DOI .34660/INF.2026.79.37.103

ON THE WELL-POSEDNESS OF DISCRETE VARIATIONAL PROBLEMS WITH EDGE-AWARE REGULARIZATION

Shchetinin Evgeny Yuryevich
Doctor of Physical and Mathematical Sciences, Associate Professor, Professor
Shevchuk Andrey Andreyevich
Postgraduate
Sevastopol State University,
Sevastopol, Russia

Abstract. *Edge-aware regularization is a cornerstone of variational methods in image processing and computer vision, typically implemented by weighting discrete gradients with functions dependent on local intensity differences. This paper investigates the phenomenon of scale inconsistency in standard weighting schemes of the form* $w \propto \exp(-\beta|\Delta I|^2)$. *We demonstrate that under grid refinement (h→0), such weights degenerate to a constant at a rate of 0(h²), effectively reducing the regularizer to an isotropic Laplacian and nullifying the edge-preserving properties in the continuum limit. To address this, we propose a scale-consistent formulation utilizing discrete directional derivatives,* $\left(I_{i+e_k} - I_i\right)/h$, *within the weight argument. We provide a rigorous analysis proving the -consistency of the modified discrete functional with a weighted anisotropic Dirichlet energy and establish 0(h)-convergence of the associated variational problems. Numerical validation on synthetic data and cardiac MRI segmentation (ACDC dataset) confirms that the proposed method ensures hyperparameter stability across varying physical resolutions, eliminating the need for resolution-dependent tuning.*

Keywords: *weighted Dirichlet energy; edge-aware regularization; scale consistency; discrete-to-continuum limit; inverse problems; -convergence.*

1. Introduction

Variational methods and regularization techniques remain fundamental to applied mathematics, particularly in solving ill-posed inverse problems [1, 2, 3] and in image reconstruction [4, 5, 6]. The classical approach involves recovering an unknown field from noisy observations by minimizing an energy functional:

$$J(u) = \mathcal{D}(u, f) + \lambda\mathcal{R}(u), \tag{1}$$

where $\mathcal{D}$ is a data fidelity term and $\mathcal{R}$ is a regularizer enforcing prior knowledge, such as smoothness.

While the standard Dirichlet energy $\int |\nabla u|^2\, dx$ promotes global smoothness, it invariably blurs structural boundaries. To mitigate this, edge-aware regularization techniques were developed, suppressing smoothing across high-gradient regions. The seminal work on anisotropic diffusion by Perona and Malik [7] introduced diffusivity coefficients dependent on the image gradient magnitude. These concepts evolved into weighted Dirichlet energies and Total Variation (TV) methods [8, 9], which are now ubiquitous in both classical solvers and modern deep learning architectures [10, 11].

Despite their widespread adoption, the discretization of these functionals on grids with varying spatial resolution (pixel/voxel size) poses a subtle but critical challenge. In many implementations, the grid step h is implicitly treated as unity. Consequently, regularization weights are computed based on raw intensity differences between adjacent pixels. As we show in this work, this practice leads to scale inconsistency: models tuned on data with one physical resolution fail to generalize to data with a different resolution, a scenario common in multi-center medical imaging studies [12].

2. Scientific Problem

The central issue addressed in this study is the mathematical degeneracy of standard discrete edge-aware weights under grid refinement.

Consider a domain $\Omega \subset R^d$ discretized by a uniform grid Ω_h with mesh size *h*. Let *I* be a fixed continuous image (or guidance signal). In standard discrete implementations [13, 14, 15], the weight $w_{i,k}$ assigned to the edge between nodes *i* and $i{+}e_k$ is typically defined as a function of the absolute intensity difference:

$$w_{i,k}^{\mathrm{std}} = g\left(\left|I_{i+e_k} - I_i\right|^2\right), \tag{2}$$

where $g(\cdot)$ is a decaying function, often the Gaussian kernel $g(s) = \exp(-\beta s)$. The intuition is that large differences imply an edge, reducing the weight and thus the smoothing.

However, from the perspective of numerical analysis, this formulation is ill-posed with respect to the limit *h*→0. For a differentiable function, the finite difference scales as *0 (h)*. Specifically, via Taylor expansion:

$$|I(x + he_k) - I(x)| \approx h|\partial_k I(x)|. \tag{3}$$

As *h*→0, the argument of the weight function tends to zero regardless of the local gradient magnitude, causing the weight $w_{i,k}^{\mathrm{std}}$ to converge to its maximum value *q(0)*. This implies that the “edge-awareness” of the regularizer is an artifact of the discretization step size rather than an intrinsic property of the image content. This scale dependence necessitates the recalibration of the sensitivity parameter β whenever the physical resolution changes, complicating model deployment and reproducibility.

3. Contributions

This paper provides a rigorous mathematical treatment of scale inconsistency and proposes a geometrically correct remedy. Our main contributions are:

Negative Result: We formally prove that standard edge-aware weights degenerate to a constant with an asymptotic error rate of $0(h^2)$ as $h\to 0$, transforming the anisotropic regularizer into an isotropic one in the continuum limit.

Proposed Remedy: We introduce a scale-consistent weight formulation based on discrete directional derivatives (normalized differences). This modification restores the correct continuum behavior.

Theoretical Analysis: We establish the consistency of the proposed discrete functional with the continuous weighted Dirichlet energy ($0(h)$ accuracy) and assert the -convergence of the respective variational problems, ensuring convergence of minimizers.

Asymptotic Characterization: We analyze the joint limits of the grid step h and the sensitivity parameter σ, classifying the regimes of isotropic, anisotropic, and TV-like behavior.

Numerical Validation: The theoretical findings are corroborated by synthetic resolution-sweep tests and a real-world cardiac MRI segmentation experiment, demonstrating superior robustness to resolution changes compared to the baseline.

4. Materials and Methods

4.1. Preliminaries and Notation

Let $\Omega \subset R^d \ (d = 2,3)$ be a bounded Lipschitz domain. We introduce a sequence of uniform grids Ω_h with spacing $h>0$. Grid nodes are denoted by $x_i = ih$, $i \in Z^d$.

For a grid function $u: \Omega_h \to R$, the forward difference operator in the direction of the basis vector e_k is defined as:

$$D_k^+ u_i = \frac{u_{i+e_k} - u_i}{h}. \tag{4}$$

The discrete L^2-norm is defined as $|u|_{2,h}^2 = h^d \sum_i |u_i|^2$.

We consider the discrete weighted Dirichlet energy:

$$E_h[u] = h^d \sum_i \sum_{k=1}^{d} w_{i,k} |D_k^+ u_i|^2, \tag{5}$$

where $w_{i,k}$ are non-negative weights derived from a guidance image.

4.2. Analysis of Standard Weight Degeneration

The standard weight formulation is given by:

$$w_{i,k}^{\mathrm{std}} = \exp\left(-\beta\left|I_{i+e_k} - I_i\right|^2\right) + w_{min}, \quad w_{min} > 0. \tag{6}$$

Proposition 1 (Asymptotic Degeneration). Assume $I \in C^3(\overline{\Omega})$. Then, for any edge $(i,\ i+e_\iota)$, the following expansion holds as $h \to 0$:

$$w_{i,k}^{\mathrm{std}} = (1 + w_{min}) - \beta h^2 |\partial_k I(\bar{x})|^2 + O(h^4), \tag{7}$$

where $\bar{x}$ is the midpoint of the edge. Consequently, $w_{i,k}^{\mathrm{std}} \to 1 + w_{min}$ uniformly.

Proof.

We expand I around $\overline{x_{\iota,k}} = x_i + \frac{h}{2} e_k$:

$$I_{i+e_k} - I_i = h\, \partial_k I\left(\overline{x_{\iota,k}}\right) + \frac{h^3}{24} \partial_k^3 I(\xi). \tag{8}$$

Squaring this expression yields:

$$\left|I_{i+e_k} - I_i\right|^2 = h^2 \left|\partial_k I\left(\overline{x_{\iota,k}}\right)\right|^2 + O(h^4). \tag{9}$$

Let $t = \beta\left|I_{i+e_k} - I_i\right|^2$. Since $t = O(h^2)$, we use the expansion $e^{-t} = 1 - t + O(t^2)$:

$$w_{i,k}^{\mathrm{std}} = \left(1 - \beta h^2 |\partial_k I|^2 + O(h^4)\right) + w_{min}. \tag{10}$$

This completes the proof.

This result implies that the contrast of the weights — the difference between weights at edges and in flat regions — vanishes quadratically: $\Delta w \sim O(h^2)$.

4.3. The Proposed Remedy: Scale-Consistent Weights

To ensure a non-trivial continuum limit, the argument of the weight function must approximate the gradient magnitude $|\nabla I|^2$, not the difference $|\Delta I|^2$. Furthermore, to handle noise and discontinuities rigorously in the continuum limit, we adopt the standard practice of using a mollified guidance image $I_\rho = G_\rho * I$, where G_ρ is a Gaussian kernel with fixed width ρ>0.

We define the **scale-consistent weights** as:

$$w_{i,k}^{*} = \exp\left(-\beta\left|D_k^+ I_{\rho,i}\right|^2\right) + w_{min} = \exp\left(-\beta\left|\frac{I_{\rho,i+e_k} - I_{\rho,i}}{h}\right|^2\right) + w_{min}. \tag{11}$$

Theorem 2 (Consistency). Let $I_\rho \in C^2(\overline{\Omega})$. The discrete functional $E_h^*[u]$ with weights $w_{i\,k}^*$ is consistent with the continuous anisotropic energy

$$E[u] = \int_\Omega \sum_{k=1}^{d} w_k^*(x)|\partial_k u(x)|^2 dx\,, \quad \text{with } w_k^*(x) = \exp\left(-\beta\left|\partial_k I_\rho(x)\right|^2\right) + w_{min}, \tag{12}$$

satisfying $|E_h^*[u_h] - E[u]| = O(h)$ for any $u \in C^2(\overline{\Omega})$, where is the grid restriction of *u*.

Proof Sketch.

Since is smooth, the discrete derivative $D_k^+ I_{\rho,i} = \partial_k I_\rho(\bar{x}) + O(h)$converges to the continuous derivative:. Since the exponential function is Lipschitz continuous on bounded intervals, $\left|w_{i,k}^* - w_k^*(\bar{x})\right| = O(h)$. The convergence of the Riemann sum to the integral introduces another *0(h)* error, resulting in a total error of *0(h)*.

Γ-Convergence.

Beyond pointwise consistency, variational correctness requires Γ-convergence [16, 17, 18, 19]. Under coercivity assumptions (e.g., Dirichlet boundary conditions) [20], it can be shown that the sequence of functionals E_h^* Γ-converges to *E* in the $L^2(\Omega)$ topology. This ensures that minimizers u^* of the discrete problems u_h^* converge to the minimizer of the continuous problem.

4.4. Two-Parameter Asymptotic Analysis

In practice, the sensitivity parameter is often denoted as σ, where $\beta = 1/\sigma^2$. We analyze the joint behavior of *h* and σ.

Regime 1: Fixed Sensitivity (σ = const independent of *h*). This corresponds to the correct anisotropic limit derived in Theorem 2. Regime 2: Vanishing Sensitivity (σ→0). As σ→0, the weights approximate a characteristic function indicating edges. This approaches a Total Variation-like regime but may introduce numerical stiffness. Regime 3: Standard Scaling (σ ∝ 1). If one uses standard weights without normalization, effectively $\beta_{eff} \propto h^2$, which corresponds to σ→∞. In this regime, weights converge to unity, and the model becomes isotropic. Recommendation: The parameter must have physical units (e.g., intensity/length) and be fixed independently of the grid step *h*.

5. Numerical Experiments

5.1. Synthetic "Resolution Sweep" Test

Setup:

We generated a 1D synthetic edge profile *I*(*x*) = tanh(*x*/δ) with width δ = 0.1 on domain [-1,1].

We generated grids with spacings $h \in \{2^{-4}, 2^{-5}, \ldots, 2^{-9}\}$.

We computed the weight contrast $\Delta w = \overline{w_{\text{flat}}} - \overline{w_{\text{edge}}}$, where $\overline{w_{\text{edge}}}$ and $\overline{w_{\text{flat}}}$ are average weights in the flat and transition regions, respectively.

5.2. Cardiac MRI Segmentation (Physical Scale Transfer)

Dataset:

We utilized the ACDC (Automated Cardiac Diagnosis Challenge) dataset [21, 22], comprising 100 patient MRI scans. A key feature of this dataset is the heterogeneity in physical resolution: pixel spacing varies from 1.37 mm to 1.68 mm.

Methodology:

A U-Net architecture [23] was trained to segment the left ventricle. The loss function included a regularization term on the softmax probability maps. We compared two strategies:

Baseline + StdWeight: Standard weights based on raw differences. fixed across all images.

Baseline + ScaleCons: Proposed scale-consistent weights (normalized by pixel size from metadata). fixed in physical units (mm^2).

Training used the Adam optimizer [24] with fixed hyperparameters. Evaluation metrics included the Dice Similarity Coefficient (DSC) and Expected Calibration Error (ECE) [25].

6. Results

6.1. Synthetic Test Results

Table 1 illustrates the dependence of weight contrast on grid resolution.

The standard weights exhibit a rapid collapse in contrast (approximately factor of 4 per halving of h, consistent with $0(h^2)$). At fine resolutions, the standard regularizer fails to distinguish edges from flat regions. In contrast, the proposed method maintains a stable contrast ≈ 0.416, demonstrating scale invariance.

6.2. MRI Segmentation Results

Performance on the ACDC test set (5-fold cross-validation):

Baseline (No Reg): Dice = 0.864 ± 0.011.

+ StdWeight: Dice = 0.869 ± 0.012. The improvement is statistically insignificant ($p = 0.09$) due to inconsistent regularization strength across patients with different resolutions.

+ ScaleCons: Dice = 0.886 ± 0.008. The proposed method yields a statistically significant improvement ($p < 0.001$) and reduces the Expected Calibration Error from 0.047 to 0.038.

7. Discussion

The results highlights a fundamental flaw in naive discretizations of variational models: the conflation of discretization parameters (grid step) with model parameters (edge sensitivity).

Interpretation of Degeneration:

The quadratic decay $\Delta w \sim O(h^2)$ explains a common practical frustration: hyperparameters tuned on low-resolution prototype data often perform poorly on high-resolution production data, leading to “oversmoothing.” The model effectively degenerates into isotropic diffusion as resolution increases.

The Role of Mollification:

While our theoretical analysis assumes $I_\rho \in C^2$, this is not merely a technical convenience. In applied settings, computing gradients on raw noisy data is unstable. Pre-smoothing (guidance) is standard in edge-preserving filters (e.g., Guided Filter [13]). Our analysis shows that this step is also necessary for the existence of a well-defined continuum limit.

8. Conclusion

This work presented a comprehensive analysis of scale inconsistency in discrete edge-aware regularization.

We proved that standard weights based on raw intensity differences degenerate as $O(h^2)$ under grid refinement.

We proposed a theoretically sound remedy using discrete directional derivatives, ensuring $O(h)$ -consistency with the anisotropic continuum model.

Experimental validation confirmed that the proposed scale-consistent weights provide superior robustness and accuracy in multi-resolution medical imaging tasks.

We recommend that all implementations of weighted Dirichlet energies or similar variational regularizers explicitly normalize intensity differences by the grid step to ensure mathematical well-posedness and practical robustness.

References

1. Engl HW, Hanke M, Neubauer A. Regularization of Inverse Problems. Dordrecht: Kluwer Academic Publishers; 1996. doi: 10.1007/978-94-009-1740-8

2. Vogel CR. Computational Methods for Inverse Problems. Philadelphia: SIAM; 2002. doi: 10.1137/1.9780898717570

3. Tikhonov AN, Arsenin VY. Solutions of Ill-Posed Problems. Washington: Winston; 1977.

4. Chambolle A. An algorithm for total variation minimization and applications. J Math Imaging Vis. 2004;20(1-2):89-97. doi: 10.1023/B: JMIV.0000011325.36760.1e

5. Natterer F, Wübbeling F. Mathematical Methods in Image Reconstruction. Philadelphia: SIAM; 2001. doi: 10.1137/1.9780898718324

6. Scherzer O, Grasmair M, Grossauer H, Haltmeier M, Lenzen F. Variational Methods in Imaging. New York: Springer; 2009. doi: 10.1007/978-0-387-69277-7

7. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell. 1990;12(7):629-39. doi: 10.1109/34.56205

8. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. Physica D. 1992;60(1-4):259-268. doi: 10.1016/0167-2789(92)90242-F

9. Ambrosio L, Fusco N, Pallara D. Functions of Bounded Variation and Free Discontinuity Problems. Oxford: Oxford University Press; 2000.

10. Litjens G, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88. doi: 10.1016/j.media.2017.07.005

11. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas; 2016. p. 770-8. doi: 10.1109/CVPR.2016.90

12. Modersitzki J. Numerical Methods for Image Registration. Oxford: Oxford University Press; 2004.

13. He K, Sun J, Tang X. Guided image filtering. IEEE Trans Pattern Anal Mach Intell. 2013;35(6):1397-409. doi: 10.1109/TPAMI.2012.213

14. Kornprobst P, Deriche R, Aubert G. Image Sequence Analysis via Partial Differential Equations. Journal of Mathematical Imaging and Vision. 1999;11:5-26. DOI: 10.1023/A:1008318126505

15. Singer A, Wu HT. Spectral convergence of the connection Laplacian from random samples. Inf Inference. 2017;6(1):58-123. doi: 10.1093/imaiai/iaw016

16. Braides A. -Convergence for Beginners. Oxford: Oxford University Press; 2002.

17. Dal Maso G. An Introduction to -Convergence. Boston: Birkhäuser; 1993. doi: 10.1007/978-1-4612-0327-8

18. Bauschke HH, Combettes PL. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Cham: Springer; 2017. doi: 10.1007/978-3-319-48311-5

19. Braides A, Defranceschi A. Homogenization of Multiple Integrals. Oxford: Oxford University Press; 1998.

20. Evans LC. Partial Differential Equations. Providence: American Mathematical Society; 2010.

21. Bernard O, et al. Deep learning techniques for automatic MRI cardiac segmentation. IEEE Trans Med Imaging. 2018;37(11):2319-31. doi: 10.1109/TMI.2018.2837502

22. ACDC Challenge. Automated Cardiac Diagnosis Challenge [Internet]. [cited 2026 Feb 14]. Available from: https://www.creatis.insa-lyon.fr/Challenge/acdc/.

23. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Cham: Springer; 2015. p. 234-41. doi: 10.1007/978-3-319-24574-4_28

24. Kingma DP, Ba J. Adam: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR); 2015 May 7-9; San Diego, CA. 2015.

25. Guo C, Pleiss G, Sun Y, Weinberger KQ. On calibration of modern neural networks. In: Proceedings of the 34th International Conference on Machine Learning (ICML). PMLR; 2017. p. 1321-30.