the cameraman is much better preserved in MMTV and D-MMTV compared to other methods as shown in the close-up images in Figure 7.

2) Error metrics: To compare the restoration results quantitatively, we use four error measures including three standard error metrics peak signal to noise ratio (PSNR), improvement in signal to noise ratio (ISNR), Mean structural similarity (MSSIM), which are widely used in image processing literature and a new metric (PSNR$_E$) based on gradient edge maps [18] which uses the PSNR of the edge maps (EM), EM$(u) = 2 - (2/(1 + k |\nabla G_\sigma \star u|^2)$ with $k = 0.0025, \sigma = 0.5$.

$$\text{PSNR}_E(u) = 20 * \log_{10} \left(\frac{\max \text{EM}(u_O) - \min \text{EM}(u_O)}{\text{MSE}_E} \right) dB$$

where $\text{MSE}_E = (mn)^{-1} \sum \sum (\text{EM}(u) - \text{EM}(u_O))^2$. Higher PSNR$_E$ indicates the scheme performs better edge preservation by way of matching the derivatives.

Tables I-II show MSSIM, SNR, PSNR (dB) values for contemporary denoising filters and adaptive, exponent based
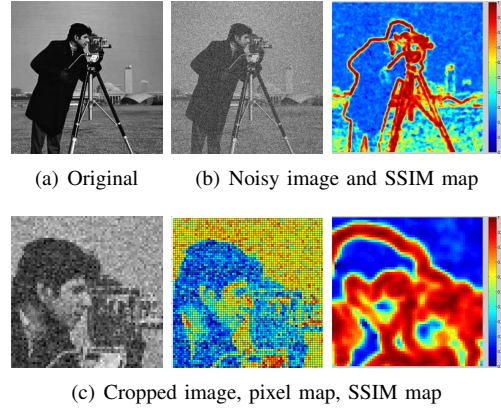

(a) Original   (b) Noisy image and SSIM map

(c) Cropped image, pixel map, SSIM map

Fig. 6. Noisy $Cameraman$ gray scale image used in our experiments and their SSIM maps. (a) Original image (b) Noisy image obtained by adding Gaussian noise $\sigma_n = 30$ and its SSIM map (c) Zoomed region with detail (left) image (middle) pixel map (right) SSIM map.

TABLE I
MSSIM/SNR(DB) VALUES CORRESPONDING TO FIGURE 5.

Fig.	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
MSSIM	0.3378	0.6276	0.4040	0.7858	0.7213	0.7287	0.7575	0.6520	0.7189
SNR	7.64	11.69	8.42	12.11	7.27	7.00	13.49	7.12	11.24
(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)	(s)	(t)
0.8244	0.6460	0.7494	0.7853	0.7237	0.6850	0.8364	0.7702	0.8012	**0.8318**
11.93	10.65	9.31	10.06	7.78	8.66	11.69	10.79	10.47	**12.64**

variational-PDE schemes corresponding to Figures 5-7 respectively. Tables III-IV show PSNR/PSNR$_E$ (dB), ISNR (dB)/MSSIM values for synthetic and other test images for different schemes[8]. Overall our scheme performs better and higher MSSIM values indicate we preserve salient structures and PSNR improvements further support our claim of efficient noise removal. Following [26], [41] we use the normalized step difference energy (NSDE) at every iteration,

$$\text{NSDE} = \frac{|u(\cdot, t) - u(\cdot, t-1)|^2}{u(\cdot, t)^2} \qquad (18)$$

as measure of convergence and Figure 8 shows a comparison with [26] for the restoration of noisy synthetic test image from Figure 5(b) and $Cameraman$ image from Figure 6(b). Our scheme converges faster and few iterations are usually enough to obtain good denoising results. Figure 9 shows PSNR$_E$ values for different noise levels on standard test images and our D-MTTV outperforms all the schemes overall.

C. Further applications

1) Thresholding based segmentation: Improved denoising leads to improved threshold estimation for image binarization and foreground-background region partitioning. Our nonlinear D-MTTVsmoothing can be used as a pre-processing step

---

[8]For the restoration of Gaussian noise corrupted ($\sigma_n = 30$) for different schemes are given here. We provide error metrics for the restoration results of 17 standard test images from the USC-SIPI Miscellanies data-set in a supplementary file. Further results for different noise levels, stopping criteria, data-sets, error metrics, SSIM maps, and denoised images are available at the project website: http://cell.missouri.edu/pages/mttv

TABLE III
PSNR/PSNR$_E$ (DB) COMPARISON OF VARIOUS SCHEMES FOR SYNTHETIC AND BIOMEDICAL TEST IMAGES. NOISY IMAGE IS OBTAINED BY ADDING GAUSSIAN NOISE OF STRENGTH $\sigma_n = 30$ TO THE ORIGINAL IMAGE. EACH COLUMN INDICATES THE PSNR/PSNR$_E$ VALUES FOR DIFFERENT TEST IMAGES. BEST RESULTS ARE IN **BOLDFACE**.

Images	TIREG $p = 2$	TVREG $p = 1$	VAREG [14]	IGEDV [17]	ATVDC [26]	LAADE [16]	APM [18]	DPM [18]	D-MTTV
Corner	25.61 / 10.60	34.56 / 24.92	34.62 / 21.66	36.81 / 24.31	37.18 / 25.75	34.77 / 21.37	36.07 / 22.34	33.83 / 19.15	**44.78** / **34.61**
Steps	27.33 / 15.59	35.85 / 25.28	37.76 / 26.92	36.01 / 26.81	38.04 / 27.98	38.23 / 26.99	36.66 / 24.60	36.21 / 24.69	**43.44** / **32.12**
Circles	14.33 / 6.13	30.51 / 20.87	31.13 / 20.77	30.18 / 20.39	31.02 / 20.99	31.40 / 20.66	29.95 / 20.13	29.89 / 20.35	**36.35** / **28.49**
Line-C.	21.31 / 11.96	34.98 / 23.20	35.36 / 21.51	34.01 / 22.82	35.46 / 22.97	36.12 / 20.98	33.51 / 21.41	33.29 / 21.51	**43.13** / **30.27**
Kiki	25.27 / 12.73	33.18 / 21.52	33.85 / 21.77	33.15 / 21.95	34.13 / 22.13	34.35 / 21.85	34.20 / 21.65	33.56 / 21.04	**36.23** / **23.83**
Shapes1	14.69 / 7.18	30.80 / 22.07	32.64 / 22.73	32.46 / 23.13	33.36 / 23.66	32.74 / 22.54	36.17 / 23.63	35.59 / 23.60	**37.16** / **29.43**
Shapes2	12.10 / 4.61	29.59 / 21.02	30.98 / 21.73	30.06 / 21.17	30.81 / 21.85	31.30 / 21.76	29.02 / 20.20	28.97 / 20.41	**36.41** / **29.19**
Bacteria1	17.27 / 4.20	25.49 / 10.00	25.16 / 9.26	25.88 / 10.60	25.70 / 9.99	24.48 / 8.74	23.64 / 8.47	24.02 / 8.69	**28.55** / **12.50**
Bacteria2	19.19 / 6.65	29.68 / 14.80	30.08 / 14.85	30.32 / 16.24	30.57 / 16.00	28.71 / 13.62	25.90 / 11.45	27.08 / 12.25	**34.51** / **19.58**
Mammo.	31.57 / 20.93	31.49 / 21.37	31.91 / 21.43	31.44 / 21.72	31.86 / 21.54	31.86 / 21.16	31.82 / 21.08	31.81 / 21.37	**32.78** / **22.16**
Brain	15.23 / 2.37	23.96 / 9.92	23.50 / 9.18	24.71 / 11.11	24.40 / 10.43	22.99 / 8.51	22.69 / 7.95	22.88 / 8.19	**27.96** / **13.76**
Nemacb3	24.37 / 9.24	26.91 / 12.16	26.48 / 11.03	27.15 / 12.96	26.99 / 12.06	25.78 / 10.39	25.77 / 10.40	26.14 / 10.61	**28.64** / **14.78**

TABLE IV
TABLE2 ISNR (DB)/MSSIM COMPARISON OF VARIOUS SCHEMES FOR SYNTHETIC AND BIOMEDICAL TEST IMAGES. NOISY IMAGE IS OBTAINED BY ADDING GAUSSIAN NOISE OF STRENGTH $\sigma_n = 30$ TO THE ORIGINAL IMAGE. EACH COLUMN INDICATES THE ISNR/MSSIM VALUES FOR DIFFERENT TEST IMAGES. BEST RESULTS ARE IN **BOLDFACE**.

Images	TIREG $p = 2$	TVREG $p = 1$	VAREG [14]	IGEDV [17]	ATVDC [26]	LAADE [16]	APM [18]	DPM [18]	D-MTTV
Corner	6.53 / 0.7244	15.48 / 0.9756	15.54 / 0.9546	17.72 / 0.9766	18.09 / 0.9781	15.69 / 0.9553	16.99 / 0.9777	14.74 / 0.9507	**25.70** / **0.9960**
Steps	8.77 / 0.8966	17.28 / 0.9327	19.20 / 0.9527	17.45 / 0.9211	19.48 / 0.9539	19.68 / 0.9584	18.09 / 0.9500	17.65 / 0.9416	**24.89** / **0.9827**
Circles	−4.27 / 0.2975	11.90 / 0.8337	12.52 / 0.8587	11.57 / 0.8183	12.41 / 0.8563	12.79 / 0.8911	11.34 / 0.8470	11.28 / 0.8319	**17.74** / **0.9410**
Line-C.	2.75 / 0.7658	16.42 / 0.8412	16.80 / 0.8661	15.52 / 0.8032	16.90 / 0.8626	17.56 / 0.9322	14.95 / 0.8822	14.73 / 0.8418	**24.57** / **0.9679**
Kiki	6.56 / 0.8488	14.62 / 0.9177	15.29 / 0.9327	14.59 / 0.9044	15.57 / 0.9332	15.79 / 0.9415	15.64 / 0.9369	15.00 / 0.9251	**17.66** / **0.9515**
Shapes1	−3.87 / 0.4241	12.23 / 0.7783	14.08 / 0.8139	13.90 / 0.7634	14.80 / 0.8164	14.18 / 0.8747	17.61 / 0.8402	17.03 / 0.8001	**18.60** / **0.9480**
Shapes2	−6.45 / 0.3671	11.03 / 0.8823	12.43 / 0.9092	11.50 / 0.8717	12.26 / 0.9056	12.75 / 0.9378	10.46 / 0.8929	10.41 / 0.8862	**17.85** / **0.9707**
Bacteria1	−1.30 / 0.2925	6.93 / 0.6407	6.60 / 0.6129	7.32 / 0.6658	7.13 / 0.6421	5.92 / 0.5735	5.07 / 0.5665	5.46 / 0.5886	**9.99** / **0.7454**
Bacteria2	0.61 / 0.5367	11.10 / 0.8527	11.51 / 0.8611	11.75 / 0.8687	12.00 / 0.8724	10.14 / 0.8287	7.33 / 0.7811	8.51 / 0.8147	**15.94** / **0.9133**
Mammo.	13.00 / 0.8336	12.93 / 0.8273	13.34 / 0.8436	12.87 / 0.8200	13.30 / 0.8426	13.29 / 0.8378	13.26 / 0.8410	13.24 / 0.8400	**14.22** / **0.8768**
Brain	−3.31 / 0.3156	5.42 / 0.7120	4.96 / 0.7138	6.17 / 0.7193	5.85 / 0.7435	4.45 / 0.7234	4.14 / 0.6788	4.33 / 0.6668	**9.42** / **0.8857**
Nemacb3	5.81 / 0.3739	8.35 / 0.4897	7.92 / 0.4564	8.59 / 0.5101	8.43 / 0.4864	7.22 / 0.4180	7.21 / 0.4228	7.57 / 0.4450	**10.08** / **0.5642**

TABLE II
MSSIM/PSNR(DB) VALUES CORRESPONDING TO FIGURE 7.

Fig.	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
MSSIM	0.7741	0.7923	0.7954	0.7607	0.7674	0.7689	0.8359	**0.8673**
PSNR	25.49	26.58	26.30	25.05	25.59	25.67	28.06	**29.29**

for thresholding based bi-modal segmentation and can be of potential use in machine vision and biomedical imagery [52], [53]. For comparison with ground-truth (GT) we use the Dice coefficient between two binary images,

$$D(A, B) = \frac{2\,|A \cap B|}{|A| + |B|} \qquad (19)$$

where $|A|$ denotes the number of elements in the set. Dice coefficient closer to one indicate that automatic segmentation is closer to ground truth. Figure 10 shows segmentation results for synthetic $Spiral$ image used in [54] for different high Gaussian noise levels. As can be seen from corresponding histograms, we obtain good binary segmentations even at higher noise levels. For example, at severe noise $\sigma_n = 150$ the bi-modality of the histogram is lost (Figure 10(c)) and after smoothing with our D-MTTV method the histogram (Figure 10(g)) cleary shows a separation of peaks and thresholding works quite well. Additionally, comparing with the segmentation model tested in [54] our scheme is accurate as shown by Dice values. Figure 11 shows some sample images used in the non-destructive evaluation (NDT) field and their corresponding smoothing based thresholding results. As can be seen, our D-MTTV smoothing based thresholding obtains good segmentations compared to ground truth results.
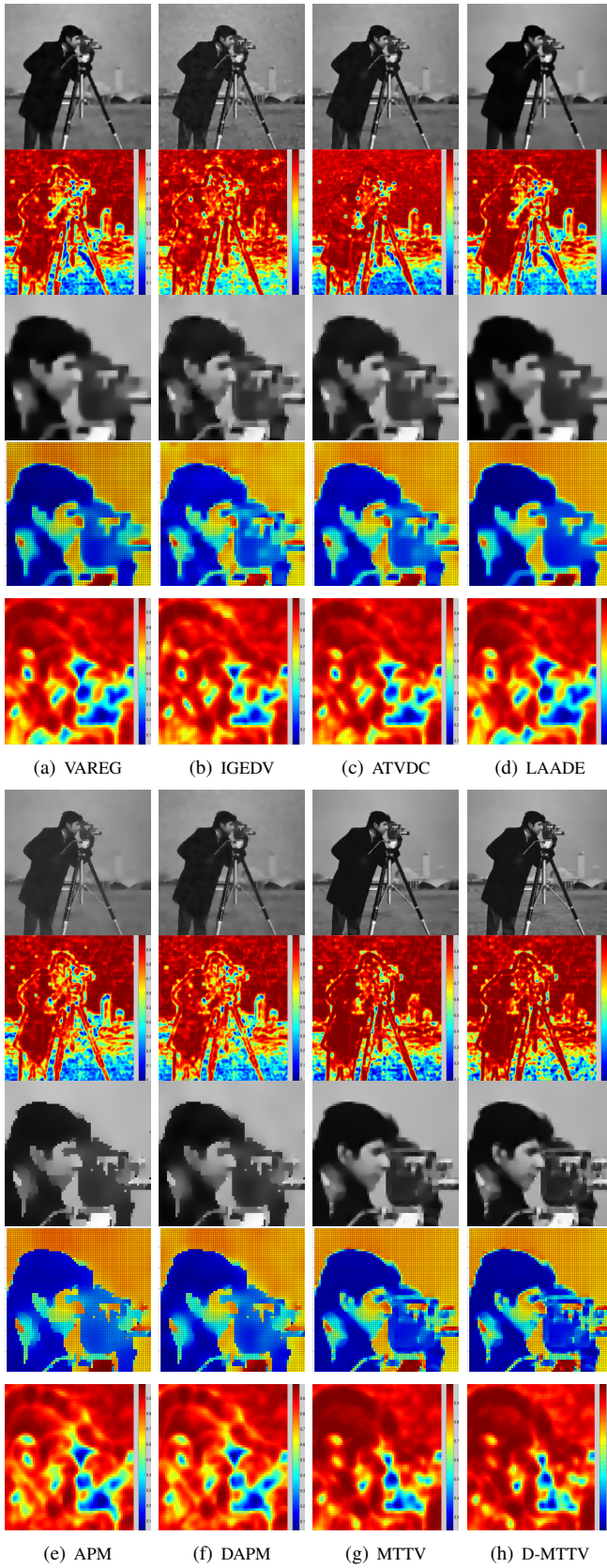
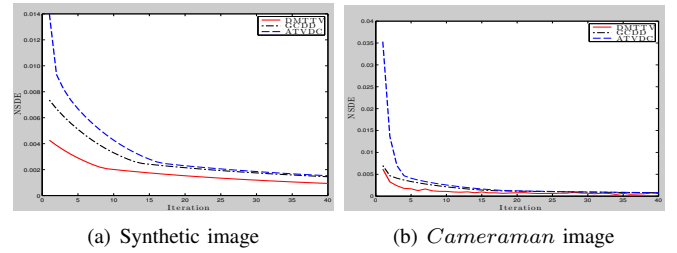(a) Synthetic image   (b) $Cameraman$ image

Fig. 8. Iteration versus NSDE for the restoration of (a) noisy synthetic image in Figure 5(b) and (b) noisy $cameraman$ image in Figure 6(b) with different schemes. Convergence curves for our D-MTTV, GCDD [41] and ATVDC [26] using NSDE (18).
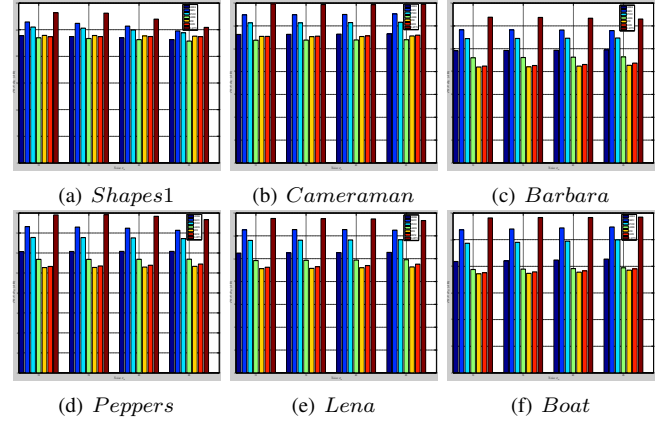

(a) $Shapes1$ (b) $Cameraman$ (c) $Barbara$

(d) $Peppers$ (e) $Lena$ (f) $Boat$

Fig. 9. Noise level $\sigma_n = \{10, 15, 20, 25\}$ along the x-axis versus $PSNR_E$ (dB) values for six different images to evaluate seven different denoising algorithms VAREG, IGEDV, ATVDC, LAADE, APM, DAPM and D-MTTV (colored bars).
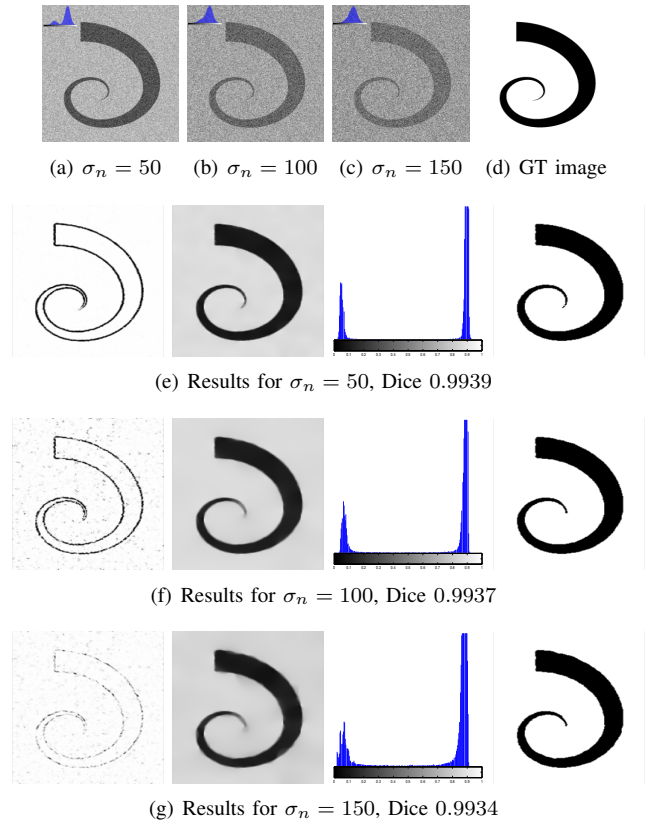

(a) VAREG   (b) IGEDV   (c) ATVDC   (d) LAADE

(a) $\sigma_n = 50$ (b) $\sigma_n = 100$ (c) $\sigma_n = 150$ (d) GT image

(e) Results for $\sigma_n = 50$, Dice 0.9939

(f) Results for $\sigma_n = 100$, Dice 0.9937

(e) APM   (f) DAPM   (g) MTTV   (h) D-MTTV

Fig. 7. Comparison results for additive Gaussian noise ($\sigma_n = 30$) corrupted $Cameraman$ gray scale image with adaptive methods. Top to bottom: Image, SSIM map, cropped image, pixel map, SSIM map. Best results of: (a) [14] (b) [17] (c) [26] (d) [16] (14) (e) [18] (15) with (16) (f) [18] (15) with (17) (g) Our MTTV (7) with (12) computed from noisy image u_0 (h) Our dynamic MTTV (7) with (12) computed from updated image $u(x,t)$.

(g) Results for $\sigma_n = 150$, Dice 0.9934

Fig. 10. Synthetic $Spiral$ image segmentation by D-MTTV smoothing based thresholding method. (a-c) Different Gaussian noise corrupted images with their histogram at top left corner (d) Ground truth binary image (e-f) show $p(x)$ map, restoration, histogram, and thresholded results for each case.

Dice 0.9990

Dice 0.9375

Dice 0.9628

Dice 0.8789

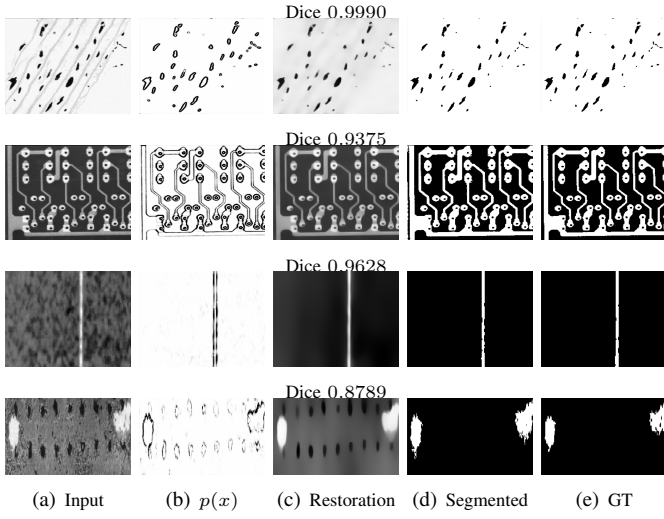(a) Input   (b) $p(x)$ (c) Restoration   (d) Segmented   (e) GT

Fig. 11. Non-destructive evaluation (NDT) image segmentation with our proposed D-MTTV smoothing based segmentation. (a) Input image (b) Computed $p(x)$ exponent from the input image (c) Restored image using our scheme (d) Simple thresholding on the resultant denoised image in (d) results in good segmentation (e) Ground truth (GT) for comparison.
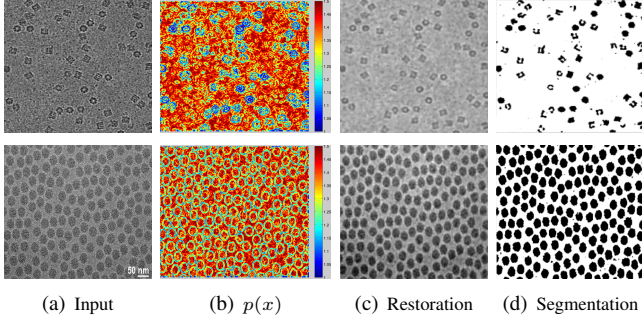

(a) Input   (b) $p(x)$ (c) Restoration   (d) Segmentation

Fig. 12. Cryo-electron microscopy images denoised and segmented based on our adaptive scheme. Top row: GroEL suspended in vitreous ice (Image Courtesy of Dr. Neil R. Voss, The Scripps Research Institute, USA). Bottom row: Tomato Bushy Stunt Virus (Image Courtesy of Dr. Ricardo Aramayo, Université Pierre et Marie Curie, France). (a) Input image (b) Computed $p(x)$ from the input image. (c) Restored image using our scheme. (d) Simple thresholding on the resultant denoised image in (b) results in reliable segmentation of compact viruses.

The applicability of image denoising using our D-MTTV scheme on biomedical cryo-electron micrograph of bacteria and viruses is illustrated in Figure 12. After smoothing out severe multiplicative noise degradation using our D-MTTV (20), a simple thresholding of grayscale values suffices to show a clear separation of foreground-background. Note that the amount of noise level is unknown and the segmentation shows separation of bacteria, and viruses against a non-homogeneous background. Thus, the proposed D-MTTV scheme can be used for effective denoising based segmentation processes.

2) Multiplicative noise removal: Next, we consider the multiplicative noise model for images and the corresponding MTTV scheme based restoration,

$$\min_u E(u) = \int_\Omega |\nabla u|^{p(x)}\, dx + \frac{\mu}{2} \int_\Omega \left(\log u + \frac{u_0}{u} \right)\, dx, \quad (20)$$

where the fidelity term is important and the fidelity parameter μ determines the amount of noise removed. We compare with a recent work of Liu et al [55] which uses a scalar $p(x) \equiv p$ based scheme (with μ values set appropriately). Figure 13

(a) Noisy u_0 (b) Noise, $\sigma_n^2 = 1/9$ (c) $p = 1$, $\mu = 40$, SNR $= 8.34$ dB

(d) $p = 1.2$, $\mu = 75$, SNR$= 8.52$ dB (e) $p = 1.5$, $\mu = 225$, SNR$= 8.50$ dB

(f) $p = 1.8$, $\mu = 605$, SNR$= 8.78$ dB (g) D-MTTV, $\mu = 120$, SNR$= 9.23$ dB

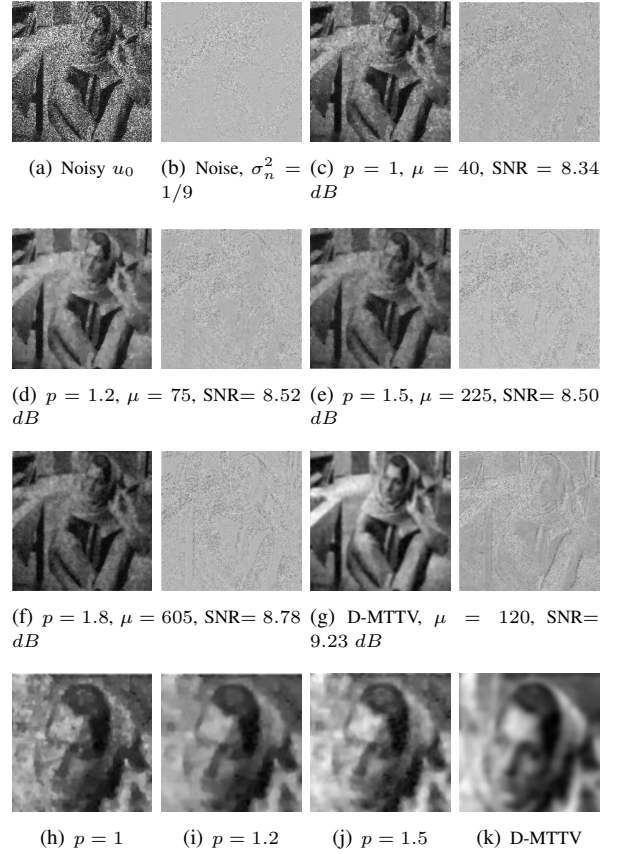(h) $p = 1$ (i) $p = 1.2$ (j) $p = 1.5$ (k) D-MTTV

Fig. 13. Multiplicative noise removal results for the $Barbara$ grayscale image based on highest possible SNR (dB). (a) Noisy image with multiplicative Gamma noise with mean 1 and deviation $1/3$. (b) Amount of noise level $(u_0 - u_O)$, where u_O is the original image. (c-f) Result of [55] with different scalar exponent p values. (g) Result based on our scheme, see Eqn. (20). The right side image in each sub-figure is residue $(u_0 - u)$ image indicating the amount of noise and details removed by each scheme. (h)-(k) Comparison with close-up showing noticeable improvements obtained with our D-MTTV scheme.

shows restoration results of $Barbara$ 256×256 image with heavy multiplicative Gamma noise[9] with mean 1 and variance $1/9$. Note that noise level depends on the intensity, that is the noise is larger in bright areas, see Figure 13 (b) which shows the amount of noise added. Figure 13 (c)-(f) show results of with increasing scalar p values in Eqn. (20) with our adaptive $p(x)$ based scheme. To compare the performance of the schemes we use signal to noise ratio (SNR)[10]. As can be seen using the corresponding residue $(u_0 - u)$ images for different p values, the proposed approach using adaptive exponent leads to improved SNR values compared to the method studied in [55].

3) The multichannel case: Extending the variable exponent model to the multichannel (color, multi spectral) images is an interesting option and is currently our focus of attention. We let $\mathbf{u} : \Omega \to \mathbb{R}^N$ be the multichannel image $\mathbf{u} =$

[9]Multiplicative noise model is given by $u_0(x) = n(x) \times u(x)$. Noise $n(\cdot)$ follows a Gamma law of mean 1 and variance σ_n^2, i.e., $n(x) \sim (b^a \Gamma(a))^{-1} x^{a-1} e^{-x/b}$, $ab = 1$, $ab^2 = \sigma_n^2$. We used the following MATLAB command to generate noise: n = gamrnd(1/σ²ₙ, σ²ₙ, 256, 256).

[10]This was the metric used in [55] to present their restoration results. Note that PSNR values were similar to SNR values and our scheme outperformed [55] in PSNR values as well.

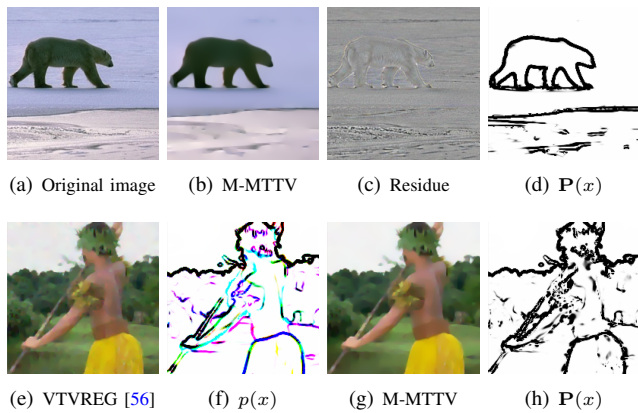| (a) Original image | (b) M-MTTV | (c) Residue | (d) $\mathbf{P}(x)$ |
| (e) VTVREG [56] | (f) $p(x)$ | (g) M-MTTV | (h) $\mathbf{P}(x)$ |

Fig. 14. Color image smoothing of noise free image 100007 (top row) and restoration of a noisy image 101087 (bottom row) from BSDS500 with different schemes (MSE values). (a) Original noise free image (b) Denoising with M-MTTV (c) Residue $\|\mathbf{u}_0 - \mathbf{u}\|_2$ scaled to $[0, 1]$ for visualization (e) VTVREG [56] (f) Channel-wise $p(x)$ exponent (h) Multichannel SST (21) based $\mathbf{p}(x)$ exponent. Better viewed online and zoomed in.

$(u^1, u^2, \ldots, u^N)$. Then the definition of the multichannel SST is given by,

$$\mathbf{K}_\sigma = G_\sigma \star \sum_{i=1}^{N} \nabla u^i (\nabla u^i)^T \qquad (21)$$

where as before we compute the variable exponent $\mathbf{p}(x)$ (see Eqn. (12)) with the largest and smallest eigenvalues of $\mathbf{K}_\sigma$ as Λ_+, Λ_-. Then the following regularization is used

$$\min_{\mathbf{u}} E(\mathbf{u}) = \sum_{i=1}^{N} \int_\Omega |\nabla u^i|^{\mathbf{p}(x)} \, dx + \frac{\mu}{2} \int_\Omega (u^i - u_0^i)^2 \, dx, \quad (22)$$

where we assume the noisy image $\mathbf{u}_0 = (u_0^1, u_0^1, \ldots, u_0^N)$ is obtained by adding adding Gaussian noise of standard deviation $\sigma_n = 30$ to each channel. Figure 14(a) shows smoothing a noise free RGB color image (from the Berkeley Segmentation Data Set (BSDS500), image id 101087) with our multichannel MTTV (M-MTTV) scheme. The result in Figure 14(b) shows strong piecewise smoothing occurring with edge preservation, and Figure 14(c) shows the amount of texture removed by the scheme. Figure 14(d) shows the variable exponent map. Figure 14(bottom row) shows an example comparison in color image restoration for a noisy (image id 101087) of size $201 \times 201 \times 3$. We compare with Bresson and Chan [56] vectorial total variation (VTVREG), and our MTTV applied with channel-wise $p(x)$ (C-MTTV) estimation, and multichannel SST (21) exponent $\mathbf{P}(x)$ based scheme (M-MTTV). Note that the regularization couples different color gradients using the multichannel SST (21) based variable exponent $\mathbf{P}(x)$ (Figure 14(h)) whereas the channel-wise uses the exponent function (8) in each channel separately. As can be seen, M-MTTV model provides better restoration result without staircasing artifacts when compared to VTVREG. Table V shows PSNR(dB) and MSSIM values obtained with our schemes against VTVREG method. We obtain better structural similarity scores indicating better edge preservation. Other approaches can also be incorporated and the implicit coupling of regularization in each channel can further improve denoising results [57]–[59].

TABLE V
PSNR/MSSIM COMPARISON OF OUR SCHEMES CHANNEL-WISE MTTV (C-MTTV), MULTICHANNEL MTTV (M-MTTV) WITH VTVREG [56] FOR RGB IMAGES. NOISY IMAGES WERE OBTAINED BY ADDING GAUSSIAN NOISE OF STRENGTH $\sigma_n = 30$ TO ORIGINAL IMAGES. BEST RESULTS ARE IN **BOLDFACE**.

Images	Noisy	VTVREG [56]	C-MTTV	M-MTTV
Baboon	18.59	19.09	19.10	**19.33**
	0.4339	0.5052	0.5372	**0.6383**
Barbara	18.59	17.04	20.39	**24.27**
	0.7416	0.7400	0.7884	**0.8360**
Boat	18.57	16.16	19.01	**23.69**
	0.5338	0.6584	0.6971	**0.7530**
House	18.57	14.79	17.92	**24.72**
	0.3493	0.7256	0.7718	**0.8141**
Lena	18.59	15.36	17.84	**22.41**
	0.6884	0.7744	0.8102	**0.8681**
Peppers	18.59	18.28	21.97	**28.83**
	0.6884	0.8574	0.8926	**0.9303**

IV. CONCLUSIONS

In this paper we developed a novel adaptive multiscale PDE regularization-based denoising method using smoothed structure tensor variable exponents that preservers coherent image structures. The proposed non-linear anisotropic operator incorporates the local structure tensor eigenvalues for coherent feature adaptation. Wellposedness of the proposed model is studied using the theory of variable exponent spaces and we obtain existence and long time behavior of solutions to the proposed model. The MMTV method incorporates elements of recently successful filtering approaches including nonparametric operators, adaptation to signal content and repeated filtering with aggregation. Extensive experiments on noisy synthetic, machine vision and biomedical imagery demonstrated the effectiveness of the proposed MMTV method and its variants compared to other denoising methods especially those related to variational-PDEs. The proposed adaptive multiscale exponent provides robust fine structure maps that enables smoothing without generating staircasing artifacts typically associated with point-wise estimators. We show analytically that the proposed method satisfies the maximum-minimum principle ensuring that artifact edge structures are not created in the denoising process. Extensions to the MMTV method readily handle multiplicative noise, RGB or multispectral images and lead to improved binary segmentation results. Extending the multiscale exponent to handle other types of noise and image decompositions defines our future work in this area.

APPENDIX

ANALYSIS OF THE VARIATIONAL PDE MODEL

A. Existence of minimizer

We recall the basic notions of variable exponent spaces, for further details we refer to the recent monograph [24]. Let $\mathcal{M}(\Omega)$ be the set of measurable functions in Ω. Though in imaging problems $\Omega \subset \mathbb{R}^2$ is a rectangle, we can study general open domain $\Omega \subset \mathbb{R}^N$ for any $N > 1$ as well. The Orlicz-Lebesgue spaces $L^{p(x)}(\Omega)$ are a generalization of the classical L^p spaces to the variable exponent case $p(x)$. Let

$$p_- = ess \inf_\Omega p(x), \quad \text{and} \quad p_+ = ess \sup_\Omega p(x).$$

The exponent sets are

$$\mathcal{P}(\Omega) = \{p \in \mathcal{M}(\Omega) : 1 \leq p_- \leq p_+ < \infty\}$$

$$\mathcal{P}^o(\Omega) = \{p \in \mathcal{M}(\Omega) : 1 < p_- \leq p_+ < \infty\}.$$

We next recall the variable exponent spaces.

Let $p \in \mathcal{M}(\Omega)$, $\rho(u) = \int_\Omega |u(x)|^{p(x)} \, dx$. Then

$$L^{p(x)}(\Omega) := \{u \in \mathcal{M}(\Omega) | \exists \lambda > 0 : \rho(u/\lambda) < \infty\}.$$

The *Luxemburg-Nakano norm* on $L^{p(x)}$ is $\|u\|_{p(\cdot)} := \inf\{\lambda > 0 : \rho(u/\lambda) \le 1\}$. The Orlicz-Sobolev space is defined as[11]

$$W^{m,p(x)}(\Omega) := \{u \in L^{p(x)}(\Omega) : D^\alpha u \in L^{p(x)}(\Omega), \forall |\alpha| \le m\}$$

with $\|u\|_{m,p(\cdot)} := \sum_{|\alpha| \le m} \|D^\alpha u\|_{p(\cdot)}$. We first recall a general existence result for a variable exponent $p(x)$ that is bounded away from 1 in the variable Sobolev space $W^{1,p(x)}(\Omega)$.

Lemma 1. *The regularization functional $E(u)$ given in (7) is lower semi-continuous in $W^{1,p(x)}(\Omega)$, i.e., for every sequence $u_j \to u$ in $W^{1,p(x)}(\Omega)$ we have,*

$$\int_\Omega |Du|^{p(x)} \, dx \le \liminf_{j \to \infty} \int_\Omega |Du_j|^{p(x)} \, dx.$$

Proof: This follows from the facts that the space $W^{1,p(x)}(\Omega)$ is a Banach space with the norm $\|u\|_{1,p(\cdot)} := \|u\|_{p(\cdot)} + \|Du\|_{p(\cdot)}$, and the semi-norm part of it, $|Du|_{1,p(x)} = \|Du\|_{p(\cdot)} \, dx$ is sequentially lower semi-continuous, and from the following inequality

$$\int_\Omega |Du|^{p(x)} \, dx \le \max\{\|Du\|_{p(\cdot)}^{p_+}, \|Du\|_{p(\cdot)}^{p_-}\}.$$

■

We recall the following fact from functional analysis (see [60, Sec. 25]).

Lemma 2. *Let B be a reflexive Banach space. If $F : B \to [0, \infty)$ is convex, lower semicontinuous, and coercive, then there is an element in B which minimizes F.*

Theorem 1. *Let $p \in \mathcal{P}^o(\Omega)$. Then*

$$F(u) = \int_\Omega |Du|^{p(x)} \, dx + \frac{\mu}{2} \int_\Omega |u - u_0|^2 \, dx$$

has a unique minimizer in $W^{1,p(x)}(\Omega)$.

Proof: From Lemma 1, the first term is lower semicontinuous in $W^{1,p(x)}(\Omega)$. The second term is strictly convex and continuous in u. Since $p_- > 1$ the function F is coercive as $|Du|^{p(x)}$ satisfies $|Du|^{p(x)} \ge |Du|^{1+\delta}$ and the space $W^{1,p(x)}$ is reflexive. Thus from Lemma 2 we have a minimizer in $W^{m,p(x)}(\Omega)$. Since $\xi \to |\xi|^{p(x)}$ is strictly convex if $p_- > 1$, the minimizer is unique. ■

Our proposed variable exponent function (12) satisfies $p(x) \in [1, 1.5]$, hence Theorem 1 is not applicable. Therefore, it requires a careful treatment as in [37] and it represents the border case of total variation [6] regularization. This also means we leave the realm of reflexive Sobolev spaces and need a new space with the properties of the traditional functions of bounded variation space $BV(\Omega)$. Before extending the results to the variable bounded variation space $BV^{p(x)}(\Omega)$ we require the following notations and preliminary results.

Recall [61] the definition of $BV(\Omega)$ and its norm defined in terms of test functions: $u \in BV(\Omega)$ if $u \in L^1(\Omega)$, and its *total variation*

$$TV(u) := \sup\left\{\int_\Omega u \, div\phi : \phi \in C_0^1(\Omega), \|\phi\|_{L^\infty} \le 1\right\}$$

is finite; the norm in $BV(\Omega)$ is given by

$$\|u\|_{BV} := \|u\|_1 + TV(u).$$

The distributional gradient Du is in fact a vector-valued Radon measure, and its "modulus" can also be interpreted as a Radon measure, which is called the *total variation measure*, and is denoted by $\|Du\|$. Note that

$$\|Du\|(\Omega) = TV(u).$$

For more details and results about the bounded variation space we refer to [61]–[63].

Let $p(x) \in \mathcal{P}(\Omega)$ be a lower semicontinuous function. Similarly to $W^{1,p(x)}(\Omega)$ spaces, we can define $BV^{p(x)}(\Omega)$. Let $Y(\Omega) := \{x \in \Omega : p(x) = 1\}$, that is the *critical* set where the exponent takes the value 1.

We now put $BV^{p(x)}(\Omega) := BV(\Omega) \cap W^{1,p(x)}(\Omega \setminus Y(\Omega))$, and define the modular

$$\Gamma(u) := TV(u) + \int_{\Omega \setminus Y(\Omega)} |Du|^{p(x)} \, dx.$$

Then the norm in $BV^{p(x)}$ is

$$\|u\|_{BV^{p(x)}(\Omega)} := \|u\|_{p(\cdot)} + \inf\{\lambda > 0 : \Gamma(u/\lambda) \le 1\}.$$

One can get a similar existence result like that of Theorem 1 in the $BV^{p(x)}$ space.

Theorem 2 (see [37]). *Assume that Ω has Lipschitz boundary, and $p(x) \in \mathcal{P}(\Omega)$ is lower semicontinuous. Then the functional*

$$\tilde{F}(u) = \int_{\Omega \setminus Y(\Omega)} |Du|^{p(x)} \, dx + \|Du\|(Y(\Omega)) + \frac{\mu}{2} \int_\Omega |u - u_0|^2 \, dx$$

has a minimizer in $BV^{p(x)}(\Omega)$.

B. Wellposedness of the PDE flow

To derive the PDE flow associated with the energy minimization in Eqn. (7) we need to use the weight $\frac{1}{p(x)}$, i.e., $\int_\Omega \frac{|\nabla u|^{p(x)}}{p(x)} \, dx$, the derivation is straightforward using the Euler-Lagrange formulation [4] and is omitted here for brevity;

$$\mu(u(x,t) - u_0(x)) = \triangle_{p(x)}$$
$$:= div\left(|\nabla u(x,t)|^{p(x)-2} \nabla u(x,t)\right), \quad x \in \Omega. \quad (23)$$

This is the $p(x)$-Laplacian which generalizes the traditional p-Laplacian and has been found useful in various scenarios, see the recent monograph [24]. We next turn our attention to the corresponding dynamic time-dependent version of the PDE given in Eqn. (23),

$$\frac{\partial u(x,t)}{\partial t} = div\left(|\nabla u(x,t)|^{p(x)-2} \nabla u(x,t)\right)$$
$$- \mu(u(x,t) - u_0(x)) \quad \text{in } \Omega \times (0,\infty), \quad (24)$$
$$u(x,t) = 0 \quad \text{on } \partial\Omega \times (0,\infty), \quad (25)$$
$$u(x,0) = u_0(x) \quad \text{in } \Omega. \quad (26)$$

---

[11]We revert to the notation of Du for the generalized gradient function. Note that if $u \in W^{1,2}(\Omega)$ then $Du = \nabla u$.

This is an inhomogeneuous $p(x)$-Laplacian PDE, and the existence result may be proven via a subdifferential approach, cf. [64]. The above time dependent PDE flow paves the way to relate to the concept of scale-space $\{u(x,t)\}_{t=0}^{\infty}$ and is related to the classical Perona-Malik [1] anisotropic diffusion paradigm [65].

To prove the existence result, we define the following subspaces of L^2:

$$L^{2,p(x)}(\Omega) = \left\{ u \in L^2(\Omega) : \frac{\partial u}{\partial x_i} \in L^{p(x)}(\Omega), \ i = 1, \ldots, N \right\}$$

with the norm $\|u\|_{L^{2,p(x)}(\Omega)} = \left(\|u\|_2^2 + \|\nabla u\|_{p(x)}^2 \right)^{1/2}$, and

$$L_0^{2,p(x)}(\Omega) = L^{2,p(x)}(\Omega) \cap W_0^{1,p_-}(\Omega).$$

Now the main existence result can be stated as follows.

Theorem 3. *Let $p \in \mathcal{P}^o(\Omega)$. Then the following hold:*

(i) *For $u_0 \in L^2(\Omega)$, the problem (24-26) admits a unique solution*

$$u \in W_{loc}^{1,2}(0, \infty, L^2(\Omega)) \cap C((0, \infty), L_0^{2,p(x)}(\Omega)).$$

(ii) *For $u_0 \in L_0^{2,p(x)}(\Omega)$ and $T > 0$,*

$$u \in W^{1,2}(\Omega)(0, T; L^2(\Omega)) \cap C([0, T]; L_0^{2,p(x)}(\Omega)).$$

(iii) *The solution u continuously depends on the original image u_0:*

$$\|u_1(t) - u_2(t)\|_2 \le (t+1) \|u_{0,1} - u_{0,2}\|_2, \ \forall t \ge 0,$$

where u_1 and u_2 are the two solutions of (24-26) with $u_0 = u_{0,1}$ and $u_{0,2}$, resp.

To prove the above theorem, we note that the functional $\varphi_{p(x)} : L^2(\Omega) \to [0, \infty]$ defined by

$$\varphi_{p(x)}(\psi) = \begin{cases} \int_\Omega \frac{\mu}{2} \psi^2(x) + \frac{1}{p(x)} |\nabla \psi(x)|^{p(x)} \ dx & \text{if } \psi \in L_0^{2,p(x)}(\Omega), \\ \infty & \text{otherwise,} \end{cases} \tag{27}$$

is proper, lower semicontinuous and convex functional on $L^2(\Omega)$ (cf. [64]). Then, the subdifferential[12] satisfies

$$\partial \varphi_{p(x)}(\psi) = \mu \psi - \triangle_{p(x)} \psi.$$

Therefore, the parabolic problem (24-26) is equivalent to an abstract Cauchy problem,

$$\frac{du}{dt}(t) + \partial \varphi_{p(x)}(u(t)) = \mu u_0 \quad \text{in } L^2(\Omega), \quad u(0) = u_0.$$

Then the proof follows from the results of [66] and [64].

Remark 5. *The critical case $p \in \mathcal{P}(\Omega) \setminus \mathcal{P}^o(\Omega)$ seems to be much more involved, and existence of any kind of solution to (24-26) is an open problem in this case.*

---

[12] The subdifferential of the functional $\varphi_{p(x)}$ is defined as usual: $\partial \varphi_{p(x)}(u) = \left\{ \rho \in L^2 : \varphi_{p(x)}(v) - \varphi_{p(x)}(u) \ge \int_\Omega (\rho, v - u), \ \forall v \in L_0^{2,p(x)} \right\}$.

Assume in addition that p satisfies the *log-Hölder* condition. Then one can obtain decay of solutions if we have a strong growth in our exponent:

Theorem 4 (see [64]). *Let $p_- > 2$, $u_0 \in L_0^{2,p(x)}(\Omega) \setminus \{0\}$, and $u = u(x,t)$ be the solution of (24-26) with $\mu = 0$. Then a constant $C > 0$ such that,*

$$\frac{(t+1)^{-1/(p_- - 2)}}{C} \le \|u(t)\|_2 \le C (t+1)^{-1/(p_+ - 2)}, \quad \forall t \ge 0.$$

One can also obtain extinction of solution if we restrict our exponent to less than a quadratic:

Theorem 5 (see [64]). *Let $\frac{2N}{N+2} \le p_- \le p_+ < 2$, $u_0 \in L_0^{2,p(x)}(\Omega) \setminus \{0\}$, and $u = u(x,t)$ be the solution of (24-26) with $\mu = 0$. Then there exists a finite time $t_* > 0$ and a constant $C > 0$ such that*

$$\frac{(t_* - t)_+^{1/(2-p_-)}}{C} \le \|u(t)\|_2 \le C (t_* - t)_+^{1/(2-p_+)}, \quad \forall t \ge 0.$$

Thus, the solution $u = u(x,t)$ vanishes at time t_* (extinction time of u).

REFERENCES

[1] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[2] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 221–232, 1992.

[3] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: Partial differential equation and calculus of variations*. New York, USA: Springer-Verlag, 2006.

[4] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*. New York, USA: Springer-Verlag, 2009.

[5] A. Tikhonov and V. Aresenin, *Solutions of Ill-posed Problems*. New York, NY, USA: John Wiley, 1997.

[6] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[7] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, 2006.

[8] A. Kuijper, "P-Laplacian driven image processing," in *IEEE ICIP*, San Antonio, TX, 2007, pp. 257–260.

[9] A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete p-laplacian driven image and manifold processing," *Comptes Rendus Mécanique*, vol. 336, no. 5, pp. 428–433, 2008.

[10] B. Zhou and C.-L. Mu, "Level set evolution for boundary extraction based on a p-Laplace equation," *Applied Mathematical Modelling*, vol. 34, no. 12, pp. 3910–3916, 2010.

[11] G. Chung and L. Vese, "Image segmentation using a multilayer level-set approach," *Computing and Visualization in Science*, vol. 12, no. 6, pp. 267–285, 2009.

[12] P. Blomgren, T. F. Chan, P. Mulet, and C. K. Wang, "Total variation image restoration: numerical methods and extensions," in *IEEE International Conference on Image Processing (ICIP)*, Santa Barbara, CA, USA, October 1997, pp. 384–387.

[13] E. M. Bollt, R. Chartrand, S. Esodoglu, P. Schultz, and K. R. Vixie, "Graduated adaptive image denoising: local compromise between total variation and isotropic diffusion," *Advances in Computational Mathematics*, vol. 31, no. 1–3, pp. 61–85, 2009, special isse: Mathematical Methods for Image Processing.

[14] Y. Chen, S. Levine, and M. Rao, "Variable exponent, linear growth functionals in image restoration," *SIAM Journal on Applied Mathematics*, vol. 66, no. 4, pp. 1383–1406, 2006.

[15] A. Chambolle and P. L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[16] B. Zhou, C.-L. Mu, J. Feng, and W. Wei, "Continuous level anisotropic diffusion for noise removal," *Applied Mathematical Modelling*, vol. 36, no. 8, pp. 3779–3786, 2012.

[17] F. Li, Z. Li, and L. Pi, "Variable exponent functionals in image restoration," *Applied Mathematics and Computation*, vol. 216, no. 3, pp. 870–882, 2010.

[18] Z. Guo, J. Sun, D. Zhang, and B. Wu, "Adaptive Perona–Malik model based on the variable exponent for image denoising," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 958–967, 2012.

[19] D. Strong, "Adaptive total variation minimizing image restoration," Ph.D. dissertation, UCLA Mathematics Department, USA, August 1997.

[20] V. B. S. Prasath and A. Singh, "Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising," *Journal of Applied Mathematics*, vol. 2010, p. 14pp, 2010, article ID 763847.

[21] V. B. S. Prasath, "A well-posed multiscale regularization scheme for digital image denoising," *International Journal of Applied Mathematics and Computer Science*, vol. 21, no. 4, pp. 769–777, 2011.

[22] V. B. S. Prasath and A. Singh, "An adaptive anisotropic diffusion scheme for image restoration and selective smoothing," *International Journal of Image and Graphics*, vol. 12, no. 1, p. 18pp, 2012.

[23] V. B. S. Prasath and D. Vorotnikov, "Weighted and well-balanced anisotropic diffusion scheme for image denoising and restoration," *Nonlinear Analysis: Real World Applications*, vol. 17, pp. 33–46, 2013.

[24] L. Diening, P. Harjulehto, P. Hästö, and M. Ružička, *Lebesgue and Sobolev spaces with variable exponents*. Heidelberg: Springer-Verlag, 2011.

[25] T. Brox, R. Van Den Boomgaard, F. Lauze, J. Van De Weijer, J. Weickert, P. Mrázek, and P. Kornprobst, "Adaptive structure tensors and their applications," in *Visualization and processing of tensor fields*. Springer Berlin Heidelberg, 2006, pp. 17–47.

[26] Q. Chen, P. Montesinos, Q. S. Sun, P. A. Heng, and D. S. Xia, "Adaptive total variation denoising based on difference curvature," *Image and Vision Computing*, vol. 28, no. 3, pp. 298–306, 2010.

[27] M. Nikolova, "Weakly constrained minimization: application to the estimation of images and signals involving constant regions," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 2, pp. 155–175, 2004.

[28] M. Nitzberg and T. Shiota, "Nonlinear image filtering with edge and corner enhancement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 8, pp. 826–833, 1992.

[29] P. Milanfar, "A tour of modern image filtering," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 106–128, 2013.

[30] J. Weickert, *Anisotropic diffusion in image Processing*. Stuttgart, Germany: B.G. Teubner-Verlag, 1998.

[31] W. Forstner and E. Gulch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *ISPRS Intercommission Workshop*, 1987.

[32] J. A. Noble, "Finding corners," *Image and Vision Computing*, vol. 6, no. 2, pp. 121–128, 1988, 3rd Alvey Vision Meeting.

[33] ——, "Descriptions of image surfaces," Ph.D. dissertation, Department of Engineering Science, Oxford University, UK, August 1989.

[34] C. G. Harris and M. J. Stephens, "A combined corner and edge detector," in *4th Alvey Vision Conference*, Manchester, UK, 1988, pp. 147–151.

[35] T. Lindeberg and L. Bretzner, "Real-time scale selection in hybrid multi-scale representations," in *Scale Space Methods in Computer Vision*, 2003, pp. 148–163, springer LNCS Volume 2695/2003 Eds.: L. D. Griffin, M. Lillholm.

[36] T. Lindeberg and D. Fagerström, "Scale-space with casual time direction," in *European Conference on Computer Vision*, 1996, pp. 229–240, springer LNCS Volume 1064/1996 Eds.: B. Buxton, R. Cipolla.

[37] P. Harjulehto, P. Hästö, V. Latvala, and O. Toivanen, "Critical variable exponent functionals in image restoration," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 56–60, 2013.

[38] T. Lindeberg, *Scale-Space Theory in Computer Vision*. Springer, 1994.

[39] J. Weickert, B. M. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 398–410, 1998.

[40] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *IEEE International Conference on Computer Vision*, 1998, pp. 59–66.

[41] S.-H. Lee and J. K. Seo, "Noise removal with Gauss curvature-driven diffusion," *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 904–909, 2005.

[42] G. Gilboa, N. Sochen, and Y. Zeevi, "Image enhancement and denoising by complex diffusion processes," *IEEE Transactions on Pattern Analysis and Machines Intelligence*, vol. 26, no. 8, pp. 1020–1036, 2004.

[43] K. Chen, "Adaptive smoothing via contextual and local discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1552–1567, 2005.

[44] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising methods, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2006.

[45] ——, "The staircasing effect in neighborhood filters and its solution," *IEEE Transactions on image Processing*, vol. 15, no. 6, pp. 1499–1505, 2006.

[46] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421–432, 1998.

[47] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Variational denoising of partly textured images by spatially varying constraints," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2281–2289, 2006.

[48] Y.-L. You and M. Kaveh, "Fourth–order partial differential equation for noise removal," *IEEE Transactions on Image Processing*, vol. 9, pp. 1723–1730, 2000.

[49] F. Li, C. Shen, J. Fan, and C. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *Journal of Visual Communication and Image Representation*, vol. 18, no. 4, pp. 322–330, 2007.

[50] K. Ito and K. Kunisch, "BV-type regularization methods for convoluted objects with edge, flat and grey scales," *Inverse Problems*, vol. 16, no. 4, pp. 909–928, 2000.

[51] C. Schnörr, "Unique reconstruction of piecewise-smooth images by minimizing strictly convex nonquadratic functionals," *Journal of Mathematical Imaging and Vision*, vol. 4, no. 2, pp. 189–198, 1994.

[52] R. Pelapur, F. Bunyak, O. Glinskii, V. Glinskii, V. Huxley, and K. Palaniappan, "Multi-focus image fusion using epifluorescence microscopy for robust vascular segmentation," in *36th IEEE/EMBS International Conference*, Chicago, USA, August 2014.

[53] V. B. S. Prasath, R. Pelapur, O. V. Glinskii, V. V. Glinsky, V. H. Huxley, and K. Palaniappan, "Multi-scale tensor anisotropic filtering of fluorescence microscopy for denoising microvasculature," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2015.

[54] G. Lathen, J. Jonasson, and M. Borga, "Blood vessel segmentation using multi-scale quadrature filtering," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 762–767, 2010.

[55] Q. Liu, L. Xia, and T. Gao, "A nondivergence p-Laplace equation in a removing multiplicative noise model," *Nonlinear Analysis: Real World Applications*, vol. 14, no. 5, pp. 2046–2058, 2013.

[56] X. Bresson and T. F. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems and Imaging*, vol. 2, no. 4, pp. 455–484, 2008.

[57] D. Tschumperle and R. Deriche, "Vector-valued image regularization with PDE's: A common framework for different applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 506–517, 2005.

[58] V. B. S. Prasath and A. Singh, "Multichannel image restoration using combined channel information and robust M-estimator approach," *International Journal Tomography and Statistics*, vol. 12, no. F10, pp. 9–22, 2010.

[59] ——, "Multispectral image denoising by well-posed anisotropic diffusion scheme with channel coupling," *International Journal of Remote Sensing*, vol. 31, no. 8, pp. 2091–2099, 2010.

[60] E. Zeidler, *Nonlinear Functional Analysis and its Applications II*. Berlin, Germany: Springer-Verlag, 1990.

[61] L. C. Evans and R. Gariepy, *Measure theory and fine properties of functions*. CRC press, 1991.

[62] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces: Applications to PDEs and optimization*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics (SIAM), 2006.

[63] F. Giusti, *Minimal Surfaces and Functions of Bounded Variation*. Basel, Switzerland: Birkhauser, 1984.

[64] G. Akagi and K. Matsuura, "Nonlinear diffusion equations driven by the p (·)-laplacian," *Nonlinear Differential Equations and Applications NoDEA*, vol. 20, no. 1, pp. 37–64, 2013.

[65] L. Alvarez, F. Guichard, P. L. Lions, and J.-M. Morel, "Axioms and fundamental equations of image processing," *Archive for Rational Mechanics and Analysis*, vol. 123, no. 3, pp. 199–257, 1993.

[66] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. Amsterdam, The Netherlands: North-Holland Publishing Company, 1973.