Adaptive color transfer with relaxed optimal transport — **Source link** ↗

Julien Rabin, Sira Ferradans, Nicolas Papadakis

Institutions: University of Caen Lower Normandy, Duke University, Centre national de la recherche scientifique

Related papers:

- Sinkhorn Distances: Lightspeed Computation of Optimal Transport

- Optimal Transport for Domain Adaptation

- Color transfer between images

- Wasserstein Generative Adversarial Networks

- Regularized Discrete Optimal Transport

Adaptive Color Transfer With Relaxed Optimal Transport

Julien Rabin, Sira Ferradans, Nicolas Papadakis

HAL Id: hal-01002830
https://hal.archives-ouvertes.fr/hal-01002830

Submitted on 6 Jun 2014

Adaptive color transfer with relaxed optimal transport

J. Rabin (GREYC, Univ. Caen) S. Ferradans*(Duke University) N. Papadakis†(CNRS, IMB)

June 6, 2014

Abstract

This paper studies the problem of color transfer between images using optimal transport techniques. While being a generic framework to handle statistics properly, it is also known to be sensitive to noise and outliers, and is not suitable for direct application to images without additional post-processing regularization to remove artifacts. To tackle these issues, we propose to directly deal with the regularity of the transport map and the spatial consistency of the reconstruction. Our approach is based on the relaxed and regularized discrete optimal transport method of [8]. We extend this work by (i) modeling the spatial distribution of colors within the image domain and (ii) tuning automatically the relaxation parameters. Experiments on real images demonstrate the capacity of our model to adapt itself to the considered data.

1 Introduction

A large class of Image Processing problems involve probability densities estimated from local or global image features. To compare these densities, most distances from information theory (*e.g.* Bhattacharya distance or Kullback-Leibler divergence) give fast and accurate estimations when the densities are similar. However, these tools are not adapted when the density modes are not located spatially at the same place. Furthermore, for some classes of problems, the mapping between densities itself is more important than the value of the distance, whereas classic metrics do not define correspondence functions between densities.

Optimal Transport in Imaging   As a consequence, the application of the Optimal Transport (OT) framework for Image Processing problems has been widely studied in the last decade. The OT problem consists in estimating the minimal cost of transferring a source distribution onto a target one.

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 discrete statistical descriptors, see for instance [20]. It can therefore deal perfectly well with discrete densities that have either similar or different supports. As a byproduct of the OT distance computation, the minimization of the transport cost also provides the *optimal plan* between the two distributions. For discrete densities represented by cloud of points, the problem amounts to finding the optimal bijective assignment between the clouds, which is a NP-hard problem (see for instance [22][4]). While the estimation of the transport and the optimal path are explicit in 1-D domain, the existing algorithms for multidimensional data (linear programming, Hungarian or auction algorithms) are limited to small sets of points (typically less than 10^4).

Color transfer   The goal of color transfer is to impose the 'image look and feel' [19] of one image onto another exemplar image. More precisely, given two input images, color transfer algorithms impose the color palette of the first image onto a second one. Initiated by [19], they propose to transfer some simple statistical descriptors (the mean and standard deviation) in the $l\alpha\beta$ color space. This idea was then extended to any color space by [23]. More precise methods for statistic transfer propose to match the whole color distribution. For grayscale images, histogram matching corresponds to the application of the 1-D OT plan to an image, as noted by [6] for the case of histogram equalization. In [10], a reduced color index was defined, so that the 1-D discrete OT can be easily computed. Approximate methods for 3D OT were proposed by [13, 18] for color images, and more recently for videos [3].

However, as pointed out by [16, 14], the perfect transfer of color is not satisfying in real applications given that color densities may have very different shapes and outliers generally appear in the final images. In fact, as the transfer is performed in the color space, it does not take into account the fact that coherent colors should be transferred to neighboring pixels, making obvious JPEG compression blocks, or increasing noise. As a consequence, methods have been proposed to consider the spatial nature of images and model some regularity priors on the image domain. Similarly to previous works on image enhancement (see *e.g.* [2, 11]), the color transfer may be formalized [12, 17] as an energy minimization problem in the image domain, which allows to directly incorporate a spatial regularization of colors. The post-regularization of the image color obtained by OT has also been proposed in [13, 15].

With these last approaches, pixels that were originally close in the color space can nevertheless be very different after the color transfer. In [21], an EM approach is used to estimate a Gaussian mixture model in both color and pixel space, showing that taking into account the pixel location helps preserve the image geometry. A major limitation of this method is that, contrary to OT

methods which match point clouds in order to minimize some global cost function, the clusters are matched using a greedy approach based on nearest-neighbor criterion. To avoid contrast inversion, they used only luminance instead of full color information, so that there is no control over the color distortion.

In the case of OT, the introduction of regularity priors on the transport map has been introduced in [8] to transfer 3D color histograms. This method relaxes the assignment problem but it does not deal with the spatial distorsion introduced by the color mapping.

Contributions In this paper, we reformulate the discrete relaxed and regularized OT method of [8] in order to tackle the aforementioned flaws. We first define the relaxed optimal transport between densities and present the regularization of the transport map on a non regular grid taking into account the spatial distribution of colors. Hence, we propose to automate the setting of relaxation parameters. Numerical results show the relevance of this approach for color transfer problems.

2 Relaxed, Regularized and Weighted OT

Problem Formulation We consider the bipartite weighted matching problem between two histograms. Let μ_X and ν_Y be two discrete and normalized distributions of features $X \in \mathbb{R}^{N_X \times d}$ and $Y \in \mathbb{R}^{N_Y \times d}$, respectively. Their corresponding weights are $\mu \in \mathbb{R}^{N_X}$ and $\nu \in \mathbb{R}^{N_Y}$, so that: $\mu_X = \sum_{i \in I_X} \mu_i \delta_{X_i}$ and $\nu_Y = \sum_{j \in I_Y} \nu_j \delta_{Y_j}$, where:

- $I_X = \{1, \ldots, N_X\}$ and $I_Y = \{1, \ldots, N_Y\}$ denote the two index sets;
- $X = \{X_i \in \mathbb{R}^d\}_{i \in I_X}$ and $Y = \{Y_j \in \mathbb{R}^d\}_{j \in I_Y}$ are respectively two sets of features in a d dimensional space;
- $\delta_X(.) = \delta(. - X)$ is the Dirac delta function located at X;
- μ and ν are non-negative and normalized vectors, *i.e.* $\mu, \nu \geq 0$ and $\sum_{i \in I_X} \mu_i = \sum_{j \in I_Y} \nu_j = 1$.

We look for a mapping $T : \mathbb{R}^d \mapsto \mathbb{R}^d$ between X and Y, referred to as *transfer* map in the following. In order to transport μ_X towards ν_Y (a.k.a. the "Monge's Problem" [22]), we define $T\#\mu_X = \sum_i \mu_i \delta_{T(X_i)}$. In practice, depending on the targeted application, the mapping should have some of the following properties: *(i)* $T\#\mu$ should be close to or match the target distribution ν, *(ii)* T should be regular in some sense, *(iii)* T should take into account the weights μ and ν. We now describe how modeling such constraints.

Optimal Transport When looking for a perfect match, an effective way of addressing this problem (see *e.g.* [22]) is to define a local matching cost function $c(X_i, Y_j) = (C_{XY})_{i,j}$ between features X and Y, and then finding the *optimal transport* plan π. This optimal plan gives the perfect match between μ and ν, such that $\pi\#\mu_X = \nu_Y$, and is estimated through the matrix minimizing the following linear program:

$$P^* \in \underset{P \in \mathcal{P}(\mu,\nu)}{\mathrm{argmin}} \ \{\langle C_{XY}, P \rangle = \sum_{i \in I_X, j \in I_Y} (C_{XY})_{i,j} P_{i,j}\} \quad (1)$$

where the set of admissible solutions is defined as $\mathcal{P}(\mu,\nu) = \{P \in [0,1]^{N_X \times N_Y} \backslash P\mathbb{I}_{N_Y} = \mu, \ P^*\mathbb{I}_{N_X} = \nu\}$, $\mathbb{I}_N$ being a column vector of N elements equal to 1 and P^* corresponding to the transpose of P. The first constraint of $\mathcal{P}(\mu,\nu)$ implies that $(P\mathbb{I}_{N_Y})_i = \sum_{j \in I_Y} P_{ij} = \mu_i$. Thus $\mathbb{I}^*_{N_X} P \, \mathbb{I}_{N_Y} = \sum_{i \in I_X} \mu_i = 1$. Thanks to the second constraint, the whole "mass" from the source X is transferred to Y. The values P_{ij} of matrix P then corresponds to the proportion of the mass of cluster X_i that is transferred to the cluster Y_j.

The advantage of such a formulation is that it is a convex problem that can easily be solved using linear program methods, provided that the size of the problem is small (typically $N_X N_Y \leq 10^6$). The main limitation of this framework is that π is a multi-valued mapping function, meaning that each cluster X_i may be matched to several clusters Y_j: $\pi\#\mu_X = \sum_{i,j} P_{i,j} \delta_{Y_j}$. For color transfer application, this mixing can therefore lead to the creation of spatial irregularities. Moreover, OT suffers from high sensibility to weight inconsistency [16], which is known to be a strong limitation for the target we have in mind. To address those issues, we propose to consider a relaxed and regularized convex problem.

Relaxed Weighted OT Following [8], we can tackle the problem of weight inconsistencies by relaxing the matching constraint towards the target distribution ν_Y on the matrix P. We introduce the set of acceptable transfer maps $\mathcal{P}_\kappa(\mu,\nu)$:

$$\mathcal{P}_\kappa(\mu,\nu) = \{P \in [0,1]^{N_X \times N_Y} \backslash P \, \mathbb{I}_{N_Y} = \mu, \ P^*\mathbb{I}_{N_X} \leqslant \kappa \otimes \nu\}$$

where $\kappa \in \mathbb{R}^{N_Y}$, $\kappa \geq 0, \kappa \otimes \nu = (\kappa_j \, \nu_j)_{j \in I_Y}$ and s.t. $\langle \kappa, \nu \rangle \geq 1$.

Observe here that the second constraint involves now a set of local relaxation parameters $\kappa = (\kappa_j)_{j \in I_Y}$. Such a relaxation κ_j parameter allows the model to increase or decrease the capacity value of the corresponding feature Y_j of the target histogram ν_Y; the maximum amount of mass from μ_X that Y_j can receive is now bounded by $\kappa_j \, \nu_j$. This relaxed modeling will allow defining a model that locally adapt itself to the data. If $\kappa = \mathbb{I}_{N_Y}$, we get the previous *OT* problem.

Average transport map Here, we resort to the following definition adapted from [8] which may be seen as a *Posterior mean* estimate to define a one-to-one transfer function T:

$$T(X_i) = \overline{Y_i} = \frac{1}{\sum_{j \in I_Y} P^\star_{ij}} \sum_{j \in I_Y} P^\star_{ij} Y_j, = \frac{1}{\mu_i} \sum_{j \in I_Y} P^\star_{i,j} Y_j. \quad (2)$$

One thus have a linear relation between the transport matrix and the average transport map: $T(X) = D_\mu PY$, where $D_\mu = diag(\mu^{-1})$.

Regularity of a transport map We consider a structured representation of X through the weighted graph $\mathcal{G}_X = (I_X, E_X)$, where I_X is the set of vertices, and $E_X \subseteq$

I_X^2 is the set of edges [5]. We refer to $w_{ij} \geq 0$ as the weight of edge $(i,j) \in E_X$ between features X_i and X_j. The divergence of a vector field $V = \{(V_i^\ell)_{1 \leq \ell \leq n}\}_{i \in I_X} \in \mathbb{R}^{N_X \times n}$ on the graph $\mathcal{G}_X$ is defined at point X_i, for each dimension ℓ, as the (opposite) adjoint operator of the gradient operator

$$\forall i \in I_X, \forall \ell \leq n \quad (div_X V)_i^\ell := \sum_{j \in I_X} \omega_{ij} \left(V_i^\ell - V_j^\ell\right) \tag{3}$$

In the following, we will measure the regularity of the flow V by considering the L^1 norm of the divergence

$$J(V) = \|div_X V\|_1 := \sum_{i \in I_X} \sum_{1 \leq \ell \leq n} \left|(div_X V)_i^\ell\right|. \tag{4}$$

Convex formulation When considering color transfer, we would like the transport to be piecewise constant, since we do not want to penalize color shift. To that end, we consider the regularization of the average transport displacement by defining $V = A(D_\mu PY - X)$ in the penalization model (4). Matrix $A \in \{0,1\}^{n \times d}$ selects features components that should be regularized. For color transfer, we only consider $n = 3$ (RGB coordinates) among $d = 5$ dimensional features (color attributes along with additional spatial dimensions). Moreover, in order to control that the relaxed and regularized transported color palette is close to the target, we add an additional term $\|\kappa - \mathbb{I}_{N_Y}\|$ to check that the relaxation is tight.

Introducing the regularization parameter $\lambda \geqslant 0$ and the fidelity one $\rho \geqslant 0$, our model becomes the following convex optimization problem with linear constraints:

$$\{P^\star, \kappa^\star\} \in \underset{\substack{P \in \mathcal{P}_\kappa(\mu, \nu) \\ \kappa \in \mathbb{R}^{N_Y}, \kappa \geq 0, \langle \kappa, \nu \rangle \geq 1}}{\operatorname{argmin}} \langle C_{XY}, P \rangle + \lambda J(A(D_\mu PY - X)) + \rho \|\kappa - \mathbb{I}\|_1. \tag{5}$$

Notice that this energy can be minimized with a linear solver by introducing auxiliary variables to deal with the L^1 norm (see [7] for more details). This formulation generalizes the OT model proposed in [8] on cloud of points to handle distributions. We will also show that it permits to include spatial regularization for color transfer purposes. Observe that the transport is here automatically relaxed, since the we do not require the user to tune the capacity vector κ. Note also that, unlike [8], we consider the divergence operator instead of Total Variation in order to reduce the time complexity.

3 Application to color transfer

In this section we consider the application of our model to color-transfer (or color "grading"). Let u be the source image and v an exemplar image with the desired color distribution. We denote as $u : \Omega_u \subset \mathbb{Z}^2 \mapsto \Sigma \subset \mathbb{R}^3$, where Ω_u is the regular pixel grid of u and Σ is the quantized RGB color space. The problem is to find a new image w whose geometry is as close as possible to the source image u and whose color distribution is close to the one of the exemplar image v.

Let us now detail how the framework presented in the previous section can be applied to color transfer (see illustration in Figure 1). The first step consists in defining

the source and target sets X, Y, which involves spatio-color clustering on the input images u, v, respectively. These clusters are then used to build a weighted graph $(\omega_{i,j})$ and define the transport cost matrix C_{XY} that are involved in the local adaptive functional (5) to minimize. The color transfer is finally applied using the estimated relaxed and regularized transport map.

Spatio-color clustering To cope with the dimensional issue, we perform a joint clustering on the pixel grid and on the color space ($d = 5$). The images u and v are quantized to generate X and Y with N_X, N_Y clusters, respectively. We denote $X_i = (x_i, U_i) \in \mathbb{R}^5$, to specify the spatial component ($x_i \in \Omega_u$) and the color component ($U_i \in \Sigma$). These values (x_i, U_i) represent the mean spatial and color values of the set of pixels in u (and equivalently for v) assigned to cluster X_i. Here, we computed this assignment using the fast super-pixels method [1], with the default regularization parameter 0.02 and a raw 10×10 seed initialization. Note that each cluster X_i has an assigned group of pixels $\{\mathcal{N}_i \subset \Omega_u \times \Sigma\}_{i \in I_X}$ that can have a different amount $|\mathcal{N}_i|$ of pixels. We define the cluster weight as $\mu_i = |\mathcal{N}_i|/|\Omega_u|$.

Thus, we consider from now the measure of interest $\mu_X : X = (x, U) \mapsto \sum_{i \in I_X} \mu_i \delta_{X_i}(X) = \sum_{i \in I_X} \mu_i \delta_{x_i}(x) \delta_{U_i}(U)$, and similarly for the cluster set Y obtained from image v, ν_Y.

Transport cost Considering only color characteristics for transfer, the ground cost is taken as $(C_{XY})_{i,j} = \|U_i - V_j\|^2$.

Graph construction As stated in the previous section, to consider a regularization term, we need to define the concept of neighborhood on the data set. In this work, a NN graph is computed through k-NN search based on the L^2-distance between clusters. The weight $\omega_{i,j}$ of the edge that links clusters X_i to X_j in (3) is defined as $\omega_{i,j} = \exp\left(-\frac{1}{2}\|X_i - X_j\|_{V_i}^2\right)$, where $\|X\|_\Sigma^2 = X^* \Sigma^{-1} X$ is the Mahalanobis distance and

$$V_i = \frac{1}{\mu_i} \sum_{E, F \in \mathcal{N}_i} S(E - X_i)(F - X_i)^* S^* \in \mathbb{R}^{5 \times 5}$$

is the weighted empirical covariance matrix of cluster X_i, using diagonal weighting matrix $S = diag(\sigma_x, \sigma_y, \sigma_r, \sigma_g, \sigma_b)$. Note that, the matrix V_i^{-1} is in practice a pseudo-inverse matrix (Moore-Penrose) to avoid numerical problem when dealing with uniform clusters.

Color Palette Transfer Once the optimization problem is solved and the relaxed OT matrix $P^\star$ is computed, we define the *color transfer* as $\mathcal{T}(U) = D_\mu P^\star V$, *i.e.*

$$\mathcal{T} \# \mu(x, U) = \sum_{i \in I_X} \mu_i \delta_{x_i} \delta_{\mathcal{T}(U_i)}. \tag{6}$$

Image Synthesis As we work at a super-pixel scale to speed-up the OT computation, the last step of the proposed approach is to synthesize a new image w from the source image u using the new color palette $\mathcal{T} \# \mu_X(x, U)$.

Figure 1: Color transfer process. (a) source image u and target color image v. (b) Super-pixel segmentation of u and v. Raw color transfer without (c) and with (d) relaxation (i.e. $\rho = \infty$ and $\rho = 10^3$). (e) Final image w with post-processing. *Photo credit: Nicolas Le Dilhuit*

Like [21], we use maximum likelihood estimation to incorporate geometrical information from the source image u into the synthesis process. For each pixel $X = (x, u(x))$ of the source image, the idea is to use the likelihood to all clusters $\{X_i\}_i$ through the weights $\{\omega_i\}_i$ in order to compute a linear combination of the estimated transferred colors $\mathcal{T}(X_i)$. Using (6), we set $\forall x \in \Omega_u$, $w(x) = \frac{1}{W(X)} \sum_{i \in I_X} w_i(X)\mathcal{T}(U_i)$, with $w_i(\cdot) = \exp\left(-\frac{1}{2}\|\cdot - X_i\|_{V_i}^2\right)$, and the normalizing factor $W(X) = \sum_i w_i(X)$.

Post-processing In order to restore the sharp details from the original image that may have been lost in the process, we make use of the NLMR filter from [15]. Note that this step can be approximated and speeded up using the real time guided filter proposed by [9].

Experiments We first present in Figure 1 the full pipeline including the super-pixel segmentation (b) of the images (a). The raw color transfers obtained without (c) and with (d) the automatic tuning of the relaxation parameters are presented as well as the result of the post-processing (e). In all our experiments, we used the following parameters: $n = 10$ neighbors in the graph, a regularization parameter $\lambda = 10^{-1}$, a capacity fidelity $\rho = 10^3$ and the covariance rescaling $\sigma_x = \sigma_y = 10$, $\sigma_r = \sigma_g = \sigma_b = 1$, to give more importance to the color information. A comparison with state-of-the art methods [13, 12, 8] is given in Figure 2. Notice that [13] can be improved with a post-processing [15]. Thanks to the segmentation, our algorithm only requires a few seconds whereas the methods of the literature last a few minutes. Contrary to [13, 12], we do not match exactly the target color palette, since our model adapt itself to the source image data. With respect to [8], the improvement of the spatial regularization allows removing artifacts that appear for instance in the background. We finally present in Figure 3 some examples of color transfer. The reader may find more results and comparisons at http://www.math.u-bordeaux1.fr/~npapadak/Color_transfer.html.

4 Conclusion and perspectives

In this paper, we have proposed a method for transferring color between images using relaxed and regularized weighted optimal transport. Our convex minimization approach involves a spatial representation of colors and allows an automatic parameterization of the relaxation parameters.

Further improvements of the method may reside in

the incorporation of image details in the model to avoid the final post-processing step and the preservation of contrast.

Figure 2: Comparison of out approach with [13] (without and with the post-processing of [15]), [12], and [8]. *Photo credit: Elektrofisch*
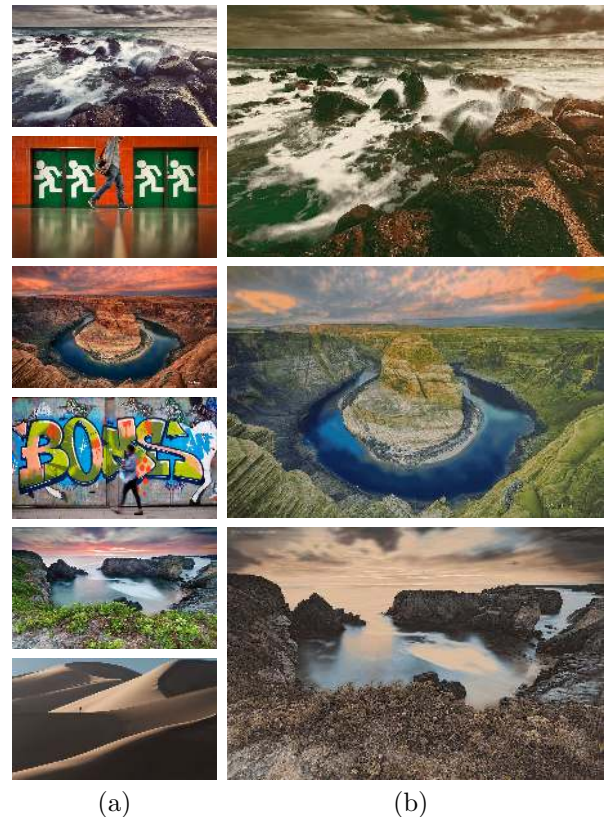

Figure 3: Examples of color transfer. (a) source image u and target color image v. (b) Final image w with post-processing. *Photo credit: 'Waves' photo by Daniele Zedda, 'Grafiti' and 'fire exits' by Thomas Leuthard, 'Canyon' by Chen Su, 'Desert' by Steve Yabek, and 'Creek' by Nicolas Le Dilhuit.*

References

[1] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence 34*, 11 (2012), 2274–2282. 3

[2] Ballester, C., Caselles, V., Igual, L., Verdera, J., and Rougé, B. A variational model for p+xs image fusion. *International Journal of Computer Vision 69*, 1 (2006), 43–58. 1

[3] Bonneel, N., Sunkavalli, K., Paris, S., and Pfister, H. Example-based video color grading. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2013) 32*, 4 (2013). 1

[4] Burkard, R., Dell'Amico, M., and Martello, S. *Assignment Problems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009. 1

[5] Chung, F. R. K. *Spectral Graph Theory.* American Mathematical Society, 1997. 3

[6] Delon, J. Midway image equalization. *Journal of Mathematical Imaging and Vision 21* (2004), 119–134. 1

[7] Ferradans, S., Papadakis, N., Peyré, G., and Aujol, J.-F. Regularized discrete optimal transport. *ArXiv e-prints 1307.5551* (2013). 3

[8] Ferradans, S., Papadakis, N., Rabin, J., Peyré, G., and Aujol, J.-F. Regularized discrete optimal transport. In *SSVM'13* (2013), Scale Space and Variational Methods in Computer Vision, pp. 428–439. 1, 2, 3, 4

[9] He, K., Sun, J., and Tang, X. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence 35*, 6 (2013), 1397–1409. 4

[10] Morovic, J., and Sun, P.-L. Accurate 3d image colour histogram transformation. *Pattern Recognition Letters 24*, 11 (2003), 1725–1735. 1

[11] Palma-Amestoy, R., Provenzi, E., Bertalmío, M., and Caselles, V. A perceptually inspired variational framework for color enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence 31*, 3 (2009), 458–474. 1

[12] Papadakis, N., Provenzi, E., and Caselles, V. A variational model for histogram transfer of color images. *IEEE Transactions on Image Processing 20*, 6 (2011), 1682–1695. 1, 4

[13] Pitié, F., Kokaram, A. C., and Dahyot, R. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding 107* (2007), 123–137. 1, 4

[14] Pouli, T., and Reinhard, E. Progressive color transfer for images of arbitrary dynamic range. *Comput. Graph. 35*, 1 (Feb. 2011), 67–80. 1

[15] Rabin, J., Delon, J., and Gousseau, Y. Removing artefacts from color and contrast modifications. *IEEE Transactions on Image Processing 20*, 11 (2011), 3073–3085. 1, 4

[16] Rabin, J., Delon, J., and Gousseau, Y. Transportation distances on the circle. *Journal of Mathematical Imaging and Vision 41*, 1-2 (2011), 147–167. 1, 2

[17] Rabin, J., and Peyré, G. In *IEEE International Conderence on Image Processing (ICIP'11)*, pp. 1541–1544. 1

[18] Rabin, J., Peyré, G., Delon, J., and Bernot, M. Wasserstein barycenter and its application to texture mixing. In *SSVM'11* (2011), vol. 6667 of *Scale Space and Variational Methods in Computer Vision*, pp. 435–446. 1

[19] Reinhard, E., Adhikhmin, M., Gooch, B., and Shirley, P. Color transfer between images. *IEEE trans. on Computer Graphics and Applications 21*, 5 (2001), 34–41. 1

[20] Rubner, Y., Tomasi, C., and Guibas, L. A metric for distributions with applications to image databases. In *ICCV'98* (1998), International Conference on Computer Vision, pp. 59–66. 1

[21] Tai, Y.-W., Jia, J., and Tang, C.-K. Local color transfer via probabilistic segmentation by expectation-maximization. In *IEEE Conference on Computer Vision and Pattern Recognition* (2005), CVPR '05, pp. 747–754. 1, 4

[22] Villani, C. *Topics in Optimal Transportation.* American Mathematical Society, 2003. 1, 2

[23] Xiao, X., and Ma, L. Color transfer in correlated color space. In *ACM International Conference on virtual reality continuum and its applications* (2006), VRCIA '06, ACM, pp. 305–309. 1