

Sliced \mathcal{L}_2 Distance for Colour Grading

Hana Alghamdi
Computer Science Dpt.
Umm Al-Qura University
Mecca, Saudi Arabia
hamghamdi@uqu.edu.sa

Rozenn Dahyot
Computer Science Dpt.
Maynooth University
Maynooth, Ireland
ORCID:0000-0003-0983-3052

Abstract—We propose a new method with \mathcal{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 mapping, 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— \mathcal{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 for registering directly multidimensional Gaussian mixtures capturing target and source datasets using the \mathcal{L}_2 distance [8]. Grogan & Dahyot \mathcal{L}_2 framework [8] relies on a parametric formulation with Thin Plate Splines (TPS) of the mapping function $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that pushes source image colours toward the target image ones. Our recent research efforts have extended IDT-SWD projection based algorithms using image patches as samples in higher dimensional spaces than the RGB triplets in \mathbb{R}^3 [9], taking advantage of correspondences between source and target patches [10] and replacing OT 1D solution by the Nadaraya-Watson estimator [11]. This paper proposes to extend this projective based strategy using the more robust \mathcal{L}_2 distance [12], [13] to infer mapping functions suitable with and without correspondences [8], [14]. Using the 1D projective strategy also allows us to tackle the inherent limitation of TPS parametric formulation that does not scale well in high dimensional spaces [8].

II. THE PROPOSED METHOD

Our Sliced \mathcal{L}_2 Distance (SL2D) approach for distribution transfer is inspired by the projection-based approach utilized

This work is partly supported by a scholarship from Umm Al-Qura University, Saudi Arabia, and the ADAPT Centre for Digital Content Technology that is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

by the IDT [5] and SWD [7] algorithms that have been originally proposed to solve OT in N -dimensional spaces.

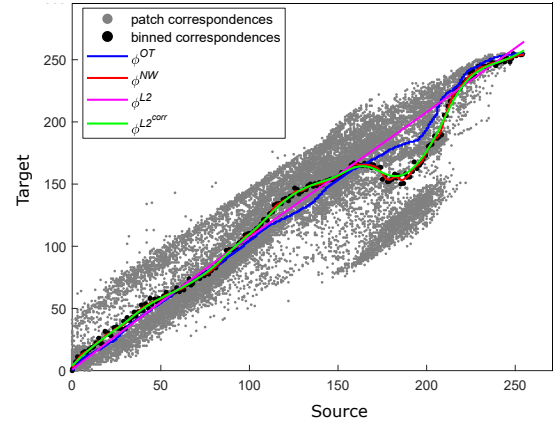


Fig. 1. Mappings $\phi^{\mathcal{L}_2^{corr}}$, $\phi^{\mathcal{L}_2}$, in comparison to ϕ^{NW} [11] and ϕ^{OT} [5].

A. \mathcal{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 $\phi^{OT}(u) = G_v^{-1} \circ F_u(u)$ (cf. Fig. 1). We propose to replace the non-parametric $\phi^{OT}(u)$ by the following parametric non-rigid 1D transformation model:

$$\phi_{\theta}^{\mathcal{L}_2}(u) = c_0 + c_1 u + \sum_{l=1}^r w_l \varphi(\|u - u_l\|_2) \quad (1)$$

$\{u_l\}_{l=1}^r$ are the control points, $\varphi(\|u - u_l\|_2) = \|u - u_l\|_2$ is the radial basis function, and $\theta = \{c_0, c_1, w_1, w_2, \dots, w_r\}$ the parameters that control the transformation estimated with:

$$\hat{\theta} = \arg \min_{\theta} \left[\mathcal{L}_2(f_{u|\theta}, g_v) = \|f_{u|\theta} - g_v\|^2 \right] \quad (2)$$

The probability density function (pdf) $f_{u|\theta}$ is a 1D Gaussian mixture model fitted to source samples $\{\phi_{\theta}^{\mathcal{L}_2}(u_i)\}_{i=1}^n$, 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 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 θ 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 $\{(u_k, v_k)\}_{k=1, \dots, K}$, the binned correspondences procedure proposed in [11] is used (c.f. Fig. 2).

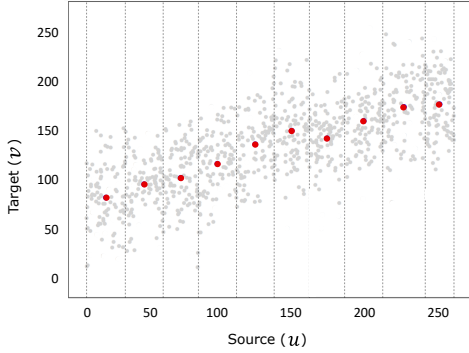


Fig. 2. An illustration of creating binned correspondences $\{(\bar{u}_i, \bar{v}_i)\}_{i=1}^K$ from $\{(u_i, v_i)\}_{i=1}^q$. The red dots representing the mean of the corresponding target values for the source observations falling in each bin.

In order to estimate the parameters of the transformation ϕ_θ in Eq. (1) that minimizes the cost function in Eq. (2), we used Quasi-Newton method [15], which is a gradient-based numerical optimization technique.

C. Convergence

The \mathcal{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 ($\phi^{\mathcal{L}_2^{corr}}$) and without correspondences ($\phi^{\mathcal{L}_2}$). As we can see from the figure, $\phi^{\mathcal{L}_2}$ at iteration 30 is not yet matching the target distribution in comparison to $\phi^{\mathcal{L}_2^{corr}}$ 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 $\phi^{\mathcal{L}_2^{corr}}$ when taking into account the correspondences, and $\phi^{\mathcal{L}_2}$ without correspondences to tackle

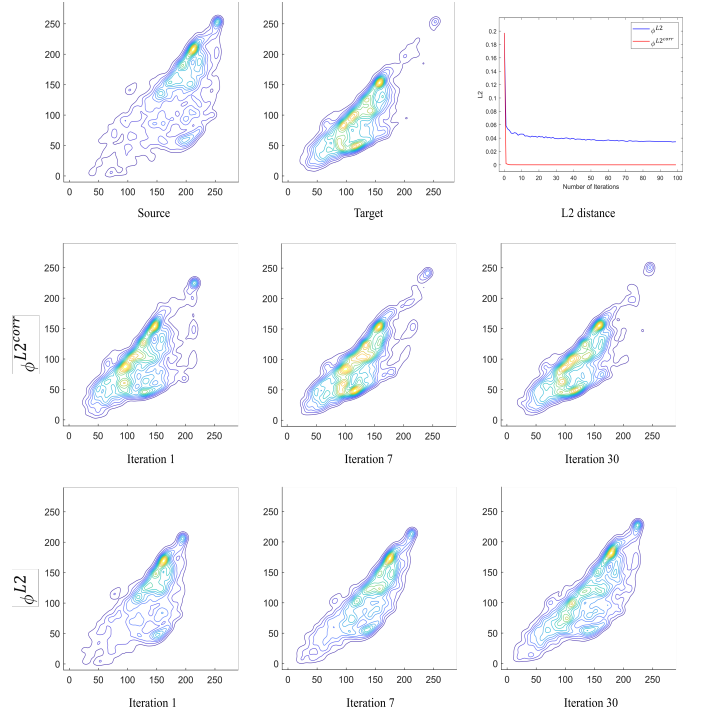


Fig. 3. Example of pdf of the transferred source patches projected in 2D space (RG). The patch size chosen is 1×1 and only the colour information is used $N = 3$ (space RGB). The standard \mathcal{L}_2 distance is computed after each iteration to illustrate the convergence of the source distribution to the target one by our transfer methods. We note that $\phi^{\mathcal{L}_2}$ at iteration 30 is not yet matching the target distribution in comparison to $\phi^{\mathcal{L}_2^{corr}}$ that is able to match the target and converge faster by iteration 7.

semi-supervised situations where correspondences are only partially available [8], [14]:

$$\forall \lambda \in [0, 1], \quad \phi_\lambda^{\mathcal{L}_2^{int}} = (1 - \lambda) \phi_{\hat{\theta}_1}^{\mathcal{L}_2^{corr}} + \lambda \phi_{\hat{\theta}_2}^{\mathcal{L}_2} \quad (3)$$

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 8).

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_c^{corr}), and the 4th with correspondences using colour patches with pixel location (SL2D_{cp}^{corr}). 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

for correspondences were also tested for comparison but are not reported as PMLS has been shown to outperform these two [8]. We have also compared our approach with [21], [22] that do not account for correspondences, however Grogan et al. [8] already reported that IDT is superior, hence IDT is the one reported here for ease of comparison.

We use the RGB colour space and we found a patch size of 3×3 captures enough of a pixel's neighbourhood. For our $SL2D_{cp}$ and $SL2D_{cp}^{corr}$ versions, each pixel is represented by its 3D RGB colour values and its 2D pixel position (i.e 5D). The patches with combined colour and spatial features create a vector in 45 dimensions ($N = 5 \times 3 \times 3 = 45$). For $SL2D_c$ and $SL2D_c^{corr}$, pixel position is not accounted for, and only RGB colours are used, which create patch vectors in 27 dimensions ($N = 3 \times 3 \times 3 = 27$).

The numerical results for PSNR, SSIM, CID and FSIMc are shown in box plots shown in Figures 4 to 7 (the means shown as red dots in the plots, and the medians shown as horizontal black lines). Sliced \mathcal{L}_2 with correspondences ($SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$) significantly outperforms the Sliced \mathcal{L}_2 solutions without correspondences ($SL2D_c$ and $SL2D_{cp}$). Incorporating colour with spatial information ($SL2D_{cp}^{corr}$) improves the performance over using colour information only ($SL2D_c^{corr}$). Moreover, the iterative projection approach with \mathcal{L}_2 (colour only $SL2D_c^{corr}$ and combined colour and position $SL2D_{cp}^{corr}$) outperform the iterative projection approach with OT solution (IDT algorithm). The medians reported for techniques B-PMLS, L2, PMLS, OT_NW, INWDT, $SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$ are not statistically different (95% confidence level) indicating equivalent quantitative performance. For qualitative performance, see Fig. 9 showing $SL2D$ performance against State of the Art. Multiple projections and 1D pdf registrations in our algorithm can be performed in parallel in the same fashion as in IDT and SWD.

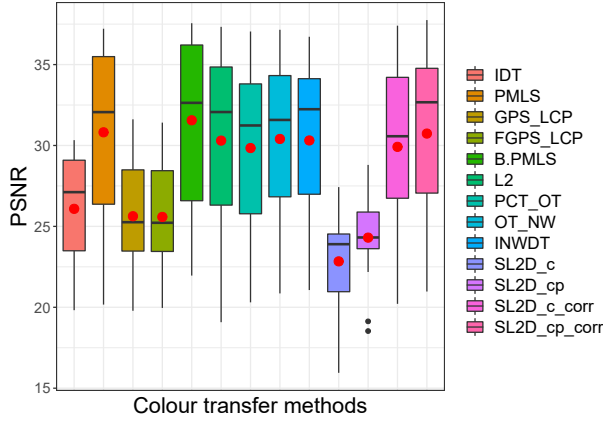


Fig. 4. Comparing our algorithms $SL2D_c$, $SL2D_{cp}$, $SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$ with the state of the art colour transfer methods using PSNR metric [23] (higher values are better, best viewed in colour and zoomed in).

IV. CONCLUSION

We have introduced the Sliced \mathcal{L}_2 distance for registering two high dimensional probability density functions, that allows

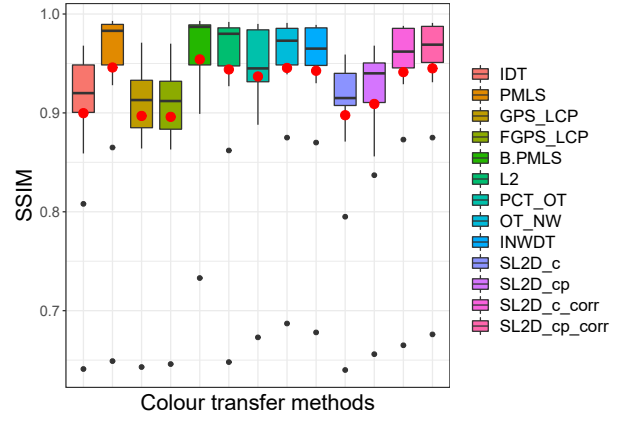


Fig. 5. Comparing $SL2D_c$, $SL2D_{cp}$, $SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$ with the state of the art colour transfer methods using SSIM metric [24] (higher values are better, best viewed in colour and zoomed in).

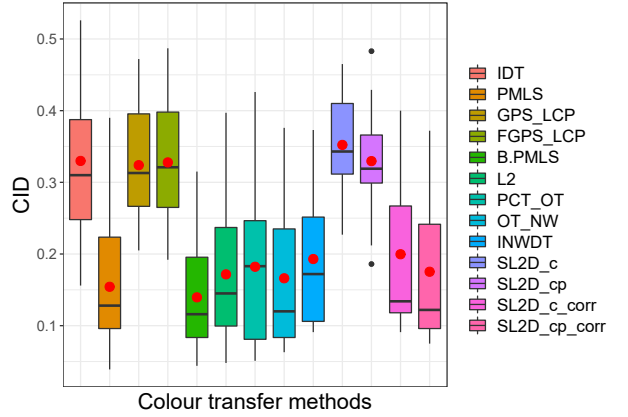


Fig. 6. Comparing $SL2D_c$, $SL2D_{cp}$, $SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$ with the state of the art colour transfer methods using CID metric [25] (lower values are better, best viewed in colour and zoomed in).

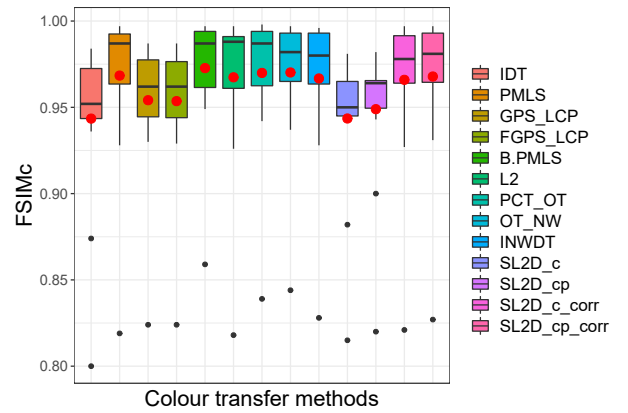


Fig. 7. Comparing $SL2D_c$, $SL2D_{cp}$, $SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$ with the state of the art colour transfer methods using FSIMc metric [26] (higher values are better, best viewed in colour and zoomed in).

correspondences to be taken into account when these are available. Our $SL2D$ technique applied to colour transfer extends

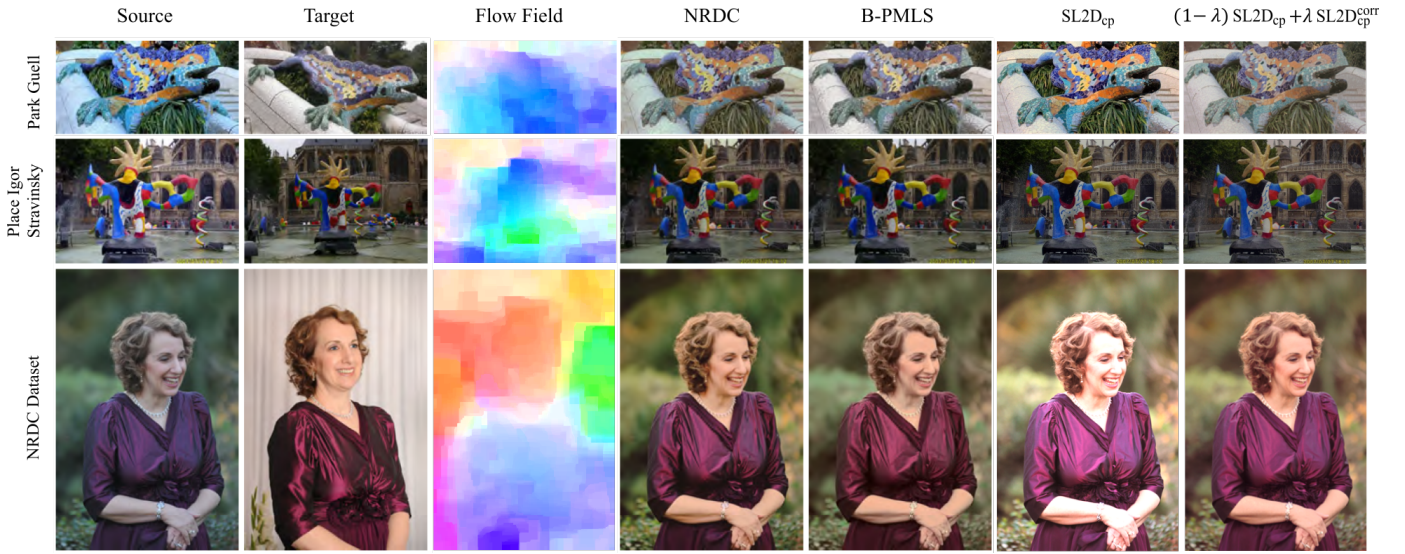


Fig. 8. Colour transfer results of interpolating between our solutions $SL2D_{cp}^{corr}$ and $SL2D_{cp}$ with comparisons to the colour transfer results in [16], [27].

Grogan et al. \mathcal{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. 6816–6826.
- [3] C. Meng, Y. Ke, J. Zhang, M. Zhang, W. Zhong, and P. Ma, “Large-scale 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 sift 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 recolouring 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, “Fsim: 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, “Non-rigid 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.

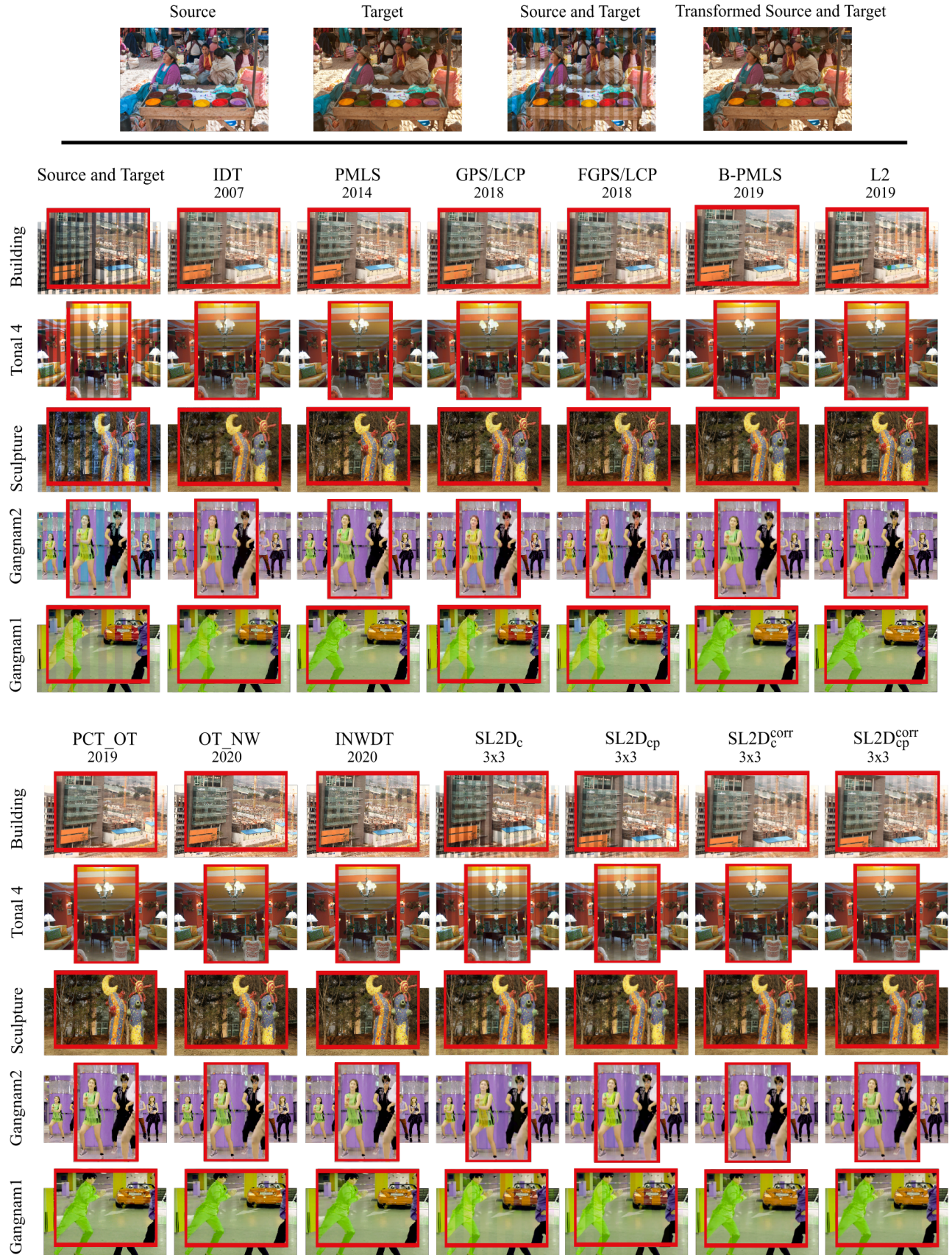


Fig. 9. A close up look at some of the results generated using the IDT [5], PMLS [18], GPS/LCP and FGPS/LCP [17], B-PMLS [16], L2 [8], PCT_OT [9], OT_NW [10], INWDT [11] and our algorithms using correspondences (SL2D_c^{corr} and SL2D_{cp}^{corr}) and without using correspondences (SL2D_c and SL2D_{cp}) - best viewed in colour and zoomed in.