Sliced $L_2$ Distance for Colour Grading

Hana Alghamdi  
Computer Science & Statistics  
Trinity College Dublin  
Dublin, Ireland  
alghamdh@tcd.ie

Rozenn Dahyot  
Computer Science & Statistics  
Trinity College Dublin  
Dublin, Ireland  
ORCID:0000-0003-0983-3052

Abstract—We propose a new method with $L_2$ distance that maps one $N$-dimensional distribution to another, taking into account available information about correspondences. We solve the high-dimensional problem in 1D space using an iterative projection approach. To show the potentials of this method, we apply it to colour transfer between two images that exhibit overlapped scenes. Experiments show quantitative and qualitative competitive results as compared with the state of the art colour transfer methods.

Index Terms—$L_2$, Colour transfer, Colour correction

I. INTRODUCTION

Optimal Transport (OT) has been successfully used as a way for defining cost functions for optimization when performing colour distribution transfer [1], and have been used more recently to solve machine learning problems [2]–[4]. Popular OT algorithms such as Iterative Distribution Transfer (IDT) [5], [6] and Sliced Wasserstein Distance (SWD) [7] use iteratively the OT explicit solution available for mapping 1D projected source dataset onto a 1D projected target dataset. More recently an efficient framework for colour transfer was proposed [8]–[10] using IDT algorithm [5] projects two multidimensional independent datasets $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^m$ sampled for two random vectors $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^N$ with respective distributions $f_x$ and $g_y$, to 1D subspace. This projection creates two 1D datasets $\{u_i\}_{i=1}^n$ and $\{v_j\}_{j=1}^m$ with corresponding marginals $f_u$ and $g_v$ whose cumulative distributions $F_u$ and $G_v$ are matched using the 1D optimal transport solution $\phi^{OPT}(u) = G_v^{-1} \circ F_u(u)$ (cf. Fig. 1). We propose to replace the non-parametric $\phi^{OPT}(u)$ by the following parametric non-rigid 1D transformation model:

$$\phi^L_2(u) = c_0 + c_1 u + \sum_{l=1}^r w_l \varphi(||u - u_l||_2)$$

$$\hat{\theta} = \arg\min_{\theta} \left[ L_2(f_{u|\theta}, g_v) = ||f_{u|\theta} - g_v||^2 \right]$$

A. $L_2$ based solution vs Optimal Transport solution in 1D

IDT algorithm [5] projects two multidimensional independent datasets $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^m$ sampled for two random vectors $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^N$ with respective distributions $f_x$ and $g_y$, to 1D subspace. This projection creates two 1D datasets $\{u_i\}_{i=1}^n$ and $\{v_j\}_{j=1}^m$ with corresponding marginals $f_u$ and $g_v$ whose cumulative distributions $F_u$ and $G_v$ are matched using the 1D optimal transport solution $\phi^{OPT}(u) = G_v^{-1} \circ F_u(u)$ (cf. Fig. 1). We propose to replace the non-parametric $\phi^{OPT}(u)$ by the following parametric non-rigid 1D transformation model:

$$\phi^L_2(u) = c_0 + c_1 u + \sum_{l=1}^r w_l \varphi(||u - u_l||_2)$$

$$\hat{\theta} = \arg\min_{\theta} \left[ L_2(f_{u|\theta}, g_v) = ||f_{u|\theta} - g_v||^2 \right]$$

The probability density function (pdf) $f_{u|\theta}$ is a 1D Gaussian mixture model fitted to source samples $\{\phi^L_2(u_i)\}_{i=1}^n$, and the pdf $g_v$ is a 1D Gaussian mixture fitted to the target samples $\{v_j\}_{j=1}^m$. We experimented with different numbers $r$ of control points.
points and we found $r = 125$ control points on regular intervals spanning the entire range of the 1D projected dataset give best results. As a consequence the dimension of the latent space that needs to be explored when estimating $\theta$ is:

$$\dim(\theta) = (125 \times 1) + 1 + 1 = 127$$

with $\dim(u_l) = 1$, $\dim(c_0) = 1$, and $\dim(c_1) = 1$ (Eq. 1).

**B. Optimisation**

Following [8], when no correspondences are available between source and target datasets, K-means algorithm is applied in $\mathbb{R}^N$ to reduce their cardinalities $n$ and $m$ to cardinality $K$. The projections of these K-means in 1D subspace define the means of the Gaussian Mixture models $f_u$ and $g_v$. A data-driven bandwidth value $h$ is selected to control the isotropic variances of the Gaussians [13]. When $K$ correspondences in 1D are available $\{(u_k, v_k)\}_{k=1}^K$, the binned correspondences procedure proposed in [11] is used (c.f. Fig. 2).

$$K \text{means algorithm is applied to reduce their cardinalities } n, m \text{ to } K.$$  

$$\text{The projections of these } K \text{-means in } 1D \text{ subspace define the means of the Gaussian Mixture models } f_u \text{ and } g_v.$$  

$$\text{A data-driven bandwidth value } h \text{ is selected to control the isotropic variances of the Gaussians [13].}$$  

$$\text{When } K \text{ correspondences in } 1D \text{ are available } \{(u_k, v_k)\}_{k=1}^K, \text{the binned correspondences procedure proposed in } [11] \text{ is used (c.f. Fig. 2).}$$  

**C. Convergence**

The $L_2$ distance between the $N$-dimensional source and target probability distributions [13] is used as a measure to quantify how well the transformed distribution $f$ matches the target pdf $g$ after each iteration $k$ of the algorithm. Figure 3 illustrates several iterations $k$ of our algorithm visualized in 2D space, using correspondences ($\phi^{corr}$) and without correspondences ($\phi^L$). As we can see from the figure, $\phi^L$ at iteration 30 is not yet matching the target distribution in comparison to $\phi^{corr}$ that is able to match the target and converge faster by iteration 7.

**D. Interpolation between solutions**

New transformations can be interpolated in each iteration between the solutions $\phi^{corr}$ when taking into account the correspondences, and $\phi^L$ without correspondences to tackle semi-supervised situations where correspondences are only partially available [3, 14]:

$$\forall \lambda \in [0, 1], \quad \phi^L \lambda = (1 - \lambda) \phi^L + \lambda \phi^{corr} \quad (3)$$

This strategy can be useful when no correspondence can be found locally between areas of the target and source images (e.g. non-overlapped areas or occlusions, Figure 5).

**III. EXPERIMENTAL ASSESSMENT**

Quantitative evaluations have been carried out with our techniques: the 1st without correspondences using colour patches only (SL2D$_c$), the 2nd without correspondences using colour patches with pixel location information (SL2D$_{cp}$), the 3rd with correspondences using colour patches only (SL2D$_{corr}$), and the 4th with correspondences using colour patches with pixel location (SL2D$_{corr}^{cp}$). We compare our methods with different state-of-the-art colour transfer methods noted by B-PMLS [16], L2 [8], GPS/LCP and FGPS/LCP [17], PMLS [18], IDT [5], PCT_OT [10], OT_NW [11], and INWDT [11].

Hwang et al. dataset [18] is used for evaluation and it includes registered pairs of images (source and target) taken with different cameras, different in-camera settings, and different illuminations and recolouring styles. Results shown for PMLS [18] and B-PMLS [16] were provided by the authors. Two other recent techniques [19, 20] that account...
IV. CONCLUSION

We have introduced the Sliced $L_2$ distance for registering correspondences to be taken into account when these are available. Our SL2D technique applied to colour transfer extends two high-dimensional probability density functions, that allows the state of the art colour transfer methods using FSIMc metric [26] (higher values are better, best viewed in colour and zoomed in).

The patches with combined colour and spatial features create a vector in 45 dimensions ($N = 3 \times 3 \times 3 = 27$). The numerical results for PSNR, SSIM, CID and FSIMc are shown in box plots shown in Figs. 4-7 (the means shown as red dots in the plots, and the medians shown as horizontal black lines). Sliced $L_2$ with correspondences (SL2D) significantly outperforms the iterative projection approach with $L_2$ colour only (SL2Dc). Moreover, the iterative projection approach with combined colour and position (SL2Dcp) outperforms the Sliced $L_2$ solution (SL2D) algorithm. The medians reported for each group (SL2D, SL2Dc, SL2Dcp) are not statistically different (95% confidence intervals). Our SL2D technique applied to colour transfer extends the state of the art colour transfer methods using FSIMc metric [26] (higher values are better, best viewed in colour and zoomed in).

The numerical results for PSNR, SSIM, CID and FSIMc are shown in box plots shown in Figs. 4-7 (the means shown as red dots in the plots, and the medians shown as horizontal black lines). Sliced $L_2$ with correspondences (SL2D) significantly outperforms the iterative projection approach with $L_2$ colour only (SL2Dc). Moreover, the iterative projection approach with combined colour and position (SL2Dcp) outperforms the Sliced $L_2$ solution (SL2D) algorithm. The medians reported for each group (SL2D, SL2Dc, SL2Dcp) are not statistically different (95% confidence intervals). Our SL2D technique applied to colour transfer extends the state of the art colour transfer methods using FSIMc metric [26] (higher values are better, best viewed in colour and zoomed in).
Grogan et al. $L_2$ approach [8], [14] to higher dimensional spaces, and performs well against the state of the art. Future work will look at applying SL2D for shape registration [13], [28].

**REFERENCES**

[1] G. Peyré and M. Cuturi, “Computational optimal transport: With applications to data science,” Foundations and Trends® in Machine Learning, vol. 11, no. 5-6, pp. 355–607, 2019.

[2] A. Tanaka, “Discriminator optimal transport,” in Adv. Neural Inf. Process. Syst., 2019, pp. 6818–6826.

[3] C. Meng, Y. Ke, J. Zhang, M. Zhang, W. Zhong, and P. Ma, “Largescale optimal transport map estimation using projection pursuit,” in Adv. Neural Inf. Process. Syst., 2019, pp. 8118–8129.

[4] B. Muzellec and M. Cuturi, “Subspace detours: Building transport plans that are optimal on subspace projections,” in Adv. Neural Inf. Process. Syst., 2019, pp. 6917–6928.

[5] F. Pitie, A. C. Kokaram, and R. Dahyot, “Automated colour grading using colour distribution transfer,” Comput. Vis. Image Underst., vol. 107, no. 1, pp. 123 – 137, 2007.

[6] F. Pitie, A. C. Kokaram, and R. Dahyot, “N-dimensional probability density function transfer and its application to color transfer,” in IEEE Int. Conf. Comput. Vis. (ICCV), vol. 2, 2005, pp. 1434–1439.

[7] N. Bonneel, J. Rabin, G. Peyré, and H. Pfister, “Sliced and radon wasserstein barycenters of measures,” J. Math. Imaging Vis., vol. 51, no. 1, pp. 22–45, Jan 2015.

[8] M. Grogan and R. Dahyot, “L2 divergence for robust colour transfer,” Comput. Vis. Image Underst., vol. 181, pp. 39 – 49, 2019.

[9] H. Alghamdi, M. Grogan, and R. Dahyot, “Patch-based colour transfer with optimal transport,” in Proc. Eur. Signal Process. Conf. (EUSIPCO), Sep. 2019, pp. 1–5.

[10] H. Alghamdi and R. Dahyot, “Patch based colour transfer using sifting flow,” in Proc. Irish Mach. Vis. Image Process. conf. (IMVIP), 2020.

[11] ——, “Iterative nadaraya-watson distribution transfer for colour grading,” in Proc. IEEE Int. Workshop Multimed. Signal Process. (MMSP). Ieee, 2020, pp. 1–6.

[12] D. W. Scott, “Parametric statistical modeling by minimum integrated square error,” Technometrics, vol. 43, no. 3, pp. 274–285, 2001.

[13] B. Jian and B. C. Vemuri, “Robust point set registration using gaussian mixture models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 8, pp. 1633–1645, 2011.

[14] M. Grogan, R. Dahyot, and A. Smolic, “User interaction for image recoloring using l2,” in Proc. Eur. Conf. on Visual Media Production (CVMP), 2017, p. 10.

[15] D. F. Shanno, “Conditioning of quasi-newton methods for function minimization,” Math. Comp., vol. 24, no. 111, pp. 647–656, 1970.

[16] Y. Hwang, J.-Y. Lee, I. S. Kweon, and S. J. Kim, “Probabilistic moving least squares with spatial constraints for nonlinear color transfer between images,” Comput. Vis. Image Underst., vol. 180, pp. 1 – 12, 2019.

[17] F. Bellavia and C. Colombo, “Dissecting and reassembling color correction algorithms for image stitching,” IEEE Trans. Image Process., vol. 27, no. 2, pp. 735–748, 2018.

[18] Y. Hwang, J. Lee, I. S. Kweon, and S. J. Kim, “Color transfer using probabilistic moving least squares,” in Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR), 2014, pp. 3342–3349.

[19] M. Xia, J. Y. Renping, X. M. Zhang, and J. Xiao, “Color consistency correction based on remapping optimization for image stitching,” in IEEE Int. Conf. Comput. Vis. Workshops, Oct. 2017, pp. 2977–2984.

[20] J. Park, Y. Tai, S. N. Sinha, and I. S. Kweon, “Efficient and robust color consistency for community photo collections,” in Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR). June 2016, pp. 430–438.

[21] N. Bonneel, G. Peyré, and M. Cuturi, “ Wasserstein barycentric coordinates: Histogram regression using optimal transport,” ACM Trans. Graph. (TOG), vol. 35, no. 4, Jul. 2016.

[22] S. Ferradans, N. Papadakis, J. Rabin, G. Peyré, and J.-F. Aujol, “Regularized discrete optimal transport,” in Int. Conf. on Scale Space and Variational Methods in Computer Vision. Springer Berlin Heidelberg, 2013, pp. 428–439.

[23] D. Salomon, Data Compression: The Complete Reference. Springer New York, 2004.

[24] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004.

[25] J. Preiss, F. Fernandes, and P. Urban, “Color-image quality assessment: From prediction to optimization,” IEEE Trans. Image Process., vol. 23, no. 3, pp. 1366–1378, 2014.

[26] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “Fsrm: A feature similarity index for image quality assessment,” IEEE Trans. Image Process., vol. 20, no. 8, pp. 2378–2386, 2011.

[27] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski, “Nonrigid dense correspondence with applications for image enhancement,” ACM Trans. Graph. (TOG), vol. 30, no. 4, Jul. 2011.

[28] M. Grogan and R. Dahyot, “Shape registration with directional data,” Pattern Recognition, vol. 79, pp. 452 – 466, 2018.
Fig. 9. A close up look at some of the results generated using the IDT [5], PMLS [18], GPS/LCP and FGPS/LCP [17], B-PMLS [16], L2 [8], PCT_OT [9], OT_NW [10], INWDT [11] and our algorithms using correspondences (SL2D_c and SL2D_cp) and without using correspondences (SL2D_c and SL2D_cp) - best viewed in colour and zoomed in.