Regularized Discrete Optimal Transport

Sira Ferradans, Nicolas Papadakis, Julien Rabin, Gabriel Peyré, Jean-François Aujol

Regularized Discrete Optimal Transport

Sira Ferradans[1⋆], Nicolas Papadakis[2⋆⋆], Julien Rabin[3],
Gabriel Peyré[1] and Jean-François Aujol[5]

[1] Ceremade, Univ. Paris-Dauphine sira.ferradans@ceremade.dauphine.fr
[2] IMB, Université Bordeaux 1 Nicolas.Papadakis@imag.fr
[3] ENSICAEN, Université de Caen julien.rabin@unicaen.fr
[4] Ceremade, Univ. Paris-Dauphine grabriel.peyre@ceremade.dauphine.fr
[5] IMB, Université Bordeaux 1. Jean-Francois.Aujol@math.u-bordeaux1.fr

Abstract. This article introduces a generalization of discrete Optimal Transport that includes a regularity penalty and a relaxation of the bijectivity constraint. The corresponding transport plan is solved by minimizing an energy which is a convexification of an integer optimization problem. We propose to use a proximal splitting scheme to perform the minimization on large scale imaging problems. For un-regularized relaxed transport, we show that the relaxation is tight and that the transport plan is an assignment. In the general case, the regularization prevents the solution from being an assignment, but we show that the corresponding map can be used to solve imaging problems. We show an illustrative application of this discrete regularized transport to color transfer between images. This imaging problem cannot be solved in a satisfying manner without relaxing the bijective assignment constraint because of mass variation across image color palettes. Furthermore, the regularization of the transport plan helps remove colorization artifacts due to noise amplification.

Key words: Optimal Transport, color transfer, variational regularization, convex optimization, proximal splitting, manifold learning

1 Introduction

A large class of Image Processing problems involves probability densities estimated from local or global image features. In contrast to most distances from information theory (e.g. the Kullback-Leibler divergence), Optimal Transport takes into account the spacial localization of the modes of the densities [1]. Furthermore, it also provides as a by-product a warping (the so-called transport plan) between the densities. This plan can be used to perform image modifications such as color transfer. However, an important flaw of this Optimal

⋆ This work has been supported by the European Research Council (ERC project SIGMA-Vision)

⋆⋆ Nicolas Papadakis acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-11-BS01-014-01.

Transport plan is that it is usually highly irregular, thus introducing unwanted artifacts in modified images. In this article, we propose a variational formalism to relax and regularize the transport. This novel regularized Optimal Transport improves visually the result for color image modification.

1.1 Optimal Transport and Imaging

Discrete Optimal Transport. The discrete Optimal Transport (OT) is the solution of a convex linear program originally introduced by Kantorovitch. It corresponds to the convex relaxation of a combinatorial problem when the densities are sums of the same number of Diracs. This relaxation is tight (i.e. the solution of the linear program is an assignment) and extends the notion of Optimal Transport to arbitrary sum of weighted Diracs, see for instance [1]. Although there exists dedicated linear solvers (transportation simplex) and combinatorial algorithms (such as the Hungarian and auction algorithms), computing Optimal Transport is still a challenging task for densities composed of thousands of Dirac masses.

Optimal transport distance. The OT distance (also known as the Wasserstein distance or the Earth Mover distance) has been shown to produce state of the art results for the comparison of statistical descriptors, see for instance [2].

Optimal transport map. Another line of applications of OT makes use of the transport plan to warp an input density on another one. Optimal transport is strongly connected to fluid dynamic partial differential equations [3]. These connexions have been used to perform Image Registration [4]. Color transfer between images is a challenging problem, and has been tackled by computing non-linear mappings between color spaces, see for instance [5,6,7]. For grayscale images, the usual histogram equalization algorithm corresponds to the application of the 1-D Optimal Transport plan to an image, see for instance [8]. It thus makes sense to consider the 3-D Optimal Transport as a mathematically-sound way to perform color palette transfer, see for instance [9] for an approximate transport method.

1.2 Regularized and relaxed transport

Removing transport artifact. The Optimal Transport map between complicated densities is usually irregular. Using directly this transport plan to perform color transfer creates artifacts and amplifies the noise in flat areas of the image. Since the transfer is computed over the 3-D color space, it does not take into account the pixel-domain regularity of the image. The visual quality of the transfer is thus improved by denoising the resulting transport using a pixel-domain regularization either as a post-processing [10] or by solving a variational problem [10,11].

Transport regularization. A more theoretically grounded way to tackle the problem of colorization artifacts should use directly a regularized Optimal Transport. This corresponds to adding a regularization penalty to the Optimal Transport energy. This however leads to difficult non-convex variational problems, that have not yet been solved in a satisfying manner either theoretically or numerically. The only theoretical contribution we are aware of is the recent work of Louet and Santambrogio [12]. They show that in 1-D the (un-regularized) Optimal Transport is also the solution of the Sobolev regularized transport problem.

Quadratic assignment problems. Regularized transport shares similarities with regularized graph matching, which is a quadratic assignment problem, known to be NP-hard to solve. This class of problems have been convexified using an SDP relaxation of the quadratic assignment problem [13]. Such a relaxation is only tractable for small size problems, and cannot be used for imaging applications. We propose here to use a simpler convexification that works well in practice for imaging problems.

Graph regularization. For imaging applications, we use regularizations built on top of a graph structure connecting neighboring points in the input density. This follows ideas introduced in manifold learning [14], that have been applied to various Image Processing problems [15]. Using these regularizations enables us to design regularizations that are adapted to the geometry of the input density, that often has a manifold-like structure.

Transport relaxation. The result of Louet and Santambrogio [12] is deceiving from the applications point of view, since it shows that, in 1-D, no regularization is possible if one maintains a 1:1 assignment between the two densities. This is our first motivation for introducing a relaxed transport which is not a bijection between the densities. Another (more practical) motivation is that relaxation is crucial to solve imaging problems such as color transfer. Indeed, the color distributions of natural images are multi-modals. An ideal color transfer should match the modes together. This cannot be achieved by classical Optimal Transport because these modes often do not have the same mass. A typical example is for two images with strong foreground and background dominant colors (thus having bi-modal densities) but where the proportion of pixels in foreground and background are not the same. Such simple examples cannot be handle properly with Optimal Transport. Allowing a controlled variation of the matched densities thus requires an appropriate relaxation of the bijective matching constraint.

1.3 Contributions

In this paper, we generalize the discrete formulation of Optimal Transportation to tackle the two major flaws that we just mentioned: i) the lack of regularity of the transport and ii) the need for a relaxed matching between densities. Our main contribution is the integration of these two properties in a unified convex

variational problem. This problem can be solved using standard convex optimization procedures such as proximal splitting methods. Our second contribution is the application of this framework to the problem of color transfer. Numerical results show the relevance of this approach to this particular imaging problem.

2 Discrete Optimal Transport

Optimal assignment. We consider discrete measures in $\mathbb{R}^d$ with a fixed number N of points, that we write as $\mu_X = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$ where δ_a is the Dirac at position $a \in \mathbb{R}^d$, and $X = (X_i)_{i=1}^N \in \mathbb{R}^{N\times d}$, is the position of the N points supporting the distribution.

The Optimal Transport between two such distributions solves the optimal assignment

$$W(\mu_X, \mu_Y)^2 = \sum_i C_{i,\sigma^\star(i)} \quad \text{where} \quad \sigma^\star \in \operatorname*{argmin}_{\sigma\in\mathcal{S}_1} \sum_i C_{i,\sigma(i)} \tag{1}$$

where $\mathcal{S}_1$ is the set of permutation of N indexes.

A usual choice is to consider the L^α Wasserstein distance for some $\alpha > 0$, so that the cost is $C_{i,j} = \|X_i - Y_j\|^\alpha$ where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$.

Convex relaxation. The Kantorovitch Optimal Transport formulation uses the embedding of permutation σ as permutation matrix $\mathcal{M}(\sigma) \in \mathbb{R}^{N\times N}$

$$\mathcal{M}(\sigma)_{i,j} = \begin{cases} 1 & \text{if} \quad j = \sigma(i), \\ 0 & \text{otherwise.} \end{cases}$$

The convex hull of permutation matrices $\mathcal{M}(\mathcal{S}_1)$ is the set of bi-stochastic matrices

$$\bar{\mathcal{S}}_1 = \left\{ \Sigma \in \mathbb{R}^{N\times N} \setminus \Sigma\mathbb{I} = \mathbb{I}, \Sigma^*\mathbb{I} = \mathbb{I}, \Sigma \geqslant 0 \right\}$$

where $\mathbb{I} = (1, \ldots, 1)^* \in \mathbb{R}^N$, where A^* is the adjoint of the matrix A, which for real matrices amounts to the transpose. One can show that the relaxation is tight, i.e. there exists a solution $\Sigma^\star$ of

$$\min_{\Sigma\in\bar{\mathcal{S}}_1} \langle C, \Sigma\rangle \quad \text{where} \quad \langle C, \Sigma\rangle = \sum_{i,j=1}^N C_{i,j}\Sigma_{i,j}$$

such that $\Sigma^\star = \mathcal{M}(\sigma^\star)$ where $\sigma^\star$ is a solution of (1).

3 Relaxed Transport

For many applications in Imaging Science, it is not desirable to impose that the mapping σ between X and Y is 1:1. This is for instance the case for the colorization problem we consider in Section 5 where the ratio of similar colors across the image is not constant.

We relax this constraint by only imposing a maximum number of elements of X linked to a single element of Y

$$\mathcal{S}_\kappa = \left\{\sigma : \{1,\dots,N\} \to \{1,\dots,N\} \setminus \forall\, j = 1,\dots,N,\ |\sigma^{-1}(j)| \leqslant \kappa\right\}$$

where $\kappa \in \mathbb{N}^*$ is a maximum capacity parameter. Note that for $\kappa = 1$ one recovers the set $\mathcal{S}_1$ of permutations. The natural convex relaxation is

$$\bar{\mathcal{S}}_\kappa = \left\{\Sigma \in \mathbb{R}^{N\times N} \setminus \Sigma^*\mathbb{I} \leqslant \kappa\mathbb{I}, \Sigma\mathbb{I} = \mathbb{I}, \Sigma \geqslant 0\right\}.$$

The following proposition shows that this relaxation is tight when κ is an integer.

Proposition 1 *For $\kappa \in \mathbb{N}^*$, there exists a solution $\Sigma^\star$ of*

$$\min_{\Sigma\in\bar{\mathcal{S}}_\kappa} \langle C,\, \Sigma\rangle \tag{2}$$

such that $\Sigma^\star = \mathcal{M}(\sigma^\star)$ where $\sigma^\star$ is solution of

$$\min_{\sigma\in\mathcal{S}_\kappa} \sum_i C_{i,\sigma(i)}. \tag{3}$$

Proof. One can write $\bar{\mathcal{S}}_\kappa = \left\{\Sigma \in \mathbb{R}^{N\times N} \setminus \mathcal{A}(\Sigma) \leqslant b_\kappa\right\}$ where $\mathcal{A}$ is the linear mapping $\mathcal{A}(\Sigma) = (-\Sigma, \Sigma\mathbb{I}, \Sigma^*\mathbb{I}) \in \mathbb{R}^{N\times N} \times \mathbb{R}^N \times \mathbb{R}^N$ and $b_\kappa = (0_{N\times N}, \mathbb{I}, \kappa\mathbb{I})$. A standard result shows that $\mathcal{A}$ is a totally unimodular matrix [16]. For any $\kappa \in \mathbb{N}$, the vector b_κ has integer coefficients, and thus the polytope $\bar{\mathcal{S}}_\kappa$ has integer vertices. Since there is always a solution of the linear program (2) which is a vertex of $\bar{\mathcal{S}}_\kappa$, it has coefficients in $\{0,1\}$.

Note however that this relaxation is in general not tight when κ is an arbitrary real number. When $\kappa \geqslant N$, there is no restriction on the map $\sigma \in \mathcal{S}_\kappa$. The solution of (3) is then the nearest neighbor assignment,

$$\forall\, i = 1,\dots,N, \quad \sigma^\star(i) = \operatorname*{argmin}_{0\leqslant j\leqslant N} C_{i,j}. \tag{4}$$

3.1 Numerical Illustrations

In Fig. 1, we show a simple example to illustrate the properties of the method proposed so far. Given a set of points X (in blue) we compute the mapping with the set of points Y (in red), that is a solution of (2). For all the mappings between X_i and Y_j with a value $\Sigma_{i,j} > 0$, we draw a line, solid if $\Sigma_{i,j} = 1$, and dashed otherwise.

As we pointed out in last section, for non integer values of κ, the mappings $\Sigma_{i,j}$ are in $[0,1]$ while for integer values of κ, $\Sigma_{i,j} \in \{0,1\}$. Note that as we increase the values of κ (Fig. 1, right), the points in X tend to be mapped to the closer points in Y, as defined in (4).

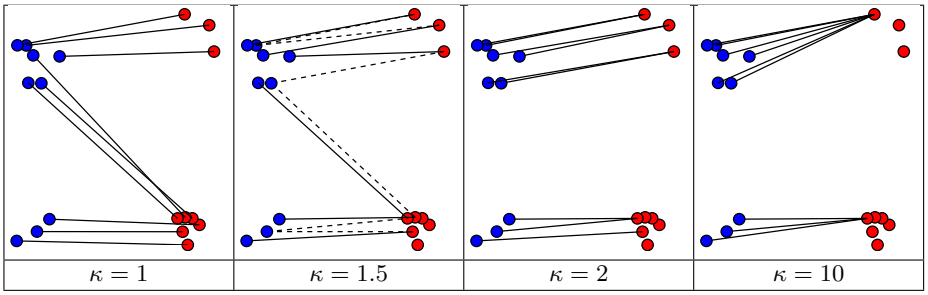


Fig. 1. Relaxed transport computed between X (blue dots) and Y (red dots) for different values of κ. Note that $\kappa = 1$ corresponds to classical OT. A dashed line between X_i and Y_j indicates that $\Sigma_{i,j}$ is not an integer.

4 Discrete Regularized Transport

4.1 Gradient on Graphs

One can view a relaxed assignment $\sigma \in \mathcal{S}_\kappa$ as a vector field $X_i \mapsto V_i = Y_{\sigma(i)} - X_i$ defined on the point cloud X. A usual way to impose regularity of such a map V is by measuring the amplitude of its derivatives GV where $G : \mathbb{R}^{N\times d} \to \mathbb{R}^{P\times d}$ is a discrete differential operator. A natural way to define a gradient is by imposing a graph structure defined by $\mathcal{G}_X \subset \{1,\dots,N\}^2$, and where $P = |\mathcal{G}_X|$. This graph structure is application dependent, and one can think of it as some sort of nearest neighbor graph. Section 5 gives an example of such a construction for the color transfer problem. The gradient measures the (weighted) difference along the edges of the graph

$$GV = (w_{i,j}(V_i - V_j))_{(i,j)\in\mathcal{G}_X} \in \mathbb{R}^{P\times d}.$$

where $w_{i,j} > 0$ is some weight. A classical choice, to ensure consistency with the directional derivative, is to choose $w_{i,j} = \|X_i - X_j\|^{-1}$.

4.2 Convex Formulation

The regularity of a transport map $V \in \mathbb{R}^{N\times d}$ is then measured according to some norm of GV, that we choose here for simplicity to be the vectorial ℓ^p norm $J_p(GV)$

$$J_p(GV) = \sum_{(i,j)\in\mathcal{G}_x} \|w_{i,j}(V_i - V_j)\|_2^p.$$

The case $p = 2$ corresponds to a graph-based Sobolev H^1 norm, whereas the case $p = 1$ corresponds to a graph-based total variation norm, see for instance [15] for applications of these functional to imaging problem regularization.

The relaxed and regularized optimal assignment problem thus reads

$$\min_{\sigma\in\mathcal{S}_\kappa} \sum_i C_{i,\sigma(i)} + \lambda J_p(G(X - Y\circ\sigma)) \tag{5}$$

where $Y\circ\sigma = (Y_{\sigma(i)})_{i=1}^N \in \mathbb{R}^{N\times d}$. To introduce a convexified regularized energy, we replace the relaxed assignment $\sigma \in \mathcal{S}_\kappa$ by $\Sigma \in \bar{\mathcal{S}}_\kappa$, and consider $X - \Sigma Y$ in place of the mapping $X - Y\circ\sigma$. We consider the following relaxed and regularized convex formulation

$$\Sigma^\star \in \min_{\Sigma\in\bar{\mathcal{S}}_\kappa} \langle C,\, \Sigma\rangle + \lambda J_p(G(X - \Sigma Y)) \tag{6}$$

The case $(\kappa,\lambda) = (1,0)$ corresponds to the usual Optimal Transport, and $\lambda = 0$ corresponds to the un-regularized formulation (3).

4.3 Minimization Algorithm

Problem (6) is a convex minimization. In the case $p = 2$, it corresponds to a quadratic minimization, whereas in the case $p = 1$ it can be cast as a conic optimization problem. They can be solved for medium-scale problem using standard interior point methods. An alternative solution is to use first order proximal scheme (see for instance [17]), that are well tailored for such highly structured problems.

Proximal Splitting. Problem (6) can be reformulated as

$$\min_{\Sigma} \langle C,\, \Sigma\rangle + \lambda J_p(G(X - \Sigma Y)) + \iota_{\mathcal{D}_1}(\Sigma) + \iota_{\mathcal{C}_\kappa}(\Sigma^*) \tag{7}$$

where ι_C is the indicator function of a convex set C and we introduced the constraint sets

$$\mathcal{C}_\kappa = \left\{\Sigma \setminus \Sigma \geqslant 0,\ \Sigma\mathbb{I} \leqslant \kappa\mathbb{I}, \Sigma \in \mathbb{R}^{N\times N}\right\}.$$

and $\iota_{\mathcal{D}_1}$ is the indicator function of the convex set $\mathcal{D}_1$ defined as:

$$\mathcal{D}_1 = \left\{\Sigma \setminus \Sigma \geqslant 0,\ \Sigma\mathbb{I} = \mathbb{I}, \Sigma \in \mathbb{R}^{N\times N}\right\},$$

where every line of $\Sigma \in \mathcal{D}_1$ belongs to the N-dimensional simplex. Problem (7) can be re-casted as a minimization of the form

$$\min_{\Sigma\in\mathbb{R}^{N\times N}} F(K(\Sigma)) + H(\Sigma)$$

$$\text{where} \quad \begin{cases} K(\Sigma) = (\Sigma, G\Sigma Y), \quad K^*(\Sigma,U) = \Sigma + G^*UY^*, \\ F(\Sigma,U) = \lambda J_p(U - GX) + \iota_{\mathcal{D}_1}(\Sigma), \\ H(\Sigma) = \langle C,\, \Sigma\rangle + \iota_{\mathcal{C}_\kappa}(\Sigma^*). \end{cases}$$

Orthogonal projection on constraint sets. The proximal operators read

$$\begin{aligned}\mathrm{Prox}_{\gamma F}(\Sigma, U) &= (\mathrm{Proj}_{\mathcal{D}_1}(\Sigma), \lambda G X + \mathrm{Prox}_{\gamma J_p(\cdot)}(U - \lambda G X)) \\ \mathrm{Prox}_{\gamma H}(\Sigma) &= \left(\mathrm{Proj}_{\mathcal{C}_\kappa}((\Sigma - \gamma C)^*)\right)^*\end{aligned}$$

where the orthogonal projection on $\mathcal{C}_\kappa$ is computed for each line Σ_ℓ of a matrix Σ as

$$\tilde{\Sigma} = \mathrm{Proj}_{\mathcal{C}_\kappa}(\Sigma) \quad \text{where} \quad \tilde{\Sigma}_\ell = \left\{ \begin{array}{ll} \max(0, \Sigma_l) & \text{if } \max(0, \Sigma_l)\mathbb{I} \leqslant \kappa \\ \mathrm{Proj}_{\mathcal{D}_\kappa}(\Sigma_l) & \text{otherwise,} \end{array} \right\}$$

where $\mathrm{Proj}_{\mathcal{D}_\kappa}(V)$ is the projection of the line vector V on the convex set $\mathcal{D}_\kappa$

$$\mathcal{D}_\kappa = \{V \setminus V \geqslant 0,\ V\mathbb{I} = \kappa\}.$$

This last projection, as well as the projection on $\mathcal{D}_1$, can be efficiently computed as detailed for instance in [18].

Proximal Operators. Let us now recall that the proximal operator of a function F is defined as

$$\mathrm{Prox}_{\gamma F}(\Sigma) = \underset{\tilde{\Sigma}}{\mathrm{argmin}}\ \frac{1}{2}\|\Sigma - \tilde{\Sigma}\|^2 + \gamma F(\tilde{\Sigma}).$$

One can check that for $p = 1$ and $p = 2$, the proximal operator of J_p evaluated at $U \in \mathbb{R}^{N \times d}$ can be computed in closed form as:

$$\mathrm{Prox}_{\lambda\gamma J_1(\cdot)}(U)_i = \max\left(0, 1 - \frac{\lambda\gamma}{\|U_i\|}\right) U_i \quad \text{and} \quad \mathrm{Prox}_{\lambda\gamma J_2(\cdot)}(U)_i = \frac{U_i}{1 + 2\gamma\lambda}$$

Note that being able to compute the proximal mapping of F is equivalent to being able to compute the proximal mapping of F^*, thanks to Moreau's identity

$$\Sigma = \mathrm{Prox}_{\gamma F^*}(\Sigma) + \gamma\, \mathrm{Prox}_{F/\gamma}(\Sigma/\gamma).$$

Primal-dual splitting scheme. The primal-dual algorithm of [17] applied to our problem finally reads

$$\begin{aligned}\Gamma^{k+1} &= \mathrm{Prox}_{\mu F^*}(\Gamma^k + \mu K(\tilde{\Sigma}^k)), \\ \Sigma^{k+1} &= \mathrm{Prox}_{\tau H}(\Sigma^k - \tau K^*(\Gamma^{k+1})), \\ \tilde{\Sigma}^{k+1} &= \Sigma^{k+1} + \theta(\Sigma^{k+1} - \Sigma^k),\end{aligned}$$

where $\theta \in [0; 1]$ and the two other parameters should satisfy $\mu\tau\|K\|^2 < 1$ where $\|K\|$ is the spectral norm of the operator K. Under these conditions, it is shown in [17] that Σ^k converges to a solution $\Sigma^\star$ of (6).

4.4 Numerical Illustrations

In Fig. 2, we can observe, on a synthetic example, the influence of the parameters κ and λ (see equation (6)). Given two sets X (in blue) and Y (in red), we compute the mapping an optimal regularized transport $\Sigma^\star$ solving eq. 6 with the minimization algorithm proposed in Section 4.3, and plot in green the set $\Sigma^\star Y$. Points X_i and Y_j are connected by a line if $\Sigma^\star_{i,j} > 0$, which is dashed if $\Sigma^\star_{i,j} \in]0,1[$ or solid if $\Sigma^\star_{i,j} = 1$. For $\lambda = 0$ and $\kappa = 1$, one obtains the classical OT solution. The influence of regularization can be observed as we increase λ: the influence of the two outliers in Y (top of the figures) in the mappings is reduced.

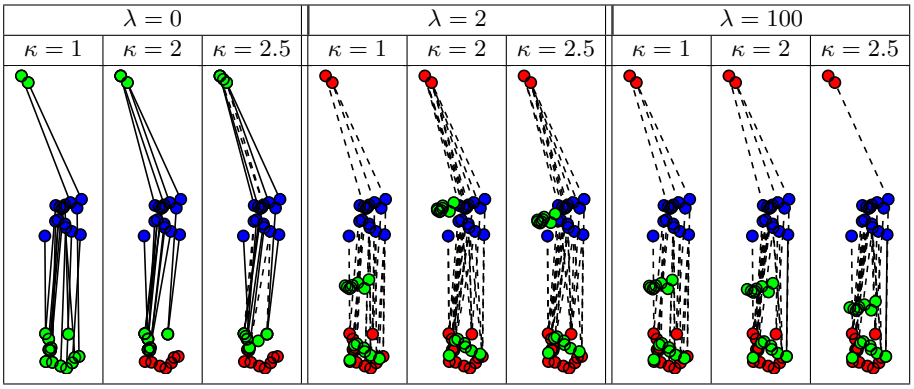


Fig. 2. Given two sets of points X (in blue) and Y (in red), we show in ΣY (in green), and the mappings $\Sigma_{i,j}$ as lines connecting X_i and Y_j, which are dashed if $\Sigma_{i,j} \in]0,1[$ and solid if it is an integer.

5 Application to Color Transfer

The color transfer problem consists in modifying an input image $X^0 \in \mathbb{R}^{N_0 \times d}$ (here $d = 3$ for RGB color image) to obtain $\tilde{X}^0 = T(X^0)$ whose color palette (its pixel empirical distribution) $\mu_{\tilde{X}^0}$ is equal or close to a target color distribution μ_{Y^0}. This target distribution is here chosen as the empirical distribution of a second image Y^0.

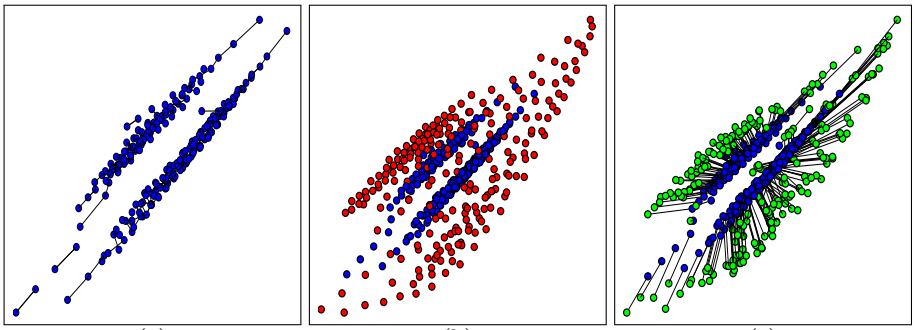
Nearest-neighbors transport interpolation. We compute T as an interpolation of a regularized Optimal Transport map between sub-sampled point clouds $X, Y \in \mathbb{R}^{N \times d}$. These two points clouds X, Y are computed by applying the K-means algorithm to the input clouds X^0, Y^0.

The regularized Optimal Transport matrix $\Sigma^\star \in \mathbb{R}^{N \times N}$ is obtained by solving (6). The quantized regularized transport then maps X to $U = \Sigma Y$. It is then extended to the whole space by a nearest neighbor interpolation

$$\forall\, x \in \mathbb{R}^d, \quad T(x) = U_{i(x)} \quad \text{where} \quad i(x) = \operatorname*{argmin}_{1 \leqslant i \leqslant N} \|x - X_i\|.$$

This transport is then applied to the input image to obtain the new pixel values $(\tilde{X}^0)_i = T(X_i^0)$.

Graph and G operator. As exposed in Section 4.1, computing a regularized transfer requires the user to design a graph structure $\mathcal{G}_X$ and weights $w_{i,j}$ that reflects the geometry of the input cloud X that supports the distribution μ_X. Inspired by several recent works on manifold learning (see Section 1.2), we use here a K-nearest neighbor graph, where K is the number of edges adjacent to each vertex, i.e. $|\{j \setminus (i,j) \in \mathcal{G}_x\}| = K$.



(a) (b) (c)

Fig. 3. In blue, the empirical distribution of the "Wheat" image pair (Fig. 4 second column), projected on the Red-Green plane. **(a)** Nearest neighbor graph $\mathcal{G}_X$ with $K = 1$, **(b)** Y in red, and **(c)** $U = \Sigma^\star Y$ in green with the lines connecting X_i to U_i.

Comparison with the state of the art. In Fig. 4, we show some results and compare them with the methods of Pitie et al. [9] and Papadakis et al. [10]. The goal of the experiment is to transfer the color palette of the images in the second row to the image on the first row. Note that the methods in the state of the art introduce color aberration (in the first column there is violet outside the flower, and in the second column the wheat is blueish), which can be avoided with the proposed method by an appropriate choice of λ and κ.

Implementation details. The results shown in this paper were obtained setting $w_{i,j} = \|X_i - X_j\|^{-1}$ and $N = 400$. The set of parameters (λ, κ, K, p) used in Fig. 4 are, by column from left to right (1400,1.05,2,1), (1000,1.2,1,1), and (1000,1,1,1).

Conclusion

In this paper, we have proposed a generalization of discrete Optimal Transport that enables to regularize the transport map and to relax the bijectivity constraints. We show how this novel class of transports can be applied to color transfer. Regularization is crucial to reduce noise amplification artifacts, while relaxation enables to cope with mass variation of the modes of the color palettes.

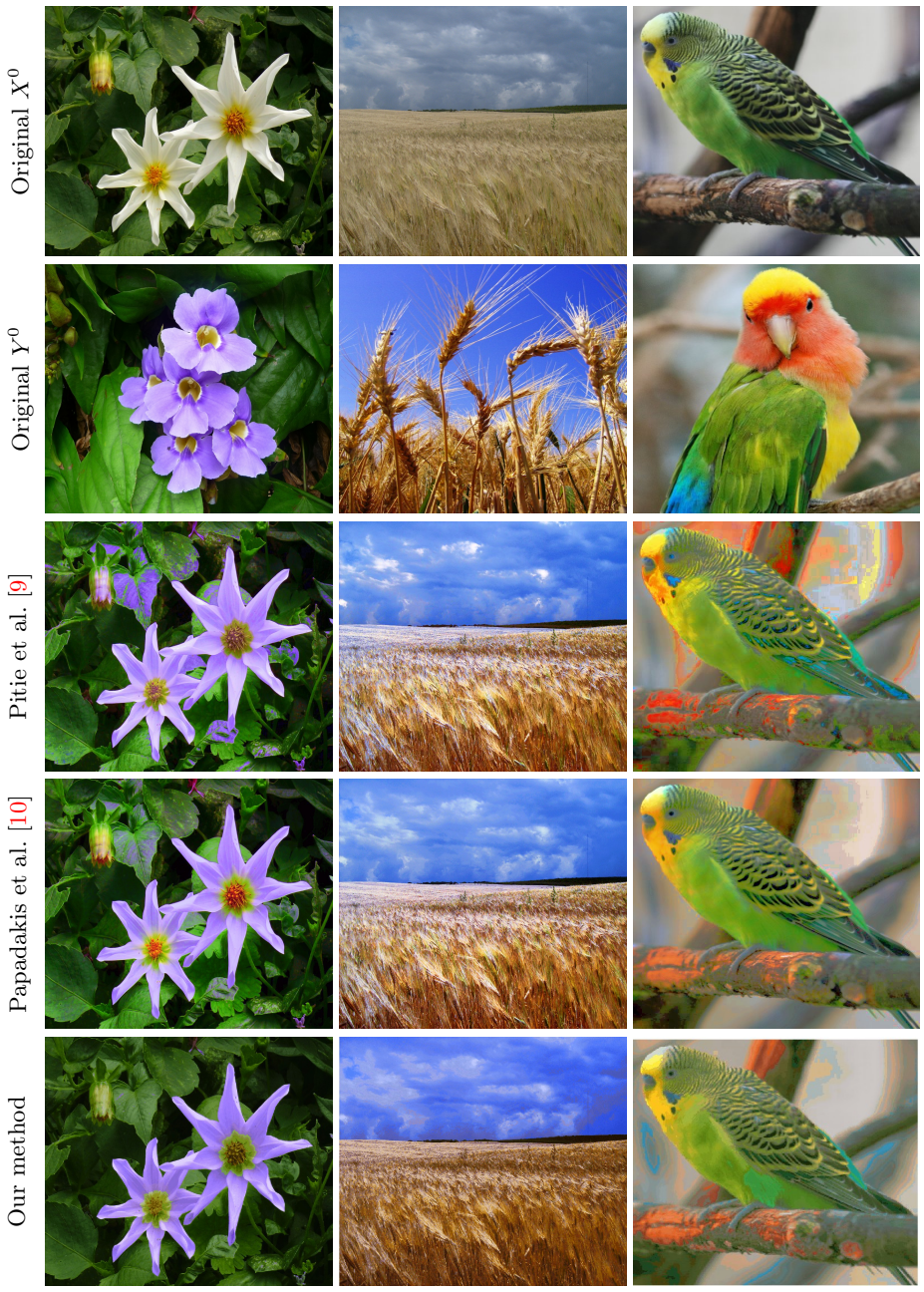


Fig. 4. Comparison between the results obtained with our method and with the methods of [9] and [10].

References

1. Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics Series. American Mathematical Society (2003)
2. Rubner, Y., Tomasi, C., Guibas, L.: A metric for distributions with applications to image databases. In: International Conference on Computer Vision (ICCV'98). (1998) 59–66
3. Benamou, J.D., Brenier, Y.: A computational fluid mechanics solution of the monge-katorovich mass transfer problem. Numerische Mathematik **84** (2000) 375–393
4. Haker, S., Zhu, L., Tannenbaum, A., Angenent, S.: Optimal mass transport for registration and warping. International Journal of Computer Vision **60** (2004) 225–240
5. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE transactions on Computer Graphics and Applications **21** (2001) 34 –41
6. Morovic, J., Sun, P.L.: Accurate 3d image colour histogram transformation. Pattern Recognition Letters **24** (2003) 1725–1735
7. McCollum, A.J., Clocksin, W.F.: Multidimensional histogram equalization and modification. In: International Conference on Image Analysis and Processing (ICIAP'07). (2007) 659–664
8. Delon, J.: Midway image equalization. Journal of Mathematical Imaging and Vision **21** (2004) 119–134
9. Pitié, F., Kokaram, A.C., Dahyot, R.: Automated colour grading using colour distribution transfer. Computer Vision and Image Understanding **107** (2007) 123–137
10. Papadakis, N., Provenzi, E., Caselles, V.: A variational model for histogram transfer of color images. IEEE Transactions on Image Processing **20** (2011) 1682–1695
11. Rabin, J., Peyré, G.: Wasserstein regularization of imaging problem. In: IEEE International Conderence on Image Processing (ICIP'11). (2011) 1541–1544
12. Louet, J., Santambrogio, F.: A sharp inequality for transport maps in via approximation. Applied Mathematics Letters **25** (2012) 648 – 653
13. Schellewald, C., Roth, S., Schnörr, C.: Evaluation of a convex relaxation to a quadratic assignment matching approach for relational object views. Image Vision Comput. **25** (2007) 1301–1314
14. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science **290** (2000) 2319–2323
15. Elmoataz, A., Lezoray, O., Bougleux, S.: Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing. Trans. Img. Proc. **17** (2008) 1047–1060
16. Schrijver, A.: Theory of linear and integer programming. John Wiley & Sons, Inc., New York, NY, USA (1986)
17. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision **40** (2011) 120–145
18. Duchi, J., Shalev-Shwartz, S., Singer, Y., Chandra, T.: Efficient projections onto the ℓ^1-ball for learning in high dimensions. In: Proc. ICML'08. ICML'08, New York, NY, USA, ACM (2008) 272–279