

Nonlinear DIP-DiracVTV model for color image restoration

Natalia Huitzil Santamaría¹, Thomas Batard¹, y Carlos Brito Loeza²

¹Centro de Investigación en Matemáticas {natalia.huitzil,thomas.batard}@cimat.mx

² Universidad Autónoma de Yucatán carlos.brito@correo.uady.mx

1. Resumen

Los modelos variacionales para problemas inversos se basan principalmente en la elección del regularizador, cuyo objetivo es asignar alguna propiedad o característica que deba cumplir la solución. La Variación Total, es uno de los regularizadores más populares para la restauración de imágenes, está inducida por el operador de gradiente euclidiano.

En este trabajo [2], introducimos un nuevo regularizador para la restauración de imágenes en color, con base en el operador de Dirac, debido a que fomenta que los gradientes de las tres componentes de color de las soluciones estén alineados, lo cual es una propiedad de las imágenes naturales [3].

Se comparan diferentes regularizadores para la restauración de imágenes en color combi-nándolos con el modelo Deep Image Prior (DIP) [5], asumiendo este último que la imagen restaurada es la salida de una red neuronal. Los experimentos de denoising y deblurring muestran que el operador de Dirac propuesto proporciona mejores resultados que el operador de gradiente euclidiano.

2. Modelo propuesto

Consideremos el siguiente modelo de degradación

$$u_0 = Au^* + n$$

donde u_0 es la imagen observada degradada, A es un operador de degradación aplicado a la imagen limpia u^* , y n algo de ruido aditivo Gaussiano.

El modelo DIP-Reg [1], [2], [4] supone que u^* puede aproximarse mediante la salida \underline{u} de una red neuronal definido de la siguiente forma:

$$\begin{cases} \underline{\theta} = \arg \min_{\theta} \frac{1}{2} \|AT_{\theta}(z) - u_0\|_{L^2}^2 + \lambda \text{Reg}(T_{\theta}(z)) & \text{para } \lambda > 0 \\ \underline{u} = T_{\underline{\theta}}z \end{cases} \quad (1)$$

donde T_{θ} es la red neuronal (se usa una U-net con conexiones de salto entre las partes codificadora y decodificadora) cuya entrada z es una imagen aleatoria, y Reg es un regularizador. En la práctica, el modelo se resuelve mediante un descenso gradiente detenido después de un cierto número de iteraciones.

En [1], se propuso el modelo **DIP-VTV**, donde el regularizador es la Variación Total Vectorial (VTV) definida por

$$\text{VTV}(u) = \int_{\Omega} \sqrt{\sum_{k=1}^3 \|\nabla u_k\|^2} d\Omega \quad (2)$$

para $u = (u_1, u_2, u_3): \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, ∇ el operador de gradiente, y $\|\cdot\|$ la norma euclidiana.

El modelo propuesto **DIP-DiracVTV^{NL}** tiene como regularizador

$$\text{DiracVTV}^{NL}(u) = \int_{\Omega} \sqrt{2 \sum_{k=1}^3 \|\nabla u_k\|^2 + 2 \sum_{1 \leq i < j \leq 3} |\nabla u_i \times \nabla u_j|} d\Omega, \quad (3)$$

donde los terminos

$$\sum_{1 \leq i < j \leq 3} |\nabla u_i \times \nabla u_j|$$

determinan el alineamiento de los gradientes usando el producto cruz \times y el valor absoluto $||$.

3. Resultados

Se probaron y compararon diferentes modelos en denoising y deblurring: DIP, DIP-VTV, DIP-DiracVTV^{NL}, y DIP-RiemannVTV que usa un gradiente riemanniano.

Probamos los modelos en el conjunto de datos Kodak, que contiene 24 imágenes naturales (<http://r0k.us/graphics/kodak/>). Para la eliminación de ruido, consideramos las imágenes en un espacio de color oponente y en sus tamaños originales (768×512 o 512×768) corrompidas con ruido gaussiano blanco aditivo de varianza 25. Para la eliminación de desenfoque, consideramos las imágenes en el espacio de color RGB, reducimos el tamaño de las imágenes a la mitad, y corrompimos las imágenes primero con desenfoque uniforme de tamaño 9×9 y luego con ruido gaussiano blanco aditivo de varianza $\sqrt{2}$.

Se tienen los siguientes resultados

Resultados	Modelo	DIP	DIP-VTV	DIP-RiemannVTV	DIP-DiracVTV ^{NL}
Valor optimo λ		-	0.01	0.02	0.03
Máx PSNR		30.76	30.85	30.86	30.96
Numero de iteraciones		4818	5125	5869	6834
Criterio de parada (Numero de iteraciones promedio)		4800	5000	6000	6750
PSNR		30.56	30.66	30.69	30.83

Cuadro 1: Resultados promedio de diferentes modelos DIP-Reg en la eliminación de ruido en el conjunto de datos Kodak. Para cada modelo, informamos de el máximo PSNR que se detuvo después de 8K iteraciones, el número de iteraciones en las que se alcanza el máximo PSNR, y el valor λ que proporciona el máximo PSNR.

Resultados	Modelo	DIP	DIP-VTV	DIP-RiemannVTV	DIP-DiracVTV ^{NL}
Valor optimo λ		-	10^{-4}	10^{-4}	10^{-4}
Máx PSNR		28.66	29.14	29.17	29.30
Numero de iteraciones		10795	17766	19785	23356
Criterio de parada (Numero de iteraciones promedio)		10000	18000	20000	23000
PSNR		28.6	29.08	29.10	29.22

Cuadro 2: Resultados promedio de los distintos modelos DIP-Reg para el desenfoque en el conjunto de datos Kodak. Para cada modelo, informamos de el máximo PSNR que se obtiene después de 30K iteraciones, y el número de iteraciones en las que se alcanza el máximo PSNR.

Referencias

- [1] T. Batard, G. Haro, and C. Ballester. Dip-vbvtv: A color image restoration model combining a deep image prior and a vector bundle total variation. *SIAM Journal on Imaging Sciences*, 14(4):1816–1847, 2021.
- [2] N. Huitzil Santamaría, T. Batard, and C. Brito-Loeza. Nonlinear dip-diracv tv model for color image restoration. In *Proceedings of the 22nd Mexican International Conference on Artificial Intelligence MICAI 2023*.
- [3] R. Kimmel, R. Malladi, and N. Sochen. Images as embedding maps and minimal surfaces: Movies, color, texture, and volumetric medical images. *International Journal of Computer Vision*, 39(2):111–129, 2000.
- [4] J. Liu, X. Xu, and U.S. Kamilov. Image restoration using total variation regularized deep image prior. In *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7715–7719, 2019.
- [5] D. Ulyanov, A. Vedaldi, and V. Lempitsky. Deep image prior. *International Journal of Computer Vision*, 128(7):1867–1888, 2020.