

Transport optimal régularisé semi-déséquilibré pour la restauration d’images

Simon MIGNON¹, Bruno GALERNE¹, Moncef HIDANE², Cécile LOUCHET¹, Julien MILLE²

¹Institut Denis Poisson
Université d’Orléans, Université de Tours, CNRS

²LIFAT
INSA Centre Val de Loire, Université de Tours
{simon.mignon, bruno.galerie, cecile.louchet}@univ-orleans.fr,
{moncef.hidane, julien.mille}@insa-cvl.fr

Résumé – Nous étudions dans cet article l’utilisation d’une pénalité basée sur le transport optimal afin de régulariser des problèmes inverses en traitement d’images. L’approche proposée est formulée dans un cadre variationnel et vise à favoriser des images dont les patches ont une distribution proche de celle apprise par un modèle génératif, ou d’exemples non dégradés disponibles. À l’aide d’illustrations numériques, nous montrons la nécessité d’adopter une forme non symétrique de déséquilibre dans la formulation du transport optimal. Nous donnons ensuite les détails permettant de calculer et de différentier cette formulation. Enfin, nous détaillons son application à un problème particulier de super-résolution : la complétion de zoom.

Abstract – In this paper we consider the use of a penalty based on optimal transport in order to regularize inverse problems in imaging. The proposed approach is formulated in a variational setting and aims at promoting images whose patch distribution is close either to the one learned by a generative model, or to available uncorrupted patches. With the aid of numerical illustrations, we argue in favor of adopting an asymmetric form of unbalanced transport. We then provide details concerning the computation and the differentiation of the proposed penalty. Finally, we detail the application of our approach to a particular super-resolution setting: the image zoom completion problem.

1 Introduction

Les problèmes de restauration constituent, aujourd’hui encore, un thème de recherche important en traitement d’images et couvrent un large champ d’applications. Ils correspondent à l’estimation d’une image cible u à partir d’une observation y dégradée par un opérateur linéaire souvent non inversible ou mal conditionné A :

$$y = Au + \eta,$$

où η correspond à une réalisation d’une variable aléatoire centrée. Ces problèmes ont été traditionnellement traités dans un cadre variationnel [19] où une estimation de u est obtenue en minimisant une fonction de coût composée d’une pénalité permettant d’induire des propriétés structurelles (parcimonie de la représentation ou de la réponse à des filtres passe-haut, redondance de patches) et d’un terme d’attache aux données.

Les avancées importantes en apprentissage profond ont donné naissance à des approches dans lesquelles la restauration est traitée comme un problème d’apprentissage discriminant, c’est-à-dire dans lesquelles on modélise, de bout-en-bout, la loi conditionnelle de u sachant y [14, 15]. Les performances empiriques de ces méthodes sont bonnes,

mais leur manque d’interprétabilité ainsi que la nécessité de ré-entraîner le réseau dès que le modèle de dégradation (A, η) change représentent des limitations importantes. Il est aussi à noter que des travaux récents [8] ont mis en lumière l’instabilité de ce type d’approches, en particulier lorsque l’opérateur A n’est pas utilisé explicitement par le réseau de neurones.

Depuis peu, on assiste à l’émergence d’approches variationnelles de restauration d’images dans lesquelles le terme de régularisation est appris, parfois de manière implicite, par un réseau profond [13, 17]. Ainsi, l’apprentissage peut être découplé du modèle de dégradation considéré. Des travaux récents ont aussi proposé d’employer un modèle génératif profond, entraîné de manière non-supervisée, pour définir le terme de régularisation [1, 7] ou une loi a priori au sein d’une approche bayésienne [11].

L’approche que nous adoptons dans cet article se situe dans le cadre des méthodes variationnelles qui emploient un modèle génératif ou des exemples non dégradés pour définir le terme de régularisation. Afin de garantir l’indépendance de l’approche proposée vis-à-vis de la dimension des images, la modélisation se fera à l’échelle des patches. L’objectif est de concevoir une pénalité qui favorise des

images dont les patches sont consistants avec un modèle génératif appris ou avec un ensemble de patches non dégradés disponibles au moment de la restauration. Pour cela, nous proposons de tirer profit des avancées récentes en transport optimal numérique [2, 5, 16, 20] et proposons une pénalité basée sur le coût d’un transport optimal non-équilibré entre la distribution empirique des patches de l’image restaurée et celle d’un modèle génératif disponible au moment de la restauration. À l’aide d’illustrations numériques, nous montrons la nécessité d’adopter une forme non symétrique de déséquilibre, dans la suite appelée transport semi-déséquilibré. Nous donnons les détails permettant de calculer et de différencier cette formulation. Enfin, nous détaillons son application à un problème particulier de super-résolution : la complétion de zoom [3, 10].

Notons enfin que le transport optimal dans l’espace des patches d’images a été récemment employé pour la synthèse de texture dans [6, 12] et pour une application de super-résolution [9] proche de celle que nous considérons dans cet article.

2 Pénalité basée sur le transport optimal équilibré

Nous adoptons dans cette section les notations et définitions de [5, 16]. Nous considérons des mesures de probabilité discrètes à support dans \mathbb{R}^n . Notons $\alpha = \sum_{i=1}^N a_i \delta_{x_i}$ et $\beta = \sum_{j=1}^M b_j \delta_{y_j}$ deux telles mesures, où $x_i, y_j \in \mathbb{R}^n$. Notons également $a = (a_i)_{i \in \llbracket N \rrbracket} \in \Sigma_N$ et $b = (b_j)_{j \in \llbracket M \rrbracket} \in \Sigma_M$ les vecteurs de probabilités associés, Σ_p désignant le simplexe de probabilité dans \mathbb{R}^p . La mesure produit de α et β sera notée $\alpha \otimes \beta$. On considère la matrice $C \in \mathbb{R}^{N \times M}$ telle que $c_{i,j} = c(x_i, y_j) = \|x_i - y_j\|_2^2$. Pour toute matrice $\pi \in \mathbb{R}_+^{N \times M}$ on notera $\pi_1 = \pi \mathbf{1}_M$ et $\pi_2 = \pi^T \mathbf{1}_N$. On note ι_E la fonction qui vaut 0 sur E et $+\infty$ sur E^c , et on note $\iota_v = \iota_{\{v\}}$ si l’ensemble est réduit à un élément.

Définition 1 (Transport optimal régularisé [5, 16]). *Soit $\varepsilon > 0$ fixé. On définit le transport optimal régularisé entre deux mesures de probabilité discrètes α et β par*

$$\text{OT}_\varepsilon(\alpha, \beta) = \min_{\pi \in \mathbb{R}_+^{N \times M}} \langle C, \pi \rangle + \varepsilon \text{KL}(\pi | \alpha \otimes \beta) + \iota_\alpha(\pi_1) + \iota_\beta(\pi_2), \quad (1)$$

où $\text{KL}(\alpha | \beta) = \langle a, \log(\frac{a}{b}) \rangle$ est la divergence de Kullback-Leibler.

Le problème de minimisation (1) est associé à un problème dual de maximisation concave en deux vecteurs $f \in \mathbb{R}^N$ et $g \in \mathbb{R}^M$. Dans [16, Prop. 4.4] et [5] sont présentés cette version duale du transport optimal ainsi qu’un algorithme de montée de gradient par blocs équivalent à celui de Sinkhorn en variables logarithmiques. Les gradients de $\text{OT}_\varepsilon(\alpha, \beta)$ par rapport à a , b , $(x_i)_{i \in \llbracket N \rrbracket}$ et $(x_j)_{j \in \llbracket M \rrbracket}$

peuvent être obtenus par différentiation implicite [4, p. 124].

Comme expliqué dans [5], le transport optimal équilibré régularisé est biaisé. Rigollet et al. [18] montrent que la projection, au sens OT_ε , d’une mesure empirique sur une classe de mesures vérifiant une hypothèse dite de fermeture sous domination, correspond à un estimateur du maximum de vraisemblance dans un modèle de déconvolution gaussien dont l’écart-type est précisément ε .

Néanmoins, pour des valeurs de ε faibles, OT_ε permet de comparer des distributions et est calculable et différentiable à l’aide de l’algorithme de Sinkhorn. Ces arguments en font un bon candidat pour définir une pénalité dans des méthodes variationnelles de restauration d’images. Considérons en particulier le cas où l’on dispose d’un modèle génératif de patches, représenté par une mesure de probabilité β . Ce modèle peut être appris en amont sur un ensemble de patches non dégradés, ou correspondre à la donnée d’un ensemble de patches similaires à ceux que l’on souhaite restaurer, auquel cas β sera une mesure empirique. La restauration pénalisée par le transport optimal régularisé est posée comme le problème suivant :

$$\min_x \frac{\lambda}{2} \|Ax - y\|^2 + \mathbb{E}_{\tilde{\beta}_M \sim \beta} [\text{OT}_\varepsilon(\alpha_x, \tilde{\beta}_M)], \quad \lambda, \varepsilon > 0. \quad (2)$$

Dans (2), $\alpha_x = \frac{1}{N} \sum_{n=1}^N \delta_{P_n x}$ est la distribution empirique associée aux patches $(P_n x)_{n \in \llbracket N \rrbracket}$ extraits de x et $\tilde{\beta}_M$ est la distribution empirique associée à un tirage de M patches générés selon β . Ainsi, dans le cas où β est à support continu, l’approche que nous proposons ne nécessite pas de connaître sa densité, mais simplement de pouvoir l’échantillonner.

Illustrons la résolution de ce problème avec un exemple de débruitage, c’est-à-dire avec $A = \mathcal{I}$, sur une texture paramétrique u (voir figure 1) dont les patches sont générés à partir d’une gaussienne \mathcal{N}_1 . Le modèle génératif β choisi est un mélange de densité $0.8\mathcal{N}_1 + 0.2\mathcal{N}_2$, où \mathcal{N}_2 est une gaussienne qui génère des patches avec un bord horizontal, non présents dans u . Le problème (2) est minimisé par descente de gradient stochastique dans laquelle un échantillon de taille $M = 24000$ selon β est tiré à chaque itération. Avec les paramètres $\lambda = 0.3$ et $\varepsilon = 10^{-4}$, on obtient le résultat OT_ε de la figure 1. On remarque la création de patches aberrants, causée par la contrainte forte sur β . Le manque de souplesse du transport optimal régularisé nous conduit à introduire ci-après le transport optimal régularisé semi-déséquilibré.

3 Transport optimal semi-déséquilibré

Le transport que nous proposons relâche la contrainte sur la distribution cible β . L’objectif est d’accroître la robustesse vis-à-vis du modèle β en autorisant l’algorithme de restauration à ignorer une partie de cette distribution.

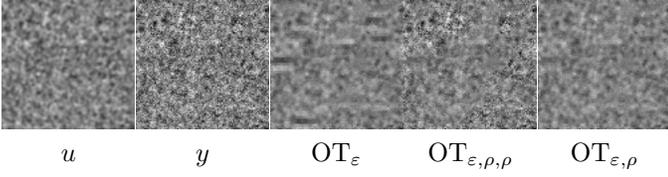


FIGURE 1 – Débruitage avec transport optimal régularisé équilibré, déséquilibré et semi-déséquilibré ($\lambda = 0.3, \varepsilon = 10^{-4}, \rho = 0.1$). Les PSNR sont respectivement 24.81, 21.72 et 25.48.

Définition 2 (Transport optimal régularisé semi-déséquilibré). Soient $\varepsilon > 0$ et $\rho > 0$ fixés. On définit le transport optimal régularisé semi-déséquilibré entre les mesures α et β par

$$\text{OT}_{\varepsilon, \rho}(\alpha, \beta) = \min_{\pi \in \mathbb{R}_+^{N \times M}} \langle C, \pi \rangle + \varepsilon \text{KL}(\pi, \alpha \otimes \beta) + \iota_\alpha(\pi_1) + \rho \text{KL}(\pi_2, \beta). \quad (3)$$

Ce nouveau transport s’inspire de la version déséquilibrée de Séjourné et al. [20], notée $\text{OT}_{\varepsilon, \rho, \rho}$ dans la suite, où les deux contraintes sont remplacées par ρKL . Comme l’illustre le résultat $\text{OT}_{\varepsilon, \rho, \rho}$ de la figure 1, obtenu en remplaçant dans (2) OT_ε par $\text{OT}_{\varepsilon, \rho, \rho}$, la relaxation KL sur la distribution α_x produit des images dont certaines zones ne sont pas restaurées. Dans (3), seule la seconde contrainte est remplacée par ρKL . Le résultat de débruitage $\text{OT}_{\varepsilon, \rho}$ de la figure 1, obtenu en remplaçant dans (2) OT_ε par $\text{OT}_{\varepsilon, \rho}$, ne fait pas apparaître de bords provenant des 20% de patches non adaptés du modèle β grâce à la relaxation KL sur β . De plus, le débruitage est homogène sur toute l’image grâce à la contrainte d’égalité sur α_x .

Nous détaillons ci-dessous comment une formulation duale permet de calculer la fonctionnelle $\text{OT}_{\varepsilon, \rho}(\alpha, \beta)$ et son gradient [5, 20].

Proposition 1. Pour $\varepsilon > 0$ fixé, le transport optimal régularisé semi-déséquilibré entre deux mesures de probabilités discrètes α et β admet la version duale

$$\text{OT}_{\varepsilon, \rho}(\alpha, \beta) = \max_{(f, g) \in \mathbb{R}^N \times \mathbb{R}^M} \langle a, f \rangle - \langle b, \Phi^*(-g) \rangle - \varepsilon \langle a \otimes b, \exp\left[\frac{f \oplus g - C}{\varepsilon}\right] - 1 \rangle \quad (4)$$

avec $\Phi^*(q) = \rho(\exp[\frac{q}{\rho}] - 1)$, appliquée à chaque composante.

(4) est un problème de maximisation concave qui peut être résolu par maximisation alternée de la fonctionnelle en f et g comme précisé par le théorème suivant.

Théorème 1 (Algorithme de Sinkhorn [20]). En partant de f^0 quelconque, l’algorithme suivant converge vers la solution du transport optimal régularisé semi-déséquilibré (4) :

$$\begin{cases} g_j^{t+1} = -\frac{\varepsilon}{1+\varepsilon} \log \left(\sum_{i=1}^N a_i \exp \left[\frac{f_i^t - C_{ij}}{\varepsilon} \right] \right), & j \in \llbracket M \rrbracket, \\ f_i^{t+1} = -\varepsilon \log \left(\sum_{j=1}^M b_j \exp \left[\frac{g_j^{t+1} - C_{ij}}{\varepsilon} \right] \right), & i \in \llbracket N \rrbracket. \end{cases}$$

Proposition 2. Soit $F(f, g)$, la fonction à maximiser dans (4). La suite de vecteurs (f^t, g^t) définie par l’algorithme de Sinkhorn vérifie :

$$F(f^t, g^t) = \langle a, f^t \rangle - \langle b, \Phi^*(-g^t) \rangle. \quad (5)$$

Utiliser la fonctionnelle $\text{OT}_{\varepsilon, \rho}$ pour résoudre un problème variationnel de restauration nécessite de pouvoir différencier par rapport aux coordonnées x_i la mesure $\alpha = \sum_i a_i \delta_{x_i}$. Ceci est possible en supposant que la convergence de l’algorithme de Sinkhorn est atteinte [5] : si (f^*, g^*) est une solution de (4), alors

$$\partial_{x_i} \text{OT}_\varepsilon \left(\sum_i a_i \delta_{x_i}, \beta \right) = a_i \nabla \varphi(x_i) \quad (6)$$

où $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ a pour expression

$$\varphi(x) = -\varepsilon \log \left(\sum_{j=1}^M b_j \exp \left[\frac{g_j^* - c(x, y_j)}{\varepsilon} \right] \right). \quad (7)$$

4 Application à la complétion de zoom

La complétion de zoom est un problème particulier de restauration d’images [10]. Le but est de restaurer une image u étant données sa version basse résolution $y^{\text{BR}} = SHu$, où H est un opérateur de convolution, S de sous-échantillonnage, et une sous-partie $y^{\text{PHR}} = \mathcal{M}u$ connue, \mathcal{M} étant un opérateur de masquage. Afin de résoudre ce problème nous considérons le problème de minimisation

$$\min_x \frac{\lambda}{2} \|SHx - y^{\text{BR}}\|^2 + \text{OT}_{\varepsilon, \rho}(\alpha_x, \beta) + \iota_{y^{\text{PHR}}}(\mathcal{M}x), \quad (8)$$

où $\alpha_x = \frac{1}{N} \sum_{i=1}^N \delta_{P_i x}$ et $\beta = \frac{1}{M} \sum_{j=1}^M \delta_{P_j y^{\text{PHR}}}$ sont respectivement les distributions empiriques associées aux patches de x et de y^{PHR} .

On résout numériquement ce problème par une descente de gradient en projetant sur la contrainte à chaque étape de la descente. La descente de gradient est faite par l’optimiseur Adam. La figure 2 montre les résultats que nous obtenons. Le cadre rouge désigne la portion connue de la haute résolution, dont le complémentaire est estimé par notre méthode. On remarque une supériorité du transport optimal semi-déséquilibré sur le transport optimal équilibré du fait de sa robustesse. En effet, en relaxant la seconde contrainte par ρKL , seuls les patches pertinents sont utilisés. Même si les résultats obtenus par transport régularisé semi-équilibré sont pour certains meilleurs visuellement que la méthode comparative [10], la quantification par PSNR et SSIM est moins bonne.

5 Conclusion

Nous avons présenté un nouveau cadre méthodologique pour la restauration d’images, basé sur une notion asymptotiquement déséquilibrée du transport optimal. À travers deux expériences, nous avons montré que le transport

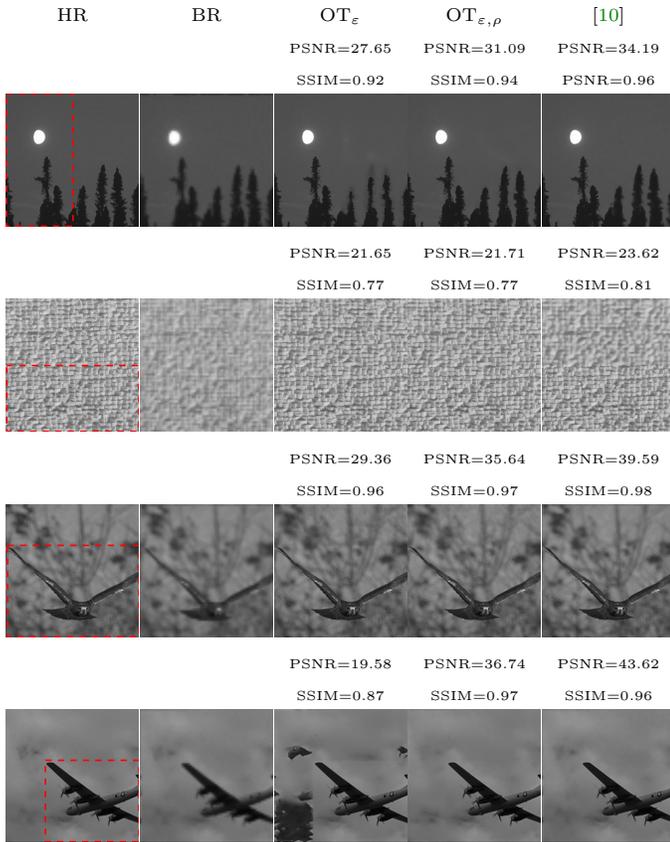


FIGURE 2 – Application à la complétion de zoom. De gauche à droite : haute résolution (avec masque), basse résolution, restauration par OT_ϵ , $OT_{\epsilon,\rho}$ et [10]. Résultats obtenus avec $\lambda = 0.3$, $\rho = 0.05$ pour les images 1, 3, 4, 5 et $\lambda = 1$, $\rho = 0.256$ pour l'image 2. Les PSNR sont été calculés sur la partie restaurée loin des bords et le SSIM sur toute l'image.

régularisé semi-déséquilibré est un outil prometteur pour la restauration d'images par patches, contrairement aux formulations usuelles équilibrée et déséquilibrée. Ce travail ouvre la porte à plusieurs perspectives comme l'étude d'une version débiaisée du transport régularisé appelée divergence de Sinkhorn [5], ou l'extension du transport semi-déséquilibré dans un cadre semi-discret.

Remerciements : Ce travail a bénéficié du soutien des projets ANR PostProdLEAP (ANR-19-CE23-0027-01) et REGETO (AAP GdR ISIS 2021).

Références

[1] A. Bora, A. Jalal, E. Price, and A.G Dimakis. Compressed sensing using generative models. In *ICML*, pages 537–546, 2017.

[2] B. Charlier, J. Feydy, J. Glaunès, F.-D. Collin, and G. Durif. Kernel operations on the gpu, with autodiff, without memory overflows. *Journal of Machine Learning Research*, 22(74) :1–6, 2021.

[3] M. El Gheche, J.-F. Aujol, Y. Berthoumiou, and C.-A. Deledalle. Texture reconstruction guided by a high-resolution patch. *IEEE TIP*, 26(2) :549–560, 2016.

[4] J. Feydy. *Analyse de données géométriques, au delà des convolutions*. PhD thesis, Université Paris-Saclay, 2020.

[5] J. Feydy, T. Séjourné, F. Vialard, S. Amari, A. Trouvé, and G. Peyré. Interpolating between optimal transport and MMD using Sinkhorn divergences. *AiSTATS*, pages 2681–2690, 2019.

[6] B. Galerne, A. Leclaire, and J. Rabin. A texture synthesis model based on semi-discrete optimal transport in patch space. *SIAM Journal on Imaging Sciences*, 11(4) :2456–2493, 2018.

[7] M. González, A. Almansa, and P. Tan. Solving inverse problems by joint posterior maximization with autoencoding prior. *arXiv preprint :2103.01648*, 2021.

[8] N.M. Gottschling, V. Antun, B. Adcock, and A.C. Hansen. The troublesome kernel : why deep learning for inverse problems is typically unstable. *arXiv preprint :2001.01258*, 2020.

[9] J. Hertrich, A. Houdard, and C. Redenbach. Wasserstein patch prior for image superresolution. *arXiv preprint :2109.12880*, 2021.

[10] M. Hidane, M. El Gheche, J.-F. Aujol, Y. Berthoumiou, and C.-A. Deledalle. Image zoom completion. *IEEE TIP*, 25(8) :3505–3517, 2016.

[11] M. Holden, M. Pereyra, and K.C. Zygalakis. Bayesian imaging with data-driven priors encoded by neural networks : Theory, methods, and algorithms. *arXiv preprint :2103.10182*, 2021.

[12] A. Houdard, A. Leclaire, N. Papadakis, and J. Rabin. Wasserstein generative models for patch-based texture synthesis. In *SSVM*, pages 269–280, 2021.

[13] R. Laumont, V. De Bortoli, A. Almansa, J. Delon, A. Durmus, and M. Pereyra. On maximum-a-posteriori estimation with plug & play priors and stochastic gradient descent. *arXiv preprint :2201.06133*, 2022.

[14] A. Lucas, M. Iliadis, R. Molina, and A.K. Katsaggelos. Using deep neural networks for inverse problems in imaging : beyond analytical methods. *IEEE Signal Processing Magazine*, 35(1) :20–36, 2018.

[15] V. Monga, Y. Li, and Y.C. Eldar. Algorithm unrolling : Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2) :18–44, 2021.

[16] G. Peyré and M. Cuturi. Computational optimal transport : With applications to data science. *Foundations and Trends in Machine Learning*, 11(5–6) :355–607, 2019.

[17] J. Prost, A. Houdard, A. Almansa, and N. Papadakis. Learning local regularization for variational image restoration. *arXiv preprint :2102.06155*, 2021.

[18] P. Rigollet and J. Weed. Entropic optimal transport is maximum-likelihood deconvolution. *Comptes Rendus Mathématique*, 356(11-12) :1228–1235, 2018.

[19] O. Scherzer. *Handbook of mathematical methods in imaging*. Springer Science & Business Media, 2015.

[20] T. Séjourné, J. Feydy, F. Vialard, A. Trouvé, and G. Peyré. Sinkhorn divergences for unbalanced optimal transport. *arXiv preprint :1910.12958*, 2019.