Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Probability Density Function Transfer in High Dimensional Spaces and its Application to Colour Transfer

Hana Alghamdi

This thesis is submitted to the University of Dublin, Trinity College
in fulfillment of the requirements for the degree of Doctor of Philosophy in
Computer Science

School of Computer Science and Statistics

Supervisor: Prof. Rozenn Dahyot

January 2021

Declaration

I declare that this thesis has not previously been submitted as an exercise for a degree at this, or any other University, and it is my own work.

_____

Hana Alghamdi

November 11, 2020

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This thesis addresses the topic of example-based colour transfer from the domain of image processing. Colour transfer is often recast as a distribution transfer problem in which the actual probability density function of a given target colour samples is transferred to the input colour samples. Existing colour transformations rely on solving the problem only in colour space and do not scale well to high-dimensional spaces due to the increase in the computational complexity associated with the increase of the data dimensionality. This poses a challenging problem to current colour transfer methods and limits the potential of encoding and fusing different types of relevant information that could enhance image descriptors and guide the transfer process.

The aim of this thesis is to propose transfer functions that are scalable to high-dimensional spaces and suitable to parallel computation. We extend the colour problem space to high-dimensional space by constructing pixel descriptors that encode colour, spatial and local structure information to guide and enhance the performance of the transfer function. We propose to solve the high-dimensional distribution transfer problem in 1D space using an iterative projection approach with three statistical methods: Optimal Transport, $\mathcal{L}_2$ inference, and Kernel Regression, taking into account available information about correspondences between the input and the target distributions. Our high-dimensional construction implies a new reconstruction step since each recoloured pixel benefits from the contribution of several estimated candidates using an averaging solution that allows denoising and artifact removal as well as colour transfer. The extensive experiments and analysis conducted in this thesis show quantitative and qualitative competitive results compared with the leading state of the art methods of colour transfer.

ACKNOWLEDGEMENTS

At the end of this pleasant journey, my gratitude goes first to my wonderful family, parents and sisters (we are three sisters finishing our PhD's at the same time); without their love, support and encouragement, I would barely finish this journey. Therefore, I proudly dedicate this thesis and every success in my life to them.

Second, with love I express my gratitude and appreciation to my supervisor Prof. Rozenn Dahyot for her stimulating discussions, advice, and encouragement she gave me throughout the course of this research. I have always been impressed by her deep understanding and her ability to fluently demonstrate interdisciplinary insight throughout the research journey. Finally, I would like to offer my sincere gratitude to the wonderful people and friends I met in the Graphics, Vision and Visualization (GV2) group.

HANA ALGHAMDI

University of Dublin, Trinity College
November 2020

ABBREVIATIONS

B-PMLS   Bilateral weights for the Probabilistic Moving Least Squares

CDF   Cumulative Distribution Functions

CID   Colour Image Difference

CNN   Convolution Neural Network

EM   Expectation-Maximization

EMD   Earth Mover's Distance

FDA   Functional Data Analysis

FGPS/LCP   Fast coarse-to-fine Gradient Preserving Spline/Linear Colour Propagation

FSIMc   Feature SIMilarity Index with Chrominance

GMM   Gaussian Mixtures Models

GPS/LCP   Gradient Preserving Spline/Linear Colour Propagation

IDT   Iterative Distribution Transfer

INWDT   Iterative Nadaraya-Watson Distribution Transfer

KDE   Kernel Density Estimate

KL   Kullback-Leibler

LF   Light Field

MGD   Multivariate Gaussian Distribution

NW   Nadaraya-Watson

NW_c	Nadaraya-Watson solution using Colour patches only
NW_cp	Nadaraya-Watson solution using Colour patches with Pixel location information
OT	Optimal Transport
OT_NW	smoothed Optimal Transport using Nadaraya-Watson
PCA	Principal Component Analysis
PCT_OT	Patch-based Colour Transfer with Optimal Transport
PDF	Probability Density Function
PMLS	Probabilistic Moving Least Squares
PSNR	Peak Signal to Noise Ratio
RBF	Radial Basis Functions
RGB	Red Green Blue colour space
RNN	Recurrent Neural Networks
SAI	Sub-Aperture Images
SE	Standard Error
SIFT	Scale Invariant Feature Transform
SL2D	Sliced $\mathcal{L}_2$ Distance
SL2D$_c$	Sliced $\mathcal{L}_2$ Distance solution using Colour patches only
SL2D$_{cp}$	Sliced $\mathcal{L}_2$ Distance solution using Colour patches with Pixel location information
SL2D$_c^{corr}$	Sliced $\mathcal{L}_2$ Distance solution with correspondences using Colour patches only
SL2D$_{cp}^{corr}$	Sliced $\mathcal{L}_2$ Distance solution with correspondences using Colour patches with Pixel location information
SWD	Sliced Wasserstein Distance
SSIM	Structural SIMilarity Index
TPS	Thin Plate Spline

INTRODUCTION

1.1 Context and Motivation

Computer vision, as an interdisciplinary scientific field, is concerned with building computers that can automate three main tasks the human vision system can do: image acquisition, image processing, and image analysis. The goal is to enable computers to reach a high-level understanding of digital images that empower them with the ability of decision making in such contexts as prediction and control, Figure 1.1.

In image acquiring task, a digital image may be acquired by one or several image sensors such as light-sensitive cameras, range sensors, tomography devices, radar, ultra-sonic cameras, etc. These devices produce different types of images: 2D image, 3D volume, or an image sequence (video). The pixel values typically correspond to light intensity in one spectral band (gray image) or several spectral bands (colour images), and can also be related to various physical measures, such as depth measurements and electromagnetic waves. Additionally, images are often captured at different device settings, different resolutions, and under different lighting conditions. These
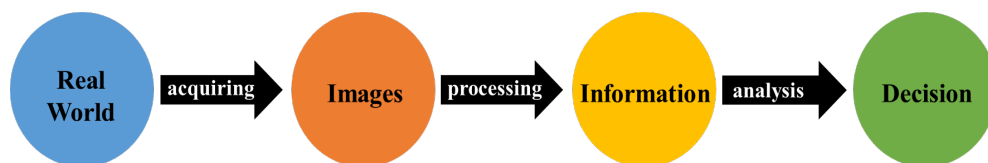
Figure 1.1: The purpose of Computer Vision is to mimic human vision in three tasks: image acquisition, image processing, and image analysis. The goal is to enable computers to reach a high-level understanding of digital images that empower them with the ability of decision making.
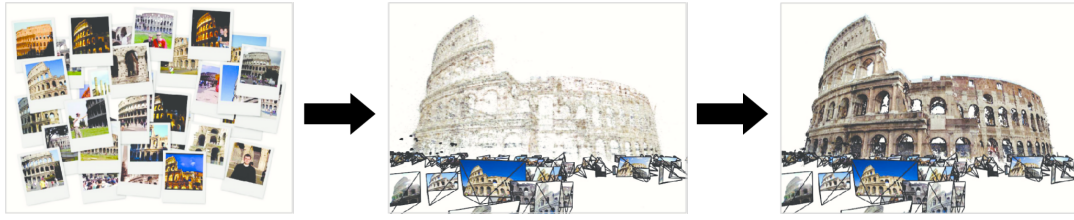
Figure 1.2: Reconstruction of the Colosseum from extremely diverse, large, and unconstrained photo collections (Agarwal et al., 2010).

different conditions make images largely heterogeneous.

Many applications involve high-level image analysis and integration of complementary and valuable information from multiple imaging resources to represent them in a single coherent informative image. An example is presented in Figure 1.2, where a diverse and large number of images distributed on the Internet that exhibit overlapped scenes are used for 3D reconstruction of buildings. Two problems arise in this integration: the geometric and the photometric alignment. The geometric alignment is commonly referred to as image registration, which is a fundamental task in image processing used to geometrically overlay two or more images of the same scene, possibly taken at different times, from different angles or different devices. In contrast, the photometric alignment refers to colour transfer, and it is the process of matching colour statistics or eliminate colour variations between images. The colour variations between images often happen due to illumination changes, using different cameras, different in-camera settings, or tonal adjustments of the users, and this poses a problem to the fusion of information from several images. Moreover, the alignment that is given by an image registration algorithm (geometric alignment) can never be accurate to the pixel level and it may contain mismatches and errors, which makes a direct colour matching of pixel to pixel unfeasible solution. Therefore, the problem of how to match the colours of images must be tackled differently. The works undertaken in this thesis aim to contribute to the domain of example-based colour transfer (photometric alignment). In the following, we briefly discuss the research problems.

1.2   Research Problems and Scope

Example-based colour transfer refers to the process of changing the colours of an input image (denoted as the source image) so that they match the colours of an example

image (denoted also as the target image). In many computer vision applications such as image mosaicking or stitching where there is a shared content, images are required to have consistent colours. In these cases, colour transfer methods can be used for harmonizing and eliminating the colour variations between images. Figure 1.3 shows examples of images with overlapped scenes that need to be harmonized. Colour transfer techniques can also be used to transfer the colour feel of one image to another when there is no shared content between them, as in Figure 1.4. This type of colour transfer problem is a fundamental process in film post-production. In this thesis, we focus on colour transfer between realistic images that exhibit overlapped scenes.

A popular approach is to formulate colour transfer as a problem of transforming statistical distributions, where the actual probability density function (pdf) of the target colour samples is transferred to the source colour samples. This problem is known as the distribution transfer problem, and it is the main interest of this thesis. Recent algorithms based on the idea of minimizing the distance between probability density functions proved to be efficient tools for a wide range of applications. Of particular interest are the methods based on the Optimal Transport cost function (Muzellec and Cuturi, 2019; Meng et al., 2019; Kolouri et al., 2018; Arjovsky et al., 2017) and $\mathcal{L}_2$ divergence cost function (Grogan and Dahyot, 2019, 2018; Ma et al., 2015; Jian and Vemuri, 2011). However, Optimal Transport has not been investigated in high dimensional spaces, and the $\mathcal{L}_2$ divergence based methods do not scale well in high dimensional spaces due to the increase in computational complexity associated with the increase of the data dimensionality. Indeed, enhancing image descriptors by fusing different information such as spatial information and colour information as well as directional information (orientations of image gradients) is interesting for the improvement of the quality of results, but that would increase the dimensionality of the data. This poses a challenging problem to current colour transfer methods and limits the potential of utilizing different data types in estimating the transformation functions. This implies the need for transfer functions that are scalable to multi-dimensional spaces and reduce computational costs.

In this perspective, to propose a new example-based colour transfer method, several questions need to be addressed:
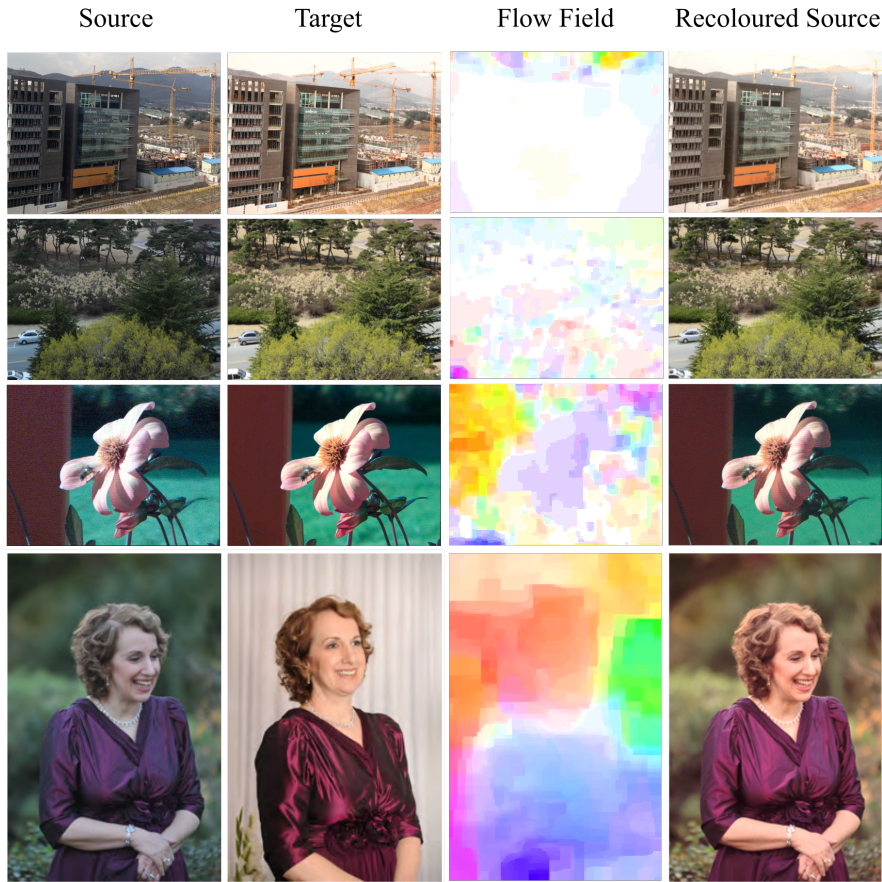
Figure 1.3: This figure shows examples of source and target images with overlapped contents. Colour transfer technique is applied with the goal of harmonizing and eliminating the colour variations between images. Flow fields show motion changes between the source and target images. In the last column, the recoloured source images in the 1st, 2nd, and 3rd rows are processed with our proposed colour transfer techniques in the same order (Alghamdi and Dahyot, 2020b; Alghamdi et al., 2019; Alghamdi and Dahyot, 2020a), the 4th row is processed with our proposed technique in Chapter 7.

- How to estimate a transformation function that matches high-dimensional probability density functions and is computationally efficient.

- Is it possible to solve the colour transfer problem using high dimensional representations that encode spatial information and colour information and preserve local structures.

- How to estimate a transformation function that is accurate enough to match the colours of the target image.

- How to estimate a transformation function that preserves most of the structures of the source image and allows denoising and artifact removal.
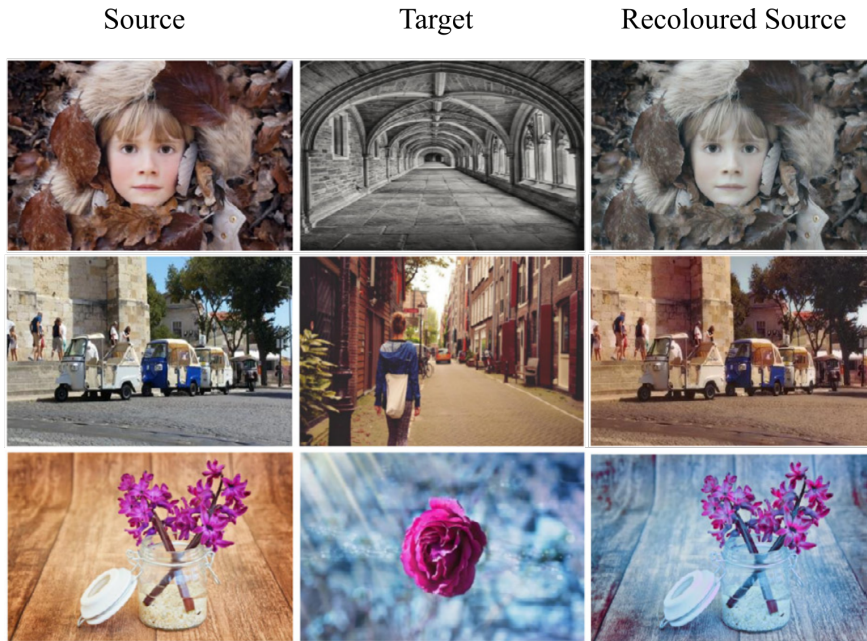
Figure 1.4: This figure shows examples of source and target images with different contents. Colour transfer technique is applied with the goal of transferring the colour feel of the target image to the source image. In the last column, the recoloured source images are processed with the `IDT` colour transfer technique (Pitie et al., 2007).

- How to account for pixel correspondences in matching high-dimensional probability density functions.

To address these questions, we extend the original problem space to high-dimensional spaces, encoding a lot of information that could guide the estimation of the transformation function. We investigate solving the high-dimensional problem using the 1D iterative projection approach with three statistical models: Optimal Transport, $\mathcal{L}_2$ inference, and Nadaraya–Watson estimator that accounts for pixel correspondences. This approach has the potential to adapt well in higher dimensions, and it is also suitable for parallel optimization.

1.3 Thesis Contributions

In 1D space, optimal transportation, inference with $\mathcal{L}_2$ divergence, and computing Nadaraya–Watson estimates are straightforward and have low computational cost. In this thesis, we utilize these efficient 1D transformation functions to tackle high dimensional distribution transfer problems using the 1D iterative projection approach.

The idea is to iteratively project dataset values originally in high dimensional space to a 1D subspaces and solve 1D transformation functions and then propagate the solution back to the original space. This operation is repeated with different directions in 1D space until convergence. The 1D transformation problems are independent of each other; hence the process is suitable for parallel optimization. We show the potentials of the proposed approaches in solving colour transfer between two images that exhibit overlapped scenes. We conduct experiments that show quantitative and qualitative improvements over the previous state of the art colour transfer methods. We now summarize the main contributions in this thesis:

- We compare and visualize how Optimal Transport, inference with $\mathcal{L}_2$ divergence, and Nadaraya–Watson behave in the 1D case. We conducted simulated experiments to compare popular approaches proposed in the literature that tackle the high-dimensional problems using 1D Optimal Transport solution.

- We propose a new colour transfer, the patch-based colour transfer with Optimal Transport, to transfer the colour of a source image to match the colour of a target image of the same scene. In order to preserve local topology and transfer coherent colours to neighboring pixels, we encode overlapping neighborhoods of pixels (patches) in high-dimensional spaces. In addition, spatial information as well as colour content of pixels are both encoded in high-dimensional feature vectors. Both overlapping patches and combining spatial and colour information improve the pixel descriptors in the 1D dimensional subspaces and this compensates for the loss of the structure of the original dimensional space. This original construction implies a new reconstruction step since each recoloured pixel benefits from the contribution of several estimated candidates using an averaging solution that allows denoising and artifact removal. Experiments show quantitative and qualitative improvements over the state of the art colour transfer in images where spaces of dimensions up to 100 have been used.

- We propose a new colour transfer method with Optimal Transport to transfer the colour of a source image to match the colour of a target image of the same scene that may exhibit large motion changes between images. By definition, Optimal Transport does not take into account any available information about correspondences when computing the optimal solution. To tackle this problem, we

propose to estimate motions flow using motion estimation and incorporate the spatial correspondences in the encoded overlapping neighborhoods of pixels. This new formulation makes Optimal Transport implicitly takes into account correspondences when computing the optimal solution. Moreover, we introduce smoothing as part of the iterative algorithms for solving optimal transport, namely Iterative Distribution Transport (IDT) and its variant, the Sliced Wasserstein Distance (SWD). We show quantitative and qualitative improvements over the state of the art colour transfer in images where spaces of dimensions up to 125 have been used.

- We propose Iterative Nadaraya-Watson Distribution Transfer for accounting for pixel correspondences in matching high-dimensional probability density functions. We demonstrate the potentials of this mapping in solving colour transfer between two images that exhibit overlapped scenes with motion changes between them. The 2D/3D problem is extended to higher dimensions by encoding overlapping neighborhoods of data points and solving the high dimensional problem in 1D space using the iterative projection approach. Nadaraya-Watson provides smooth mapping and alleviates artifacts. Quantitative and qualitative experiments show that our approach is competitive with the state of the art colour transfer in images where spaces of dimensions up to 45 have been used.

- We introduce the Sliced $\mathcal{L}_2$ Distance based approach that minimizes the $\mathcal{L}_2$ distance between two pdfs using the 1D projection approach. This approach extends the standard $\mathcal{L}_2$ inference to high-dimensional optimization problems. The $\mathcal{L}_2$ based inference is flexible as it takes advantage of correspondences that may be available and performs well in the presence of correspondence outliers. We show how to combine $\mathcal{L}_2$ solutions with correspondences and without correspondences to tackle semi-supervised situations where correspondences are only partially available.

1.4 Thesis Outline

The works undertaken in this thesis is structured into eight chapters:

- In Chapter 2, we provide the theory and applications of Optimal Transport, $\mathcal{L}_2$ inference, Nadaraya-Watson estimator, and we present the state of the art colour transfer techniques.

- In Chapter 3, we compare and visualize the models' behavior in the 1D case. Also, We compare the iterative projection approaches that tackle the high-dimensional problems using a 1D OT solution.

- In Chapter 4, we present our colour transfer approach with Optimal Transport, where we encode overlapping neighborhoods of pixels containing colour information as well as spatial information.

- In Chapter 5, we present colour transfer approach with Optimal Transport, where we incorporate spatial correspondences estimated using motion estimation. Also, we introduce smoothing as part of the iterative algorithms for solving optimal transport.

- In Chapter 6, we present colour transfer approach with Nadaraya-Watson that maps one N-dimensional distribution to another using the 1D projection-based approach, taking into account available information about correspondences.

- In Chapter 7, we extend $\mathcal{L}_2$ framework to higher dimensional problems and introduce the Sliced $\mathcal{L}_2$ Distance approach that minimizes the robust $\mathcal{L}_2$ between two N-dimensional distributions using the 1D projection-based approach.

- In Chapter 8, we summarise the work carried out in this thesis and presents possible future directions of investigation.

- Finally, Appendices provide a number of supplementary results generated by the methods proposed in this thesis.

1.5 List of Publications

A number of paper publications took place during the preparation of this thesis:

- (Under Review) Alghamdi, H., Dahyot, R., 2021. **Sliced $\mathcal{L}_2$ Distance for Colour Grading**, in: Proceedings of 29th European Signal Processing Conference (EUSIPCO).

- (Published) Alghamdi, H., Dahyot, R., 2020. **Iterative Nadaraya-Watson Distribution Transfer for Colour Grading**, in: Proceedings of the IEEE 22th International Workshop on Multimedia Signal Processing (MMSP).

- (Published) Alghamdi, H., Dahyot, R., 2020. **Patch-based Colour Transfer using SIFT Flow**, in: Proceedings of the 22th Irish Machine Vision and Image Processing Conference (IMVIP). *Won Best full-paper Award at the conference.*

- (Published) Alghamdi, H., Grogan, M., Dahyot, R., 2019. **Patch-based Colour Transfer with Optimal Transport**, in: Proceedings of 27th European Signal Processing Conference (EUSIPCO).

- (Published) Dahyot, R., Alghamdi, H., Grogan, M., 2019. **Entropic Regularisation of Robust Optimal Transport**, in: Proceedings of the 21th Irish Machine Vision and Image Processing Conference (IMVIP).

- (Published) Alghamdi, H., Dahyot, R., 2017. **IDT vs $\mathcal{L}_2$ Distance for Point Set Registration**, in: Proceedings of the 19th Irish Machine Vision and Image Processing Conference (IMVIP).

LITERATURE REVIEW: THEORY AND APPLICATIONS

2.1 Introduction

The main contributions of this thesis are to introduce high dimensional distribution transfer functions with Optimal Transport, $\mathcal{L}_2$ divergence inference, and a Kernel regression technique (Nadaraya-Watson estimator), which have been inspired by the 1D iterative approach implemented in algorithms proposed initially as solutions to Optimal Transport in N-dimensional spaces. The proposed transfer functions are guided by high dimensional point representations encoding content information (such as colour and spatial information and local structures) to produce accurate mappings. To show these mappings' potentials, we apply them to colour transfer between two images that exhibit overlapped scenes. Therefore, in this chapter, we present the background structured in a way that serves the subsequent chapters. First, we will summarize the theory and applications of Optimal Transport, $\mathcal{L}_2$ inference, and Nadaraya-Watson estimator in Section 2.2. Following this, we present a brief review of colour transfer techniques in Section 2.3. Finally, we conclude in Section 2.4.

2.2 Statistical Techniques and Applications

2.2.1 Optimal Transport

Optimal Transport (OT), also known as the Wasserstein Distance (Villani, 2003, 2008) or as the Earth Mover's Distance (EMD) (Rubner et al., 2000), is a framework concerned with the problem of finding an optimal solution with a minimum cost of transferring a source distribution to a target distribution. It was first formulated by the civil engineer Gaspar Monge (Monge, 1781) for the purpose of finding an optimal way to move a pile of soil from one site to another with minimum transportation cost. Later the problem was extensively analyzed by Kantorovich (Kantorovich, 1942) with a focus on developing techniques for the optimal economic allocation of resources. Since then, OT has appeared in numerous fields like econometrics, fluid dynamics, automatic control, transportation, and production planning as they all involve moving a group of items or distributions from a source to a target destination (Villani, 2003; Rachev and Rüschendorf, 1998).

In the last two decades, OT received significant attention from researchers in Computer Science, imaging, and data analysis. The rationale for such interest is the powerful geometric tools which provide effective means for defining distances between probability measures that may be defined in high-dimensional metric spaces. Optimal transport based techniques have been shown to produce powerful and robust results in numerous pattern recognition related problems (Kolouri et al., 2017).

In the following, we introduce an overview of key mathematical concepts and features of OT, and how it has been used as a distance metric between probability distributions, including the possible closed-form solutions in one-dimensional space, as well as how OT can be computed in N-dimensional space. Lastly, we summarize interesting applications of OT in image processing, computer vision, and learning problems. For more detailed arguments and analysis, the interested reader is referred to books that have been written on the topic by Villani (2008, 2003) for more comprehensive mathematical discussion, and by Santambrogio (2015) for a discussion targeted at applied mathematics, and more recently, a book by Peyré and Cuturi (2019) which looks at OT from a computational viewpoint centered on applications in computer vision and machine learning.
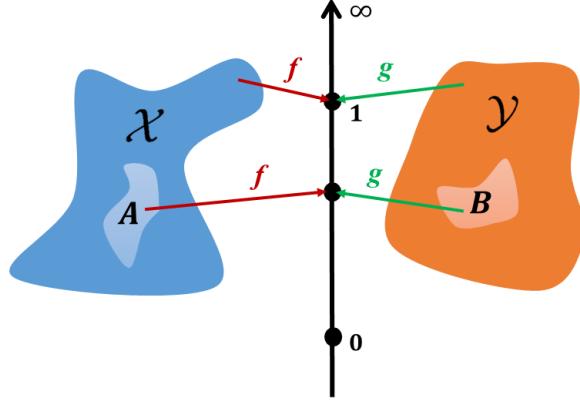
Figure 2.1: Given $(\mathcal{X}, f)$ and $(\mathcal{Y}, g)$, metric spaces paired with probabilities measures which have the same total mass=1. A mapping $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ is a measure preserving if for any measurable set $B \subset \mathcal{Y}$, $f(A) = g(B)$ such that $A = \phi^{-1}(B)$.

2.2.1.1 Theoretical foundations

Estimate a warping map ϕ. Given general metric spaces paired with probability measures $(\mathcal{X}, f)$ and $(\mathcal{Y}, g)$ that represent the source and target respectively, where the balance of mass condition (Eq. (2.1)) is satisfied, i.e., the source and target have the same total mass, generally chosen to be one (Kolouri et al., 2017):

$$\int_{\mathcal{X}} f(x)dx = \int_{\mathcal{Y}} g(y)dy = 1 \tag{2.1}$$

A mapping $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ is measure preserving if for any measurable set $B \subset \mathcal{Y}$, a local version of the condition (2.1) holds, Figure 2.1:

$$\int_{A} f(x)dx = \int_{B} g(y)dy \tag{2.2a}$$

$$\text{such that } A = \phi^{-1}(B) \tag{2.2b}$$

Monge's formulation of the optimal transport problem defines the "optimal" measure preserving map $\phi(x) = y$ that transports the source distribution into the target distri-
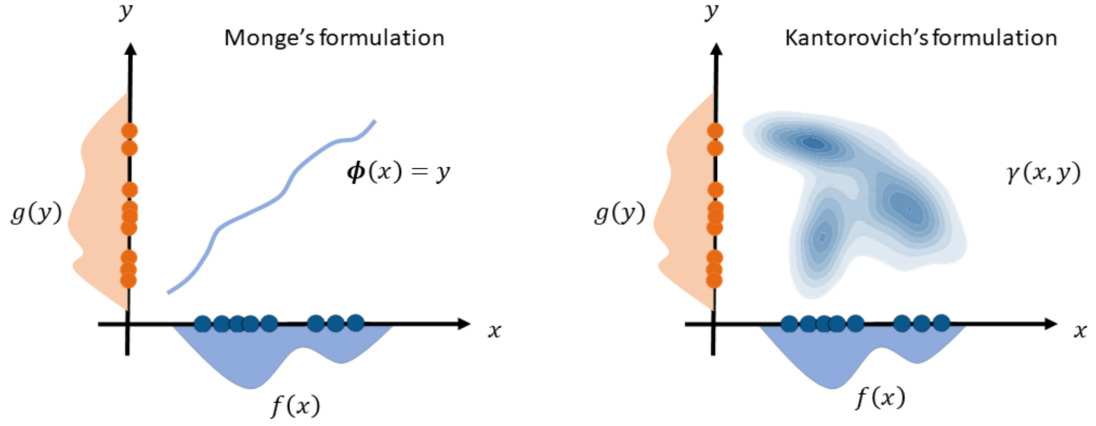
Figure 2.2: Monge's formulation seeks to find a deterministic mapping ϕ with a minimum total cost that sends the mass from a single source location x to a single target location y. Kantorovich's formulation seeks to find a joint distribution with a minimum total cost that splits the mass of a single source location and sends it to multiple target locations y.

bution with minimal total transportation cost, and it is defined as follows:

$$\underset{\phi}{\mathrm{argmin}} \quad \int_{\mathcal{X}} c(x, \phi(x)) \ f(x) dx \tag{2.3a}$$

$$\text{such that} \quad \phi \ \text{satisfies condition} \ (2.2) \tag{2.3b}$$

where $c(x, \phi(x)) = \|x - \phi(x)\|^2$ is the Euclidean distance. The cost of moving a unit of mass from source point x to target point y is induced by the product of the source point's mass $f(x)$ by the distance between the source point x and the target point y to which the source is transported. Condition (2.3b) implies that $f(x \in A) = g(y \in B)$ such that $A = \phi^{-1}(B)$. Note that Monge's map is a deterministic coupling, a mass from a single location x is being sent to a single location y, i.e., it is a one-to-one mapping, Figure 2.2.

Estimate a joint distribution γ. Kantorovich optimal transport formulation seeks to estimate a transport "plan" γ instead of a "map" ϕ that matches f into g with a minimum modification of the source as possible (Kantorovich, 1942). The amount of modification done is interpreted as the total cost and induced by the product of the transformed source mass associated with a point x by the distance between that source

point and the target point which is assigned to. Kantorovich's formulation is defined
as follows:

$$\operatorname*{argmin}_{\gamma \in \Pi(f,g)} \int_{\mathcal{X} \times \mathcal{Y}} c(x,y)\, \gamma(x,y) dx dy \quad \text{[Total cost]} \tag{2.4a}$$

$$\text{such that}: \gamma(x,y) \geq 0 \quad \text{[Positive mass]} \tag{2.4b}$$

$$\gamma(A \times \mathcal{Y}) = f(A) \quad \text{[Start at } f] \tag{2.4c}$$

$$\gamma(\mathcal{X} \times B) = g(B) \quad \text{[end at } g] \tag{2.4d}$$

For all measurable sets $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$. The set of all possible measure couplings
between f and g is denoted $\Pi(f,g)$. Kantorovich's formulation, unlike Monge's for-
mulation, has the ability to distribute the mass from one location in source distribution
to multiple locations in target distribution, i.e., splits the mass instead of sending it to
a single location, Figure 2.2.

2.2.1.2 Geometric properties: Optimal Transport as a distance metric

Optimal transport can be viewed as a tool for comparing distributions in a way that
measures the similarity between f and g. For example, the (Eq. (2.4a)) returns the
transport plan γ that gives "minimum" total transportation cost of transforming f
into g. If the minimum total transportation cost is small, then the amount of mass
needed to be moved from each source x will not travel so far to match its destination
y, i.e., they are close to each other. If the minimum transportation cost is large, then
that indicates that f and g are dissimilar, and the amount of the effort to match them
is large. The quadratic Wasserstein distance is defined as follows:

$$\mathcal{W}(f,g) = \min_{\gamma \in \Pi(f,g)} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\|^2 \, \gamma(x,y) dx dy \tag{2.5}$$

$\mathcal{W}(f,g)$ is the cost of the optimal plan that transfers the total mass of f to g, where
$\|x - y\|^2$ is the cost of transforming the mass from x to y. The quadratic Wasserstein
distance satisfies the distance properties: non-negativity, symmetry, and triangle in-
equality (see Villani (2008, 2003) for more detailed proof and a generalized version of
Wasserstein distance). In the case of a deterministic coupling $y = \phi(x)$, the quadratic
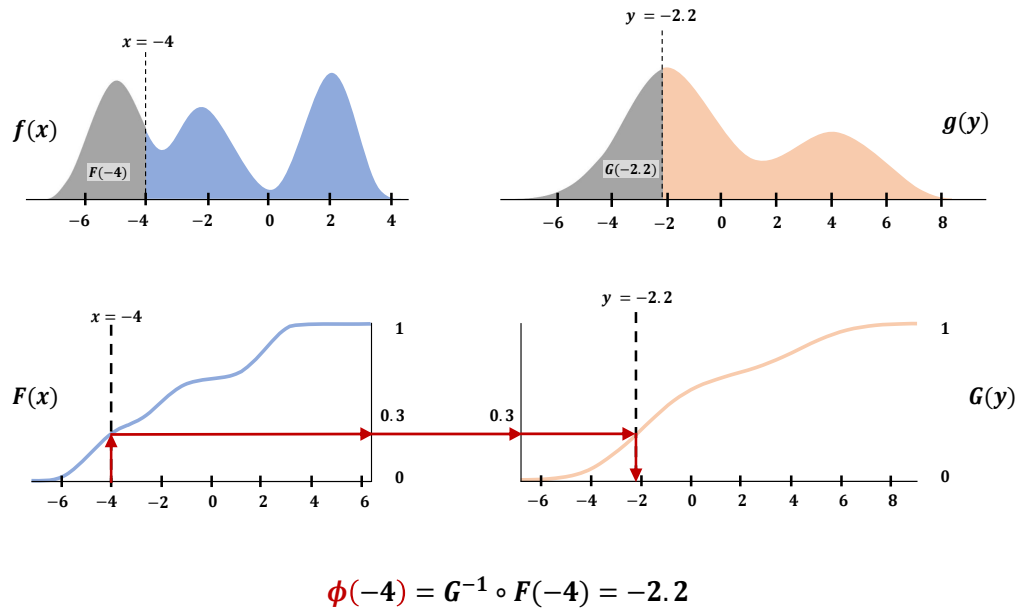
Figure 2.3: Illustration of CDF-matching approach. f and g are the source and target pdfs respectively. G and F are the corresponding cumulative functions. ϕ maps the source and target pdfs using CDF-matching approach.

Wasserstein distance is defined as follows:

$$\mathcal{W}(f, g) = \min_{\phi} \int_{\mathcal{X}} \|x - \phi(x)\|^2 \, f(x)dx \tag{2.6}$$

2.2.1.3 Closed-form solution for ϕ in one-dimensional space

In one-dimensional space (1D), there are closed-form solutions for the optimal transport map ϕ; we will refer to these solutions as the CDF-matching approach and the quantile-matching approach. The simplicity and explicitness with which a solution can be obtained in that setting have been used to tackle problems in higher dimensions in different contexts as it alleviates the need for optimization, as we will see in Section 2.2.1.4. In the following we explain these solutions:

CDF-matching approach. Suppose that f and g are probability measures on $\mathbb{R}$ for 1D random variables x and y respectively, and their cumulative distribution functions (CDF) are defined as follows:
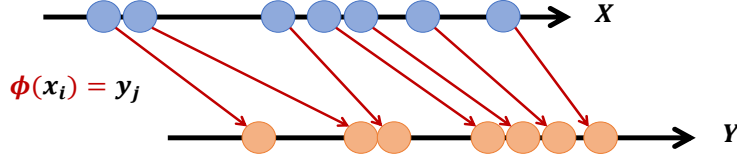
Figure 2.4: Illustration of Quantile-matching approach. x and y are random variables represent the source and target respectively. ϕ maps the two random variables using Quantile-matching approach.

$$F(x) = \int_{\infty}^{x} f(x) \ dx, \qquad G(y) = \int_{\infty}^{y} g(y) \ dy \qquad (2.7)$$

And the inverse of G is defined as:

$$G^{-1}(t) = \inf \{ \ y \in \mathbb{R}; \ \ G(y) \geqslant t\} \qquad (2.8)$$

Then the optimal map ϕ is set as:

$$y = \phi(x) = G^{-1} \circ F(x) \qquad (2.9)$$

According to the above, the data point x is sent onto y (i.e. $y = \phi(x)$), if and only if $F(x) = G(y)$, Figure 2.3. This method is referred to as the increasing rearrangement by Villani (2008). $\phi(x) = y$ defines an optimal transport map for Monge's problem that transports the mass at minimal cost in such a way that associates to each point x a single point y, i.e, no mass splitting.

Quantile-matching approach. This approach does not employ cumulative distribution; instead, it is based on matching two sorted arrays, Figure 2.4. More specifically, it sorts the n observations $\{x_i\}$ and $\{y_j\}$ of the source and target distributions respectively to define quantiles with regular increments of size $\frac{1}{n}$ between 0 and 1 for both distributions, and then the optimal solution ϕ is a direct assignment between the sorted datasets. This approach was first introduced by Rabin et al. (2012), where the authors formulated the problem as a permutation problem. Let σ_X and σ_Y be the permutations that order the points of the source and the target data sets respectively:
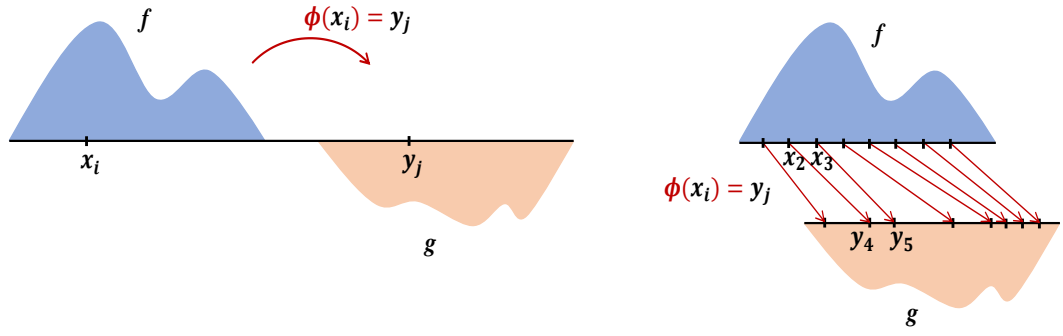
Figure 2.5: Optimal transport ϕ maps the source f into g. The optimal transport map has the property of being monotone in the sense that if $x_1 \leqslant x_2$ then $\phi(x_1) \leqslant \phi(x_2)$.

$$\forall i, \ x_{\sigma_X(i)} \leqslant x_{\sigma_X(i+1)} \tag{2.10a}$$

$$\text{and} \quad \forall j, \ y_{\sigma_Y(j)} \leqslant y_{\sigma_Y(j+1)} \tag{2.10b}$$

The optimal permutation σ is computed as follows:

$$\sigma(i) = \sigma_Y \circ \sigma_X^{-1}(i) \tag{2.11}$$

And then the deterministic coupling ϕ can be rewritten as follows:

$$\phi(x_i) = y_{\sigma(i)} \tag{2.12}$$

Both approaches CDF-matching and quantile-matching provide monotone rearrangements that preserve the ordering of the points after transformation (Figure 2.5) such that:

$$\text{if} \ \ x_1 \leqslant x_2 \ \ \text{then} \ \ \phi(x_1) \leqslant \phi(x_2) \tag{2.13}$$

Quantile-matching and CDF-matching are approaches used for histogram specification and histogram equalization that have many applications in image processing (Morovic et al., 2002; Rolland et al., 2000). Histogram specification is a process where a given source time series or an image is modified such that its histogram matches

that of target time series or an image, and hence the source cumulative density function (CDF) matches that of a target CDF. Histogram equalization (normalization) is a special case, where the target histogram is a uniform random distribution (the CDF is a straight line). The effect of histogram equalization is to spread the data values over the range of data.

However, in terms of the computational complexity, the CDF-matching can be computed in $\mathcal{O}(n)$ where n is the number of samples processed (Pitie et al., 2005b). Quantile-matching can be computed in $\mathcal{O}(n \log(n))$ operations using a fast sorting algorithm (Rabin et al., 2012). When a small number of observations are available, using quantile-matching is best, but with a very large number of observations, CDF-matching is more efficient.

2.2.1.4   Computational Optimal Transport

The high computational cost induced by solving for the optimal transport due to the huge amount of data that modern applications deal with raised the need for developing fast algorithms for optimal transport. In the following, we give a brief overview of available computational solutions for optimal transport.

Linear programming.   The optimal transport problem (Eq. (2.4a)) is a basic formulation of a linear program which consists of minimizing a cost function that is linear with the unknown variables (a value of γ for each pair (x, y)) with linear inequality and equality constraints. Linear programming is a classical problem in operations research, and there exist algorithms to solve linear programs, including the Simplex algorithm developed by Dantzig (1990), the Auction algorithm developed by Bertsekas (1988), and the Hungarian algorithm (Kuhn, 1955). These algorithms can be used to solve transportation problems for small or medium sized problems but can suffer from scaling issues, adding that the computational complexity with linear program solvers increases in multidimensional spaces (Santambrogio, 2015; Villani, 2003).

Entropic regularization.   Motivated by the large scale transport problems that using Linear programming to induce solutions is expensive in terms of computational complexity, a regularized version of optimal transport is proposed, which means that the problem is slightly modified in order to make it easier and more tractable. The

optimal transport in Eq. (2.4a) is modified as follows:

$$\underset{\gamma \in \Pi(f,g)}{\operatorname{argmin}} \left[\int_{\mathcal{X} \times \mathcal{Y}} c(x,y)\, \gamma(x,y) dx\, dy \; - \; \lambda\, H(\gamma(x,y)) \right] \tag{2.14}$$

$H(\gamma(x,y))$ computes the entropy of the transport plan γ and it is defined as:

$$H(\gamma(x,y)) = \int_{\mathcal{X} \times \mathcal{Y}} -\gamma(x,y)\, log\, \gamma(x,y)\, dx\, dy \tag{2.15}$$

When λ=0, we recover the original optimal transport Eq. (2.4a). When λ increases, the optimization objective favors minimizing the second term, which equals minus the entropy of the transport plan γ and hence large values of λ increase the entropy of the transport plan. Adding the regularization term leads to a fast algorithm by introducing a solution of the transport plan γ in the following form:

$$\gamma(x,y) = u(x)e^{(\frac{-c(x,y)}{\lambda})}v(y) \tag{2.16}$$

In the original optimal transport Eq. (2.4a), solving for $\gamma(x,y)$ involves quadratic number of variables to be optimized (one value of γ for each (x,y) pair), but with the regularized Eq. (2.14) we have linear number of variables ($v(x)$ and $u(y)$ for every x and y) to solve for. The Sinkhorn algorithm is developed to estimate u and v jointly. With the regularized formulation of optimal transport and the Sinkhorn algorithm, an approximation of the solution to transport problems can be obtained without the need to use complex and specialized algorithms. We refer the interested reader to Peyré and Cuturi (2019) for more detailed treatment.

Efficient approximate mass transport solvers. The Iterative Distribution Transfer algorithm (IDT) proposed by Pitie et al. (2007) is considered an approximate mass transport solver. IDT uses the 1D Optimal Transport solution Eq. (2.9) to tackle problems in multidimensional spaces. More specifically, IDT proposed to iteratively project the observations $\{x_i\}$ and $\{y_j\}$ originally in $\mathbb{R}^N$ to a 1D subspace and solve the 1D OT using Eq. (2.9) in this 1D subspace and then propagate the solution back to the original space $\mathbb{R}^N$. This operation is repeated with different directions in 1D space until convergence. This strategy was inspired by the idea of the Radon Transform (Pitie

et al., 2007), which states the following proposition: if the target and source distributions are aligned in all possible 1D projective spaces, then matching is also achieved in $\mathbb{R}^N$ space. IDT can be considered as a form of quantile matching but with irregular quantile increments derived from the cumulative histograms of the source and target distributions - as source and target quantiles do not match exactly, interpolation can be used to compute the solution (Pitie et al., 2007).

Another approximate mass transport solver called the Sliced Wasserstein Distance (SWD) algorithm, which follows from the iterative projection approach of IDT but computes the 1D OT solution with quantile matching (Eq. (2.11)) instead of cumulative histogram matching. Moreover, SWD uses a stochastic gradient descent method to derive the solution (Bonneel et al., 2015; Rabin et al., 2012).

2.2.1.5 Optimal Transport applications

Nowadays, many modern computer vision, statistical, and machine learning problems can be recast as finding the optimal transport map (or plan) between two probability distributions. One of the earliest applications of the optimal transport problem was in image retrieval. Rubner et al. (2000) employed the Wasserstein Distance (also denoted as the Earth Mover's Distance) to measure the dissimilarity between image histograms. They compared the Wasserstein metric with common metrics such as Jeffrey's divergence, χ^2 statistics, $\mathcal{L}_1$ distance, and $\mathcal{L}_2$ distance in an image retrieval task; and it was shown that the Wasserstein metric achieves the highest performance amongst all. A recent popular application in machine learning is domain adaptation, which is designed to compensate for the difference between training and test data distribution. Optimal Transport based techniques have been utilized in designing a stable domain adaptation algorithms (Redko et al., 2019; Muzellec and Cuturi, 2019; Courty et al., 2016; Courty et al., 2014). Other popular applications nowadays are deep generative models (Goodfellow et al., 2014), which map an input distribution, such as Gaussian or uniform distribution, to the underlying target distribution of a real dataset. Wasserstein distance has been utilized as a metric between the target distribution and the generator distribution to enhance and improve the stability of the learned model (Meng et al., 2019; Tanaka, 2019; Sanjabi et al., 2018; Arjovsky et al., 2017). In statistics, Wasserstein Distances provide a powerful tool to carry out statistical tasks such as: provide goodness-of-fit testing (Panaretos and Zemel, 2019),

and as statistical inference on a distribution registration model for general deformation functions (Del Barrio et al., 2019), or as a tool for statistical analysis of datasets whose elements are random histograms by computing Principal Component Analysis of histograms with respect to the Wasserstein distance between probability density functions (Cazelles et al., 2018). Chen et al. (2018) proposed optimal mass transport framework to compare, interpolate and average Gaussian mixture models which are widely used in statistical inference. In Computer Vision, there are many applications, for example, shape deformation (Peyré and Cuturi, 2019; Bonneel et al., 2016), shape matching and comparison (Su et al., 2015), and colour transfer applications (Bonneel and Coeurjolly, 2019; Shu et al., 2017; Bonneel et al., 2016, 2015; Rabin et al., 2014; Ferradans et al., 2013; Rabin and Peyré, 2011; Freedman and Kisilev, 2010). More details about colour transfer application with optimal transportation are presented in Section 2.3. A more detailed review of Optimal Transport applications in signal processing and machine learning can be found in (Kolouri et al., 2017).

2.2.2 Inference with $\mathcal{L}_2$

Another approach for solving the distribution transfer problem is estimating a transformation function ϕ that minimizes the divergence between two probability density functions. In the shape registration context, several probabilistic methods recast the shapes matching problem as a matching process between probability density functions. In this case, the source and target datasets are modeled as probability density functions, and then the source pdf is transformed to match the target pdf by minimizing a distance metric between them. Each density matching method differs in how to model these pdfs and what distance metric is used. As for modeling pdfs, Kernel Density Estimate (KDE) is a popular non-parametric method used to estimate a density function in density matching methods. As for distance metrics, several measures have been proposed to compute the similarity between two probability density functions, including $\mathcal{L}_2$ divergence (Jian and Vemuri, 2011), Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951), Cauchy-Schwartz divergence (Hasanbelliu et al., 2011), Shanon Divergence (Wang et al., 2006), Bregman divergence (Liu et al., 2010) and Havrda-Charvt divergence (Chen et al., 2010). Of these divergences, $\mathcal{L}_2$ has been shown to be a robust metric (Jones et al., 2001; Basu et al., 1998). Besides robustness, $\mathcal{L}_2$ has another attractive feature of having a closed-form solution when density func-

tions are modeled as Gaussian Mixtures Models (GMM). Of particular interest is the successful work of Jian and Vemuri (2011), who proposed to model the datasets using GMMs and used an optimization process that minimizes $\mathcal{L}_2$ distance between the pdfs to estimate the parameters of the transformation function ϕ that matches the source and target shapes (Jian and Vemuri, 2011). The GMM can be obtained by applying a kernel density estimate with a Gaussian kernel. Jian and Vemuri's work inspired successive works in shape registration and colour transfer applications (see Section 2.2.2.2).

In the following, we introduce an overview of key mathematical concepts of $\mathcal{L}_2$ divergence between probability density functions and parameter inference with $\mathcal{L}_2$, and then we summarize interesting applications of inference with $\mathcal{L}_2$ in computer vision. For more detailed arguments and analysis, the interested reader is referred to (Jian and Vemuri, 2011; Scott, 2001), and to recent theses (Grogan, 2017; L. Arellano Vidal, 2014) for generalized $\mathcal{L}_2$ inference in colour transfer and shape registration, respectively.

2.2.2.1 Theoretical foundations

Unlike the OT model which relies on two random variables x and y, only one random variable x appears explicitly in $\mathcal{L}_2$ model, and both the source pdf f and target pdf g are candidate pdfs for the random variable x. A parametric model proposed by this approach to solving distribution transfer is $f(x|\theta)$ for the source distribution such that the parameter θ controls the mapping function ϕ. The $\mathcal{L}_2$ is the Euclidean distance between the density functions f and g, and can be defined as follows:

$$\mathcal{L}_2(f(x), g(x)) = \int f^2(x)dx - 2\int f(x)g(x)dx + \int g^2(x)dx \tag{2.17}$$

The parameter θ that controls the transformation ϕ that transforms the pdf f into g is estimated by minimizing $\mathcal{L}_2$ distance between the the moving source distribution $f(x|\theta)$ and the target $g(x)$ as follows:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left[\mathcal{L}_2(f(x|\theta), g(x)) = \int f^2(x|\theta)dx - 2\int f(x|\theta)g(x)dx \right] \tag{2.18}$$

Since the target is fixed and independent of the transformation estimation, the term $\int g(x)^2 dx$ is removed from the cost function (2.18). The transfer function ϕ can be

conveniently defined as a rigid, affine or non-rigid parametric function such as Thin Plate Splines (TPS) (Grogan and Dahyot, 2019, 2018; Jian and Vemuri, 2011).

Let us assume that the two probability density functions f and g representing source and target samples are Gaussian mixtures as follows:

$$f(x) = \sum_{i=1}^{n} w_i \mathcal{N}(x; \mu_i, \Sigma_i) \tag{2.19}$$

and

$$g(x) = \sum_{j=1}^{m} w_j \mathcal{N}(x; \mu_j, \Sigma_j) \tag{2.20}$$

where $\mathcal{N}(x; \mu_*, \Sigma_*)$ is the normal probability density function of mean μ_* and covariance Σ_*, and $0 \leq w_* \leq 1$ is its weight in the mixture such that $\sum_{*=1}^{**} w_* = 1$ (with $* = i, j$ and $** = n, m$). The $\mathcal{L}_2$ distance can be computed explicitly as there exists a closed-form expression for the $\mathcal{L}_2$ distance between Gaussian mixtures due to the following formula:

$$\int \mathcal{N}(x; \mu_1, \Sigma_1) \mathcal{N}(x; \mu_2, \Sigma_2) = \mathcal{N}(0; \mu_1 - \mu_2, \Sigma_1 + \Sigma_2) \tag{2.21}$$

2.2.2.2 $\mathcal{L}_2$ inference applications

As mentioned earlier, Jian and Vemuri (2011) proposed to register two point sets by fitting Gaussian mixture models to each dataset and minimizing the $\mathcal{L}_2$ distance between them to estimate ϕ, where isotropic covariance are used for the GMMs. A generalized $\mathcal{L}_2$ approach is used by Wang et al. (2009) for groupwise point-sets registration. Ma et al. (2015) also proposed a similar method which uses the approximation $\mathcal{L}_2 E$ to estimate ϕ. Roy et al. (2007) proposed to use Expectation-Maximization algorithm (EM) to fit the GMMs enforcing identical isotropic covariance matrices and uniform occupancy probabilities and uses the $\mathcal{L}_2$ distance as part of the optimization process. Arellano and Dahyot (2016) extended Jian and Vemuri's framework by adding gradient directional information to the model, capturing the normals' angles of the shapes and showed that it gives improved results for ellipse detection, where they modeled the bandwidths of the density functions as non-isotropic to reduce the number of com-
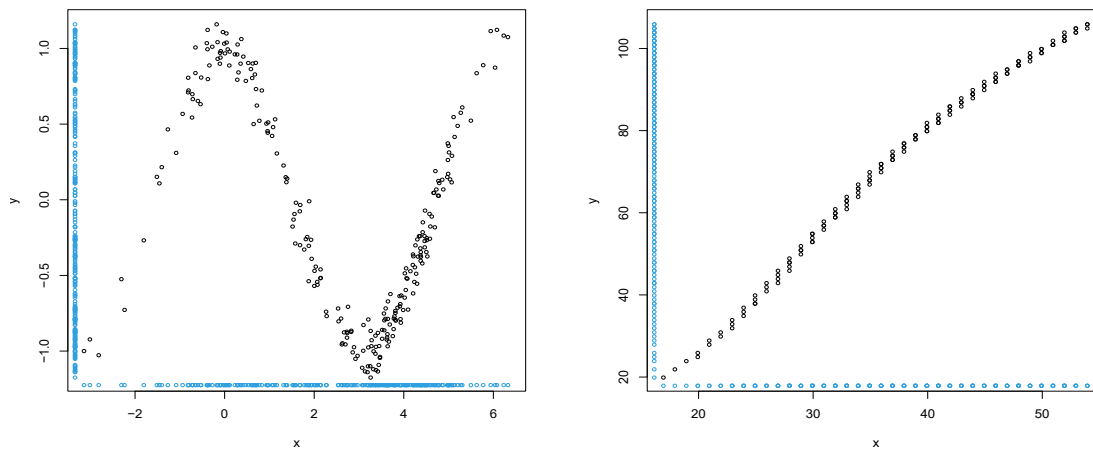
Figure 2.6: Simulated data show examples of functional data. The left panel shows discrete observations reflecting a smooth variation appears to be nicely sinusoidal, and on the right panel there is a tendency for the data to show exponential growth.

ponents needed in the Gaussian Mixture. Grogan and Dahyot (2018) proposed to use von Mises-Fisher kernels to model the normal vectors of the shape, adding them as an extra dimension to the model previously proposed by Jian and Vemuri (2011).

Inspired by the success of Jian and Vemuri's $\mathcal{L}_2$ framework for shape registration context, Grogan et al. (2015) proposed a similar approach in colour transfer applications, which models the colour distributions of the source and target images using GMMs, and match them by minimizing the robust $\mathcal{L}_2$ distance between the mixtures. The parametric transformation has been shown that can be also applied to video sequences (Grogan and Dahyot, 2015). Moreover, the $\mathcal{L}_2E$ framework has been extended further to let the users engage in the colour transfer process by choosing correspondences and refining the recolouring (Grogan et al., 2017). The subsequent extension by Grogan and Dahyot (2019) showed that $\mathcal{L}_2$ framework takes advantage of colour correspondences that may be available and performs well in the occurrence of correspondence outliers. $\mathcal{L}_2$ framework has been shown to be useful for many colour transfer applications such as correcting light fields' colour discrepancies (Matysiak et al., 2020; Grogan and Smolic, 2019).

2.2.3   Functional Data Analysis: Nadaraya-Watson

Functional data analysis (FDA) refers to a collection of methods for analyzing data over a curve, surface, or continuum. "Functional" data is usually data observed as a
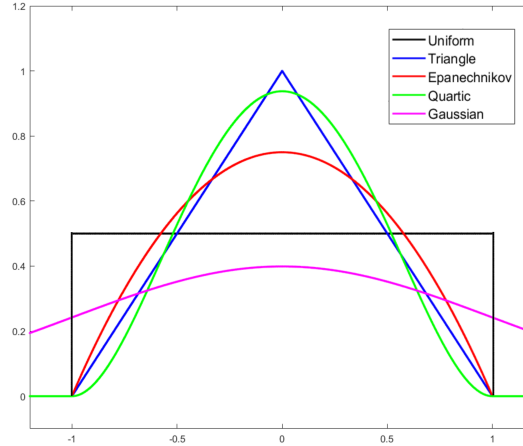
Figure 2.7: Several types of kernel functions are commonly used: uniform, triangle, epanechnikov, quartic, and gaussian.

sequence of discrete pairs (correspondences) $\{(x_i, y_i)\}_{i=1}^n$ over a continuum. The continuum can take different forms of continua such as time, spatial position, frequency, weight etc. These discrete observations reflect a smooth variation and are assumed that they are observations of some underlying function ϕ that maps x_i to y_i. An important assumption in FDA is that the underlying function ϕ has to be smooth, in the sense that two samples x_i and x_{i+1} that are similar should still be mapped to loosely similar observations y_i and y_{i+1}, i.e., the function ϕ generally is well-behaved and not too "wiggly" or locally variable.

The goal of Functional data analysis is to develop methods to turn the raw observations $\{(x_i, y_i)\}$ into smooth functions, i.e to estimate the corresponding smooth functions that are presumed to generate them. Because the raw observation may exhibit some noise, functional data analysis considers combining noise model with smooth functional representation in order to produce the observed data, and this can be expressed in the following model given the dataset of discrete observations $\{(x_i, y_i)\}$:

$$y_i = \phi(x_i) + \epsilon_i \tag{2.22}$$

Here the standard model for the error assumes that the errors ϵ_i are independently and identically distributed with mean zero and constant variance σ^2. Figure 2.6 shows examples of functional data: on the left panel, the data show a pattern that appears to be nicely sinusoidal, and on the right panel there is a tendency for the data to show

Kernel function	Equation
Gaussian	$W(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$
Triangular	$W(x) = (1 - |x|)$ support: $|x| \leqslant 1$
Epanechnikov	$W(x) = \frac{3}{4}(1 - x^2)$ support: $|x| \leqslant 1$
Quartic	$W(x) = \frac{15}{16}(1 - x^2)^2$ support: $|x| \leqslant 1$
Uniform	$W(x) = \frac{1}{2}$ support: $|x| \leqslant 1$

Table 2.1: This table shows the most popular kernel functions that can be used with the Nadaraya-Watson estimator.

overall exponential growth. These data require a smooth curve fitting, and one of the classical methods functional data analysis provides for this purpose is the non-parametric Nadaraya Watson estimator (NW).

2.2.3.1 Theoretical foundations

Nadaraya Watson estimator is considered to be part of a family of non-parametric Kernel regression techniques, and it was invented independently by Nadaraya (Nadaraya, 1965) and Watson (Watson, 1964). Non-parametric means that the given observations are used directly to compute the function, while parametric techniques assume that the observed data follows a particular parametric probability distribution and the parameters of the distribution can be estimated via fitting the distribution using the observed data. The underlying function can be estimated by sampling data from the fitted distribution.

Nadaraya Watson makes use of local weights such that the estimated value of the function at a point x_i must be influenced mostly by the observations near x_i. Given the correspondences $\{(x_i, y_i)\}_{i=1}^n$ the estimator is defined as follows:

$$\phi_h(x) = \mathbb{E}[y|x] = \int y\, p(y|x) dy \tag{2.23a}$$

$$\simeq \frac{n^{-1} \sum_{i=1}^n y_i\, W_h(x - x_i)}{n^{-1} \sum_{i=1}^n W_h(x - x_i)} \tag{2.23b}$$
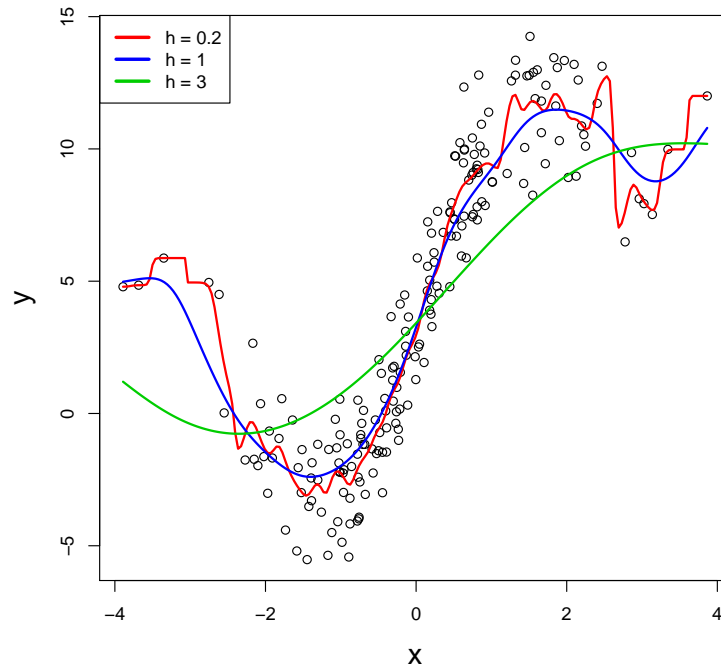
Figure 2.8: Simulated data shows an example of functional data where the discrete observations reflect a smooth variation that appears to be nicely sinusoidal. The curves show Nadaraya Watson ϕ_h solutions with different bandwidth values $h = (0.2, 1, 3)$. The selection of the smoothing parameter affects how well NW approximates the underlying function.

With this form, NW can be seen as a locally weighted average of $\{y_i\}_{i=1}^n$, using a kernel as a weighting function that defines the shape of weights of averaged values, where the bandwidth h is the hyperparameter or scale parameter of the kernel that defines the range of averaged values, the larger the value of h the more ϕ_h gets smoothed.

Table 2.1 shows popular kernel functions. As we can see, all kernels define a range for input values to be averaged, except Gaussian kernel which takes into account all neighbouring points when computing the estimation of x_i. In practice, the selection of kernel function generally influences less than the selection of the smoothing parameter h, Figure 2.8 shows an example of the effect of the value of h. Different methods proposed in the literature for estimating h can be found in (Shao and Tu, 2012; Turlach, 1993; Fan and Gijbels, 1992; Terrell and Scott, 1992; Gasser et al., 1991; Terrell, 1990).

2.3 Colour Transfer

Colour transfer techniques based on the source image only (Cohen-Or et al., 2006; Lischinski et al., 2006) often require time-consuming manual adjustments. To reduce the efforts involved in parameterization and subjective judgment, example (also denoted as reference or target) based colour transfer have been introduced. The problem can be reduced to transferring the colour of the example image to the source image by estimating a transfer function ϕ that used to recolour a source colour pixel x to $\phi(x)$.

There are many colour transfer applications in image processing problems. In general, a mosaic is created if a combination of two or more images of the same scene is needed for comparison or integration purposes. In these cases, correspondences are available and colour transfer methods can be used for harmonizing and eliminating the colour variations between the images. This process of colour transfer is also known as photometrical alignment or colour correction. Such applications range from generating colour consistent images in image mosaicing and stitching applications (Li et al., 2020; Brown and Lowe, 2007), city 3D modeling (Micusik and Kosecka, 2009), stereo images (Wang et al., 2010), to colour enhancement and colour style manipulation (Hwang et al., 2019).

Research works in the field of example based colour transfer can be classified into two categories: global methods and local methods. Global colour transfer methods calculate the colour statistics by taking into account all pixels in the images, and local colour transfer methods are guided by content information derived by considering the spatial relationships via segmentation and clustering or using colour correspondences. In this section, we give an overview of the example based colour transfer techniques explored in each area. An extended, detailed reviews about colour transfer techniques and approaches can be found in (Faridul et al., 2014; Pitie et al., 2008).

2.3.1 Global Transfer Methods

Many methods use global statistical properties to describe the colour distribution of the source and target images and propose ways to reshape the colour distribution of the source image so that it matches that of the target image. The notion of transfer statistical properties can refer to the simple process of matching the mean, variance (or

covariance) or to matching the full colour distribution. The following subsections will focus on the problem of finding a colour mapping that transfers the colour statistics of a target image to the source image and presents a brief overview of the techniques that have been explored in this context. Starting in Section 2.3.1.1 with the methods that only consider a linear transformation in which the colour distributions are modeled using the multivariate Gaussian distribution and only transfer the second statistical moments of the distribution, i.e. the mean and the covariance, and then presenting the non-linear transformation methods in Section 2.3.1.2 that can match any type of distribution.

2.3.1.1 Linear Transfer

For simplicity, many colour transfer methods assume that the probability density functions of the colours in the source and target images, denoted f and g respectively, follow a simple multivariate Gaussian distribution (MGD) model ($f \equiv \mathcal{N}(x; \mu_x, \Sigma_x)$ and $g \equiv \mathcal{N}(y; \mu_y, \Sigma_y)$). These methods consider the problem of finding a linear mapping with parametric form as follows (Faridul et al., 2014; Pitie et al., 2008):

$$\phi(x) = Ax + b \tag{2.24}$$

where A is NxN matrix and b is an offset vector. Since the Gaussian distribution is fully specified by its first two statistical moments, the mean and covariance matrix, the linear transform Eq. (2.24) is computed by matching the empirical estimates of (μ_x, Σ_x) and (μ_y, Σ_y) for source and target colour distributions as follows (Pitie et al., 2008):

$$\phi(x) = A(x - \mu_x) + \mu_y \tag{2.25a}$$

$$\text{with } A^T \Sigma_y^{-1} A = \Sigma_x^{-1} \tag{2.25b}$$

There exist several solutions proposed in the literature for the matrix A that lead to different ways of transferring the colour statistics. The first solution uses an independent transfer (Reinhard et al., 2001) that assumes multivariate distributions with independent variables, thus the covariance matrices are diagonal. In this case, the mean and

the variance of each space component (here the colour channels) are matched independently. This assumption of independence is rarely true for real images, and it is only applicable in decorrelated colour spaces such as $l\alpha\beta$ colour space (Ruderman et al., 1998). The success of these techniques depends on the chosen target images because $l\alpha\beta$ colour space cannot be guaranteed to successfully decorrelate the colour space components. To account for inter-component dependencies, colour distributions are assumed to follow MGD with non-diagonal covariance matrices, and the solution in this case for the matrix A can be found using Cholesky Decomposition (Pitie et al., 2008; Pitie and Kokaram, 2007). The disadvantage of this method is that the derived solution is dependent on the ordering of the axes components (colour channels). Another possible solution can be found using Principal Component Analysis (Abadpour and Kasaei, 2007) that shows improvements in the resulting image over Cholesky Decomposition (see evaluations by Pitie et al. (2008)). Pitie and Kokaram (2007) proposed the Linear Monge-Kantorovitch solution for multivariate Gaussian distributions, which is more reliable than Cholesky Decomposition and gives results that are slightly better than the ones of the Principal Axes method (see evaluations by Pitie et al. (2008)).

However, these linear colour transfer techniques are limited to simple affine transformations of the colour space; they perform well when a single Gaussian can well describe the colour distribution of the source and target images, but poorly perform when the colour distributions are more complex and hence fail to resynthesize the colour scheme of the target image. For detailed treatments, comparisons, advantages and disadvantages of each linear technique, the interested reader is referred to a recent comprehensive review by Pitie et al. (2008).

2.3.1.2 Non-linear Transfer

Multivariate Gaussian distributions are too simple and can not describe the complex characteristics of general distributions. Alternatively, many algorithms proposed to approximate colour distributions using histograms and propose non-parametric transfer functions. The following sections present techniques explored in the colour transfer literature that can match any colour distribution and are organized based on how the solution is approached.

Independent Transfer. Similarly to the linear transfer case, multivariate distributions with independent variables are assumed here, and histograms are used to approximate the full distribution of values in each coordinate independently, and then histogram matching is used separately for each coordinate (Grundland and Dodgson, 2005). In this case, the joint (multivariate) distribution is the product of its marginals:

$$f(x_1, x_2, ..., x_N) = f(x_1)f(x_2)...f(x_N) \tag{2.26}$$

Hence, the mapping of a sample $x \in \mathbb{R}^N$ is composed for each component independently as follows:

$$\Phi(x) = \Phi(x_1, x_2, ..., x_N) = (\phi_1(x_1), \phi_2(x_2), ..., \phi_d(x_N)) \tag{2.27}$$

Each 1D mapping $\phi_1, \phi_2, ..., \phi_N$ is found by using the corresponding pdf match:

$$
\begin{aligned}
\phi_1(x_1) &: f(x_1) \Rightarrow g(y_1) \\
\phi_2(x_2) &: f(x_2) \Rightarrow g(y_2) \\
&\vdots \\
\phi_N(x_N) &: f(x_N) \Rightarrow g(y_N)
\end{aligned}
\tag{2.28}
$$

This technique is applicable to decorrelated colour space like $l\alpha\beta$ colour space (Ruderman et al., 1998) where the joint distributions become separable. However, this full histogram transfer tends to be too harsh and may affect the appearance of the source image by producing artifacts. Therefore, recent works resolve this problem by matching histograms at different scales (Pouli and Reinhard, 2011) and by constraining the histogram matching to preserve the gradient distribution of the source image (Xiao and Ma, 2009).

Knothe Rosenblatt Rearrangement. The Knothe-Rosenblatt rearrangement was independently proposed by (Rosenblatt, 1952), and by (Knothe, 1957). This method has been applied to the colour transfer problem by Neumann and Neumann (2005). Here we assume we have dependent multivariate joint distribution and hence it is

---

defined as follows:

$$f(x_1, x_2, ..., x_N) = f(x_1)f(x_2|x_1)...f(x_N|x_1, .., x_{N-1}) \tag{2.29}$$

Each conditional marginal $f(*)$ can be treated independently from the others, and 1D histogram matching is used separately for each conditional marginal as follows (Villani, 2008):

$$\Phi(x) = \Phi(x_1, x_2, ..., x_N) = (\phi_1(x_1), \phi_2(x_2|x_1), ..., \phi_N(x_N|x_1, .., x_{N-1})) \tag{2.30}$$

Each 1D mapping $\phi_1, \phi_2, ..., \phi_N$ is found by matching the source conditional marginal with the corresponding conditional marginal in the target distribution:

$$
\begin{aligned}
\phi_1(x_1) &: f(x_1) \Rightarrow g(y_1) \\
\phi_2(x_2|x_1) &: f(x_2|x_1) \Rightarrow g(y_2|y_1) \\
&\vdots \\
\phi_N(x_N|x_1, .., x_{N-1}) &: f(x_N|x_1, .., x_{N-1}) => g(y_N|y_1, .., y_{N-1})
\end{aligned}
\tag{2.31}
$$

This method is dependent on the order of axes coordinates in which the variables are conditioned on each other which produces different results (Villani, 2008). Moreover, only a very limited number of colour samples are used to estimate the conditional marginals, as only a few pixels have exactly the same colour, which may lead to poor results (detailed evaluations can be found in (Pitie et al., 2008)).

1D Iterative Projection based Transfer. Here the entire target colour distribution is transferred to the source colour samples. The iterative approach explained in Section 2.2.1.4 is used here for colour distribution transfer, which is based on the iterative use of 1D CDF-matching (Pitie et al., 2005a,b) and its variant 1D quantile-matching (Bonneel et al., 2015) for various random projection directions in the N-dimensional space. Similar to the Knothe Rosenblatt rearrangement based transfer, the estimation process is based on the use of 1D pdfs. The difference is that the pdfs are not conditional probabilities. This implies that the method is not affected by the ordering of the axes

coordinates (colour channels). Also, the estimation of the marginals does not suffer from the data sparseness. However, when the content of the target and source images differs, grain artifacts may appear and the noise level may be amplified in the resulting image. The subsequent extension by Pitie et al. (2007) proposed a post-processing step to eliminate the artifacts and ensure that the gradient field of the resulting image matches the gradient field of the source image. Similarly, Bonneel et al. (2015) proposed using a gradient smoothing technique introduced in (Rabin et al., 2011) to reduce any artifacts that may appear in their results. Recent works used this approach in solving colour transfer problems such as (Shu et al., 2017; Bonneel and Coeurjolly, 2019).

Discrete Monge-Kantorovitch Optimal Transportation. Many researchers investigated the discrete Monge-Kantorovitch formulation in solving colour transfer problem (Rabin et al., 2014; Ferradans et al., 2013; Rabin and Peyré, 2011; Freedman and Kisilev, 2010) who proposed to relax the mass conservation constraint and introduce a regularization term in order to eliminate the transport artifacts and the amplified noise level in the resulting images.

2.3.2 Local Transfer Methods

Local colour transfer methods are guided by image content information expressed in terms of spatial relationships (using Gaussian Mixture Model segmentation or extracted dominant colours) or colour pixel correspondences between two overlapped images or semantic relationships between source and target images (using deep learning frameworks). The following gives an overview of each type.

Using Gaussian Mixture Model. Many local colour transfer methods exploit GMM for image segmentation, where the segmentation results achieved by GMM are used to derive better colour coherence in mapping results. Tai et al. (2007) proposed a solution in which the colour transfer is defined on colour segments based on GMM estimated by a modified Expectation-Maximization algorithm that allows a given pixel to fit within more than one GMM, resulting in more seamless transitions between adjacent regions. They transfer the statistics of colour distributions between each segment pair using histogram matching similar to the approach in (Reinhard et al., 2001). Jeong

and Jaynes (2008) proposed a method to harmonize the colours of object appearances from multiple disjoint camera views for subject recognition purposes. They modeled colour distributions using GMMs and minimized the Kullback–Leibler divergence between Gaussian components to estimate an affine transformation function. Grogan et al. (Grogan et al., 2015; Grogan and Dahyot, 2015) extended the affine transformation function by using a more complex colour mapping, TPS transformation, in which they modeled the colour distributions of the source and target images using GMMs, and estimated the parameters of the TPS function by minimizing the robust $\mathcal{L}_2$ divergence between the the two pdfs. Xiang et al. (2009) proposed to segment the source image and multiple target images into regions by using the GMMs fitted by EM algorithm. An affine transfer function is constructed between the regions using the multiple target regions to ensure the best target region is assigned to the source region. Hristova et al. (2015) partitioned images into Gaussian clusters by considering their main style features (either "light-based style" if the image exhibits one dominant colour, or "colour-based style" if the image exhibits at least two different and significant dominant colours) in order to improve the colour transfer process. However, small regions with insufficient numbers of pixels may be misclassified into other Gaussian components. An alternative approaches for segmenting colours into regions can be used such as mean-shift (Oliveira et al., 2015) and K-means (Shuchang Xu et al., 2005).

Using Dominant Colours. Frigo et al. (2014) proposed a colour transfer technique where colour modes are extracted from the source and target images. Then a mapping based on the Monge-Kantorovitch optimal transport between these two sets of modes is performed by using the Simplex linear programming (Dantzig, 1990). Likewise, Yoo et al. (2013) and Dong et al. (2010) proposed a dominant colour mapping methods to reduce mapping unrelated items.

Using Correspondences. Many methods assume that there are spatial correspondences $\{(x_i, y_i)\}_{i=1}^n$ to be found between the source and the target image to guide the colour transformation. Oliveira et al. (2015) proposed to extract the overlapped area from the source and target images, and from this overlapped area colour mapping functions are estimated and applied to all the pixels of the source image. The colour estimation is performed for each channel independently using fitted truncated GMM representations. Vazquez-Corral and Bertalmío (2014) proposed a method for colour

stabilization for images of the same scene, taken under the same illumination but have colours that do not exactly match. They use colour correspondences to estimate a correction matrix composed of a linear term which is a 3x3 matrix, and power terms which are gamma correction parameters. Frigo et al. (2016) proposed a colour correction model for video sequence similar to the model proposed by (Vazquez-Corral and Bertalmío, 2014) except that it is separable over colour channels and aided by motion estimation to infer large sets of correspondences between frames. The final correction is temporally weighted by the motion speed. Park et al. (2016) proposed a technique to ensure colour consistency across photo collections, estimating white balance and gamma correction across images. Their algorithm is robust to correspondence outliers and can achieve good results with fewer correspondences than similar techniques (HaCohen et al., 2013). Xia et al. (2017) proposed to process each colour channel separately by defining the transformation model as three monotonically increasing mapping curves, which are formulated as quadratic spline curves with 6 knots under colour, gradient and contrast constraints. Bellavia and Colombo (2018) proposed a compositional framework that combines local and global colour correction algorithms. More specifically, the framework employed a monotone mapping curve for each channel separately, which is formulated as a piecewise Hermite cubic spline with 4 knots to estimate the local colour transformation for the overlapped region of the source and target images. It then employed a linear propagation step to propagate the estimated colour transformation to the non-overlap region. Grogan and Dahyot (2019) extended their colour transfer model with $\mathcal{L}_2$ divergence (Grogan et al., 2015; Grogan and Dahyot, 2015) to take advantage of the available colour correspondences, and they demonstrated that it performs well in the presence of correspondence outliers. Hwang et al. (2014) proposed 3D nonlinear and nonparametric colour transformation by applying scattered point interpolation based on moving least squares algorithm combined with weights, expressed as a probability value for each correspondence pair in order to make the model robust against correspondences outliers and noise. The moving least squares transformation independently corrects each pixel's colour in the source image. it can be very costly for large images. Therefore the authors propose to add a parallel processing scheme to increase the computational efficiency. Moreover, because the algorithm corrects each pixel's colour independently, it allows nearby pixels in the resulting image to be recoloured differently, and unwanted artifacts can appear when a large number of pixel correspondences are incorrect. The

subsequent extension by Hwang et al. (2019) improved the algorithm by introducing a new weight to account for spatially varying colour transfer in such a way that colour transfer for the pixel is dominated by closer and more similar control points.

Deep Learning based Colour Transfer. Recently deep neural networks have become popular and are often more exploited in the form of "style" transfer for non-realistic rendering than colour transfer, due to the strong capacity of convolutional neural networks in capturing latent texture features. Style can be seen as a combination of colour and texture features, and style transfer is the task of transforming a source image in such a way that it mimics the style of a given target image. Deep features are usually extracted from pre-trained networks to build the relationship between the source and target images (Xiao et al., 2020; He et al., 2019, 2018; Luan et al., 2017; Liao et al., 2017). Deep learning based models usually provide excellent results, however, they require long term training and they are supposed to be trained on a large-scale target image database that contains all possible objects, which is impossible in practice; thus, they are dependent on the type of the images they are trained on, for example, if the model was trained on real images, it would fail with the synthetic image (Cheng et al., 2015) or with unusual or artistic colours (He et al., 2018).

While there are researchers exploring neural networks already, we wanted to explore a slightly different approach. The approach we chose to explore in this thesis is similar to a neural network, but we do not learn the projections (weights) in the convolutional layers; we choose them randomly instead. In fact, we learn more complex activation functions: non-parametric activation functions such as Optimal Transport and Nadaraya-Watson kernel regression and parametric activation function based on the $\mathcal{L}_2$ framework. In our approach, we need only a single pair of images to estimate the nonlinear mapping.

2.4 Conclusion

In this chapter, we have reviewed the theoretical base that will be used in this thesis. First, we presented a set of popular statistical methods in computer science and statistical communities: OT, $\mathcal{L}_2$ inference, and NW. This thesis will exploit and extend the capabilities of the presented statistical techniques to higher-dimensional space problems. We also presented the 1D iterative approach that has the potentials to scale well in high-dimensional spaces. This thesis will exploit this approach to undertake an estimation process that involves 1D marginal distributions, hence involving low computation costs. We presented an overview of the application of interest, the colour transfer, and explained the local and global approaches. Those techniques compute the transfer function considering colour space only. Fusing different available information to guide the process of the estimation would be interesting to investigate, but that would create high-dimensional spaces that would consequently cause computational complexity challenges to the current methods. In this thesis, we will extend the problem space from colour space to higher augmented space that encodes spatial features, colour features, and local structures. Our formulations are considered as global in the sense of transferring the whole target pdfs, but at the same time, they are guided by local descriptors.

The following chapters present our solutions. We compare our proposed methods to `B-PMLS` and `PMLS` (Hwang et al., 2019; Hwang et al., 2014), `L2` (Grogan and Dahyot, 2019), `GPS/LCP` and `FGPS/LCP`, (Bellavia and Colombo, 2018), `IDT` (Pitie et al., 2007). We also compared with (Xia et al., 2017; Park et al., 2016) but `PMLS` has been shown to perform better than these two (Grogan and Dahyot, 2019), so `PMLS` is the one reported in the experimental assessment sections. Moreover, we compared with (Bonneel et al., 2016; Ferradans et al., 2013) but `IDT` also has been shown to be superior (Grogan and Dahyot, 2019), hence `IDT` is the one reported in the experimental assessment sections to make the comparison clear and easy.

COMPARISONS: OPTIMAL TRANSPORT, $\mathcal{L}_2$ AND NADARAYA-WATSON

In this thesis, three models, Optimal Transport, inference with $\mathcal{L}_2$ and Nadaraya-Watson, are used for high-dimensional distribution transfer using the 1D iterative projection approach (see Section 2.3.1.2). Therefore in this chapter, we compare and visualize how each function behaves in 1D case (Section 3.1). We also compare popular approaches proposed in the literature that tackle the high-dimensional problems using 1D OT solution, which are the iterative projection approaches proposed as approximated solutions for OT, namely IDT and SWD algorithms, and the independent transfer method (Section 3.2).

3.1 OT, NW, and $\mathcal{L}_2$ inference in 1D space

To explore the differences between the three methods in 1D, we create a simple dataset $\mathcal{S}_{x,y} = \{(x_i, y_i)\}$, where $\{x_i\}$ and $\{y_i\} \in \mathbb{R}$ denoted as source and target samples respectively, such that:

$$y_i = cos(x_i) + \epsilon_i \tag{3.1a}$$

$$\text{with noise } \epsilon_i \sim \mathcal{N}(0, 1) \tag{3.1b}$$

38
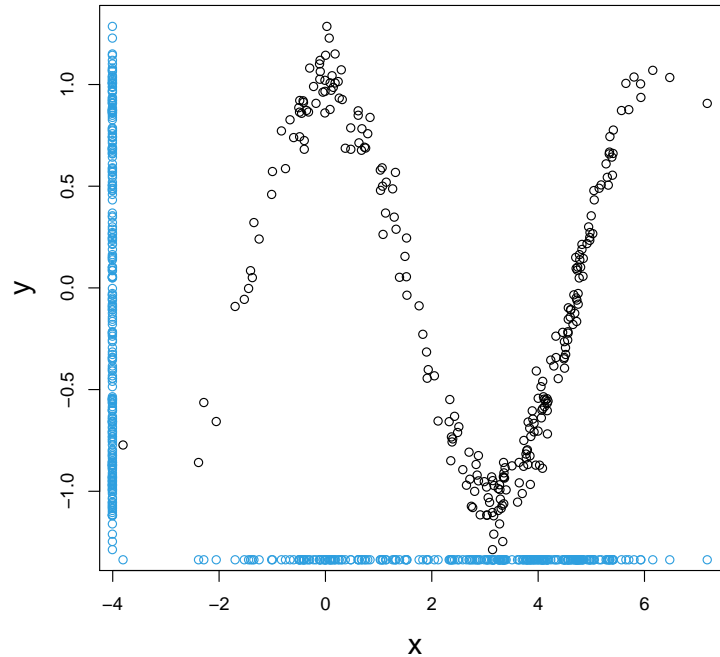
Figure 3.1: Simulated data show example of functional data. The discrete observations reflecting a smooth variation appear to be nicely sinusoidal.

In this situation we know the ground truth answer is $\phi(x) = cos(x)$. The dataset $\mathcal{S}_{x,y}$ can be visualised as a point cloud in 2D with x_i the coordinate on the x-axis and y_i the corresponding coordinate on the y-axis, Figure 3.1.

3.1.1 Nadaraya-Watson solution

The non-parametric Nadaraya-Watson estimator takes the correspondences $\{(x_i, y_i)\}$ and computes the function $\phi_h^{NW}(x)$ using the Eq. (2.23) (Section 2.2.3). As we previously explained, NW is a local weighted average of $\{y_i\}$ using a kernel as a weighting function, such that the estimated value of the function at a point x_i must be influenced mostly by the observations near x_i. The smoothing bandwidth h controls the amount of smoothing. Small bandwidths give very rough estimates while larger bandwidths give smoother estimates. If we assume deterministic coupling between variables $y = \phi(x)$, then NW provides the approximation $\hat{y} = \phi_h^{NW}(x)$ as an estimate of the conditional expectation $\mathbb{E}[y|x] = \phi(x)$, and in case of assuming non-deterministic coupling in the form $y = \phi(x) + \epsilon$, then NW provides its estimate when $\mathbb{E}[\epsilon] = 0$. The red
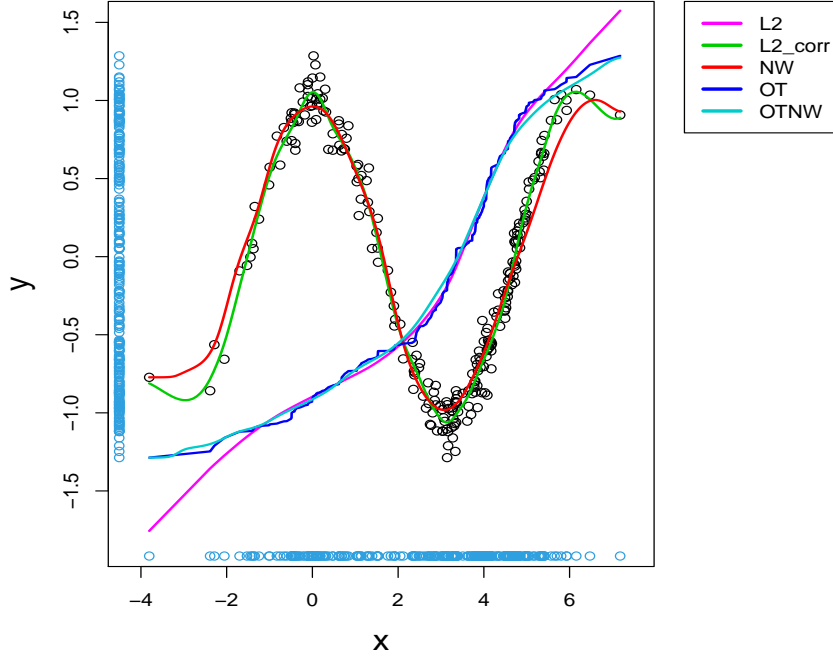
Figure 3.2: Solutions for mapping source dataset to target dataset using Optimal Transport ϕ^{OT} (OT), smoothed Optimal Transport using Nadaraya-Watson ϕ_1^{OT} (OTNW), Nadaraya-Watson ϕ_1^{NW} (NW), $\mathcal{L}_2$ estimate $\phi_\theta^{\mathcal{L}_2}$ without correspondences (L2), and with correspondences (L2_corr) .

curve in Figure 3.2 shows the Nadaraya-Watson solution $\hat{y} = \phi_1^{NW}(x)$ with Gaussian kernel and bandwidth value $h = 1$. It provides a smooth estimate of the underlying function.

3.1.2 Optimal Transport solution

If we consider that we have the datasets shown in Figure 3.1 and no pairing information (correspondences) is given, hence we only have access to separate datasets $\mathcal{S}_x = \{x_i\}$ and $\mathcal{S}_y = \{y_i\}$ and we assume an invertible deterministic mapping $y = \phi(x)$, then OT solution in 1D for the mapping ϕ is straightforward and can be defined using the non-parametric 1D mapping of the Eq. (2.9) (CDF-maching) or its variant Eq. (2.11) (quantile-matching). We estimated Optimal Transport solution $\hat{y} = \phi^{OT}$ using quantile matching with the sort algorithm and the solution is the dark blue curve visualized in Figure 3.2. Because OT does not take into account when computing the solution any available information about correspondences and the solution is

an increasing function by definition, the computed solution ϕ^{OT} for the sinusoidal behaviour type of data is irrelevant as it does not capture the pattern shown in the data, while NW solution ϕ^{NW} taking advantage of the correspondences $\{(x_i, y_i)\}$ gives a better estimate of the underlying function. The monotonic characteristic of OT mapping is useful in certain situations where there are no correspondences available and the desired mapping between the source and target datasets is increasing, such as colour transfer between images with no shared contents, as this property preserves the relative position of colours in the transformed source image. This means that areas of the source image that were bright will remain bright after the transformation, and those that were dark will remain dark (cf. Figure 1.4). Note that the result of OT is not smooth. We can compute smoother OT using NW estimator with OT estimated correspondences $\{(x_i, \phi^{OT}(x_i))\}$, and we denote it as $\phi_h^{OT}(x)$ (cf. Figure 3.2 compares dark blue curve with light blue curve for smoother OT solution).

3.1.3 Constructing $\mathcal{L}_2$ inference based solution

The $\mathcal{L}_2$ distance between probability density functions can be used to estimate the a parametric mapping ϕ_θ between the source and target datasets (cf. Section 2.2.2). The problem then is formulated as a parameter estimation:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \ \mathcal{L}_2(f(x|\theta), g(x)) \tag{3.2}$$

Unlike the OT formulation (Section 2.2.1.1), which depends on two random variables x and y (see Eq. (2.3a)), only one random variable x appears explicitly in $\mathcal{L}_2$ framework, and both the source pdf f and target pdf g are candidate pdfs for the random vector x defined here in space $\mathbb{R}$. In Section 2.2.2 we explained that there is a closed form solution for $\mathcal{L}_2$ when using Gaussian mixture models.

In the following, we first formulate the $\mathcal{L}_2$ framework in 1D space and then use it to estimate a parametric mapping ϕ_θ that maps the source to target data in the example shown in Figure 3.1:

$\mathcal{L}_2$ **between pdfs:** Let us consider that the two probability density functions f and g representing the source and target samples are Gaussian mixtures as follows:

$$f(x) = \sum_{i=1}^{n} w_i \mathcal{N}(x; \mu_i^s, h_i^2) \tag{3.3}$$

$$g(x) = \sum_{j=1}^{m} w_j \mathcal{N}(x; \mu_j^t, h_j^2) \tag{3.4}$$

The notation $\mathcal{N}(x; \mu_*, h_*^2)$ indicates a one-dimensional Gaussian distribution, with mean μ_* and variance h_*^2, and $0 < w_* < 1$ is the weight of the distribution such that $\sum_{*=1}^{**} w_* = 1$ (with $* = i, j$ and $** = n, m$). For simplicity we constructed the Gaussian mixture models explicitly as follows: 1) The number of Gaussian components is the number of samples in the dataset and all components are weighted equally, 2) for each component, the mean is given by the spatial location of each sample, and 3) all components share the same bandwidth h_*.

The $\mathcal{L}_2$ distance between Gaussian mixtures is computed as follows:

$$\mathcal{L}_2(f(x), g(x)) = \int f(x)^2 dx - 2 \int f(x)g(x)dx + \int g(x)^2 dx \tag{3.5}$$

Inference with $\mathcal{L}_2$: In order to map the source distribution $f(x)$ to the target distribution $g(x)$, $f(x)$ is changed to a parametric family of distributions $f(x|\theta)$, with parameters θ controlling the transformation ϕ_θ which maps the Gaussian mixture $f(x|\theta)$ to $g(x)$. One method to define $f(x|\theta)$ involves applying the transformation ϕ_θ to the means of the Gaussians $\{\mu_i^s\}$ (Grogan and Dahyot, 2019; Jian and Vemuri, 2011). The GMM $f(x|\theta)$ is then given by:

$$f(x|\theta) = \sum_{i=1}^{n} w_i \mathcal{N}(x; \phi_\theta(\mu_i^s), h_i^2) \tag{3.6}$$

The parameters θ are estimated by minimizing $\mathcal{L}_2$ distance between the moving source distribution $f(x|\theta)$ and the target $g(x)$ as follows:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left[\mathcal{L}_2(f(x|\theta), g(x)) = \int f(x|\theta)^2 dx - 2 \int f(x|\theta)g(x)dx \right] \tag{3.7}$$

Since the target is fixed and independent of the transformation estimation we removed the term $\int g(x)^2 dx$ from the cost function (3.7).

Closed-form solution for $\mathcal{L}_2$:   By utilizing the Eq. (2.21), the term $\int f(x|\theta)^2 dx$ in Eq. (3.7) is evaluated as follows:

$$\int f(x|\theta)^2 dx = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathcal{N}(0, \phi_\theta(\mu_i^s) - \phi_\theta(\mu_j^s), 2h^2) \tag{3.8}$$

The term $\int f(x|\theta)g(x)dx$ can be computed based on two situations: with correspondences and without taking account of the correspondences (Grogan and Dahyot, 2019). In case we do not take into account the correspondences the term is computed as follows:

$$\int f(x|\theta)g(x)dx = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathcal{N}(0, \phi_\theta(\mu_i^s) - \mu_j^t, 2h^2) \tag{3.9}$$

When the correspondences $\{(x_i^s, x_i^t)\}_{i=1}^{n}$ are given then the term $\int f(x|\theta)g(x)dx$ is computed as follows:

$$\int f(x|\theta)g(x)dx = \frac{1}{n} \sum_{i=1}^{n} \mathcal{N}(0, \phi_\theta(\mu_i^s) - \mu_i^t, 2h^2) \tag{3.10}$$

Defining the parametric transformation model ϕ_θ: We consider a parametric non-rigid 1D transformation using radial basis functions (RBF) transformation. The transformation model is decomposed into a linear part which is modeled by an affine motion in 1D, and a nonlinear part which is controlled by linear spline warping coefficients, the equation is as follows:

$$\phi(x, \theta) = c_0 + c_1 x + \sum_{l=1}^{r} w_l \ \varphi(\Gamma_l) \tag{3.11}$$

where:

$\theta = \quad \{c_0, c_1, w_1, w_2, ..., w_r\}$, parameters control the transformation to be estimated

$\varphi(\Gamma_l) = \Gamma_l$, radial basis function of linear spline type

$\Gamma_l = \quad \|x - x_l\|_2$, euclidean distance

$x_l \quad$ control points

We choose r=125 control points on regular intervals spanning the entire range of the 1D projected dataset. As a consequence the dimension of the latent space that needs to be explored when estimating θ in this case is:

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

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

Optimization: When minimizing the cost function Eq. (3.7) to estimate the parameters of the transformation ϕ_θ Eq. (3.11), we employed Quasi-Newton method (Shanno, 1970), which is a gradient-based numerical optimization technique. We computed the analytical derivative of the cost function Eq. (3.7) and passed it to the gradient ascent algorithm to speed up the optimization (Grogan and Dahyot, 2019; Jian and Vemuri, 2011). As for the implementation specifics, we used the Matlab function (*fminunc*) that implements the Quasi-Newton minimization method, which is suitable for our problem formulation. *fminunc* function returns a vector that is a local minimizer of the scalar valued function. In our case, the scalar valued function is our cost function that returns the $\mathcal{L}_2$ distance and the returned vector that gives minimum $\mathcal{L}_2$ distance is our parameters $\theta = \{c_0, c_1, w_1, w_2, ..., w_r\}$. To select the bandwidth for the Gaussians in the GMMs, we choose the fixed value $h = 1$ for the example data in Figure 3.1. In Chapter 7, we set a data-driven bandwidth value.

Estimating $\phi_\theta^{\mathcal{L}_2}$: The green curve in Figure 3.2 shows the $\mathcal{L}_2$ based solution when using the correspondences $\hat{y} = \phi_\theta^{\mathcal{L}_2}(x)$. It provides a smooth estimate of the underlying function. The pink curve in Figure 3.2 shows the $\mathcal{L}_2$ solution without taking into account the correspondences. We can see it is an increasing smooth function but it is not necessarily a strictly increasing function.

3.1.4 Wasserstein Distance comparison

Recall Wasserstein Distance $\mathcal{W}(f, g)$ (Section 2.2.1.1):

$$\mathcal{W}(f, g) = \min_{\gamma} \int_{\mathcal{X} \times \mathcal{Y}} \|x - y\|^2 \, \gamma(x, y) dx dy \tag{3.12}$$

$y = \phi(x)$	No correspondences	Correspondences
Non-parametric ϕ	Optimal Transport ϕ^{OT}	Nadaraya–Watson ϕ_h^{NW}
Parametric ϕ_θ	$\mathcal{L}_2$ rubust fitting $\phi_\theta^{\mathcal{L}_2}$	$\mathcal{L}_2$ robust fitting $\phi_\theta^{\mathcal{L}_2}$

Table 3.1: Summary of the differences between the four transfer function solutions, Optimal transport(ϕ^{OT}), Nadaraya–Watson (ϕ_h^{NW}), $\mathcal{L}_2$ based solution ($\phi_\theta^{\mathcal{L}_2}$) taking advantage of correspondences between the source and target when available, and without correspondences.

Which with a deterministic coupling reduced to:

$$\mathcal{W}(f, g) = \int_{\mathcal{X}} \|x - \phi(x)\|^2 \, f(x) dx \tag{3.13a}$$

$$= \mathbb{E}\left[\|x - \phi(x)\|^2\right] = \frac{1}{n} \sum_{i=1}^{n} \|x_i - \phi(x_i)\|^2 \tag{3.13b}$$

Which is the sum of square differences between $\phi(x)$ and x. It is interesting to calculate the Wasserstein Distance for all the transfer functions. The Wasserstein Distance results are tabulated in Table 3.2. The Optimal Transport solution ϕ^{OT} is invertible and an increasing function by definition, also the case with the smoothed OT ϕ_1^{OT}. $\mathcal{L}_2$ based solution $\phi_\theta^{\mathcal{L}_2}$, which fit linear spline function without using correspondences, is also shown to be an increasing and invertible function for the example in the Figure 3.2, but it is not guaranteed that it is always the case with the linear spline function. On the other hand, $\mathcal{L}_2$ based solution using correspondences and Nadaraya–Watson are non-monotonic functions. The Table 3.2 shows that $\mathcal{L}_2$ without correspondences gives the lowest distance, then OT is followed (Eq. (2.11) quantile-matching) followed by the smoothed OT, then comes $\mathcal{L}_2$ using the correspondences giving lower distance value than Nadaraya–Watson solution.

	OT (ϕ^{OT})	Smoothed OT (ϕ_1^{OT})	NW (ϕ_1^{NW})	L2_corr ($\phi_\theta^{\mathcal{L}_2}$)	L2 ($\phi_\theta^{\mathcal{L}_2}$)
$\mathcal{W}(f,g)$	9.657	9.682	13.754	13.659	9.622

Table 3.2: Comparing the Wasserstein Distance between the source and target after the transformation using the four transfer function solutions, Optimal transport (ϕ^{OT}), smoothed Optimal Transport (ϕ_1^{OT}), Nadaraya–Watson (ϕ_1^{NW}), $\mathcal{L}_2$ based solution ($\phi_1^{\mathcal{L}_2}$) taking advantage of correspondences between the source and target, and without correspondences. Red, blue, green, pink and cyan indicate 1st, 2nd, 3rd, 4rd, and 5rd best result in terms of minimum distance, respectively (lower values are better).

3.2 Using 1D OT in N-D space

Of the non-linear distribution transfer methods explained in Section 2.3.1.2 that tackle N-dimensional transfer problems using 1D OT pdf transfer functions (Eq. (2.9) CDF-matching, or its variant Eq. (2.11) quantile-matching), we picked the most popular ones used in the literature which are the independent transfer, Iterative Distribution Transfer (IDT) and Sliced Wasserstein Distance (SWD). In Section 2.3.1.2, we explained the independent transfer, which assumes multivariate distributions with independent variables where the histogram is used to approximate the full distribution of values in each coordinate independently. 1D non-parametric CDF-matching function is performed on each coordinate separately. On the other hand, Iterative projection-based approaches tackle the problem of estimating a N-dimensional OT by breaking down the problem into a series of subproblems, each of which finds a 1D OT solution using projected samples. IDT algorithm solve the 1D OT subproblems using CDF-matching (Eq. 2.9), and SWD algorithm solves 1D OT subproblems using quantile-matching (Eq. (2.11)) using sorting operation (see Section 2.2.1.3). These projection-based approaches provide faster calculation as well as smaller memory consumption compared with regularization-based optimal transport methods; they become useful tools when we have a large number of data samples in high-dimensional spaces. In the following we give detailed explanation for the projection-based approaches:

Denote a particular axis by its unit vector direction $e \in \mathbb{R}^N$. The projection of the N-dimensional samples x and y on the axis e is given by the 1D scalar values $u = e^T x$ and $v = e^T y$ for the source and target respectively. The projection of datasets $\{x_i\}$ and $\{y_j\}$ yields to the projection of both pdfs f and g onto the axis e, which results in two 1D marginal pdfs f_e and g_e. Using the non-linear 1D Optimal Transport mapping

---

Algorithm 1: Iterative projection-based transform using Optimal Transport

---

1: **Input:** datasets Source $\{x_i\}_{i=1}^n$ and Target $\{y_j\}_{j=1}^m$ in $\mathbb{R}^N$

2: **Initialisation:** $k \leftarrow 0$ and $\forall i,\ x_i^{(0)} \leftarrow x_i$

3: **repeat**

4:   Generate D random unit vectors in $\mathbb{R}^N$ stored in matrix R $= [e_1, ..., e_D]$

5:   **for** $r = 1$ to D **do**

6:     Compute projections $\forall i,\ u_i = e_r^T x_i^{(k)}$ and $\forall j,\ v_j = e_r^T y_j$

7:     Compute 1D transfer functions ϕ_1 to ϕ_D such that ϕ_r is the optimal transport solution (Eq. (2.9) or Eq. (2.11))

8:   **end for**

9:   Remap the source dataset

$$\Phi(x_i^{(k)}) = \text{R} \begin{pmatrix} \phi_1(e_1^T x_i^{(k)}) \\ \vdots \\ \phi_D(e_D^T x_i^{(k)}) \end{pmatrix} = \sum_{r=1}^D (\phi_r(e_r^T x_i^{(k)}))\ e_r$$

10:   Update source dataset $x_i^{(k+1)} \leftarrow (\lambda - 1)(x_i^{(k)}) + \lambda(\Phi(x_i^{(k)}))$

11:   R $\leftarrow$ Random rotation of R

12:   $k \leftarrow k + 1$

13: **until** convergence on all marginals for every possible rotation (noted $k \equiv \infty$)

14: **Result:** The final one-to-one mapping Φ in $\mathbb{R}^N$ is given by $\forall i,\ x_i \mapsto \Phi(x_i) = x_i^{(\infty)}$

---

yields a 1D mapping ϕ_e^{OT} (Eq. (2.9) or its variant Eq. (2.11)) on this axis e that matches the marginal f_e to g_e. After the transformation, the 1D sample is back-projected into the original space by the following operation:

$$x \rightarrow (\phi_e(e^T x))\ e \tag{3.14}$$

We can summarize and present both projection-based approaches IDT (Pitie et al., 2005b) and SWD (Bonneel et al., 2015) in Algorithm 1. It consists in choosing a matrix of D directions (unit vectors) R $= (e_1, ..., e_D)$, and then applying the following mapping Φ:

$$\Phi(x) = \text{R} \begin{pmatrix} \phi_1(e_1^T x) \\ \vdots \\ \phi_D(e_D^T x) \end{pmatrix} = \sum_{r=1}^D (\phi_r(e_r^T x))\ e_r \tag{3.15}$$

The ϕ_r is the corresponding 1D transfer for axis e_r and can be done independently for each axis. This manipulation is iterated over different random rotations of R matrix

over k iteration that results in a sequence of distributions $f^{(k)}$ that eventually converges to the target distribution g. Algorithm 1 converges when the overall estimated transfer function becomes the identity function: $\Phi(x^{(k)}) \simeq Id(x^{(k)}) = x^{(k)}$ (i.e the samples do not move anymore), and in practice this happens after k^t iterations where k^t can be set empirically as a stopping condition of the algorithm (Bonneel et al., 2015; Pitie et al., 2007). Pitie et al. (2005b) measured the convergence experimentally by computing the average Kullbach-Leibler divergence between pdfs, and they showed that it decreases with each iteration. The subsequent extension by Pitie et al. (2007) proposed to choose a matrix of 6 directions (unit vectors) and initialized it with optimized values for RGB space. A theoretical and numerical study of the method is developed in more depth in (Pitie et al., 2007).

In SWD (Bonneel et al., 2015), the authors introduced a regularization term λ in their model to control the transfer function with the use of the identity function. More specifically, in the update step 10 in Algorithm 1, the modified source points are obtained by taking a partial step with $\lambda{=}0.2$ such that:

$$x_i^{(k+1)} = (1 - \lambda)(x_i^{(k)}) \; + \; \lambda \; \Phi(x_i^{(k)}) \tag{3.16}$$

This regularization term λ enforces a pulling constraint on the transfer function towards the identity. The goal is to create a transfer function that moves as little as possible the source points, minimizing the displacement to eliminate the excessive stretching of the transfer function (note the increasing identity function has a minimum transport cost). On the other hand, the IDT algorithm takes the full step and can be shown in Algorithm 1 by setting $\lambda{=}1$. Figure 3.3 shows illustration of the IDT algorithm.

Simulated experiments: To explore and visualize the differences and similarities under different settings to better understand each transfer method's behavior and capabilities, we conduct a simulated experiment in $\mathbb{R}^2$. Here, the assumption regarding the joint distributions they operate on is different in terms of independent and dependent random multivariate variables. Figure 3.4 and Figure 3.5 show results of the application of the three distribution matching algorithms on multivariate distributions with independent variables, including common distributions such as uniform
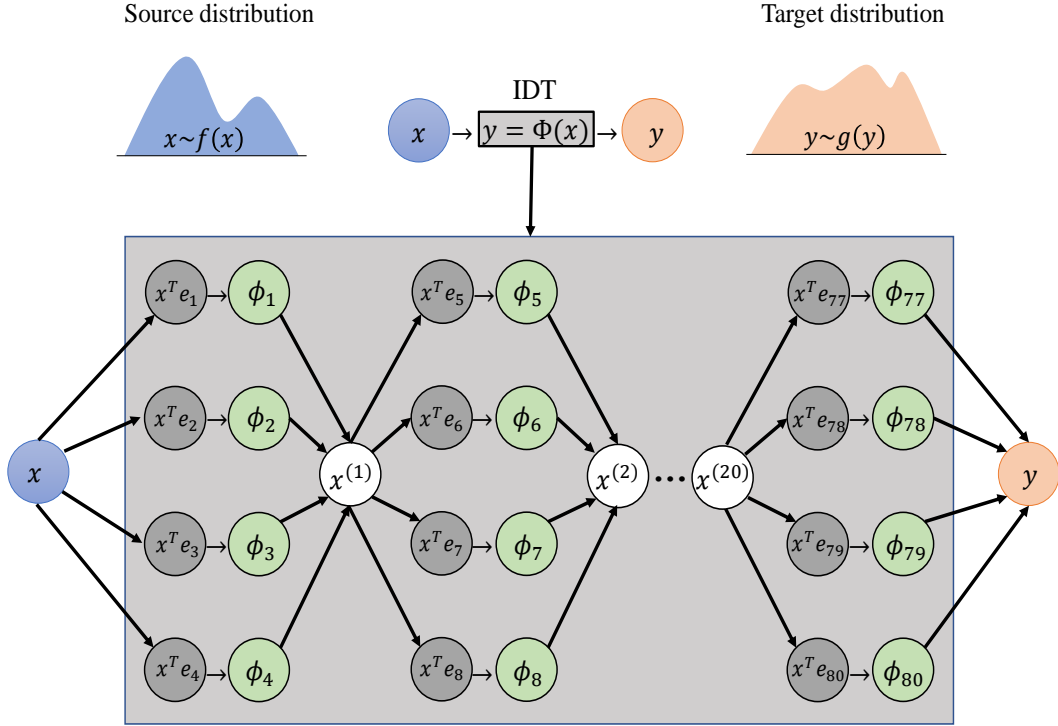
Figure 3.3: Illustration of the 1D iterative projection approach with full updating step ($\lambda = 1$) in algorithm 1 (gray boxes). The input sample x (blue node) and the output transformed sample y (orange node) with $k = 1, ..., 20$ iterations. $x^T e_r$ (gray nodes) represents the projection on 1D axis e_r. ϕ_r (green nodes) represents the estimated transformation function in the direction of the projection axis e_r. $x^{(k)}$ (white nodes) represents the back-projected transformed sample in iteration k (Best viewed in colour and zoomed in).

distribution, Gaussian distribution and Irwin-Hall distribution. Generative models such as Variational Auto Encoders (VAEs) (Kingma and Welling, 2014) and Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) are trained for a fixed prior distribution in the latent space, such as uniform or Gaussian. It is a common practice in the literature of generative models to explore the learned model behavior and outputs under various arithmetic operations on the latent space samples. The operations that are commonly used such as linear interpolation (Goodfellow et al., 2014) or vicinity sampling (Radford et al., 2016), create a distribution mismatch between the prior distribution of the latent space and the distribution of the results of these operations. Irwin hall distribution is the resulting distribution of the linear midpoint interpolation on the prior uniform distribution of the latent space. Distribution matching using the independent transfer can be used to obtain latent space operations (e.g. interpolation, vicinity sampling) that preserve the prior distribution of the latent space (Agustsson

et al., 2019). Figures 3.4 and 3.5 visualize the simulated data alongside with the contour lines of the pdfs for a better visualization. We can see from the figures that the independent transfer algorithm converges in all examples. IDT algorithm converges from the first iteration on all four examples. SWD convergence is slower than IDT due to the partial step (Eq. (3.16)).

In contrast, Figure 3.6 shows results of the application of the three algorithms on two examples of multivariate distributions with dependent variables. As we can see from the results, the independent transfer does not converge in all examples while IDT converges faster than SWD. We can conclude that using independent transfer is plausible when we are guaranteed that we are dealing with separable distributions, otherwise this approach does not produce correct results. However in most situations this setting is not realistic. IDT is considered as a generalization of the independent transfer algorithm and works well with all settings. Sliced Wasserstein Distance is slower than IDT due to the partial step (Eq. (3.16)).
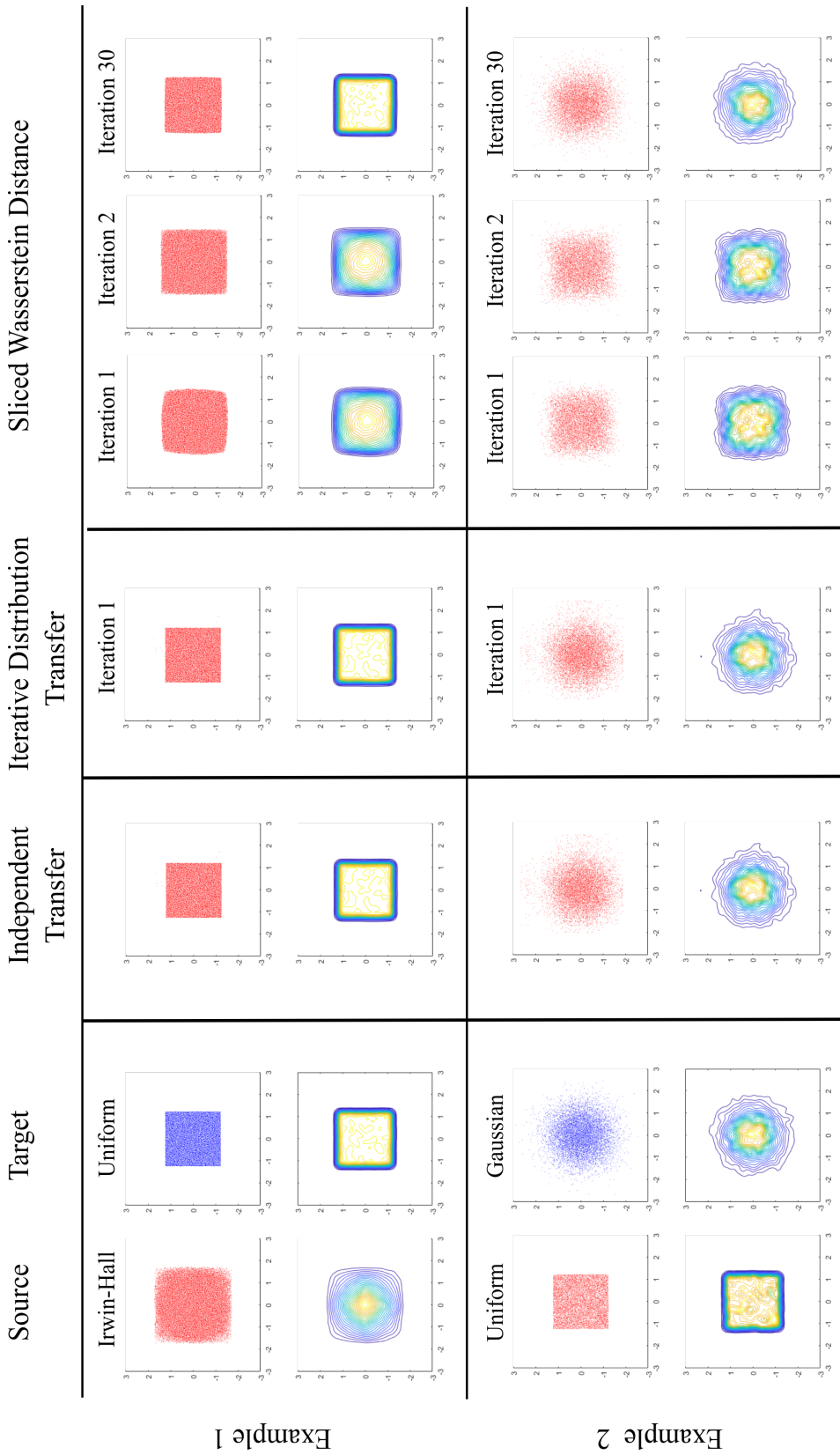
Figure 3.4: Simulation of distributions with independent variables: results of the application of the three optimal transport algorithms (Independent transfer, Iterative Distribution Transfer and Sliced Wasserstein Distance) on four multivariate distributions with independent variables. The independent transfer algorithm converges in all examples. IDT algorithm converge from the first iteration on all two examples. SWD algorithm converges after several iterations (best viewed in colour and zoomed in).
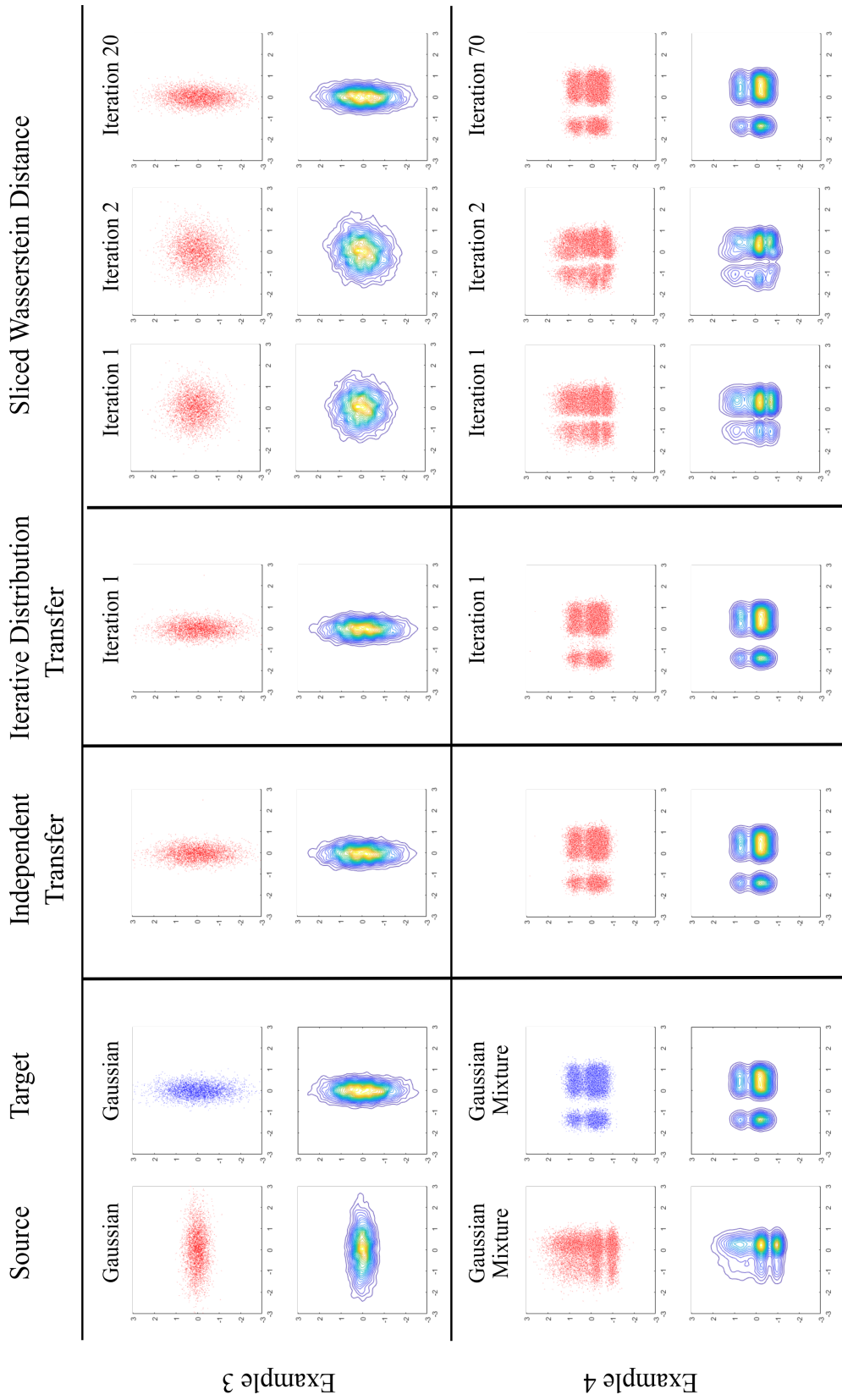
Figure 3.5: Another example of distribution transfer with independent variables. The independent transfer algorithm converges in all examples. IDT algorithm converge from the first iteration on all examples. SWD algorithm converges after several iterations (best viewed in colour and zoomed in).
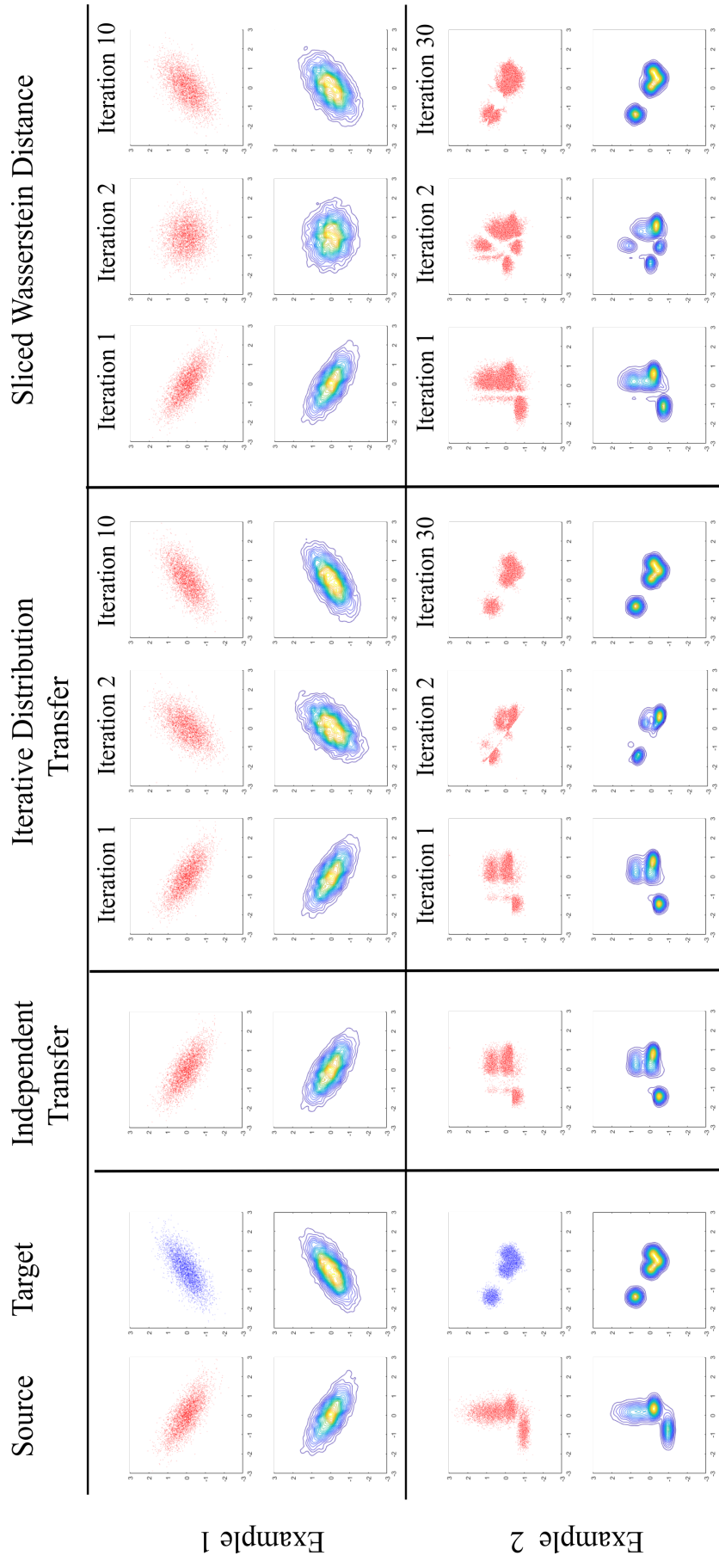
Figure 3.6: Distribution transfer with dependent variables: results of the application of the three OT algorithms (Independent transfer, Iterative Distribution Transfer and Sliced Wasserstein Distance) on 2 examples of multivariate distributions with dependent variables. The independent transfer does not converge in all examples while IDT converges faster than SWD algorithm (best viewed in colour and zoomed in).

3.3 Conclusion

In this chapter, we formulated inference with $\mathcal{L}_2$ to solve 1D problems, and compared the $\mathcal{L}_2$ based solution, Optimal Transport and Nadaraya-Watson solutions with each other in 1D space. $\mathcal{L}_2$ is a robust metric that proposes to estimate a parametric function which matches source and target distributions with a flexible framework that can take into account information about correspondences when it is available. $\mathcal{L}_2$ with non-rigid transformation functions such as splines is not scalable to high-dimensional spaces, as the complexity of the optimization increases with high dimensional spaces. Optimal Transport is a non-parametric pdf transfer function, and by definition, it does not take into account any available correspondences when computing the solution, while the ability to take advantage of the available correspondences would guide the transfer process and improve the results. Moreover, the OT solution is increasing by definition but not smooth in behavior, and that could affect the results' accuracy. Nadaraya-Watson estimator is a non-parametric kernel regression function that takes advantage of the available correspondences for computing predictions. We have also explored the capabilities of Optimal Transport based algorithms that tackle N-D problems, namely the independent transfer and the projection-based algorithms (IDT and SWD). The projection-based algorithms can match any distribution, while the independent transfer is restricted to distributions with independent variables.

This thesis explores the connections between the ideas mentioned above to overcome the limitations and solve high-dimensional space problems in colour transfer application. More specifically, we create high-dimensional representations that encode spatial information, colour information and correspondences whilst preserving local structures for better inference. These representations extend the original problem space (colour space) to higher dimensions that lead to computational complexities. To tackle the problem and avoid the computational complexities, in the next chapters, we use 1D projection-based approach to explore Optimal Transport's performance with the high-dimensional representations (cf. Chapter 4), with implicit correspondences (cf. Chapter 5), explore the application of kernel regression Nadaraya-Watson to the colour transfer problem (cf. Chapter 6) and, finally, extend the $\mathcal{L}_2$ inference (cf. Chapter 7) to high-dimensional space problems.

PATCH-BASED COLOUR TRANSFER WITH OPTIMAL TRANSPORT

4.1 Introduction

Colour variations between photographs often happen due to illumination changes, using different cameras, using different in-camera settings, or due to tonal adjustments by the users. Colour transfer methods have been developed to transform a source colour image into a specified target colour image to match colour statistics or eliminate colour variations between different photographs. There are many applications in image processing problems, ranging from generating colour consistent image mosaicing and stitching (Li et al., 2020) to colour enhancement and style manipulation (Hwang et al., 2019). In this chapter, we propose a new colour transfer method with Optimal Transport to transform the colour of a source image to match the colour of a target image of the same scene. As the transfer process is performed in the colour space, it does not consider that coherent colours should be transferred to neighboring pixels, which can create results with blocky artifacts emphasizing JPEG compression blocks or increasing noise. To tackle this problem, we propose to extend the colour problem to higher dimensional spaces by encoding overlapping neighborhoods of pixels, taking into account both their colour and spatial information. Using an iterative projection approach, we address the high dimensional distribution transfer problem in the 1D subspace where only 1D marginal distributions are used in the estimation process, thus requiring low computational costs. Since several recoloured candidates are now generated for each pixel in the source image, we define an orig-

inal procedure to efficiently merge these candidates, which allows denoising and artifact removal as well as colour transfer. Experiments show competitive results, both quantitatively and qualitatively, compared with leading state of the art colour transfer methods. Our approach can be applied to various colour transfer contexts, such as colour transfer between different camera models, camera settings, lighting conditions, and colour retouching. The work presented in this chapter has been published in (Alghamdi et al., 2019).

4.2 Optimal Transport for Colour Transfer

As described in Section 2.2.1.1, Monge's formulation of OT (Villani, 2008) defines the deterministic coupling $y = \phi(x)$ between random vectors $x \sim f(x)$ and $y \sim g(y)$ that capture the colour information of the source and target images respectively, and its solution minimizes the total transportation cost:

$$\underset{\phi}{\mathrm{argmin}} \int \|x - \phi(x)\|^2 \, f(x) \, dx \qquad (4.1.a)$$

$$\text{such that} : f(x) = g(\phi(x)) \, |\det \nabla \phi(x)| \qquad (4.1.b)$$

with f the probability density function (pdf) of x, and g the pdf of y. The solution for ϕ can be found using existing linear programming algorithms such as the Hungarian and Auction algorithms (Santambrogio, 2015). However, in practice, it is difficult to find a solution for colour images when $\dim(x) = \dim(y) = N > 1$ as the computational complexity of these solvers increases in multidimensional spaces (Villani, 2003). But as shown in Section 2.2.1.3 for dimension $N = 1$, with $x, y \in \mathbb{R}$, a solution for ϕ is straightforward and can be defined using the increasing rearrangement (Villani, 2008):

$$\phi^{OT} = G^{-1} \circ F \qquad (4.2)$$

where F and G are the cumulative distributions of the colour values in the source and target images respectively. Section 2.3.1.2 explained how the 1D OT solution (Eq. (4.2)) could be used to tackle problems in multidimensional colour spaces and of particular interest is the Iterative Distribution Transfer (IDT) algorithm for colour transfer

proposed by Pitie et al. (2005b). They proposed to iteratively project colour values $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^m$ originally in $\mathbb{R}^N$ to a 1D subspace and solve the OT using Eq. (4.2) in this 1D subspace and then propagate the solution back to $\mathbb{R}^N$ space. This operation is repeated with different directions in 1D space until convergence. This strategy was inspired by the idea of the Radon Transform (Pitie et al., 2007), which states the following proposition: if the target and source colour points are aligned in all possible 1D projective spaces, then matching is also achieved in $\mathbb{R}^N$ space. Note that the implementation of IDT approximates F and G using cumulative histograms, which can be considered as a form of quantile matching but with irregular quantile increments derived from the cumulative histograms of the source and target images - as source and target quantiles do not match exactly, interpolation can be used to compute the solution (Pitie et al., 2007). The IDT algorithm can be computed in $\mathcal{O}(nKL)$, where n is the number of samples processed, K is the number of iterations until convergence, and L is the number of projection directions considered in each iteration.

The Sliced Wasserstein Distance (SWD) algorithm was later proposed (Section 2.2.1.4), similarly using the iterative projection approach of IDT, and was applied to texture mixing applications (Rabin et al., 2012) and colour transfer (Bonneel et al., 2015). The SWD algorithm uses quantile matching (Section 2.2.1.3) instead of cumulative histogram matching to solve Eq. (4.2). More specifically, SWD sorts the n 1D projections of the source and target images respectively to define quantiles with regular increments of size $\frac{1}{n}$ between 0 and 1 for both source and target distributions. The SWD algorithm can be computed in $\mathcal{O}(n \log(n) KL)$ operations using a fast sorting algorithm (Rabin et al., 2012). When a small number of observations are available, using SWD is best, but with a very large number of observations, histogram matching with IDT is more efficient.

The warping map ϕ^{OT} derived by using the closed-form (Eq. (4.2)) has a number of important characteristics: it is parameter free, no strong hypotheses are made about the distributions (i.e. f and g). By design it is an increasing function and this is an important property for colour transfer problems since OT does not take into account available correspondences to guide the transfer process. More specifically, this increasing property preserves the relative position of colours in the transformed source image, i.e., the areas of the source image that were bright will remain bright after transformation, and those that were dark will remain dark. Also, this function provides a

tractable iterative solution for imaging applications formulated in higher dimensional spaces. However, undesirable visual artifacts can be produced in the output image and a post processing step is added to ensure that the gradient field of the resulting image matches the gradient field of the source image (Pitie et al., 2007). Similarly, Bonneel et al. (2015) use an iterative post-processing technique introduced in (Rabin et al., 2011) to regularize the transportation map.

4.3   Patch-based Colour Transfer (PCT_OT)

We propose to solve colour transfer by encoding overlapping neighborhoods of pixels, accounting for both colour information as well as pixel location, and extending the problem to higher dimensional space. We solve the transfer problem in 1D space using an iterative projection approach. In our context, we found that using the quantile matching approach that employs sorting operations is better than the histogram technique. The problem with building histograms is that the regular bin size used for every projection must be chosen, and so to alleviate the need to optimize the bin size parameter, we choose the sorting technique (Bonneel et al., 2015). Due to overlapping regions in patches, several recoloured candidates are now generated for each pixel in the source image. Therefore, we also define an original procedure to merge these candidates efficiently. Our proposed pipeline consists of the following steps:

4.3.1   Combine colour and spatial information

Shu et al. (2017) have proposed to incorporate spatial information for portrait relighting that transfers shading with an awareness of the face geometry in source and target images. Similarly, we propose to incorporate the original pixel positions in the grid coordinate of the image (global coordinates) with colour information to take into account spatial information when recolouring images with shared contents. The pixel's colour denoted x^c and its pixel position x^p are concatenated into a vector $x = (x^c, x^p)^T$ such that $\dim(x) = \dim(x^c) + \dim(x^p)$. Adding this spatial information extends Monge's formulation in Eq. (4.1.a) such that $\phi(x)$ minimizes the following cost:
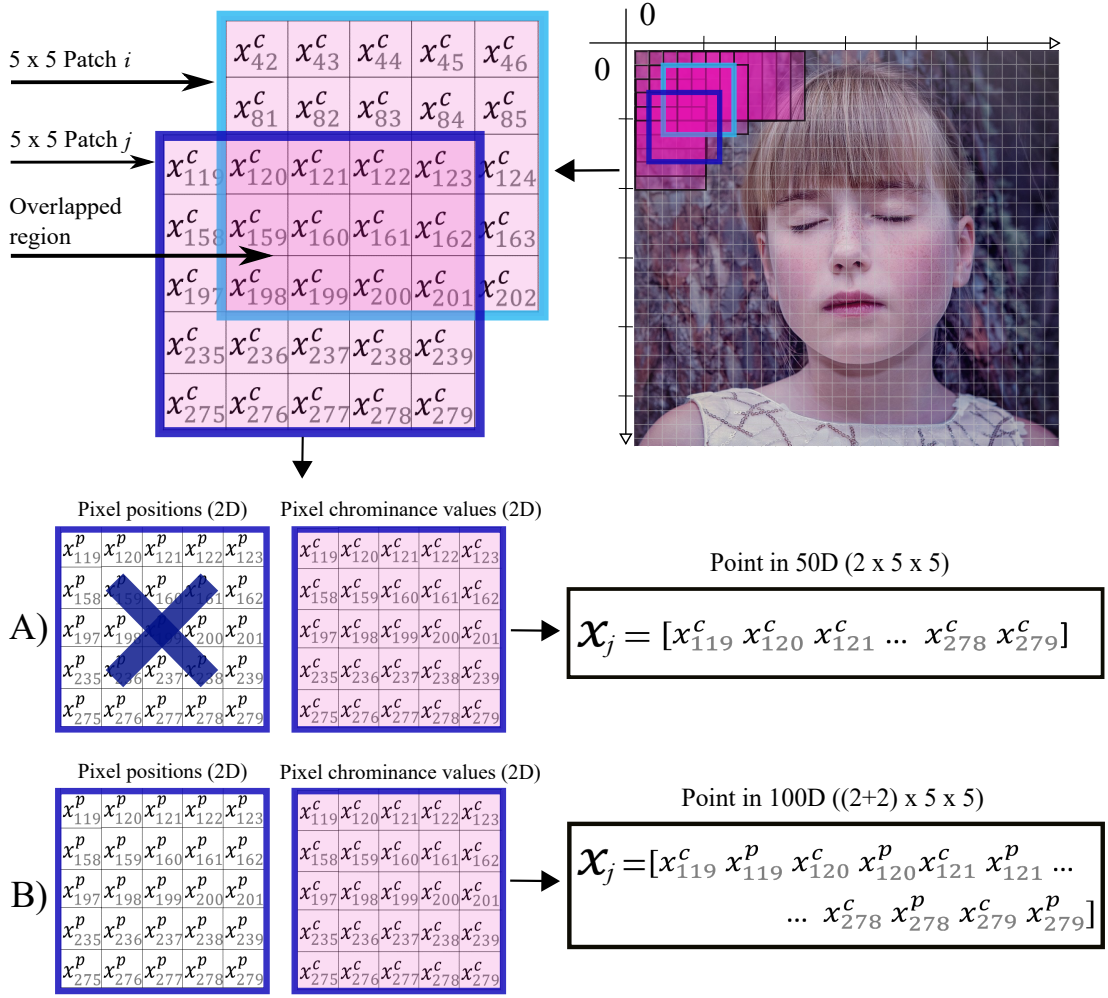
Figure 4.1: This figure shows an example of overlapping patches to create points in higher dimensions when A) using the colour information only (chrominance) and when B) combine positions and the colour information (best viewed in colour and zoomed in).

$$\int \left(\|x^c - \phi_c(x)\|^2 + \|x^p - \phi_p(x)\|^2 \right) f(x) \, dx \tag{4.3.a}$$

$$\text{such that}: f(x) = g(\phi(x)) \left| \det \nabla \phi(x) \right| \tag{4.3.b}$$

where now f denotes the pdf of x, and g the pdf of y in the high dimensional space ($\mathbb{R}^N$, where $N = \dim(x)$). ϕ_c and ϕ_p are the components of ϕ that return the transformed colour and spatial features respectively. The Eq. (4.3.a) highlights the effect of Monge's formulation on the combined colour and spatial information, which minimizes the distance in the colour space as well as in the spatial domain.

4.3.2 Data normalisation

In the case when colour and spatial information are combined, the colours have integer values from 0 to 255, and the spatial values can be anything depending on the size of the image. In order to produce consistent results regardless of the size of the image, all the colour and position coordinates are normalized to lie between 0 and 255 to create a hypercube in $\mathbb{R}^N$ (for example, $N = \dim(x) = 5$, where 5D dimensional feature space comes from the combined 3D RGB colour space and the 2D pixel's position in the image plane). We then stretch that space in the direction of the spatial coordinates by a factor w to make it harder to move the pixels in the spatial domain than in the colour domain, because since we are focusing on transferring colour between images of a same scene, we know that the scenes are overlapped.

4.3.3 Create patch vectors

We encode overlapping neighborhoods of pixels to preserve local topology information. Starting from the origin of the coordinate system of the images (upper left corner), we use a sliding window operation of window size $d \times d$ to extract overlapping patches. From each individual patch we create a high dimensional vector in $\mathbb{R}^N$, where $N = \dim(x) \times d \times d$, by concatenating all the pixels in the patch with their $\dim(x)$ dimensional features (colour values only such that $x = (x^c)^T$ Figure 4.1 (A), or colour and spatial features such that $x = (x^c, x^p)^T$ Figure 4.1 (B)). We apply this process to the source and target images to create patch vector sets $\{x_i\}$ and $\{y_j\}$ for each respectively. Note, the formulation that includes spatial information favors local mapping between patches such that it penalizes mappings between patches that are spatially distant from each other even if their colours are similar.

Note that we experimented with using local coordinates where we extract the patches first, then we consider each patch as a mini image where it has its origin of the coordinate system (upper left corner of the patch), and we calculate the pixel coordinates in the patch based on the patch's origin. We found that using local coordinates does not produce any improvements in the results, while using the image's global coordinates improves the results.

4.3.4 Iterative projection transfer

The SWD algorithm (Bonneel et al., 2015) that employs sorting operation is used as follows (Algorithm 1): let $x^{(k)}$ represents the high dimensional source point in $\mathbb{R}^N$ at iteration k. At each iteration, we generate a random orthogonal basis for the $\mathbb{R}^N$ space and project the source and target points onto these axes. Along each axis, ϕ^{OT} is estimated and then applied to the source points to create an intermediate value $\phi^{OT}(x^{(k)})$. As explained in Section 3.2, in the SWD algorithm the final modified source point at the k iteration are obtained by taking a partial step with $\lambda=0.2$ (see step 10 Algorithm 1) such that:

$$x^{(k+1)} = (1 - \lambda)(x^{(k)}) \; + \; \lambda \; \Phi^{OT}(x^{(k)}) \tag{4.4}$$

Here we have chosen a full step $\lambda=1$ to get an updating step similar to Pitie et al. (2005b), hence not enforcing any pulling constraint on the transfer function towards the identity as Bonneel et al. (2015). After convergence, we retain only the transformed colour information and discard any changes in the positions.

4.3.5 Merge recoloured candidates

Due to the overlapping regions between patches, several recoloured candidates are generated for each pixel in the source image. We compute the average colour value from the candidates and use it in conjunction with the original position to create the output recoloured image. This procedure of merging with average operation allows denoising and artifact removal as well as producing smoothed colour transfer. Our proposed pipeline is outlined in Algorithm 2.

Algorithm 2: Our pipeline for colour transfer between overlapped scene images

---

1: **Input:** Source & Target images
2: **Output:** recoloured Source image
3: Formulate the Source image & the Target image as colour samples datasets $\{x_i\}_{i=1}^n$ & $\{y_j\}_{j=1}^m$ of r.v. x and $y \in \mathbb{R}^N$, respectively. For example in RGB colour transfer, $N = 3$ and $x_i = (r_i, g_i, b_i)$ where r_i, g_i, b_i are the red, green and blue components of source pixel number i
4: **if** Combining colour and spatial information **then**
5:      The pixel's colour x^c and its pixel position in the image plane x^p are concatenated into a vector $x = (x^c, x^p)^T$ in $\mathbb{R}^N$ such that $N = \dim(x) = \dim(x^c) + \dim(x^p)$, same process is applied to the target
6:      Normalize all the colour and position coordinates to lie between 0 and 255
7:      Choose a window size $d \times d$ and use the combined colour and position features (i.e. $x = (x^c, x^p)^T$ & $y = (y^c, y^p)^T$) to create overlapping patch vectors $\{x_i\}$ & $\{y_j\}$ in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$
8: **else if** colour only **then**
9:      Choose a window size $d \times d$ and create colour only (i.e. $x = (x^c)^T$ & $y = (y^c)^T$) patch vectors $\{x_i\}$ & $\{y_j\}$ in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$
10: **end if**
11: Giving patch vectors $\{x_i\}$ & $\{y_j\}$, compute Optimal Transport mapping (Algorithm 1) where matrix R is random orthogonal basis for the $\mathbb{R}^N$ space (step 4 in Algorithm 1)
12: Merge recoloured candidates from the overlapping patches $\{x_i^{(\infty)}\}$ to create the output recoloured source image

---

4.4   Experimental Assessment

We provide quantitative and qualitative evaluations to validate our method `PCT_OT` with both of our optimal transport solutions - using colour patches only, annotated in the results as `Our_c`, and using colour patches with spatial information, annotated as `Our_cp`.

4.4.1   Colour space and parameters settings

Colour space: Here we use the YCbCr colour space and investigate transforming the luminance (Y) and chrominance (CbCr) components separately, and recombining the resulting recoloured sources (Y and CbCr) to create the final recoloured source image. Other colour spaces such as RGB, $l\alpha\beta$, CIELAB may be considered, but because full statistics (pdfs) are transferred by the proposed method, this has no implications and the method would function regardless of the colour space chosen. For `Our_cp`, when processing chrominance, each pixel is represented by its 2D chromi-

Figure 4.2: This figure shows the impact of using larger or smaller values of the stretching parameter $w = \{1, 3, 5, 10, 20, 30\}$ on the estimated mapping in the case when we combine the pixel's position with colour information. When we increase the value of w, the transfer function still maintains the structure of the source image but introduces shadows from the target image in places where there are large motion changes (best viewed in colour and zoomed in).

nance (CbCr) value and its 2D spatial position. When processing luminance, each pixel is represented by its 1D luminance (Y) value and its 2D spatial position. The spatial information corresponds to the position coordinates of the pixel in the image, with the origin of the coordinate system in the upper left corner (Figure 4.1).

Patch size: When creating patch vectors, we evaluated the impact of using different patch sizes. Figure 4.3 shows colour transfer results when using different patch sizes $d \times d$ with $d = \{1, 5, 10\}$. When we increase the value of d we get smoother colour transitions across pixels. We found that a size of 5×5 captures enough of a pixel's neighborhood. Our patches now with combined colour and spatial features create a vector in 75 dimensions ($3 \times 5 \times 5$) for the luminance and position components, and 100 dimensions ($4 \times 5 \times 5$) for the chrominance and position components. For `Our_c`, pixel position is not accounted for, and only chrominance and luminance are used. The code we used to implement our method is available online[1].

The value of the stretching parameter w**:** The w parameter controls the influence of the positions in the transfer process. We evaluate the impact of using larger or smaller values of the stretching parameter w on the estimated mapping in the case when we combine the pixel's position with colour information. Figure 4.2 shows ex-

---

[1] `https://github.com/leshep/PCT_OT`

| 1x1 | 5x5 | 10x10 |
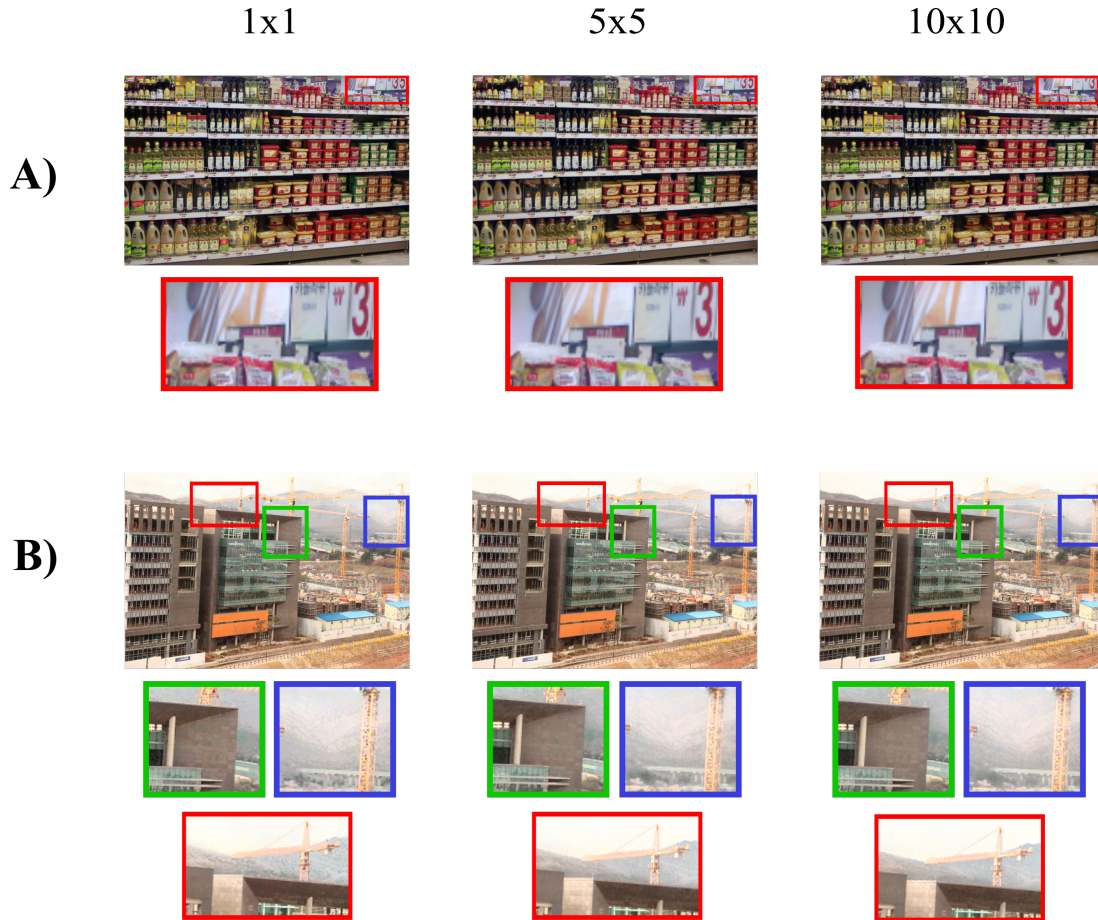
Figure 4.3: This figure shows the impact of using different patch sizes $d \times d$ with $d = \{1, 5, 10\}$. When we increase the value of d, we get smoother colour transitions across pixels (best viewed in colour and zoomed in).

amples of colour transfer results when using different values $w = \{1, 3, 5, 10, 20, 30\}$. When we increase the value of w, the transfer function still maintains the structure of the source image but introduces shadows from the target image in places where there are large motion changes. The higher the value of w, the bigger influence of pixel positions on the estimated mapping. We found that setting $w = 2.5$ is a good balance and produces the best visual results.

4.4.2   The state of the art algorithms for comparison

We compare our methods to different state of the art colour transfer methods noted by `B-PMLS` (Hwang et al., 2019), `L2` (Grogan and Dahyot, 2019), `GPS/LCP` and `FGPS/LCP`, (Bellavia and Colombo, 2018), `PMLS` (Hwang et al., 2014), `IDT` (Pitie et al., 2007). Note that the results using `B-PMLS` and `PMLS` were provided by the authors

(Hwang et al., 2019; Hwang et al., 2014). We also compared with two other recent techniques (Xia et al., 2017; Park et al., 2016) that incorporate correspondences into their framework, but `PMLS` has been shown to perform better than these two (Grogan and Dahyot, 2019), so `PMLS` is the one reported here with (Grogan and Dahyot, 2019; Bellavia and Colombo, 2018) as algorithms that incorporate correspondences in their methodologies. Moreover, we compared with (Bonneel et al., 2016; Ferradans et al., 2013) that do not take into account correspondences, but `IDT` also has been shown to be superior (Grogan and Dahyot, 2019), hence `IDT` is the one reported here to make the comparison clear and easy.

4.4.3 Evaluation dataset

To evaluate our algorithms, we use image pairs with similar content from a popular dataset provided by Hwang et al. (2014). The dataset includes 30 images that compose 15 geometrically registered pairs of images (source and target), which exhibit different colour change sources from different cameras, different in-camera settings, different illuminations, different tonal adjustments and different photo retouch styles. Some of the images contain large motion changes such as images of "building", some exhibit object changes such as "illum", some have marginal motion changes such as the case of retouch styles images (images of "gangnam") and others do not contain motion changes such as the case with tonal adjustment images (images of "tonal"). The dataset is provided in Appendix A.

4.4.4 Evaluation tools

Numerical metrics: to quantitatively assess the recolouring results, four metrics are used: peak signal to noise ratio (PSNR) (Salomon, 2004), structural similarity index (SSIM) (Zhou Wang et al., 2004), colour image difference (CID) (Preiss et al., 2014) and feature similarity index (FSIMc) (Zhang et al., 2011). These metrics are often used when considering source and target images of shared contents for which correspondences are easily available (Matysiak et al., 2020; Hwang et al., 2019; Grogan and Dahyot, 2019; Bellavia and Colombo, 2018; Oliveira et al., 2015; Hwang et al., 2014; Lissner et al., 2013). As we said earlier, the dataset used for the evaluation contains geometrically registered pairs of images, but note, the alignment that is given by the

image registration algorithm sometimes can never be accurate to the pixel level and it may contain mismatches and noise, and while the registration errors can affect the evaluation metrics, the measures for the different colour transfer methods have been calculated using the same registration information.

Descriptive statistics: we graphically depict the numerical data through box plots. The purpose of the box plots is to visually show the distribution of the numerical data and visualize differences among methods and show how close our methods are to the state of the art algorithms.

Significance analysis: we further use the non-parametric Kruskal-Wallis Rank Sum test, which is an extension of the two-sample Wilcoxon Rank Sum test, to determine if there are statistically significant differences in the experiment results between multiple methods. The advantage of this test is that it is non-parametric which means it does not rely on data belonging to any particular parametric family of probability distributions (unlike ANOVA and two-sample t-test which assume normal distribution). Also, the null hypothesis of the Kruskal-Wallis test is equal medians which makes it much less sensitive to outliers (unlike ANOVA and two-sample t-test which compare the means that are sensitive to outlier values).

Qualitative analysis: we provide visual assessments. For clarity, the results are presented in image mosaics, created by switching between the target image and the transformed source image column-wise (Figure 4.5, top row). If the colour transfer is accurate, the resulting mosaic should look like a single image (ignoring the small motion displacement between source and target images), otherwise column differences appear. The result mosaics approach makes the comparison accurate and easy. We also provide close up look results to visualize the differences between the methods that are not picked up by the numerical metrics.

4.4.5   Experimental results

Quantitative Evaluation: The numerical results of each metric with means and standard errors (SE) are shown in Table 4.2 and Table 4.3, along with box plots carrying a lot of statistical details as shown in Figures 4.6 to 4.9. The purpose of the box

plots is to visualize differences among methods and to show how close our methods are to the state of the art algorithms. By looking at the Table 4.2 and Table 4.3, all metrics show our solution that incorporates the colour and spatial information (Our_cp) outperforms the solution that uses colour features only (Our_c). Moreover, we can see the significant difference in the performance between our method Our_cp and the IDT method, as we can see that the optimal transport performance is highly improved when we extended the colour problem to high dimensional representation that encode colour, spatial, and neighborhood information (as suggested by our model Our_cp), these results can be visualized in the box plot in Figures 4.6 to 4.9. Comparing Our_cp with leading state of the art colour transfer methods, we find that Our_cp is competitive with three top methods B-PMLS, L2 and PMLS. We can see clearly in the summary statistics (box plots in Figures 4.6 to 4.9) that B-PMLS, L2, PMLS and Our_cp are greatly overlapped with each other, the median and mean values (the means shown as red dots in the plots, and the medians shown as horizontal black lines) are the highest among all algorithms and are very close in value, hence, B-PMLS, L2, PMLS and Our_cp are considered equivalent in their performances.

In conclusion, the quantitative metrics show that our method with colour and spatial information (Our_cp) performs similarly with top methods B-PMLS, L2 and PMLS, and outperforms the rest of the state of the art algorithms. It might be worth mentioning that having mean values in SSIM and FSIMC metrics (Figure 4.9 and Figure 4.7) that are located at the edge of the box plots of some methods is due to having outliers in the results (shown as black dots in the plots). However, if we removed those outliers, this will not affect our above conclusion (see Appendix A). Moreover, our conclusion is supported by statistical analysis, using Kruskal-Wallis Rank Sum test. For each individual metric results, the Kruskal-Wallis Rank Sum test compares the resulting medians of the methods. We found that with 95% confidence level, the P-values for all metrics are greater than α [2] which indicates that there is no significant difference between PMLS, L2, B-PMLS and our method Our_cp, Table 4.1.

Qualitative Evaluation: Figure 4.5 provides qualitative results. For clarity, the results are presented in image mosaics (Figure 4.5, top row). If the colour transfer is accurate, the resulting mosaic should look like a single image (ignoring the small mo-

---

[2]with confidence level = 95%, the significance level $\alpha = 0.05$. If the P-value $< \alpha$, the hypothesis test is statistically significant.

Metrics	P-values
PSNR	0.6856
SSIM	0.4614
CID	0.7293
FSIMc	0.9774

Table 4.1: This table shows results of Kruskal-Wallis Rank Sum statistical test for each metric separately. The test tests if there are significant differences between our method `Our_cp` results and the top performance methods, namely, `B-PMLS`, `L2` and `PMLS`. The null hypothesis H_o states that all medians are equal, the alternative hypothesis H_1 states that at least one median is different. The confidence level $= 95\%$, and the significance level $\alpha = 0.05$. If the P-value $< \alpha$, the null hypothesis is rejected and alternative hypothesis is accepted. Since the P-values for each metric is greater than the significance level $\alpha = 0.05$, this indicates that there is no significant difference between the aforementioned methods.

tion displacement between source and target images), otherwise column differences appear. As can be noted, Optimal Transport which by definition does not take into account the correspondences but when incorporates the colour and spatial and neighborhood information (our approach `Our_cp`) outperforms the original IDT solution in alleviating the column differences, and is better than `PMLS` and `L2` which take advantage of correspondences when estimating the colour mapping.

While `B-PMLS` and `PMLS` provide good results in terms of metrics, they introduce visual artifacts if the input images are not registered correctly (Figure 4.4), while our method is robust to registration errors. Note that although the accuracy of the PSNR, SSIM, CID and FSIMc metrics relies on the fact that the input images are registered correctly; if this is not the case, these metrics may not accurately capture all artifacts (cf. Figures 4.4 and 4.5). In addition, due to the merging step of our algorithm (cf. step 5 Section 4.3), our approach allows us to create a smooth colour transfer result, and can also alleviate JPEG compression artifacts and noise (cf. Figure 4.4 for comparison).

Our method `Our_cp` can also correctly transfer colours between images that contain moving objects, as can be seen in Figure 4.4 with examples "illum" and "mart". However, `Our_cp` can create shadow artifacts when there are large changes between target and source images (Figure 4.4 in example "building"). In this case, `Our_c` does not suffer from these shadow artifacts and creates good colour transfer results.
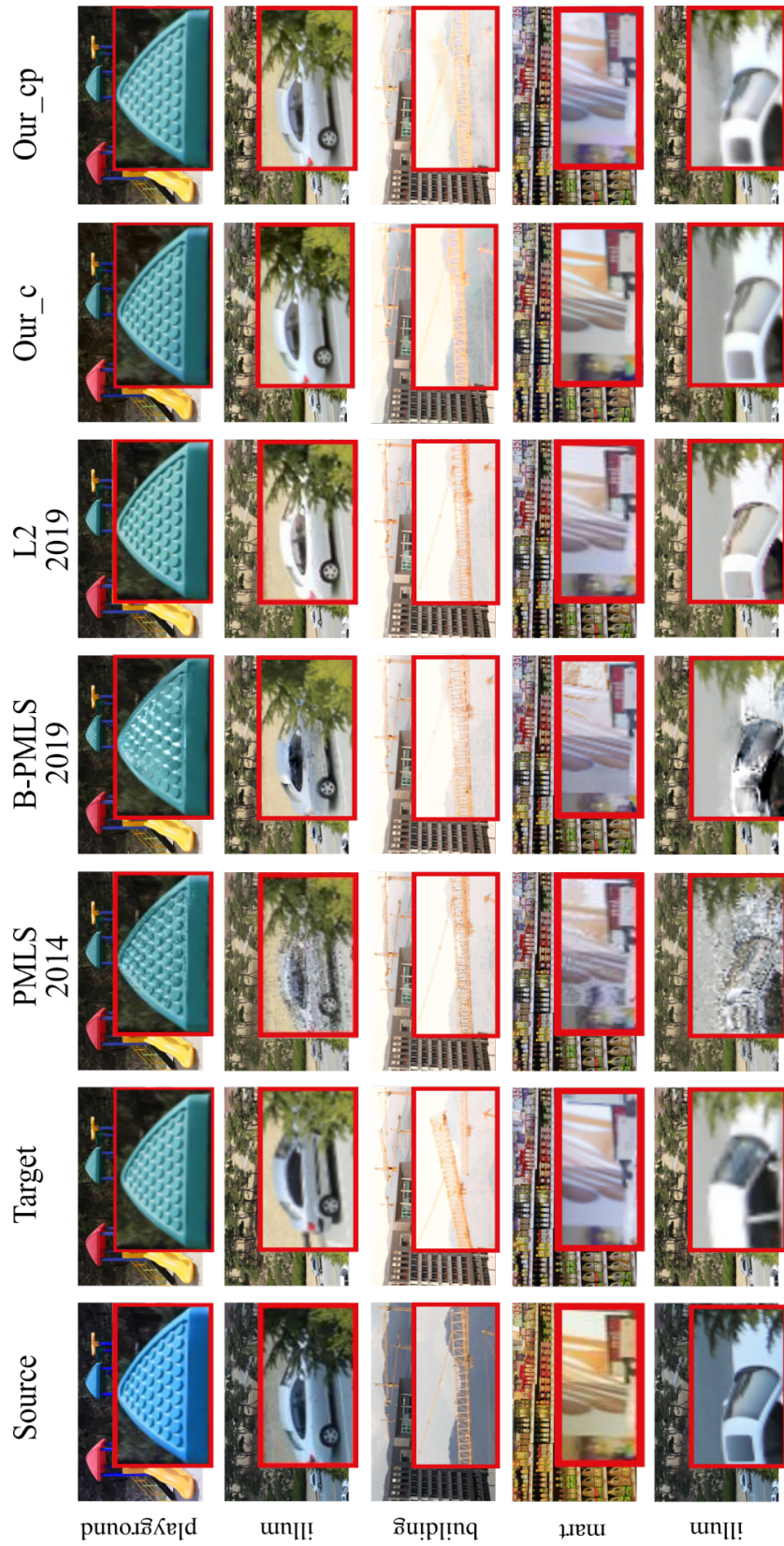
Figure 4.4: A close up look at some of the results generated using the PMLS (Hwang et al., 2014), B-PMLS (Hwang et al., 2019), L2 (Grogan and Dahyot, 2019) and our algorithms. B-PMLS and PMLS introduce clear visual artifacts that are not captured by the numerical metrics (best viewed in colour and zoomed in).

Figure 4.5: A close up look at some of the results generated using the IDT (Pitie et al., 2007), PMLS (Hwang et al., 2014), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), B-PMLS (Hwang et al., 2019), L2 (Grogan and Dahyot, 2019), and our algorithm Our_cp. The results are presented in image mosaics, created by switching between the source (or the result i.e. the transformed source) and the target image column wise, if the colour transfer is accurate, the resulting mosaic should look like a single image, otherwise column differences appear (best viewed in colour and zoomed in).

Scene	PSNR ↑								SSIM ↑							
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	Our_c	Our_cp	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP P 2018	B-PMLS 2019	L2 2019	Our_c	Our_cp
Gangnam1	25.354	35.725	24.048	23.936	36.165	35.358	25.217	31.479	0.900	0.992	0.892	0.891	0.990	0.990	0.920	0.964
Gangnam2	27.116	36.553	25.952	25.944	36.255	35.524	28.300	35.502	0.920	0.993	0.909	0.909	0.987	0.986	0.938	0.980
Gangnam3	22.372	35.007	21.908	21.913	36.022	33.284	22.373	26.393	0.859	0.991	0.873	0.864	0.987	0.980	0.849	0.930
Illum	19.822	20.167	19.785	19.960	21.960	19.079	19.370	20.306	0.641	0.649	0.643	0.646	0.733	0.648	0.657	0.673
Building	20.554	22.634	22.736	22.769	24.867	20.499	20.761	25.019	0.808	0.865	0.864	0.863	0.899	0.862	0.847	0.888
Playground	27.184	27.835	25.501	25.436	28.172	27.647	26.002	28.482	0.920	0.940	0.878	0.876	0.939	0.939	0.902	0.939
Flower1	24.238	26.981	23.765	23.706	27.194	26.857	21.661	25.186	0.909	0.967	0.913	0.912	0.968	0.966	0.874	0.926
Flower2	25.417	25.760	25.259	25.223	25.978	25.772	24.519	26.373	0.901	0.928	0.894	0.894	0.930	0.927	0.893	0.933
Tonal1	30.082	37.215	31.617	31.413	37.556	37.332	31.451	37.044	0.953	0.988	0.971	0.970	0.989	0.987	0.966	0.988
Tonal2	27.992	31.508	25.062	25.087	32.229	31.356	28.376	32.049	0.968	0.987	0.926	0.926	0.989	0.986	0.971	0.988
Tonal3	29.575	36.246	28.136	28.065	36.826	36.644	29.704	33.793	0.962	0.992	0.947	0.946	0.993	0.992	0.968	0.987
Tonal4	28.605	34.521	28.852	28.848	35.697	34.344	28.209	33.819	0.944	0.983	0.932	0.932	0.987	0.983	0.948	0.981
Tonal5	30.330	35.260	29.580	29.448	36.870	34.303	31.321	36.437	0.965	0.986	0.953	0.954	0.990	0.985	0.972	0.990
Mart	22.747	24.742	23.183	23.196	24.936	24.450	22.890	24.509	0.904	0.957	0.925	0.925	0.958	0.956	0.918	0.941
Sculpture	29.884	32.062	29.037	28.820	32.635	32.067	26.402	31.237	0.942	0.971	0.934	0.932	0.973	0.972	0.885	0.945
Mean	26.085	30.814	25.628	25.584	31.557	30.301	25.770	29.842	0.900	0.946	0.897	0.896	0.954	0.944	0.901	0.937
SE	0.905	1.459	0.841	0.821	1.408	1.518	0.983	1.306	0.022	0.023	0.020	0.020	0.017	0.023	0.021	0.020

Table 4.2: Quantitative evaluations using evaluation metrics PSNR and SSIM. Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively when applied to dataset (Hwang et al., 2014) that includes captured image pairs with different cameras, camera setups, illumination and different image recolouring styles. Higher values are better (best viewed in colour and zoomed in).

Scene	CID ↓								FSIMc ↑							
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	Our_c	Our_cp	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	Our_c	Our_cp
Gangnam1	0.252	0.040	0.226	0.222	0.044	0.048	0.189	0.085	0.936	0.986	0.944	0.943	0.986	0.985	0.946	0.972
Gangnam2	0.268	0.039	0.291	0.292	0.048	0.089	0.217	0.068	0.952	0.992	0.962	0.962	0.989	0.988	0.968	0.990
Gangnam3	0.496	0.108	0.472	0.487	0.140	0.193	0.484	0.261	0.946	0.992	0.962	0.961	0.993	0.990	0.973	0.987
Illum	0.386	0.390	0.395	0.396	0.315	0.397	0.428	0.377	0.800	0.819	0.824	0.824	0.859	0.818	0.826	0.839
Building	0.374	0.228	0.313	0.321	0.188	0.249	0.313	0.183	0.874	0.928	0.930	0.929	0.949	0.926	0.912	0.942
Playground	0.440	0.238	0.443	0.471	0.221	0.254	0.466	0.209	0.950	0.958	0.933	0.932	0.954	0.955	0.938	0.956
Flower1	0.389	0.163	0.396	0.400	0.164	0.174	0.458	0.285	0.954	0.975	0.968	0.967	0.976	0.976	0.954	0.971
Flower2	0.337	0.245	0.322	0.323	0.235	0.266	0.363	0.218	0.941	0.950	0.945	0.945	0.951	0.949	0.944	0.954
Tonal1	0.310	0.101	0.285	0.308	0.093	0.111	0.294	0.097	0.964	0.997	0.986	0.986	0.997	0.997	0.993	0.998
Tonal2	0.288	0.128	0.351	0.347	0.097	0.145	0.250	0.099	0.984	0.993	0.973	0.973	0.995	0.992	0.984	0.993
Tonal3	0.244	0.079	0.294	0.294	0.085	0.081	0.206	0.077	0.979	0.997	0.984	0.983	0.997	0.997	0.993	0.997
Tonal4	0.240	0.108	0.248	0.238	0.082	0.107	0.173	0.065	0.966	0.989	0.972	0.973	0.992	0.990	0.987	0.995
Tonal5	0.156	0.091	0.205	0.192	0.062	0.092	0.195	0.051	0.980	0.994	0.987	0.987	0.997	0.993	0.991	0.998
Mart	0.526	0.219	0.405	0.402	0.203	0.225	0.526	0.426	0.946	0.969	0.960	0.959	0.969	0.967	0.957	0.969
Sculpture	0.242	0.137	0.213	0.224	0.116	0.143	0.381	0.232	0.980	0.987	0.982	0.980	0.987	0.988	0.982	0.988
Mean	0.330	0.154	0.324	0.328	0.139	0.172	0.330	0.182	0.943	0.968	0.954	0.954	0.973	0.967	0.957	0.970
SE	0.027	0.024	0.022	0.023	0.020	0.024	0.032	0.031	0.012	0.012	0.010	0.010	0.009	0.012	0.011	0.010

Table 4.3: Quantitative evaluations using evaluation metrics CID and FSIMc. Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively when applied to dataset (Hwang et al., 2014) that includes captured image pairs with different cameras, camera setups, illumination and different image recolouring styles. Lower values are better with CID metric, and higher values are better with FSIMc metric (best viewed in colour and zoomed in).
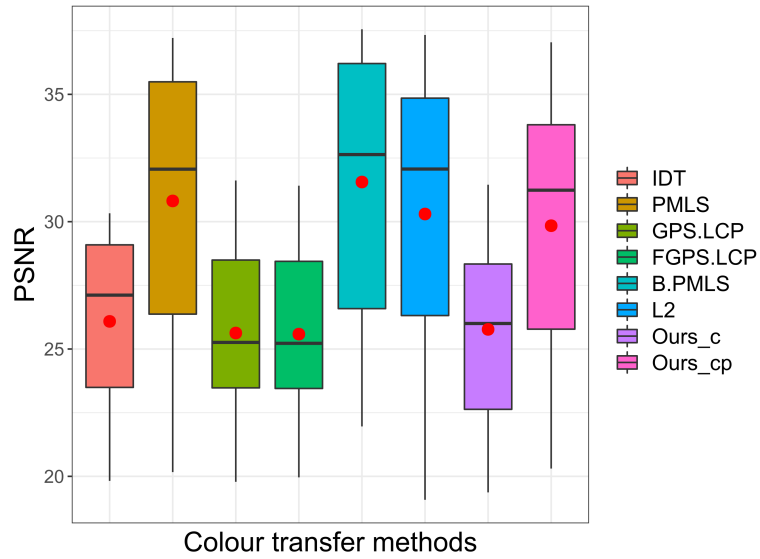
Figure 4.6: Comparing `PCT_OT` (colour only `Our_c`, colour and position `Our_cp`) with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Detailed quantitative results in Table 4.2 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).
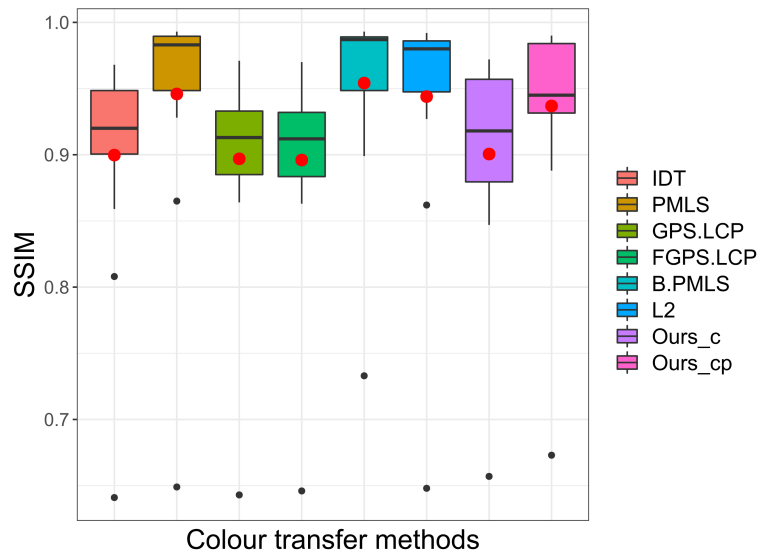
Figure 4.7: Comparing `PCT_OT` (colour only `Our_c`, colour and position `Our_cp`) with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Detailed quantitative results in Table 4.2 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).
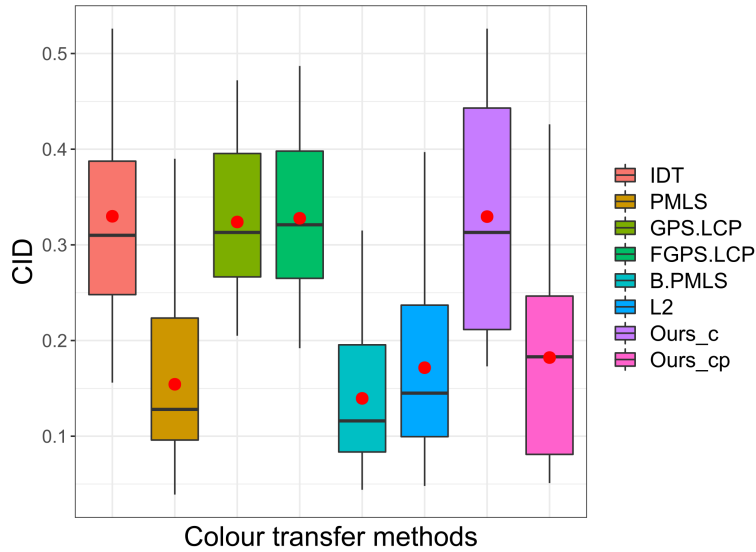
Figure 4.8: Comparing `PCT_OT` (colour only `Our_c`, colour and position `Our_cp`) with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Detailed quantitative results in Table 4.3 are summarized in a box plot (lower values are better, best viewed in colour and zoomed in).
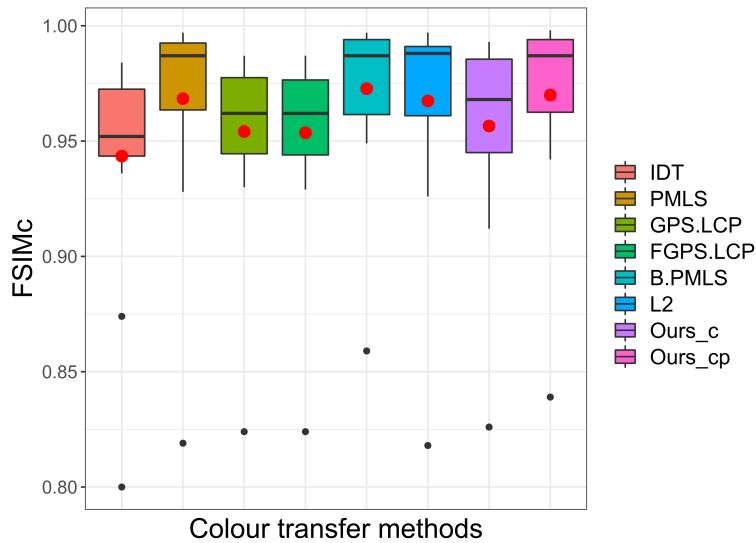
Figure 4.9: Comparing `PCT_OT` (colour only `Our_c`, colour and position `Our_cp`) with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Detailed quantitative results in Table 4.3 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

4.5 Conclusion

Several contributions to colour transfer with Optimal Transport have been made in this chapter, showing comprehensive quantitative and qualitative evaluations with leading state of the art colour transfer methods. In particular, first, neighborhoods of pixels (patches) are used with Optimal Transport algorithm in high dimensional space, and second, spatial information as well as colour content of pixels are both encoded in the high dimensional feature vectors. This original construction implies a new reconstruction step since each recoloured pixel benefits from the contribution of several estimated candidates using an averaging solution which allows denoising and artifact removal as well as smoothed colour transfer. One limitation of our proposed method is that it can create shadow artifacts when large changes exist between target and source images. In the next chapter, we tackle this problem and use a motion estimation technique (SIFT flow) to estimate spatial correspondences and encode them as well as colour content of pixels in the high dimensional feature vectors making Optimal Transport implicitly account for correspondences, then we investigate whether this formulation can improve the colour transfer results between images that may exhibit large motion changes between them. We further introduce smoothing as part of the iterative algorithms for solving Optimal Transport.

PATCH-BASED COLOUR TRANSFER USING SIFT FLOW

5.1 Introduction

Computing the colour transfer function considering only the colour information does not take into account the fact that coherent colours should be transferred to neighboring pixels, which can create results with blocky artifacts emphasizing JPEG compression blocks or increase noise. In Chapter 4, we tackled this problem and proposed the Patch-based Colour Transfer (PCT_OT) approach that encodes overlapping neighborhoods of pixels, taking into account both their colour and pixel positions (Alghamdi et al., 2019). The PCT_OT algorithm not only shows improvements over the state of the art methods but also shows limitations by creating shadow artifacts when there are large changes between target and source images. The mapping that is used in PCT_OT is Optimal Transport, and by definition OT does not take into account any available correspondences $\{(x_i, y_i)\}$ when computing the optimal solution and only has access to separate datasets $\{x_i\}$ and $\{y_j\}$. Moreover, the OT solution is invertible and an increasing function by definition but not smooth. In this chapter, we tackle these problems and propose a new colour transfer method with Optimal Transport to transfer the colour of a source image to match the colour of a target image of the same scene that may exhibit large motion changes between images. We accomplish this by improving PCT_OT method by first improving the data preparation step for defining patches thanks to SIFT flow (Liu et al., 2011). We estimate motions between images using the SIFT flow approach and incorporate the estimated spatial correspondence
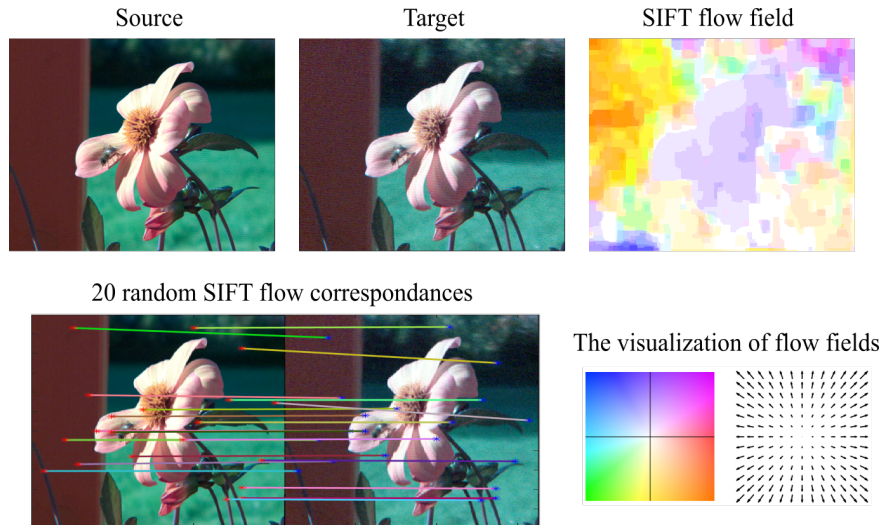
Figure 5.1: Correspondences are estimated using the SIFT flow approach that matches densely sampled, pixel-wise SIFT features between the target and source image. The flow fields on the top right show the motion changes between the two images. On the bottom right, flow field visualization is shown based on the code in (Baker et al., 2007): each pixel denotes a flow vector where the orientation and magnitude are represented by the hue and saturation of the pixel, respectively.

information in the encoded overlapping neighborhoods of pixels. This new formulation makes OT *implicitly* takes into account correspondences when computing the optimal solution. Our second contribution is to introduce smoothing as part of the iterative algorithms for solving optimal transport namely Iterative Distribution Transport (IDT) and its variant the Sliced Wasserstein Distance (SWD). Experiments show quantitative and qualitative improvements over the previous state of the art colour transfer methods, including the PCT_OT method. The work presented in this chapter has been published in (Alghamdi and Dahyot, 2020b).

5.2 PCT_OT with SIFT Flow

When target and source images are from the same scene, correspondences can be found to guide the process of recolouring. Indeed, colour transfer techniques often employ correspondences to harmonize colours across a video sequence or across multiple view images such as in image stitching or image mosaicing tasks.

The SIFT flow algorithm (Liu et al., 2011) is well suited for matching densely sampled, pixel-wise SIFT features between the two images. SIFT flow adopts the computational
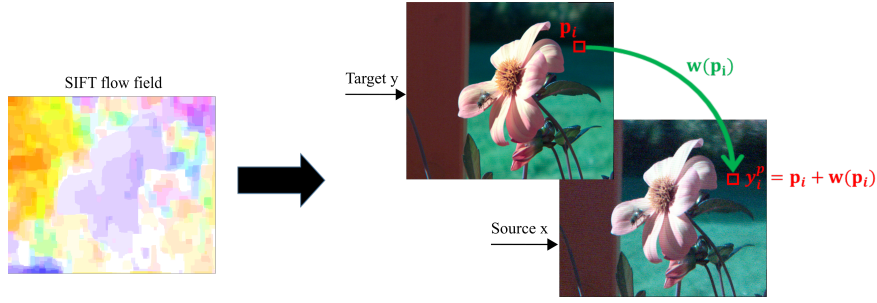
Figure 5.2: Spatial correspondences for the target image are calculated by adding the flow vectors generated by SIFT flow and the original grid coordinates of the target image and then concatenate the computed positions with the image colour channels.

framework of optical flow, but by matching SIFT descriptors instead of RGB colours or gradient that is used in optical flow (Brox et al., 2004). A discrete, discontinuity preserving, flow estimation algorithm is used to match the SIFT descriptors between two images. The use of SIFT features allows matching across different scene/object appearances and the discontinuity-preserving spatial model allows matching of objects located at different parts of the scene. Hence, SIFT flow is well suited for finding correspondences between images that exhibit colour variations and large motions, which is used in our proposed solution, Figure 5.1.

5.2.1 Combine colour and spatial correspondences

In the proposed method, the spatial information for the target image is calculated using the SIFT flow method, which estimates dense spatial correspondences (Liu et al., 2011), while in `PCT_OT` method (Chapter 4) the original pixel positions in the grid coordinate of the image are used. Using correspondences will allow colour transfer between images that contain moving objects and overcome the limitations in `PCT_OT`. More specifically, let y^p be the 2D pixel position of the target image that needs to be computed, and let $\mathbf{p} = (a, b)$ be the 2D grid coordinate of the target image and $\mathbf{w}(\mathbf{p}) = (\mathbf{w}_x(\mathbf{p}), \mathbf{w}_y(\mathbf{p}))$ be the flow vector at $\mathbf{p}$ computed using the SIFT flow method, then $y^p = \mathbf{p} + \mathbf{w}(\mathbf{p}) = (a + \mathbf{w}_x(\mathbf{p}), b + \mathbf{w}_y(\mathbf{p}))$ is the new pixel position for the target image that matches a pixel position in the source image, Figure 5.2. The target pixel's colour y^c and its new pixel position y^p are concatenated into a vector $y = (y^c, y^p)^T$ such that $\dim(y) = \dim(y^c) + \dim(y^p)$. The source image keeps the grid coordinate of the image as pixel positions, i.e $x^p = \mathbf{p}$ and similarly to the target image the pixel's

5 x 5 Patch i

5 x 5 Patch j

Overlapped region

y_{42}	y_{43}	y_{44}	y_{45}	y_{46}
y_{81}	y_{82}	y_{83}	y_{84}	y_{85}
y_{119}	y_{120}	y_{121}	y_{122}	y_{123}	y_{124}
y_{158}	y_{159}	y_{160}	y_{161}	y_{162}	y_{163}
y_{197}	y_{198}	y_{199}	y_{200}	y_{201}	y_{202}
y_{235}	y_{236}	y_{237}	y_{238}	y_{239}
y_{275}	y_{276}	y_{277}	y_{278}	y_{279}

Position (2D)

y^p_{119}	y^p_{120}	y^p_{121}	y^p_{122}	y^p_{123}
y^p_{158}	y^p_{159}	y^p_{160}	y^p_{161}	y^p_{162}
y^p_{197}	y^p_{198}	y^p_{199}	y^p_{200}	y^p_{201}
y^p_{235}	y^p_{236}	y^p_{237}	y^p_{238}	y^p_{239}
y^p_{275}	y^p_{276}	y^p_{277}	y^p_{278}	y^p_{279}

Colour (3D)

y^c_{119}	y^c_{120}	y^c_{121}	y^c_{122}	y^c_{123}
y^c_{158}	y^c_{159}	y^c_{160}	y^c_{161}	y^c_{162}
y^c_{197}	y^c_{198}	y^c_{199}	y^c_{200}	y^c_{201}
y^c_{235}	y^c_{236}	y^c_{237}	y^c_{238}	y^c_{239}
y^c_{275}	y^c_{276}	y^c_{277}	y^c_{278}	y^c_{279}

Point in 125D ((2+3) x 5 x 5)

$$y_j = [y^c_{119} \; y^p_{119} \; y^c_{120} \; y^p_{120} \; y^c_{121} \; y^p_{121} \cdots$$
$$\cdots y^c_{278} \; y^p_{278} \; y^c_{279} \; y^p_{279}]$$

Figure 5.3: An illustration shows the overlapping patches extracted from the target image, where for each patch, we concatenate all pixels information in the patch to create a point in a higher-dimensional space (best viewed in colour and zoomed in).

colour x^c and its pixel position x^p are concatenated into a vector $x = (x^c, x^p)^T$ such that $\dim(x) = \dim(x^c) + \dim(x^p)$.

5.2.2 Data normalisation

Since the colours have integer values from 0 to 255, and the spatial values can be anything depending on the size of the image, we normalize all the colour and position coordinates to lie between 0 and 255 to create a hypercube in $\mathbb{R}^N$, where $N = \dim(x) = \dim(x^c) + \dim(x^p)$, in order to produce consistent results regardless of the size of the image and better control parameters. We then stretch that space in the direction of the spatial coordinates by a weight factor w. This stretching makes the pixel positions significantly influence the estimated mapping more than the colour information in a way that forces the algorithm to match the colour of pixels that have same positions. Since we are focusing on transferring colour between images of the same scene, we know that the scenes are overlapped, and hence the more overlapped areas we have, the bigger the value of w we can set. Note, in case we have outlier correspondences, then the bigger value of w, the more influence of the outlier correspondences in the transfer process, as we will see in the experiment part. This value could be adjusted by the user, depending on the images being processed.
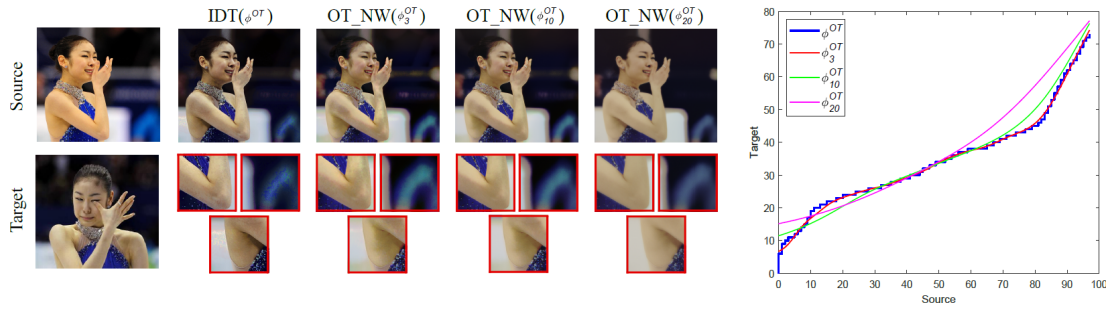
Figure 5.4: Results show the smoothed Optimal Transport solution using non-parametric Nadaraya-Watson (ϕ_h^{OT}) with different bandwidth values $h = \{3, 10, 20\}$. Nadaraya-Watson significantly reduces the grainy artifacts produced by the original Optimal Transport function (ϕ^{OT}), mapping the source patch projections to the target patch projections. The bigger h value, the more smoothed mapping. The results were processed without a post-processing step. Note that the graph is a zoom-in on 0-255 range pixel values (best viewed in colour and zoomed in).

5.2.3 Create patch vectors

In a similar way to PCT_OT (Chapter 4), we encode overlapping neighborhoods of pixels to preserve local topology information. Starting from the origin of the coordinate system of the images (upper left corner), we use a sliding window operation of window size $d \times d$ to extract overlapping patches. From each individual patch we create a high dimensional vector in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$. We apply this process to the source and target images to create patch vector sets $\{x_i\}$ and $\{y_j\}$ for each respectively, Figure 5.3.

5.3 Smoothing ϕ^{OT} with Nadaraya-Watson Estimator

As we mentioned in Section 4.2, the 1D solution ϕ^{OT} Eq. (4.2) has been used to tackle problems in multidimensional colour spaces. Of particular interest is the Iterative Distribution Transfer (IDT) algorithm for colour transfer proposed by Pitie et al. (2007) and the Sliced Wasserstein Distance (SWD) algorithm that follows from the iterative projection approach of IDT but computes the 1D solution ϕ^{OT} with quantile matching instead of cumulative histogram matching (Rabin et al., 2012; Bonneel et al., 2015). Optimal Transport does not provide an explicit expression of ϕ^{OT} but instead an estimated correspondence $(x, \phi^{OT}(x))$ for every data point x in the source dataset. We can use the OT estimated correspondences to compute a smoothed OT solution for colour

transfer using the Nadaraya Watson estimator (Section 2.2.3). Recall the general definition of Nadaraya Watson estimator assuming we are giving the correspondences $\{(x_i, y_i)\}_{i=1}^q$:

$$\mathbb{E}[y|x] = \int y \, p(y|x) \, dy = \int y \, \frac{p(y,x)}{p(x)} \, dy \simeq \frac{\sum_{i=1}^q y_i \, W_h(x - x_i)}{\sum_{i=1}^q W_h(x - x_i)} = \phi_h^{NW}(x) \qquad (5.1)$$

With this form, NW can be seen as a locally weighted average of $\{y_i\}_{i=1}^q$, using a kernel as a weighting function where the bandwidth h is the hyperparameter or scale parameter of the kernel, the larger the value of h the more ϕ_h^{NW} gets smoothed. We propose to smooth the 1D ϕ^{OT} computed in IDT or SWD by using the non-parametric Nadaraya Watson estimator: at each iteration k, following the step of projecting $\{x_i\}_{i=1}^n$ and $\{y_j\}_{j=1}^m$ originally in high dimensional space to a 1D subspace creating two 1D datasets $\{u_i\}_{i=1}^n$ and $\{v_j\}_{j=1}^m$ in the projective space (step 6 in Algorithm 4). We then calculate the optimal map ϕ^{OT} to estimate OT correspondences $\{(u_i, \phi^{OT}(u_i))\}_{i=1}^n$ (step 7 in Algorithm 4). We then choose a subset of $K = 2000$ number of correspondences $\{\phi^{OT}(u_i)\}_{i=1}^K$ spreading out the entire range of the projected dataset in the 1D subspace. Next, we feed the estimated OT correspondences $\{(u_i, \phi^{OT}(u_i))\}_{i=1}^K$ to the NW estimator to compute a smoother OT solution, denoted as ϕ_h^{OT}, defined as follows:

$$\phi_h^{OT}(u) = \frac{\sum_{i=1}^K \phi^{OT}(u_i) \, W_h(u - u_i)}{\sum_{i=1}^K W_h(u - u_i)} \qquad (5.2)$$

Figure 5.4 illustrates the effect of computing smoother OT solutions using NW with different bandwidth values on colour transfer compared with the original OT solution computed using IDT algorithm (Pitie et al., 2007). Optimal Transport solutions are suitable in situations where the function that we need to estimate must satisfy necessary side conditions, such as being strictly increasing. The non-parametric NW estimator on top of the OT solution can provide the smoothness required in the estimated function. In addition, one of the important characteristics of using OT and NW estimators is that they do not assume an explicit expression controlled by parameters on the regression function, which makes them directly employable. In the following sections, we apply OT and NW smoothing in the relevant context of colour transfer where the estimated function satisfies the condition of being an increasing function.

Algorithm 3: Our pipeline for colour transfer between overlapped scene images using SIFT flow and smoothed OT solution

---

1: **Input:** Source & Target images
2: **Output:** recoloured Source image
3: Formulate the Source image & the Target image as colour samples datasets $\{x_i\}_{i=1}^n$ & $\{y_j\}_{j=1}^m$ of r.v. x and $y \in \mathbb{R}^N$, respectively. For example in RGB colour transfer, $N = 3$ and $x_i = (r_i, g_i, b_i)$ where r_i, g_i, b_i are the red, green and blue components of source pixel number i
4: **if** Combining colour and spatial information **then**
5:     Use SIFT flow method to compute new positions y^p for the target
6:     Combine target colours y^c with the new computed positions y^p into a vector $y = (y^c, y^p)^T$ in $\mathbb{R}^N$ where $N = \dim(y) = \dim(y^c) + \dim(y^p)$, same process applied to the source colours but with positions in the grid coordinate of the image
7:     Normalize all the colour and position coordinates to lie between 0 and 255
8:     Choose a window size $d \times d$ and use the combined colour and positions features (i.e. $x = (x^c, x^p)^T$ & $y = (y^c, y^p)^T$) to create overlapping patch vectors $\{x_i\}$ & $\{y_j\}$ in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$
9: **else if** colour only **then**
10:     Choose a window size $d \times d$ and create colour only (i.e. $x = (x^c)^T$ & $y = (y^c)^T$) overlapping patch vectors $\{x_i\}$ & $\{y_j\}$ in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$
11: **end if**
12: Giving patch vectors $\{x_i\}$ and $\{y_j\}$, compute smoothed Optimal Transport mapping with Nadaraya-Watson step (Algorithm 4)
13: Merge recoloured candidates from the overlapping patches $\{x_i^{(\infty)}\}$ to create the output recoloured source image

---

Algorithm 4: Smoothed iterative projection-based OT transform

1: **Input:** datasets Source $\{x_i\}_{i=1}^n$ and Target $\{y_j\}_{j=1}^m$ in $\mathbb{R}^N$

2: **Initialisation:** $k \leftarrow 0$ and $\forall i,\ x_i^{(0)} \leftarrow x_i$

3: **repeat**

4:     Generate D random unit vectors in $\mathbb{R}^N$ stored in matrix $\mathrm{R} = [e_1, ..., e_D]$

5:     **for** $r = 1$ to D **do**

6:         Compute projections $\forall i,\ u_i^{(k)} = e_r^T x_i^{(k)}$ and $\forall j,\ v_j = e_r^T y_j$

7:         Compute 1D transfer functions $\forall i, \phi^{OT^{(r)}}(u_i^{(k)})$ such that $\phi^{OT^{(r)}}$ is the OT solution (Eq. (4.2) or Eq. (2.11)) for the projection r

8:         Choose subset of K correspondences $\{(u_i^{(k)}, \phi^{OT^{(r)}}(u_i^{(k)}))\}_{i=1}^K$ from $\{u_i^{(k)}, \phi^{OT^{(r)}}(u_i^{(k)})\}_{i=1}^n$

9:         Giving the correspondences $\{(u_i^{(k)}, \phi^{OT^{(r)}}(u_i^{(k)}))\}_{i=1}^K$, compute a smoothed OT $\forall i$ in source dataset, $\phi_h^{OT^{(r)}}(u_i^{(k)})$ such that $\phi_h^{OT^{(r)}}$ is the NW solution (Eq. (5.2))

10:    **end for**

11:    Remap the source dataset, $\forall i$

$$\Phi(x_i^{(k)}) = \mathrm{R} \begin{pmatrix} \phi_h^{OT^{(1)}}(u_i^{(k)}) \\ \vdots \\ \phi_h^{OT^{(D)}}(u_i^{(k)}) \end{pmatrix} = \sum_{r=1}^D (\phi_h^{OT^{(r)}}(u_i^{(k)}))\, e_r$$

12:    Update source dataset $x_i^{(k+1)} \leftarrow (\lambda - 1)(x_i^{(k)}) + \lambda(\Phi(x_i^{(k)}))$

13:    $\mathrm{R} \leftarrow$ Random rotation of R

14:    $k \leftarrow k + 1$

15: **until** convergence on all marginals for every possible rotation (noted $k \equiv \infty$)

16: **Result:** The final one-to-one mapping Φ in $\mathbb{R}^N$ is given by $\forall i,\ x_i \mapsto \Phi(x_i) = x_i^{(\infty)}$

5.4   Experimental Assessment

We use the evaluation tools presented in Section 4.4.4 which consist of using numerical metrics, summary descriptive statistics, statistical significance tests, and visual analysis. We provide extensive quantitative and qualitative evaluations of our approach, noted by OT_NW, with comparisons to various state of the art methods of colour transfer noted by L2 (Grogan and Dahyot, 2019), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), PMLS (Hwang et al., 2014), IDT (Pitie et al., 2007) and PCT_OT (Chapter 4). In these evaluations, we use image pairs with similar contents from an existing dataset (Section 4.4.3) provided by Hwang et al. (2014). Note that the results using PMLS were provided by the authors (Hwang et al., 2014). We also compared with two other recent techniques (Xia et al., 2017; Park et al., 2016) that account for correspondences into their framework, but PMLS has been shown to perform better than these two (Grogan and Dahyot, 2019), so PMLS is the one reported here with (Grogan and Dahyot, 2019; Bellavia and Colombo, 2018) as algorithms that incorporate correspondences in their methodologies. Moreover, we compared with (Bonneel et al., 2016; Ferradans et al., 2013) that do not take into account correspondences, but IDT also has been shown to be superior (Grogan and Dahyot, 2019), hence IDT is the one reported here for ease of comparison.

5.4.1   Colour space and parameters settings

Colour space and patch size: We use the RGB colour space where each pixel is represented by its 3D RGB colour values and its 2D spatial position. Similarly to PCT_OT (Chapter 4), we use patch size of 5×5 to capture enough of a pixel's neighborhoods. Our patches with combined colour and spatial features create a vector in 125 dimensions ($5 \times 5 \times 5$) for the RGB colours (3D) and position components (2D), Figure 5.3. Similarly to PCT_OT, other colour spaces could be considered, but since the proposed method transfers complete statistics (pdfs), this has no implications and the method would function regardless of the colour space chosen.

The value of the stretching parameter w**:**   We evaluated the impact of using larger or smaller values of the stretching parameter w (Section 5.2.2) on the estimated mapping and in the presence of outlier correspondences. Figure 5.5 (A) shows colour
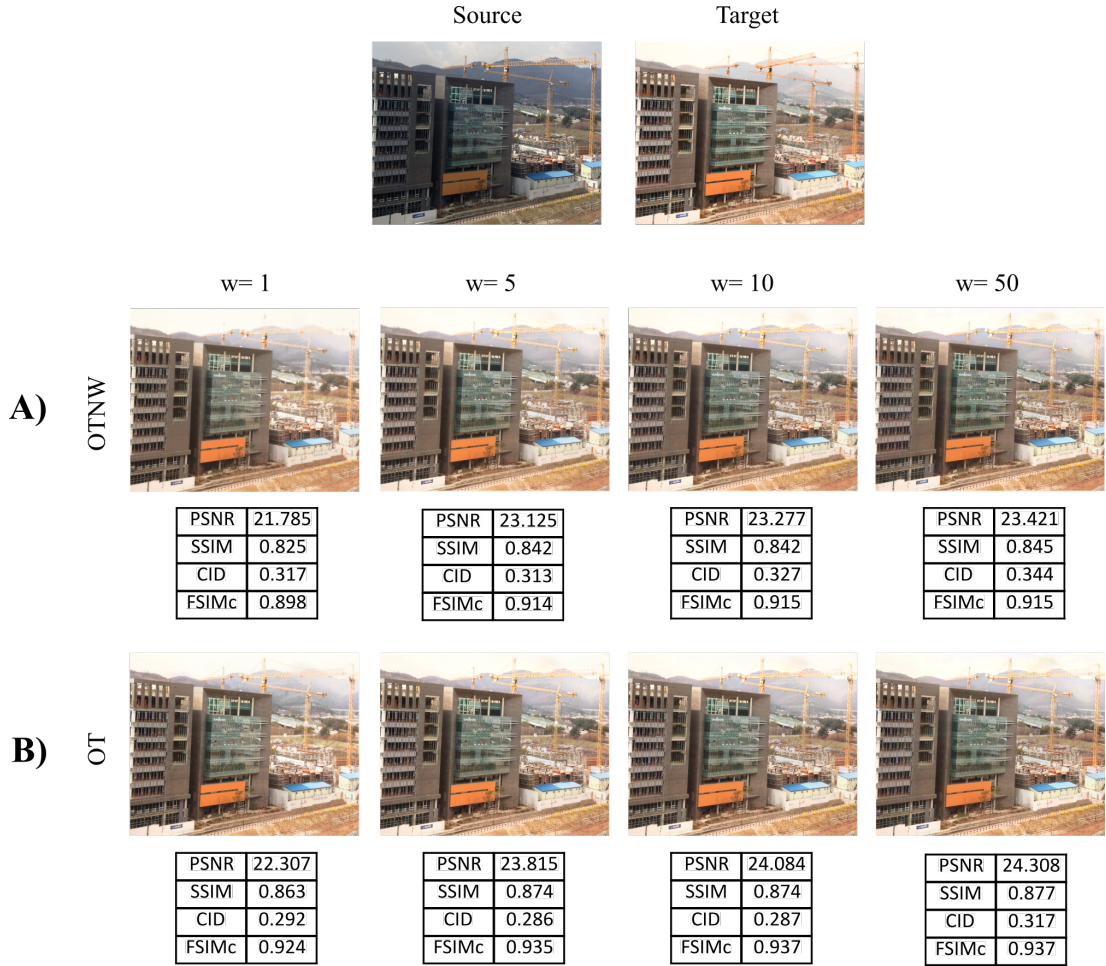
Figure 5.5: This figure shows the impact of using smaller or bigger values of the stretching parameter w. The top row is the source and target that need to be matched. The second row (A) shows results of our method `OT_NW` when using different values of $w = \{1, 5, 10, 50\}$. The third row (B) shows our method's results with motion flow but without the NW smoothing step. All results are shown with corresponding metrics results (best viewed in colour and zoomed in).

transfer results when using different values $w = \{1, 5, 10, 50\}$ with corresponding metrics results. When we increase the value of w, we get better metrics scores, but also artifacts start appearing in places where there are outlier spatial correspondences. The higher the value of w, the higher influence of pixel positions on the estimated mapping. We found that setting $w = 10$ produces the best visual results.

Assessing the method with and without the NW smoothing step: Figure 5.5 (B) shows an example of the results with motion flow only and without the NW smoothing step. We see with every w value the metrics produce better scores with comparison to adding NW smoothing step such that bandwidth $h = 10$. We found that adding NW
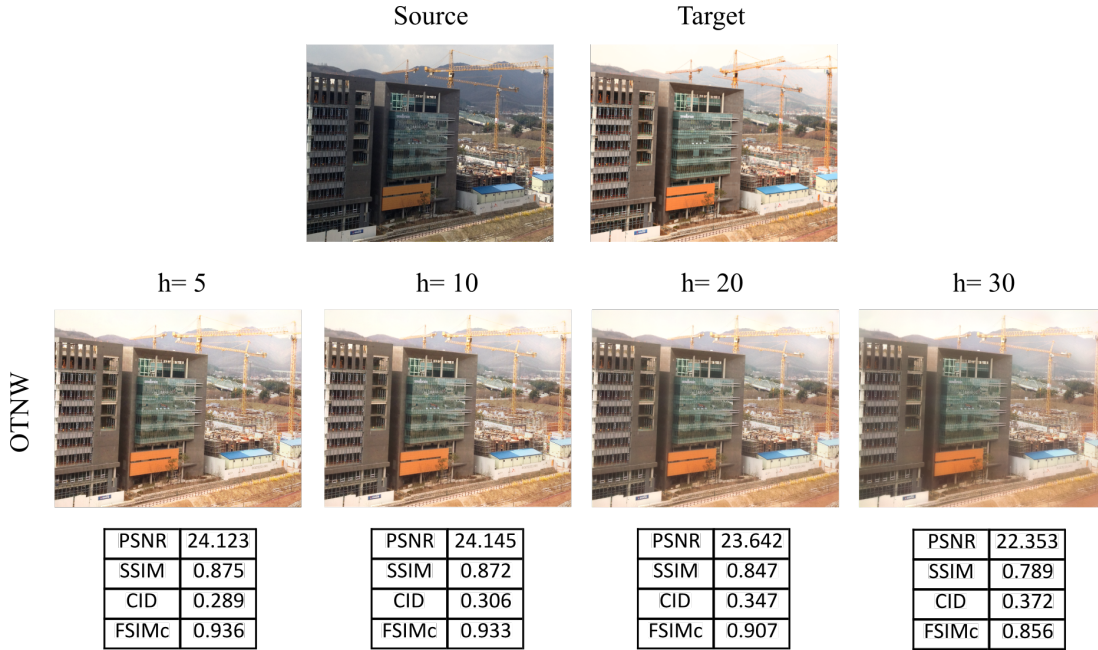
Figure 5.6: This figure shows the impact of the NW bandwidth h in our method `OT_NW` while fixing $w = 10$. The higher the values, the more smoothed colour transfer but with less accurate colour transfer. All results are shown with corresponding metrics results.

smoothing step acts as a regularization for the OT mapping and hence allows smooth colour transfer between images as well as artifact removal, but at the expense of an accurate colour transfer, hence we experimented with different values of h and we found that $h = 10$ is a good balance, producing best results. Figure 5.6 shows more results for various h values, while fixing $w = 10$.

5.4.2 Experimental results

Quantitative evaluation: The numerical results of each metric are shown in Tables 5.2 to 5.5, along with statistical measures and plots as shown in Figures 5.7 to 5.10, respectively. The purpose of the box plots is to visualize differences among methods and to show how close our method is to the state of the art algorithms. By looking at Tables 5.2 and 5.3, we can see that PSNR and SSIM metrics show that `OT_NW` method competes very well with the top methods `PMLS` and `L2`, and improves the performance over the `PCT_OT` method (Chapter 4). This conclusion is confirmed by the box plot results in Figure 5.7 and Figure 5.8, as we can see the mean (the mean shown as red dots in the plots), median (shown as horizontal black lines) and deviation of

`OT_NW` method are similar to the corresponding ones of the top methods `PMLS` and `L2`, while they indicate a better performance of the `OT_NW` method over the `PCT_OT` method. Moreover, the box plot of SSIM metric (Figure 5.8) shows that the median is higher than the mean in the `OT_NW`, `PMLS` and `L2`, which indicates left-skewness in favor of higher performance.

On the other hand, the numerical results of Tables 5.4 and 5.5 (CID and FSIMc metrics) are not certain in distinguishing the performance among the four methods. However, the box plots can provide us with better visualization. From the box plot results in Figure 5.9 and Figure 5.10, the CID and the FSIMc metrics show that `OT_NW` method still competes very well with the top methods `PMLS` and `L2`. With respect to comparing with the `PCT_OT` method, the FSIMc shows similar performance between the `OT_NW` and `PCT_OT` methods, while the CID metric indicates that the `OT_NW` outperforms the `PCT_OT`, at which there is a right-skewness in favor of higher performance (note that in CID metric, lower values indicate higher performance). In conclusion, the quantitative metrics show that our algorithm with Nadaraya Watson `OT_NW` performs similarly with top methods `PMLS`, `L2`, and improves the performance over the `PCT_OT` (Chapter 4), and outperforms the rest of the state of the art algorithms.

According to statistical significance analysis, using Kruskal-Wallis Rank Sum test (Section 4.4.4), we found that the P-values for all metrics are greater than the significance level $\alpha = 0.05$, which indicates that with confidence level $= 95\%$ there is no significant difference between `PMLS`, `L2`, `PCT_OT` and `OT_NW` method, Table 5.1.

Qualitative evaluation: Figure 5.11 provides qualitative results. For clarity, the results are presented in image mosaics, created by switching between the target image and the transformed source image column wise. If the colour transfer is accurate, the resulting mosaic should look like a single image (ignoring the small motion displacement between source and target images), otherwise column differences appear. As can be noted, our approach `OT_NW` with Nadaraya Watson step is successful in removing the column differences, improving the performance over the `PCT_OT` (Chapter 4).

Figure 5.12 shows more qualitative results. We can see that our method `OT_NW` is robust to registration errors, unlike `PMLS` which introduces visual artifacts if the input images are not registered correctly. Note that the accuracy of the PSNR, SSIM, CID and FSIMc metrics relies on the fact that the input images are registered correctly, and

Metrics	P-values
PSNR	0.9134
SSIM	0.5936
CID	0.9489
FSIMc	0.9895

Table 5.1: This table shows results of Kruskal-Wallis Rank Sum statistical test that tests if there are significant differences between `OT_NW` results and the top performance methods, namely, `PMLS`, `L2` and `PCT_OT`. The null hypothesis H_o states that all medians are equal, the alternative hypothesis H_1 states that at least one median is different. The confidence level = 95%, and the significance level $\alpha = 0.05$. If the P-value $< \alpha$, the null hypothesis is rejected and alternative hypothesis is accepted.

if this is not the case, then these metrics may not accurately capture all artifacts, and these can be visualized clearly with `PMLS` method in Figures 5.11 and 5.12 (best viewed zoomed in). `PCT_OT` creates shadow artifacts when there are large changes between target and source images (Figure 5.12, in example "building"), while `OT_NW` method that incorporates correspondences in the overlapped patches can correctly transfer colours between images that contain significant spatial differences and alleviates the shadow artifacts, as can be seen in Figure 5.12 with examples "illum", "mart" and "building". Moreover, `OT_NW` visually outperforms `PMLS` and `L2` in keeping the original structure of the source image after the transformation, and due to the Nadaraya Watson smoothing step in our algorithm, `OT_NW` allows us to create a smoother colour transfer result, and can also alleviate JPEG compression artifacts and noise (cf. Figure 5.12 for comparison between all methods).

	PSNR ↑						
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 5x5
Gangnam1	25.354	35.725	24.048	23.936	35.358	31.479	33.565
Gangnam2	27.116	36.553	25.952	25.944	35.524	35.502	33.627
Gangnam3	22.372	35.007	21.908	21.913	33.284	26.393	28.217
Illum	19.822	20.167	19.785	19.960	19.079	20.306	20.858
Building	20.554	22.634	22.736	22.769	20.499	25.019	24.039
Playground	27.184	27.835	25.501	25.436	27.647	28.482	28.491
Flower1	24.238	26.981	23.765	23.706	26.857	25.186	27.158
Flower2	25.417	25.760	25.259	25.223	25.772	26.373	26.497
Tonal1	30.082	37.215	31.617	31.413	37.332	37.044	37.151
Tonal2	27.992	31.508	25.062	25.087	31.356	32.049	31.579
Tonal3	29.575	36.246	28.136	28.065	36.644	33.793	35.014
Tonal4	28.605	34.521	28.852	28.848	34.344	33.819	35.320
Tonal5	30.330	35.260	29.580	29.448	34.303	36.437	36.616
Mart	22.747	24.742	23.183	23.196	24.450	24.509	25.189
Sculpture	29.884	32.062	29.037	28.820	32.067	31.237	32.735
Mean	26.085	30.814	25.628	25.584	30.301	29.842	30.404
SE	0.905	1.459	0.841	0.821	1.518	1.306	1.291

Table 5.2: Comparing OT_NW with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
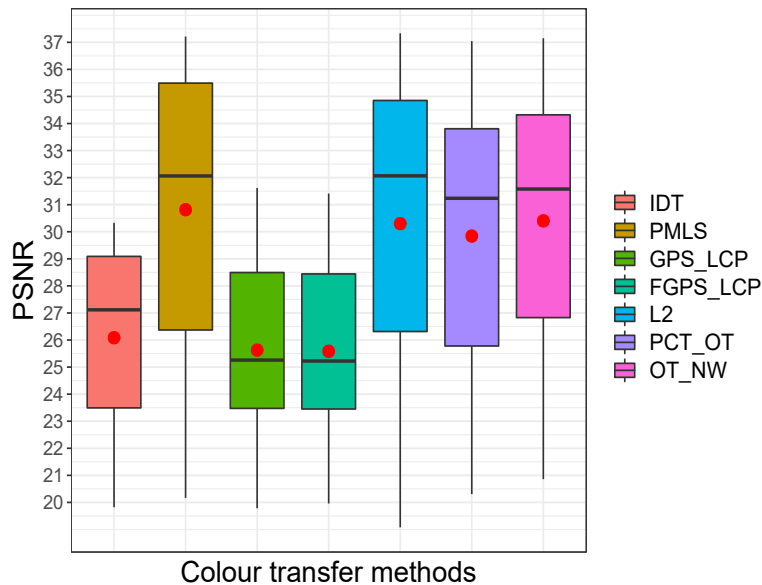
Figure 5.7: Comparing OT_NW with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Detailed quantitative results in Table 5.2 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

	SSIM ↑						
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 5x5
Gangnam1	0.900	0.992	0.892	0.891	0.990	0.964	0.973
Gangnam2	0.920	0.993	0.909	0.909	0.986	0.980	0.976
Gangnam3	0.859	0.991	0.873	0.864	0.980	0.930	0.959
Illum	0.641	0.649	0.643	0.646	0.648	0.673	0.687
Building	0.808	0.865	0.864	0.863	0.862	0.888	0.875
Playground	0.920	0.940	0.878	0.876	0.939	0.939	0.943
Flower1	0.909	0.967	0.913	0.912	0.966	0.926	0.959
Flower2	0.901	0.928	0.894	0.894	0.927	0.933	0.939
Tonal1	0.953	0.988	0.971	0.970	0.987	0.988	0.991
Tonal2	0.968	0.987	0.926	0.926	0.986	0.988	0.986
Tonal3	0.962	0.992	0.947	0.946	0.992	0.987	0.990
Tonal4	0.944	0.983	0.932	0.932	0.983	0.981	0.985
Tonal5	0.965	0.986	0.953	0.954	0.985	0.990	0.991
Mart	0.904	0.957	0.925	0.925	0.956	0.941	0.954
Sculpture	0.942	0.971	0.934	0.932	0.972	0.945	0.974
Mean	0.900	0.946	0.897	0.896	0.944	0.937	0.946
SE	0.022	0.023	0.020	0.020	0.023	0.020	0.020

Table 5.3: Comparing OT_NW with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
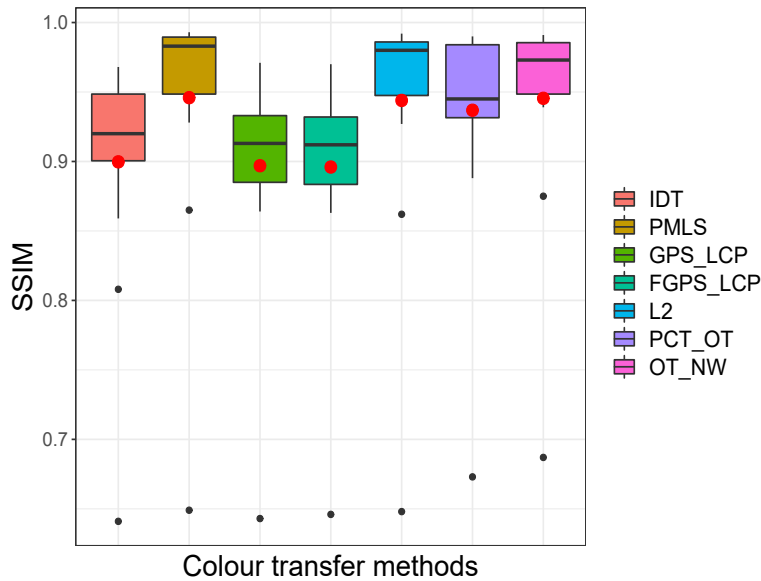
Figure 5.8: Comparing OT_NW with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Detailed quantitative results in Table 5.3 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

	CID ↓						
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 5x5
Gangnam1	0.252	0.040	0.226	0.222	0.048	0.085	0.088
Gangnam2	0.268	0.039	0.291	0.292	0.089	0.068	0.109
Gangnam3	0.496	0.108	0.472	0.487	0.193	0.261	0.267
Illum	0.386	0.390	0.395	0.396	0.397	0.377	0.376
Building	0.374	0.228	0.313	0.321	0.249	0.183	0.275
Playground	0.440	0.238	0.443	0.471	0.254	0.209	0.221
Flower1	0.389	0.163	0.396	0.400	0.174	0.285	0.194
Flower2	0.337	0.245	0.322	0.323	0.266	0.218	0.201
Tonal1	0.310	0.101	0.285	0.308	0.111	0.097	0.063
Tonal2	0.288	0.128	0.351	0.347	0.145	0.099	0.118
Tonal3	0.244	0.079	0.294	0.294	0.081	0.077	0.079
Tonal4	0.240	0.108	0.248	0.238	0.107	0.065	0.065
Tonal5	0.156	0.091	0.205	0.192	0.092	0.051	0.067
Mart	0.526	0.219	0.405	0.402	0.225	0.426	0.249
Sculpture	0.242	0.137	0.213	0.224	0.143	0.232	0.120
Mean	0.330	0.154	0.324	0.328	0.172	0.182	0.166
SE	0.027	0.024	0.022	0.023	0.024	0.031	0.025

Table 5.4: Comparing OT_NW with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (lower values are better, best viewed in colour and zoomed in).
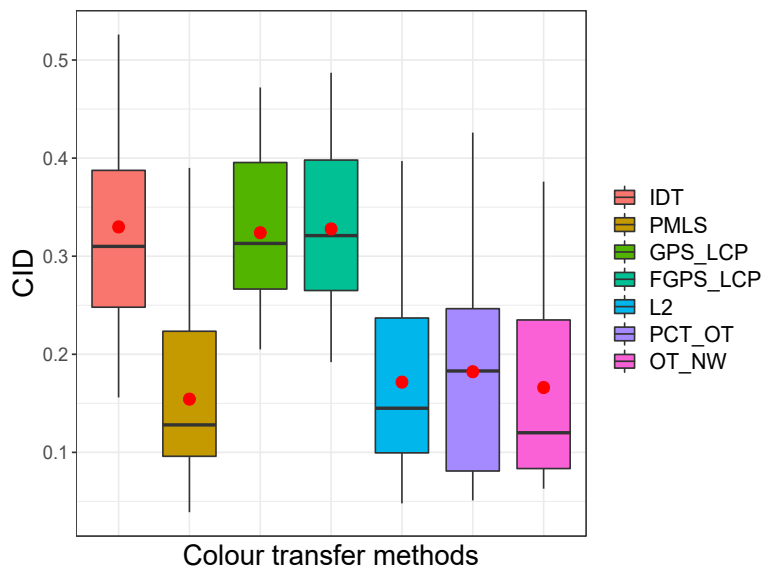
Figure 5.9: Comparing OT_NW with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Detailed quantitative results in Table 5.4 are summarized in a box plot (lower values are better, best viewed in colour and zoomed in).

	FSIMc ↑						
.	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 5x5
Gangnam1	0.936	0.986	0.944	0.943	0.985	0.972	0.979
Gangnam2	0.952	0.992	0.962	0.962	0.988	0.990	0.986
Gangnam3	0.946	0.992	0.962	0.961	0.990	0.987	0.982
Illum	0.800	0.819	0.824	0.824	0.818	0.839	0.844
Building	0.874	0.928	0.930	0.929	0.926	0.942	0.937
Playground	0.950	0.958	0.933	0.932	0.955	0.956	0.960
Flower1	0.954	0.975	0.968	0.967	0.976	0.971	0.977
Flower2	0.941	0.950	0.945	0.945	0.949	0.954	0.956
Tonal1	0.964	0.997	0.986	0.986	0.997	0.998	0.997
Tonal2	0.984	0.993	0.973	0.973	0.992	0.993	0.992
Tonal3	0.979	0.997	0.984	0.983	0.997	0.997	0.995
Tonal4	0.966	0.989	0.972	0.973	0.990	0.995	0.994
Tonal5	0.980	0.994	0.987	0.987	0.993	0.998	0.997
Mart	0.946	0.969	0.960	0.959	0.967	0.969	0.970
Sculpture	0.980	0.987	0.982	0.980	0.988	0.988	0.987
Mean	0.943	0.968	0.954	0.954	0.967	0.970	0.970
SE	0.012	0.012	0.010	0.010	0.012	0.010	0.010

Table 5.5: Comparing OT_NW with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
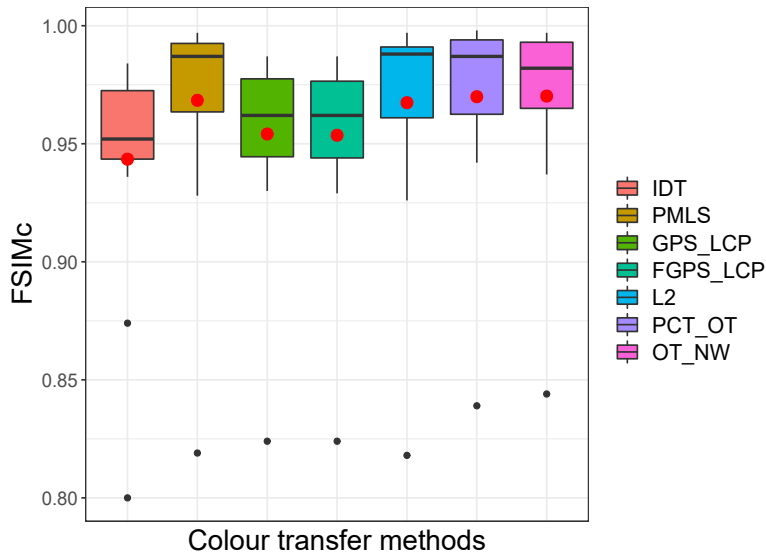
Figure 5.10: Comparing OT_NW with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Detailed quantitative results in Table 5.5 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).
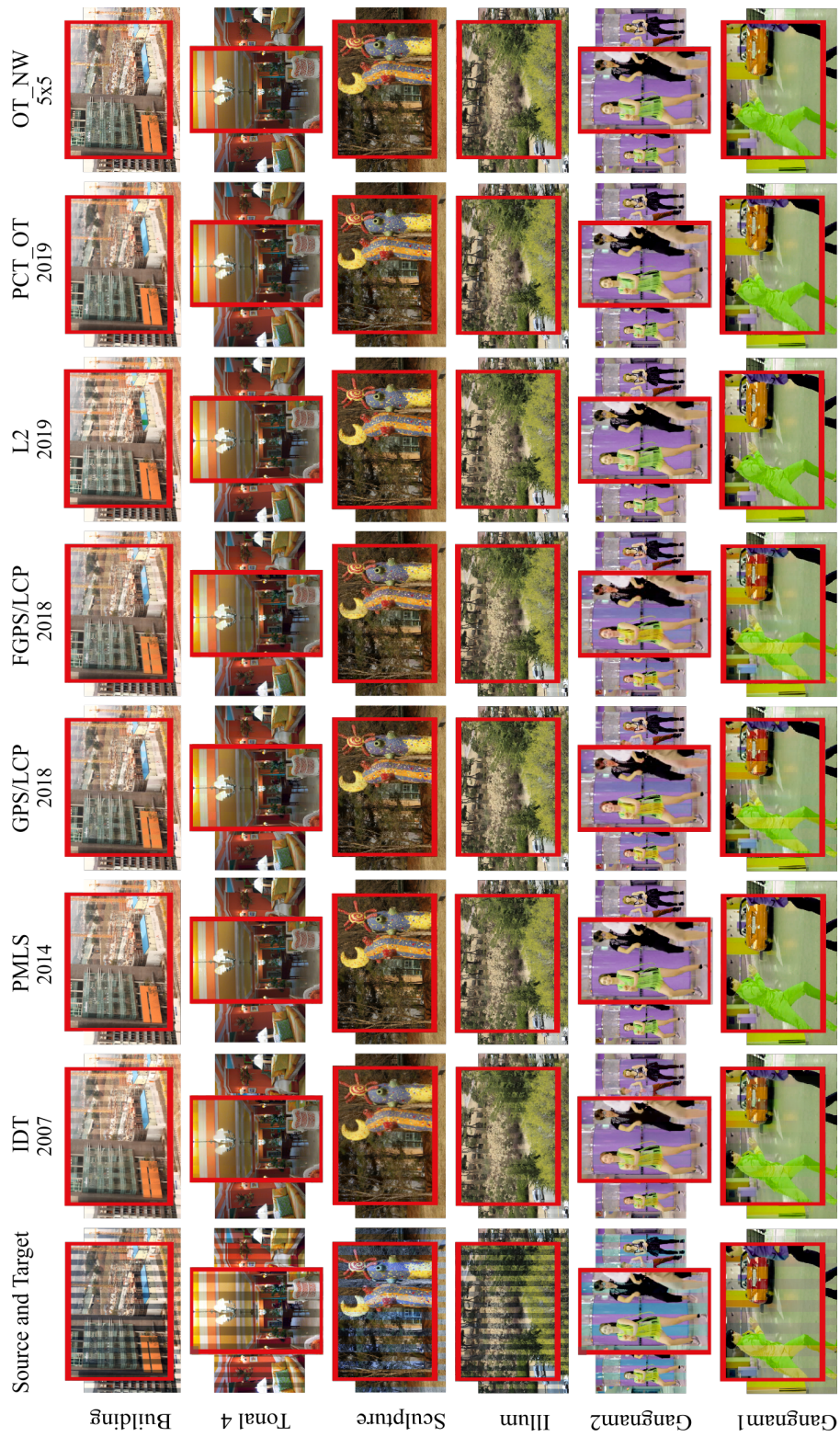
Figure 5.11: A close up look at some of the results generated using the IDT (Pitie et al., 2007), PMLS (Hwang et al., 2014), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), L2 (Grogan and Dahyot, 2019), PCT_OT (Alghamdi et al., 2019) and our algorithm OT_NW. The results are presented in image mosaics, if the colour transfer is accurate, the resulting mosaic should look like a single image, otherwise column differences appear (best viewed in colour and zoomed in).
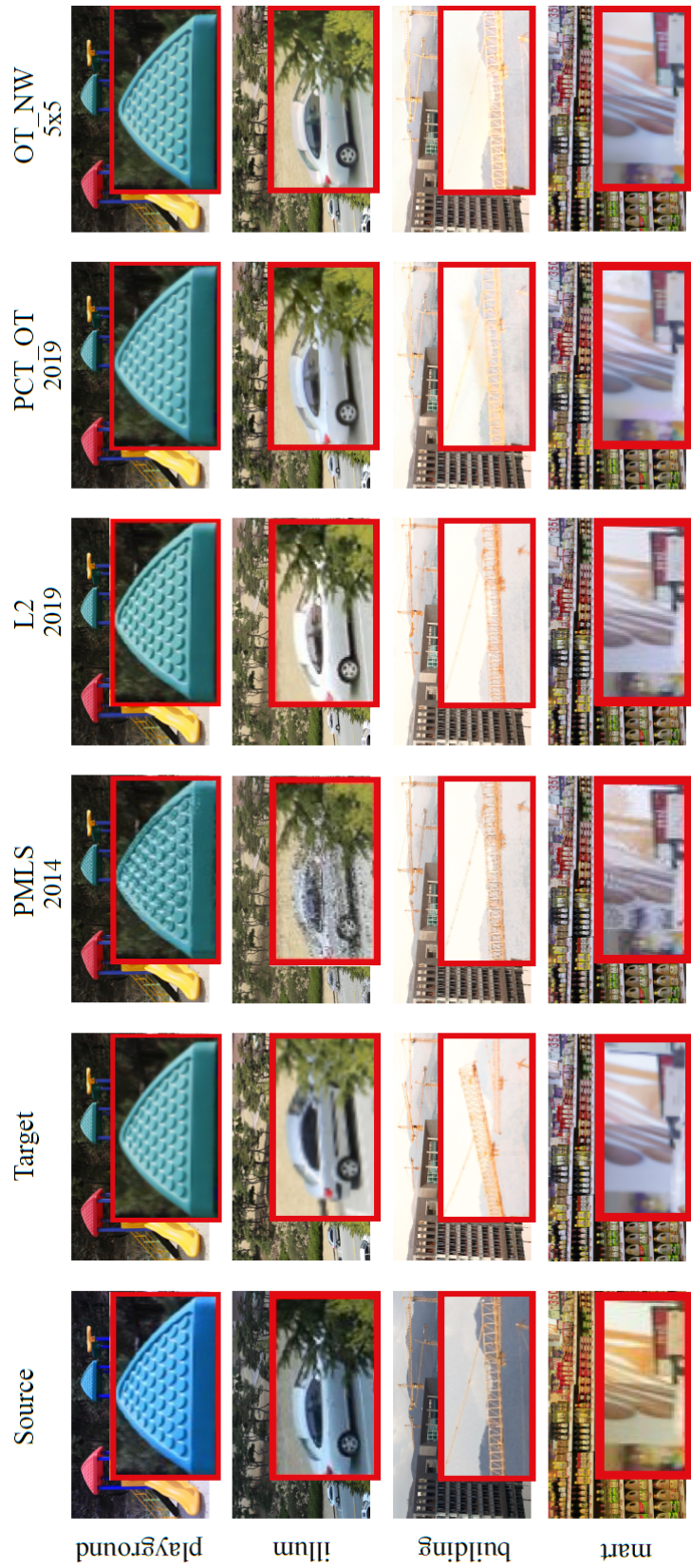
Figure 5.12: A close up look at some of the results generated using the PMLS (Hwang et al., 2014), L2 (Grogan and Dahyot, 2019), PCT_OT (Alghamdi et al., 2019) and our algorithm OT_NW (best viewed in colour and zoomed in).

5.5 Conclusion

In this chapter, several contributions to colour transfer with Optimal Transport have been made, showing comprehensive quantitative and qualitative assessments with leading state of the art methods of colour transfer. In particular, first, SIFT flows based spatial correspondence information as well as colour content of pixels are both encoded in the high dimensional feature vectors, making OT implicitly taking into account correspondences when computing the optimal solution, and second, we introduced smoothing as part of the iterative algorithms for solving optimal transport, namely Iterative Distribution Transport (IDT) and its variant the Sliced Wasserstein Distance (SWD). The algorithm allows denoising, artifact removal as well as a smoother colour transfer between images that may contain large motion changes. In the next chapter, we introduce the Nadaraya-Watson estimator with the 1D iterative projection approach taking advantage of the correspondences explicitly in its formulation, and investigate whether this method can solve the colour transfer problem.

ITERATIVE NADARAYA-WATSON DISTRIBUTION TRANSFER FOR COLOUR GRADING

6.1 Introduction

Light field (LF) is one application of colour transfer, which has recently become a significant research topic (Matysiak et al., 2020; Grogan and Smolic, 2019), Figure 6.1. The increase of this technology's popularity is due to the availability of LF camera devices such as lenslet cameras. These cameras extract sub-aperture images (SAI) with an extensive depth of field representing different viewpoints of the same scene. However, the extracted views suffer from several artifacts, such as colour discrepancies. When the source and target images are from the same scene such as light field images, correspondences can be found to guide the process of recolouring. The SIFT flow algorithm (Liu et al., 2011) (Section 5.2) is well suited for matching densely sampled pixel-wise SIFT features between the two images and is used in our proposed pipeline to extract correspondences (cf. Figure 6.3). This chapter proposes a new method with Nadaraya-Watson that maps one N-dimensional distribution to another, taking into account available information about correspondences. We extend the 3D problem to higher dimensions by encoding overlapping neighborhoods of data points, and solve the high dimensional problem in 1D subspace using the iterative projection approach. To show the potentials of this mapping, we apply it to colour transfer between two images that exhibit overlapped scene. Experiments show competitive results, both quan-
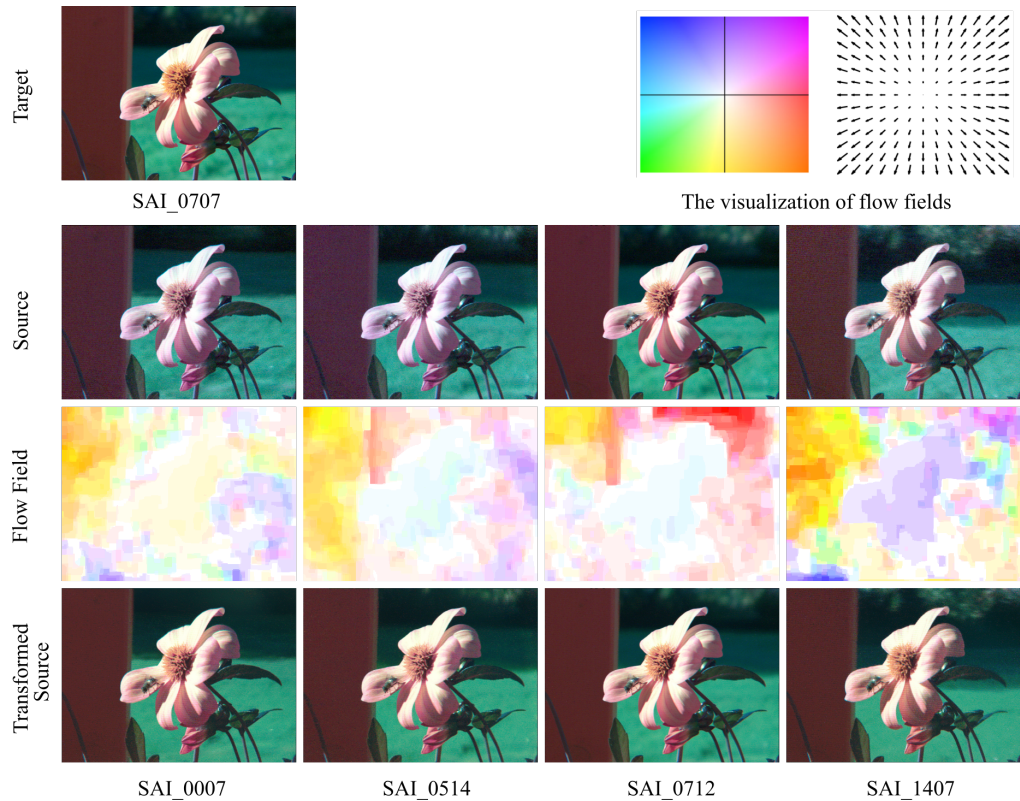
Figure 6.1: Results on SAI light fields images. Our method `NW_cp` successfully corrects colour inconsistencies in the selected outer views images using the centre view image (SAI_0707) as the target image. The flow fields show the motion changes between the target and each source image. The top row on the right shows flow field visualization based on the code in (Baker et al., 2007): each pixel denotes a flow vector where the orientation and magnitude are represented by the hue and saturation of the pixel, respectively.

titatively and qualitatively, compared with the previous state of the art colour transfer methods. The work presented in this chapter has been published in (Alghamdi and Dahyot, 2020a).

6.2   The Proposed Method

We explain our algorithm (Algorithm 6) in Section 6.2.1. It is inspired by the iterative projection approach implemented in the IDT (Pitie et al., 2007) and SWD (Bonneel et al., 2015) algorithms originally proposed to solve Optimal Transport problem in N-dimensional spaces. Our algorithm is part of our overall pipeline (Figure 6.3), which is explained in Section 6.2.2.
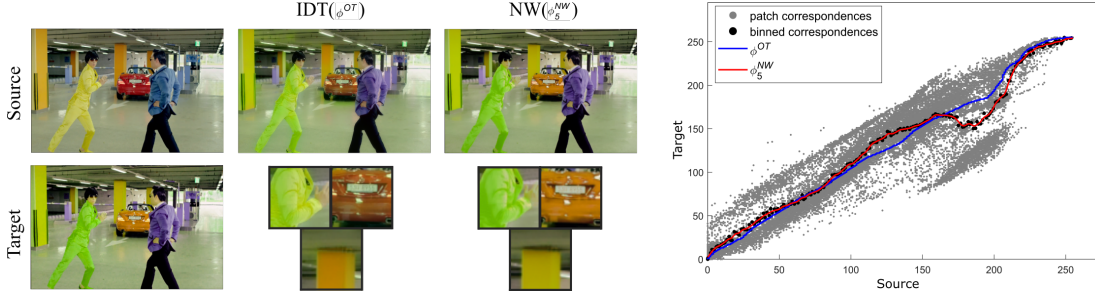
Figure 6.2: This figure shows an example result of the non-parametric Nadaraya-Watson mapping function (ϕ_5^{NW}) with bandwidth $h = 5$ that accounts for correspondences to guide the transfer process compared with the strictly increasing Optimal transport function (ϕ^{OT}) that does not take into account the correspondences.

6.2.1 Nadaraya-Watson vs Optimal Transport Solution in 1D

The IDT algorithm (Pitie et al., 2007) proposes to project 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 (cf. Algorithm 1). This projection creates two 1D datasets $\{u_i\}_{i=1}^n$ and $\{v_j\}_{j=1}^m$ with corresponding f_u and g_v whose cumulative distributions F_u and G_v are matched using the 1D optimal transport solution $\phi^{OT}(u)$ (Eq. (4.2) CDF-maching). Recall that OT by definition does not take into account any available information about the correspondences when computing the mapping, while the ability of taking advantage of the available correspondences would guide the transfer process and improve the results. Also, OT is an increasing function by definition and not smooth and locally variable. We propose to replace the 1D ϕ^{OT} by the Nadaraya-Watson estimator (cf. Algorithm 6 line 8) taking advantage of the correspondences $\{(x_i, y_i)\}_{i=1}^n$ giving correspondences $\{(u_i, v_i)\}_{i=1}^n$ in the projective space:

$$\phi_h^{NW}(u) = \frac{\sum_{i=1}^n v_i \, W_h(u - u_i)}{\sum_{i=1}^n W_h(u - u_i)} \tag{6.1}$$

The NW estimator computes the estimate of an expectation of u given v using a Gaussian kernel (Section 2.2.3.1) with bandwidth h. This bandwidth controls the smoothness of the estimated function ϕ_h^{NW}. Figure 6.2 presents the two estimates ϕ^{OT} and ϕ_h^{NW} estimated as part of one iteration of our algorithm. Not having correspondences, ϕ^{OT} is a strictly increasing function, whereas ϕ_h^{NW} provides a smooth non-monotonic mapping function u to v.
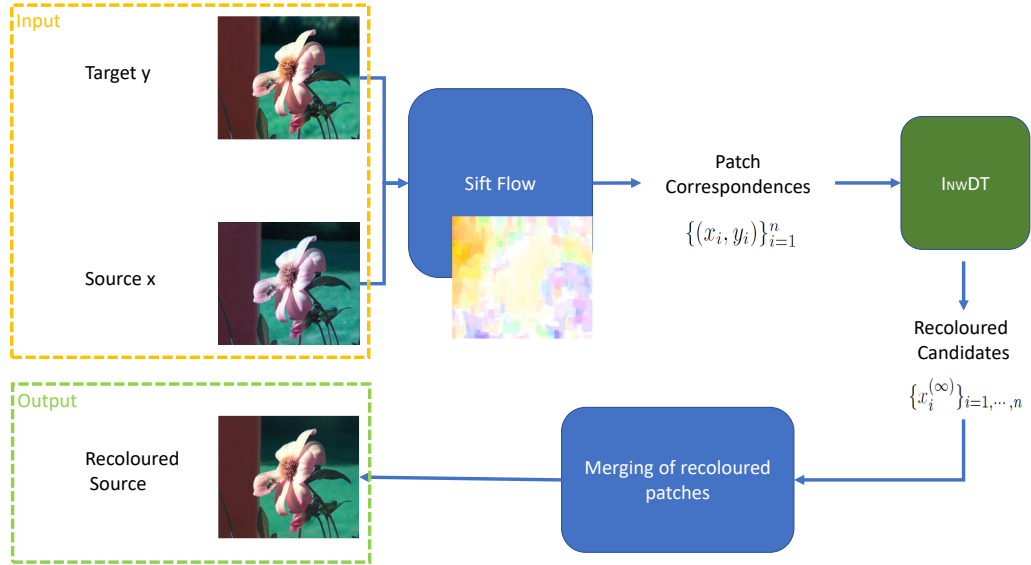
Figure 6.3: Proposed Pipeline: following OT_NW method (Chapter 5), correspondences between target and source are found using SIFT flow (Liu et al., 2011). These correspondences are used directly in our proposed INWDT algorithm (cf. Algorithm 6) to compute recoloured candidates that are then merged using the same process as PCT_OT method (Chapter 4) to compose the recoloured source image.

6.2.2 The Pipeline

We follow PCT_OT (Chapter 4) and OT_NW (Chapter 5) methods in constructing high dimensional descriptors with differences in choosing the parameters, as explained in the following:

6.2.2.1 Patch correspondences

First, giving the source colour pixels denoted by $\{x_i^c\}_{i=1}^n$ and the target colour pixels $\{y_j^c\}_{j=1}^m$, we use SIFT flow motion estimation (Liu et al., 2011) to compute motion flows (explained in Section 5.2.1) and extract the correspondences $\{(x_i^c, y_i^c)\}_{i=1}^n$ in $\mathbb{R}^N$ (for example $N = \dim(x^c) = 3$ for RGB colour space). Second, we concatenate the original pixels' positions in the grid coordinate of the images with the colour channels such that $\{(x_i, y_i) = ((x_i^c, x_i^p) , (y_i^c, y_i^p))\}_{i=1}^n$ in $\mathbb{R}^N$, where $N = \dim(x) = \dim(x^c) + \dim(x^p)$. Next, we normalize the coordinates to lie between 0 and 255, and we encode overlapping neighborhoods of pixels correspondences with patch size neighborhood $d \times d$ to create high dimensional pairs of vectors $\{(x_i, y_i)\}_{i=1}^q$ in $\mathbb{R}^N$, where $N = \dim(x) \times d \times d$. We also consider patches containing only colour informa-
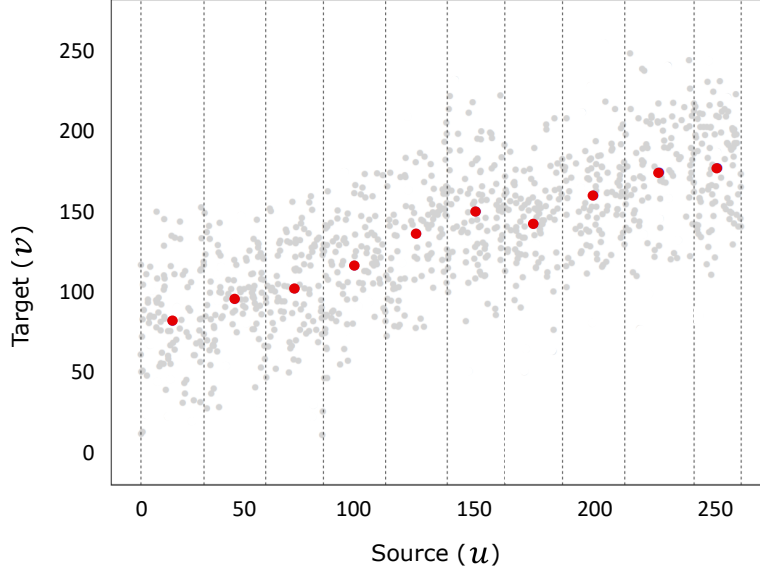
Figure 6.4: An illustration of creating binned correspondences $\{(\bar{u}_i, \bar{v}_i)\}_{i=1}^K$ from $\{(u_i, v_i)\}_{i=1}^q$. The red dots representing the mean of the corresponding target values for the source observations falling in each bin.

tion of a pixel neighborhood, i.e., $\{(x_i, y_i) = (x_i^c, y_i^c)\}_{i=1}^n$, similar to PCT_OT method (cf. Figure 4.1 (A)).

6.2.2.2 Iterative Nadaraya-Watson Distribution Transfer (INWDT)

Our INWDT algorithm outlined in Algorithm 6 is applied to our set patch correspondences $\{(x_i, y_i)\}_{i=1}^q$ (input) to compute recoloured patches $\{x_i^{(\infty)}\}_{i=1}^q$ (output). At each iteration k, we project the correspondences $\{(x_i, y_i)\}_{i=1}^q$ originally in $\mathbb{R}^N$ space, where $N = \dim(x) \times d \times d$, to a 1D subspace creating 1D datasets $\{(u_i, v_i)\}_{i=1}^q$ in the projective space (step 6 in Algorithm 6). Next, we use binned Nadaraya-Watson estimator where we create binned correspondences by dividing the support of the source points $\{u_i\}_{i=1}^q$ into $K = 200$ number of bins (subintervals), and calculate the averages of the source observations falling in each bin $\{\bar{u}_i\}_{i=1}^K$ as well as calculate the averages of the corresponding target values $\{\bar{v}_i\}_{i=1}^K$ (for example, $\bar{u}_4$ is the average of all $u_i s$ falling in the 4th bin, and $\bar{v}_4$ is the average of $v_i s$ such that the corresponding $u_i s$ is in the 4th bin), Figure 6.4 illustrates an example for this procedure. The key idea of binned implementation is to greatly reduce the number of kernel evaluations (Murphy, 2014; Fan and Marron, 1994) as we have large number of SIFT flow correspondences and many of these evaluations are the same in the projective space. Next, we feed the correspondences $\{(\bar{u}_i, \bar{v}_i)\}_{i=1}^K$ to the binned NW estimator to compute an estimate $v = \phi_h^{NW}(u)$

for every data point in the source patch dataset $\{u_i\}_{i=1}^q$ as follows:

$$\phi_h^{NW}(u) = \frac{\sum_{i=1}^K \bar{v}_i \, W_h(u - \bar{u}_i)}{\sum_{i=1}^K W_h(u - \bar{u}_i)} \tag{6.2}$$

Exploring the Convergence: To explore the convergence, the standard $\mathcal{L}_2$ distance between N-dimensional probability distributions (Jian and Vemuri, 2011) is used as a measure to quantify how well the transformed distribution f_x matches the target pdf g_y after each iteration k of the algorithm (at step 14 in Algorithm 6). We chose $\mathcal{L}_2$ distance between probability distributions over the popular Kullback-Leibler (Kullback and Leibler, 1951) divergence to assess the convergence because $\mathcal{L}_2$ distance is robust to outliers, unlike KL (Grogan and Dahyot, 2019; Sugiyama, 2013; Jian and Vemuri, 2011). As explained in Section 2.2.2.1, the $\mathcal{L}_2$ distance between probability distributions is defined as follows:

$$\mathcal{L}_2(f(x), g(x)) = \int f^2(x)dx - 2\int f(x)g(x)dx + \int g^2(x)dx \tag{6.3}$$

There is only one random variable x appears explicitly in the Eq.(6.3), and both the source pdf f and target pdf g are candidate pdfs for this random variable. We estimated the underlying pdfs using a similar setting to the one we used in Section 3.1.3, as follows:

We consider that the two probability density functions f and g representing source and target samples are Gaussian mixtures:

$$f(x) = \sum_{i=1}^n w_i \mathcal{N}(x; \mu_i^s, \Sigma_i) \quad \text{and} \quad g(x) = \sum_{j=1}^m w_j \mathcal{N}(x; \mu_j^t, \Sigma_j) \tag{6.4}$$

The notation $\mathcal{N}(x; \mu_*; \Sigma_*)$ indicates a N-dimensional Gaussian distribution, with mean μ_* and covariance matrix Σ_*, and $0 < w_* < 1$ is the weight of the distribution such that $\sum_{*=1}^{**} w_* = 1$ (with $* = i, j$ and $** = n, m$). We construct the Gaussian mixture models from the given 1D datasets in a simplified setting as follows: 1) The number of Gaussian components is the number of samples in the dataset and all components are weighted equally, 2) for each component, the mean is given by the spatial location of each sample, and 3) all components share the same spherical covariance matrix $\Sigma = h^2 I$, where I is the $N \times N$ identity matrix and h is a scalar value controls

the spherical covariance matrix. We set a data-driven bandwidth h value as follows [1] (Jian and Vemuri, 2011):

$$h = \left[\det \left(\frac{X^T X}{n} \right) \right]^{\frac{1}{(2^N)}} \tag{6.5}$$

Where X is $n \times N$ matrix denoting the source dataset. Figure 6.5 illustrates the convergence of the original distribution to the target over several iterations k of our algorithm visualized in 2D space.

6.2.2.3 Merge recoloured candidates.

Because the same pixel is present in overlapping patches $\{x_i^{(\infty)}\}_{i=1}^q$, the average colour value from all the candidates is selected for recolouring the source dataset $\{x_i\}_{i=1}^n$ similar to PCT_OT method (Chapter 4) and OT_NW method (Chapter 5). Our modified pipeline is outlined in Algorithm 5.

---

[1] https://github.com/bing-jian/gmmreg

Algorithm 5: Our pipeline for colour transfer between overlapped scene images with SIFT flow based correspondences

---

1: **Input:** Source & Target images
2: **Output:** recoloured Source image
3: Formulate the Source image & the Target image as colour samples datasets $\{x_i\}_{i=1}^n$ & $\{y_j\}_{j=1}^m$ of r.v. x and $y \in \mathbb{R}^N$, respectively. For example in RGB colour transfer, $N = 3$ and $x_i = (r_i, g_i, b_i)$ where r_i, g_i, b_i are the red, green and blue components of source pixel number i
4: Use SIFT flow to estimate dense correspondences $\{(x_i^c, y_i^c)\}_{i=1}^n \in \mathbb{R}^3$ (for 3D RGB colour space)
5: **if** Combining colour and spatial information **then**
6:     Combine spatial correspondences with colour information $\{(x_i, y_i) = ((x_i^c, x_i^p) , (y_i^c, y_i^p))\}_{i=1}^n$ in $\mathbb{R}^N$ where $N = \dim(x) = \dim(x^c) + \dim(x^p)$ (see Section 6.2.2.1)
7:     Normalize all the colour and position coordinates to lie between 0 and 255
8:     Choose a window size $d \times d$ and create patch correspondence vectors using the combined colour and positions features $\{(x_i, y_i)\}_{i=1}^q$ in $\mathbb{R}^N$ where $N = \dim(x) \times d \times d$ (for 3D RGB and $d = 3$, $\mathbb{R}^{45}$ where $N = 5 \times 3 \times 3 = 45$)
9: **else if** colour only **then**
10:     Use colour pixels correspondences (i.e. $\{(x_i, y_i) = (x_i^c, y_i^c)\}_{i=1}^n$) to create colour patch correspondence vectors $\{(x_i, y_i)\}_{i=1}^q$ in $\mathbb{R}^N$ (for 3D RGB and $d = 3$, $\mathbb{R}^{27}$ where $N = 3 \times 3 \times 3 = 27$)
11: **end if**
12: Transform $\{x_i\}_{i=1}^q$ using INWDT giving patch correspondences $\{(x_i, y_i)\}_{i=1}^q$ (cf. Algorithm 6)
13: Merge recoloured candidates from the overlapping patches $\{x_i^{(\infty)}\}_{i=1}^q$ to create the output recoloured source image

---
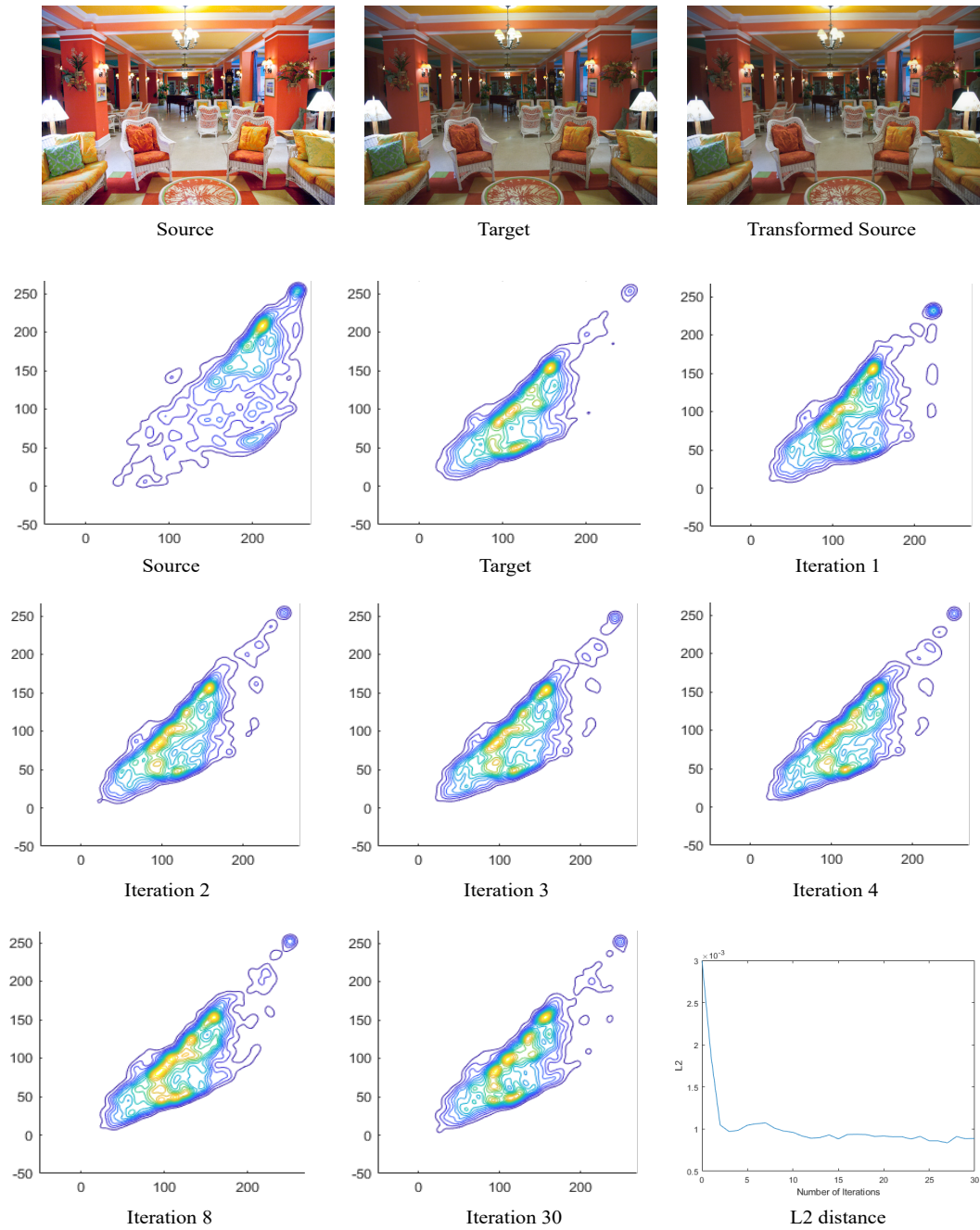
Figure 6.5: Example of pdf of the transferred source patches projected in 2D space (RG). The patch size chosen is 1×1 and only the colour information is used $N = 3$ (space RGB). The $\mathcal{L}_2$ distance (Grogan and Dahyot, 2019; Jian and Vemuri, 2011) is computed at each iteration to illustrate the convergence of the original distribution to the target one by our transfer method.

Algorithm 6: Generalized Pseudo-code for Iterative Nadaraya-Watson Distribution Transfer (INWDT)

1: **Input:** Source and Target correspondences $\{(x_i, y_i)\}_{i=1}^q$, samples of r.v. $(x, y) \in \mathbb{R}^N \times \mathbb{R}^N$

2: **Initialisation:** $k \leftarrow 0$ and $\forall i,\ x_i^{(0)} \leftarrow x_i$

3: **repeat**

4:     Generate D random unit vectors in $\mathbb{R}^N$ stored in matrix $\mathrm{R} = [e_1, ..., e_D]$

5:     **for** $r = 1$ to D **do**

6:         Compute projections $\forall i,\ u_i = e_r^T x_i^{(k)}$ and $v_i = e_r^T y_i$

7:         Choose K bins and create binned correspondences $\{(\bar{u}_i, \bar{v}_i)\}_{i=1}^K$ from $\{u_i, v_i\}_{i=1}^q$

8:         Compute 1D NW estimate $\forall i$ in $\{u_i\}_{i=1}^q$, $v_i = \phi_r^{NW}(u_i)$ with $\{(\bar{u}_i, \bar{v}_i)\}_{i=1}^K$ (Eq. 6.2)

9:     **end for**

10:    Remap the source dataset

$$\Phi(x_i^{(k)}) = \mathrm{R} \begin{pmatrix} \phi_1^{NW}(e_1^T x_i^{(k)}) \\ \vdots \\ \phi_D^{NW}(e_D^T x_i^{(k)}) \end{pmatrix} = \sum_{r=1}^D (\phi_r^{NW}(e_r^T x_i^{(k)}))\, e_r$$

11:    Update source dataset $x_i^{(k+1)} \leftarrow \Phi(x_i^{(k)})$

12:    $\mathrm{R} \leftarrow$ Random rotation of R

13:    $k \leftarrow k + 1$

14: **until** convergence $\mathcal{L}_2(f_x, g_y) \rightarrow 0$ (noted $k \equiv \infty$)

15: **Result:** With the recoloured patches $\{x_i^{(\infty)}\}_{i=1}^q$, the final one-to-one mapping Φ in $\mathbb{R}^N$ is given by $x_i \mapsto \Phi(x_i) = x_i^{(\infty)}$, $\forall i$.

6.3   Experimental Assessment

We use the evaluation tools presented in Section 4.4.4 which consist of using numerical metrics, summary descriptive statistics, statistical significance tests, and visual analysis. We provide extensive quantitative and qualitative evaluations to validate both of our NW solutions - using colour patches only, annotated in the results as `NW_c`, and using colour patches with pixel location information, annotated as `NW_cp`. We compare our methods to different state of the art colour transfer methods noted by `L2` (Grogan and Dahyot, 2019), `GPS/LCP` and `FGPS/LCP` (Bellavia and Colombo, 2018), `PMLS` (Hwang et al., 2014), `IDT` (Pitie et al., 2007), also we compare with `PCT_OT` (Chapter 4) and `OT_NW` (Chapter 5). In these evaluations, we use image pairs with similar contents from an existing dataset (Section 4.4.3) provided by (Hwang et al., 2014). Note that the results using `PMLS` were provided by the authors (Hwang et al., 2014). We also compared with two other recent techniques (Xia et al., 2017; Park et al., 2016) that account for correspondences into their framework, but `PMLS` has been shown to perform better than these two (Grogan and Dahyot, 2019), hence `PMLS` is the one reported here with (Grogan and Dahyot, 2019) as algorithms that incorporate correspondences in their methodologies. We also compared with (Bonneel et al., 2016; Ferradans et al., 2013) that do not take into account correspondences, but `IDT` has been shown to be superior (Grogan and Dahyot, 2019), hence `IDT` is the one reported here for ease of comparison.

6.3.1   Colour space and parameters settings

We use the RGB colour space and we found a patch size of 3×3 captures enough of a pixel's neighbourhood. For our `NW_cp` version, 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 ($N = 3 \times 3 \times 5 = 45$). For `NW_c`, pixel position is not accounted for, and only RGB colours are used ($N = 3 \times 3 \times 3 = 27$). For the bandwidth selection in Nadaraya-Watson Eq. (6.2), using a cross-validation method such as the popular Jackknife cross-validation criterion (Shao and Tu, 2012) is overly costly computationally. Hence, we experimented with different values, and a fixed bandwidth of $h = 5$ was selected for the best results.

6.3.2 Experimental results

Quantitative evaluation: The numerical results of each metric with means and standard errors (SE) are shown in Tables 6.2 to 6.5, along with box plots shown in Figures 6.6 to 6.9, respectively. The purpose of the box plots is to visualize differences among methods and to show how close our method is to the state of the art algorithms. By looking at the tables (Tables 6.2 to 6.5), all metrics show that both solutions with NW (colour only NW_c, and combined colour and positions NW_cp), which takes correspondences into account, significantly outperforms OT (IDT algorithm) that does not take the available correspondences into account when estimates the mapping, and this is visualized in each metric box plot, Figure 6.6 to 6.9. Moreover, the results show that NW_cp which incorporates colour and spatial information improves the performance over NW_c which uses colour only. Comparing NW_cp with the best performances we reported so far, namely, PMLS, L2 and OT_NW, we find that although NW_cp outperforms the other methods in many cases as measured by PSNR as shown in the Table 6.2, by examining all the metrics' box plots in Figures 6.6 to 6.9 we see that the methods PMLS, L2, OT_NW, and NW_cp are greatly overlapped with each other, where the values of the median and mean (the means shown as red dots in the plots, and the medians shown as horizontal black lines) are very close in value. In conclusion, the quantitative metrics show that our algorithm with Nadaraya-Watson competes very well with top methods PMLS, L2 and OT_NW (Chapter 5), and outperforms the rest of the state of the art algorithms. Moreover, our conclusion is supported by statistical significance analysis. We picked the methods that are shown to be greatly overlapped in the box plots, namely PMLS, L2, OT_NW and NW_cp, and we used Kruskal-Wallis Rank Sum test (Section 4.4.4) for each metric separately. We found that with 95% confidence level, the P-values for each metrics are greater than the significance level $\alpha = 0.05$, which indicates that there is no significant difference between PMLS, L2, OT_NW and NW_cp method, Table 6.1.

Qualitative evaluation: Figure 6.11 provides qualitative results. For clarity, the results in Figure 6.11 are presented in image mosaics, created by switching between the target image and the transformed source image column wise. If the colour transfer is accurate, the resulting mosaic should look like a single image (ignoring the small motion displacement between source and target images), otherwise column differences

Metrics	P-values
PSNR	0.9836
SSIM	0.7832
CID	0.6059
FSIMc	0.972

Table 6.1: This table shows results of Kruskal-Wallis Rank Sum statistical test for each metric separately. The test tests if there are significant differences between `NW_cp` results and the top performance methods, namely, `PMLS`, `L2`, and `OT_NW`. With 95% confidence level, the P-values for each metric is greater than the significance level $\alpha = 0.05$, which indicates that there is no significant difference between the aforementioned methods.

appear. Figure 6.11 shows that both NW solutions (colour only `NW_c`, and combined colour and positions `NW_cp`), which take correspondences into account, are successful in alleviating the column differences, where significantly outperform the original OT (IDT algorithm) that does not take the available correspondences into account (cf. Figure 6.11 in examples "Tonal 4", "Gangnam2" and "Gangnam1"). However, NW is a smooth function provides a very smooth mapping that sometimes affecting the exact colour mapping, which in this case makes `OT_NW` (Chapter 5) the best in alleviating the columns differences (cf. Figure 6.11 in examples "building" and "illum").

Figure 6.10 shows more qualitative results. We can see that both NW solutions are robust to registration errors and do not introduce visual artifacts, unlike `PMLS` which produces those artifacts although it provides good results in terms of metrics. Moreover, as can be seen in Figure 6.10 with examples 'illum', 'mart' and 'building', both NW solutions can transfer colours between images that contain significant spatial differences and eliminate the shadow artifacts better, visually, than `OT_NW`, and unlike `PCT_OT` (cf. Figure 6.10, in example 'building'). In addition, NW approach can create a smoother colour transfer result, and can also alleviate JPEG compression artifacts and noise (cf. Figure 6.10 for comparison). In general, Figure 6.10 and Figure 6.11 show that NW, which accounts for the correspondences *explicitly* in its modeling, gives similar results visually with both solutions (colour only `NW_c`, and combined colour and positions `NW_cp`), although `NW_cp` gives better numerical results in terms of metrics.

	PSNR ↑								
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 2020	NW_c 3x3	NW_cp 3x3
Gangnam1	25.354	35.725	24.048	23.936	35.358	31.479	33.565	30.353	32.241
Gangnam2	27.116	36.553	25.952	25.944	35.524	35.502	33.627	31.350	32.783
Gangnam3	22.372	35.007	21.908	21.913	33.284	26.393	28.217	27.069	29.269
Illum	19.822	20.167	19.785	19.960	19.079	20.306	20.858	20.236	21.066
Building	20.554	22.634	22.736	22.769	20.499	25.019	24.039	22.726	22.895
Playground	27.184	27.835	25.501	25.436	27.647	28.482	28.491	28.532	28.625
Flower1	24.238	26.981	23.765	23.706	26.857	25.186	27.158	26.973	27.530
Flower2	25.417	25.760	25.259	25.223	25.772	26.373	26.497	26.326	26.445
Tonal1	30.082	37.215	31.617	31.413	37.332	37.044	37.151	36.264	36.719
Tonal2	27.992	31.508	25.062	25.087	31.356	32.049	31.579	31.755	32.394
Tonal3	29.575	36.246	28.136	28.065	36.644	33.793	35.014	33.889	34.900
Tonal4	28.605	34.521	28.852	28.848	34.344	33.819	35.320	34.118	35.202
Tonal5	30.330	35.260	29.580	29.448	34.303	36.437	36.616	34.356	35.656
Mart	22.747	24.742	23.183	23.196	24.450	24.509	25.189	25.271	25.574
Sculpture	29.884	32.062	29.037	28.820	32.067	31.237	32.735	32.857	33.356
Mean	26.085	30.814	25.628	25.584	30.301	29.842	30.404	29.472	30.310
SE	0.905	1.459	0.841	0.821	1.518	1.306	1.291	1.202	1.249

Table 6.2: Comparing our algorithms NW_c and NW_cp with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
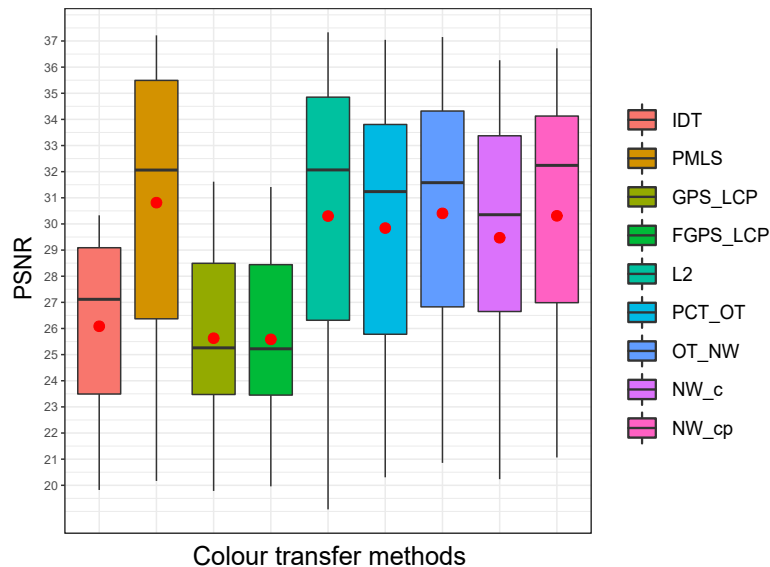
Figure 6.6: Comparing our algorithms NW_c and NW_cp with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Detailed quantitative results in Table 6.2 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

	SSIM ↑								
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 2020	NW_c 3x3	NW_cp 3x3
Gangnam1	0.900	0.992	0.892	0.891	0.990	0.964	0.973	0.947	0.965
Gangnam2	0.920	0.993	0.909	0.909	0.986	0.980	0.976	0.960	0.972
Gangnam3	0.859	0.991	0.873	0.864	0.980	0.930	0.959	0.890	0.955
Illum	0.641	0.649	0.643	0.646	0.648	0.673	0.687	0.662	0.678
Building	0.808	0.865	0.864	0.863	0.862	0.888	0.875	0.873	0.870
Playground	0.920	0.940	0.878	0.876	0.939	0.939	0.943	0.941	0.941
Flower1	0.909	0.967	0.913	0.912	0.966	0.926	0.959	0.960	0.963
Flower2	0.901	0.928	0.894	0.894	0.927	0.933	0.939	0.923	0.930
Tonal1	0.953	0.988	0.971	0.970	0.987	0.988	0.991	0.986	0.988
Tonal2	0.968	0.987	0.926	0.926	0.986	0.988	0.986	0.986	0.988
Tonal3	0.962	0.992	0.947	0.946	0.992	0.987	0.990	0.985	0.989
Tonal4	0.944	0.983	0.932	0.932	0.983	0.981	0.985	0.982	0.985
Tonal5	0.965	0.986	0.953	0.954	0.985	0.990	0.991	0.983	0.987
Mart	0.904	0.957	0.925	0.925	0.956	0.941	0.954	0.955	0.957
Sculpture	0.942	0.971	0.934	0.932	0.972	0.945	0.974	0.965	0.971
Mean	0.900	0.946	0.897	0.896	0.944	0.937	0.946	0.933	0.943
SE	0.022	0.023	0.020	0.020	0.023	0.020	0.020	0.021	0.020

Table 6.3: Comparing `NW_c` and `NW_cp` with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
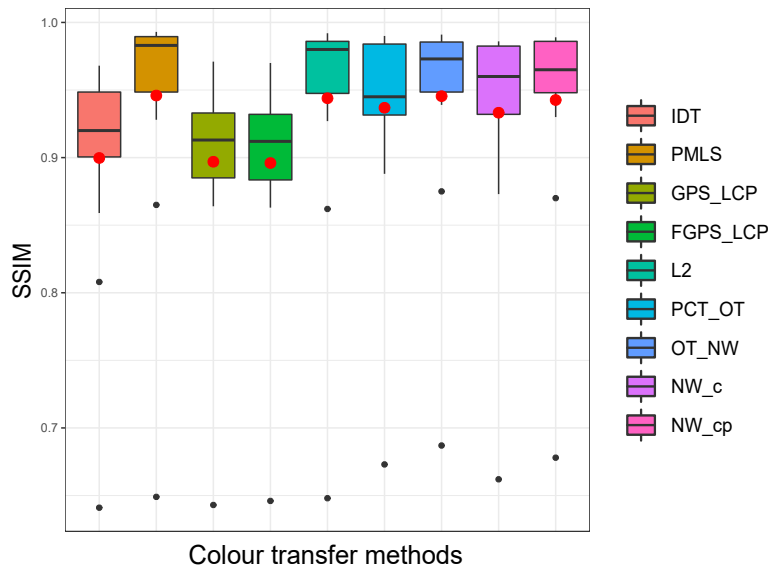
Figure 6.7: Comparing `NW_c` and `NW_cp` with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Detailed quantitative results in Table 6.3 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

	CID ↓								
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 2020	NW_c 3x3	NW_cp 3x3
Gangnam1	0.252	0.040	0.226	0.222	0.048	0.085	0.088	0.115	0.094
Gangnam2	0.268	0.039	0.291	0.292	0.089	0.068	0.109	0.140	0.123
Gangnam3	0.496	0.108	0.472	0.487	0.193	0.261	0.267	0.401	0.297
Illum	0.386	0.390	0.395	0.396	0.397	0.377	0.376	0.405	0.373
Building	0.374	0.228	0.313	0.321	0.249	0.183	0.275	0.344	0.328
Playground	0.440	0.238	0.443	0.471	0.254	0.209	0.221	0.291	0.253
Flower1	0.389	0.163	0.396	0.400	0.174	0.285	0.194	0.264	0.233
Flower2	0.337	0.245	0.322	0.323	0.266	0.218	0.201	0.300	0.250
Tonal1	0.310	0.101	0.285	0.308	0.111	0.097	0.063	0.126	0.105
Tonal2	0.288	0.128	0.351	0.347	0.145	0.099	0.118	0.157	0.129
Tonal3	0.244	0.079	0.294	0.294	0.081	0.077	0.079	0.128	0.107
Tonal4	0.240	0.108	0.248	0.238	0.107	0.065	0.065	0.132	0.091
Tonal5	0.156	0.091	0.205	0.192	0.092	0.051	0.067	0.122	0.094
Mart	0.526	0.219	0.405	0.402	0.225	0.426	0.249	0.252	0.246
Sculpture	0.242	0.137	0.213	0.224	0.143	0.232	0.120	0.221	0.172
Mean	0.330	0.154	0.324	0.328	0.172	0.182	0.166	0.227	0.193
SE	0.027	0.024	0.022	0.023	0.024	0.031	0.025	0.027	0.025

Table 6.4: Comparing NW_c and NW_cp with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (lower values are better, best viewed in colour and zoomed in).
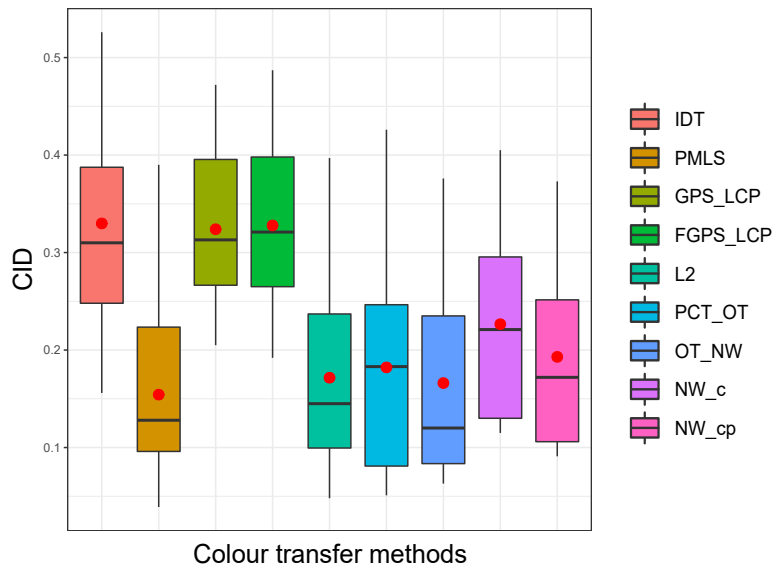
Figure 6.8: Comparing NW_c and NW_cp with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Detailed quantitative results in Table 6.4 are summarized in a box plot (lower values are better, best viewed in colour and zoomed in).

	FSIMc ↑								
	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	L2 2019	PCT_OT 2019	OT_NW 2020	NW_c 3x3	NW_cp 3x3
Gangnam1	0.936	0.986	0.944	0.943	0.985	0.972	0.979	0.967	0.975
Gangnam2	0.952	0.992	0.962	0.962	0.988	0.990	0.986	0.977	0.982
Gangnam3	0.946	0.992	0.962	0.961	0.990	0.987	0.982	0.972	0.980
Illum	0.800	0.819	0.824	0.824	0.818	0.839	0.844	0.819	0.828
Building	0.874	0.928	0.930	0.929	0.926	0.942	0.937	0.926	0.928
Playground	0.950	0.958	0.933	0.932	0.955	0.956	0.960	0.957	0.956
Flower1	0.954	0.975	0.968	0.967	0.976	0.971	0.977	0.974	0.975
Flower2	0.941	0.950	0.945	0.945	0.949	0.954	0.956	0.949	0.950
Tonal1	0.964	0.997	0.986	0.986	0.997	0.998	0.997	0.996	0.996
Tonal2	0.984	0.993	0.973	0.973	0.992	0.993	0.992	0.991	0.993
Tonal3	0.979	0.997	0.984	0.983	0.997	0.997	0.995	0.992	0.993
Tonal4	0.966	0.989	0.972	0.973	0.990	0.995	0.994	0.990	0.993
Tonal5	0.980	0.994	0.987	0.987	0.993	0.998	0.997	0.993	0.995
Mart	0.946	0.969	0.960	0.959	0.967	0.969	0.970	0.968	0.971
Sculpture	0.980	0.987	0.982	0.980	0.988	0.988	0.987	0.984	0.986
Mean	0.943	0.968	0.954	0.954	0.967	0.970	0.970	0.964	0.967
SE	0.012	0.012	0.010	0.010	0.012	0.010	0.010	0.011	0.011

Table 6.5: Comparing NW_c and NW_cp with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better, best viewed in colour and zoomed in).
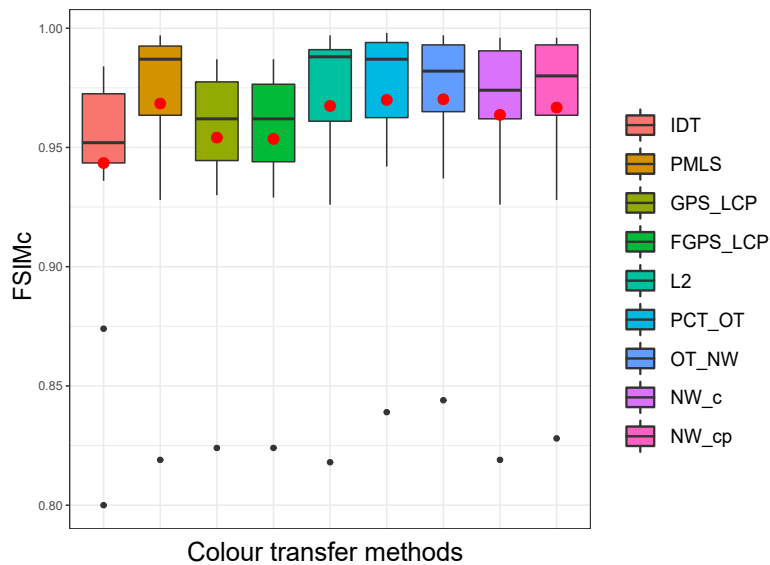
Figure 6.9: Comparing NW_c and NW_cp with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Detailed quantitative results in Table 6.5 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).
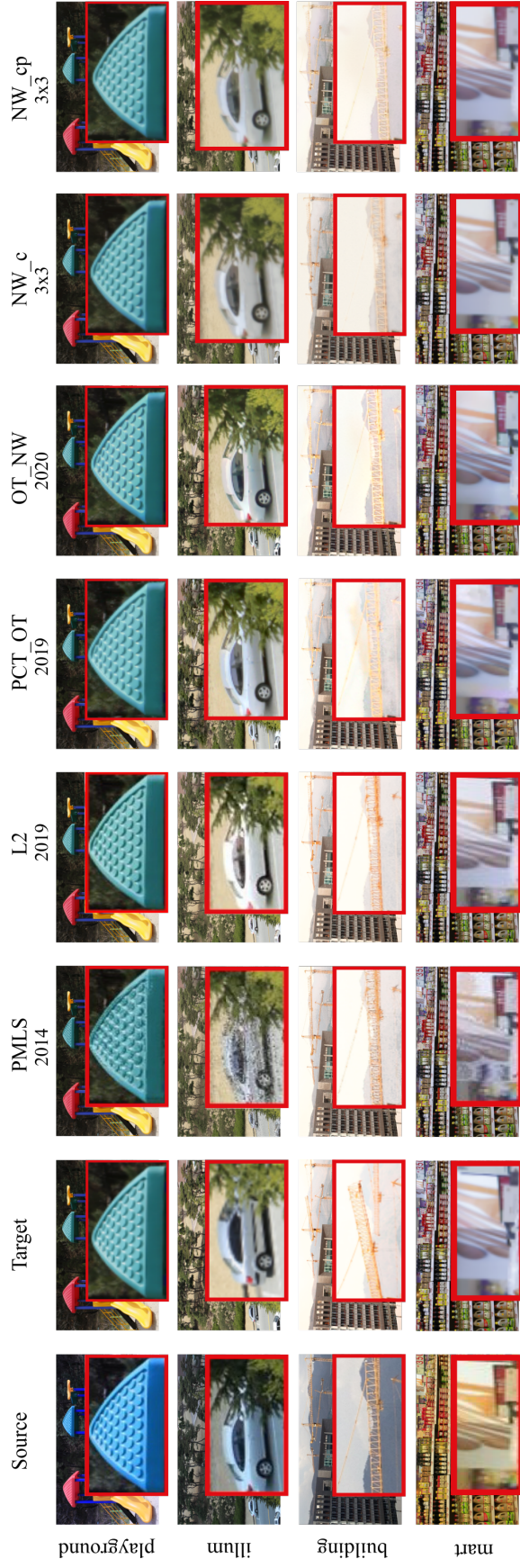
Figure 6.10: A close up look at some of the results generated using the top performance methods PMLS (Hwang et al., 2014), L2 (Grogan and Dahyot, 2019), PCT_OT (Alghamdi et al., 2019) and OT_NW (Alghamdi and Dahyot, 2020b) and our algorithms NW_C and NW_CP (best viewed in colour and zoomed in).
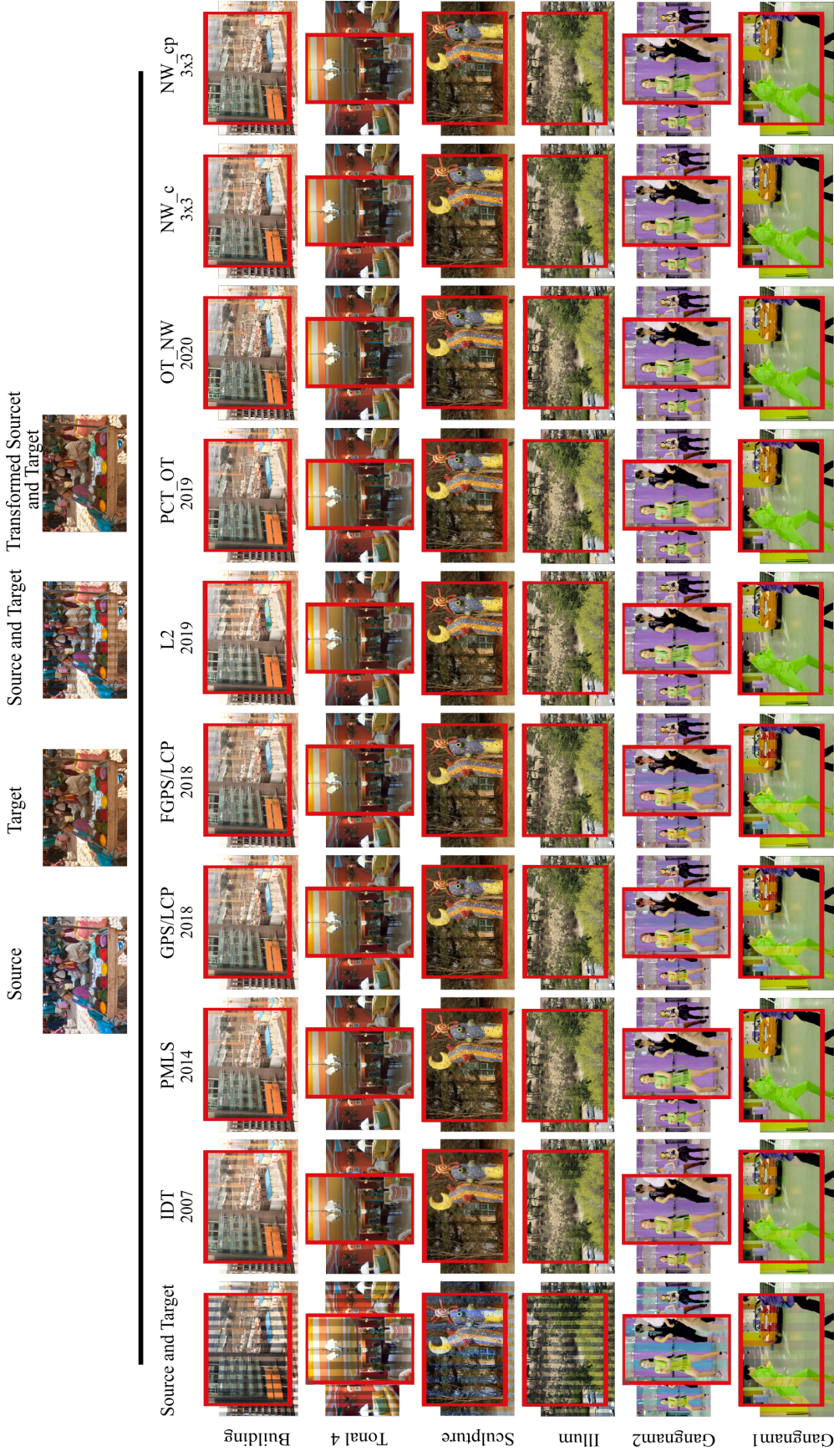
Figure 6.11: A close up look at some of the results generated using the IDT (Pitie et al., 2007), PMLS (Hwang et al., 2014), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), L2 (Grogan and Dahyot, 2019), PCT_OT (Alghamdi et al., 2019), OT_NW (Alghamdi and Dahyot, 2020b) and our algorithms NW_c and NW_cp (best viewed in colour and zoomed in).

6.4 Conclusion

We have shown how to use the Nadaraya-Watson estimator to adapt OT (the IDT algorithm) for accounting for input correspondences in matching high dimensional probability density functions. We conducted extensive quantitative and qualitative experiments and analyses. Our approach outperforms the IDT algorithm and is shown to be competitive with the leading state of the art of colour transfer in images where spaces of dimension up to 45 have been used. Future work will look into combining solution ϕ^{OT} and ϕ^{NW} to tackle semi-supervised situations where correspondences are only partially available.

Nadaraya-Watson and Optimal Transport, by definition, have no parametric formulation of the solution ϕ^{OT} and ϕ^{NW}, which limits the possibility and flexibility of applying this transformation to previously unseen values. In this context, Grogan and Dahyot (2019) proposed a parametric $\mathcal{L}_2$-based colour transfer framework that has been shown to be competitive with alternative Optimal Transport based solutions but does not scale well in high dimensional spaces due to the increase in computational complexity associated with the increase of the data dimensionality. In the next chapter, using the 1D iterative projection approach, we extend the $\mathcal{L}_2$-based framework to higher dimensions, where only 1D marginal distributions are used in the estimation process, which therefore involves low computational costs.

SLICED $\mathcal{L}_2$ DISTANCE FOR COLOUR GRADING

7.1 Introduction

Optimal Transport has been successfully used as a way for defining cost functions for optimization when performing distribution transfer (Peyré and Cuturi, 2019), which have been used more recently to solve machine learning problems (Tanaka, 2019; Meng et al., 2019; Muzellec and Cuturi, 2019). Optimal Transport solutions for colour transfer based on the 1D iterative projection approach have also been proposed, as described in Section 2.3.1.2. However, by OT definition, no parametric formulation of the solution ϕ^{OT} is available. Instead, an estimated correspondences $\{(x_i, \phi^{OT}(x_i))\}_{i=1}^{n}$ are only produced by the OT mapping, limiting the possibility and flexibility of applying this transformation to a previously unseen value x_i. In order to extend OT solution to unseen values such as recolouring multiple frames in a video sequence, Frigo et al. (Frigo et al., 2014) proposed to fit a smooth Thin Plate Spline transformation to the OT solution for colour transfer and estimate the colour transformation between target image and a key frame of a video sequence and encode it in a LUT (LookUp Table) to be applied to all frames in the video sequence.

An alternative framework was first successfully proposed by Jian and Vemuri (2011) in shape registration context, where the cost function used for inferring a parametric transfer function is defined as the robust $\mathcal{L}_2$ divergence between two probability density functions. Inspired by the success of the $\mathcal{L}_2$ framework, Grogan and Dahyot (2019) proposed an example based colour transfer approach, whereby a parametric

transformation function ϕ_θ is estimated by minimizing the $\mathcal{L}_2$ distance between two GMMs which capture the colour content of the source and target images. They have shown that $\mathcal{L}_2$ based solution is very competitive with alternative solutions based on OT for colour transfer. Unlike OT, Grogan and Dahyot's $\mathcal{L}_2$ based formulation is flexible as it allows colour correspondences $\{(x_i, y_i)\}$ that may be available to be taken into account and performs well in the presence of correspondence outliers, while OT by definition does not take into account any available information about correspondences when computing the optimal solution.

While the $\mathcal{L}_2$ based technique solves the problem directly in 2D/3D considering a smooth parametric non-linear transformation such as TPS (Grogan and Dahyot, 2019; Jian and Vemuri, 2011), TPS does not scale well in high dimensional spaces due to the increase in computational complexity associated with the increase of the data dimensionality. Hence, this chapter proposes to extend it to higher dimensions and introduce the Sliced $\mathcal{L}_2$ Distance based approach that minimizes the robust $\mathcal{L}_2$ between two N-dimensional distributions using the 1D projection-based approach. $\mathcal{L}_2$ with the projection-based approach has the potential to adapt well in N-dimensional spaces where $N > 3$. More specifically, we create high dimensional descriptors including colour and spatial features as well as neighborhoods structural information and extend the data space from 3D to higher dimensional spaces. Using the 1D iterative projection approach in solving the high dimensional problem, only 1D marginal distributions are used in the estimation process, which therefore involves low computational costs. We show how to tackle semi-supervised situations where correspondences are only partially available by combining $\mathcal{L}_2$ based solutions using correspondences and without using correspondences. To show potentials of this mapping, we apply it to the colour transfer problem between two images that exhibit overlapped scene, and conduct exhaustive experiments that show quantitative and qualitative competitive results as compared with the previous state of the art colour transfer methods.

7.2 The Proposed Method

We explain our Sliced $\mathcal{L}_2$ Distance (SL2D) for distribution transfer (Algorithm 7) in Section 7.2.1. It is inspired by the projection-based approach utilized by the IDT (Pitie et al., 2007) and SWD (Bonneel et al., 2015) algorithms that have been originally pro-

posed to solve OT in N-dimensional spaces. We modified our pipeline outlined in Algorithm 5 (Chapter 6) and incorporate SL2D as part of it, which is explained in Section 7.2.2.

7.2.1 $\mathcal{L}_2$ based solution vs Optimal Transport solution in 1D

As we explained in Section 6.2.1, the IDT (Pitie et al., 2007) 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 (cf. line 6 in Algorithm 1). 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 $\phi^{OT}(u)$ (Eq. (4.2)). We propose to replace the non-parametric $\phi^{OT}(u)$ by the $\mathcal{L}_2$ derived estimate $\phi_\theta^{\mathcal{L}_2}(u)$ (cf. line 14 & line 20 in Algorithm 7), where the parameters θ that control the mapping function $\phi_\theta^{\mathcal{L}_2}(u)$ are estimated as follows:

$$\hat{\theta} = \underset{\theta}{\mathrm{argmin}} \left[\mathcal{L}_2(f(u|\theta), g(u)) = \int f(u|\theta)^2 du - 2 \int f(u|\theta)g(u)du \right] \quad (7.1)$$

The $\mathcal{L}_2$ framework has only one random variable - denoted by u here since we are solving the problem in 1D projective spaces - appears explicitly in the modeling Eq. (7.1), and both the source pdf f and target pdf g are candidate pdfs for the random variable $u \in \mathbb{R}$. Eq. (7.1) proposes a parametric model $f(u|\theta)$ for the source distribution where θ is the parameters of the mapping function ϕ that minimizes the $\mathcal{L}_2$ divergence between the source and target marginal probability density functions. In the following, whenever needed, we distinguish the sample that is coming from the source dataset by u^s and the sample that is coming from the target dataset by u^t. In Section 3.1.3, we formulated $\mathcal{L}_2$ inference (Eq. 7.1) in 1D space. The parametric pdf $f(u|\theta)$ representing the 1D source distribution is modeled as a GMM with parameterised means $\{\phi_\theta(\mu_i^s)\}_{i=1}^K$, and the pdf $g(u)$ representing the 1D target distribution is modeled as a GMM with the means $\{\mu_j^t\}_{j=1}^K$. In colour transfer context, we select the means and variances differently, and define the parametric transformation model as described in the following sections.

7.2.1.1 Computing $\mathcal{L}_2$ without using correspondences

In the patch vectors space (see Section 7.2.2.1), we apply the K-means algorithm [1] (following Grogan and Dahyot (2019)) to create the datasets $\{x_i^s\}_{i=1}^K$ and $\{x_j^t\}_{j=1}^K$. Projecting the resulted datasets on 1D subspace (Algorithm 7 step 12) gives 1D datasets $\{u_i^s\}_{i=1}^K$ and $\{u_j^t\}_{j=1}^K$, we set the means of the GMMS as $\{\mu_i^s = u_i^s\}_{i=1}^K$ and $\{\mu_j^t = u_j^t\}_{j=1}^K$ for the source and target respectively. Both the source and target GMMs use identical variance h^2. In this case, the term $\int f(u|\theta)^2$ in Eq. (7.1) is evaluated as follows:

$$\int f(u|\theta)^2 du = \frac{1}{K^2} \sum_{i=1}^K \sum_{j=1}^K \mathcal{N}(0, \phi_\theta(\mu_i^s) - \phi_\theta(\mu_j^s), 2h^2) \tag{7.2}$$

And the term $\int f(u|\theta)g(u)$ in Eq. (7.1) is evaluated as follows:

$$\int f(u|\theta)g(u) du = \frac{1}{K^2} \sum_{i=1}^K \sum_{j=1}^K \mathcal{N}(0, \phi_\theta(\mu_i^s) - \mu_j^t, 2h^2) \tag{7.3}$$

7.2.1.2 Computing $\mathcal{L}_2$ with correspondences

Given the 1D patch correspondences $\{(u_i^s, u_i^t)\}_{i=1}^q$, the correspondence means $\{\mu_i^s, \mu_i^t\}_{i=1}^K$ are selected using the binned correspondences procedure which implemented in INWDT (Section 6.2.2.2, Figure 6.4), as follows:

We divide the support of the source points $\{u_i^s\}_{i=1}^q$ into K number of bins (subintervals), and calculate the averages of the source observations falling in each bin $\{\bar{u}_i^s\}_{i=1}^K$ as well as calculate the averages of the corresponding target values $\{\bar{u}_i^t\}_{i=1}^K$. We set the number of the Gaussians equal to the number of the binned correspondences such that $\{(\mu_i^s, \mu_i^t) = (\bar{u}_i^s, \bar{u}_i^t)\}_{i=1}^K$. With the correspondences, the term $\int f(u|\theta)g(u) du$ (in Eq. (7.1)) is computed as follows:

$$\int f(u|\theta)g(u) du = \frac{1}{K} \sum_{i=1}^K \mathcal{N}(0, \phi_\theta(\mu_i^s) - \mu_i^t, 2h^2) \tag{7.4}$$

---

[1]Note that using the K-means algorithm is equivalent to using the Expectation Maximization algorithm with constraining the algorithm to provide spherical covariance matrices.
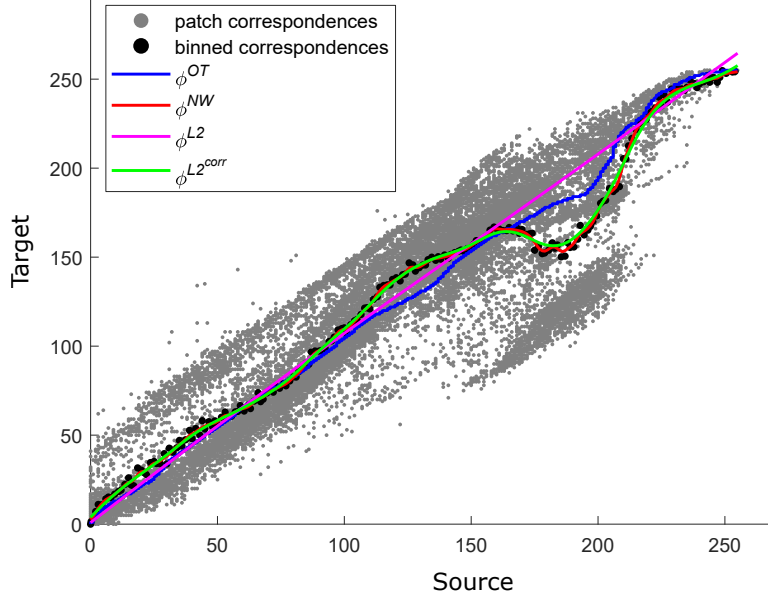
Figure 7.1: Illustration of both $\mathcal{L}_2$ based parametric functions, using correspondences $(\phi^{\mathcal{L}_2{}^{corr}})$ and without $(\phi^{\mathcal{L}_2})$, compared with the non-parametric Nadaraya-Watson mapping function (ϕ^{NW}) that accounts for correspondences and the non-parametric strictly increasing Optimal transport function (ϕ^{OT}) that does not take into account the correspondences.

7.2.1.3 Selecting the bandwidth h:

All Gaussian components share the same bandwidth h. We set a data-driven bandwidth value h using Eq. (6.5) (Jian and Vemuri, 2011), where $X = \{u_i^s\}_{i=1}^n$ is vector of our 1D projected source dataset, and $N = 1$ dimensional space.

7.2.1.4 Defining the parametric transformation model ϕ_θ:

We consider a parametric non-rigid 1D transformation model decomposed into a linear part which is modeled by an affine motion in 1D, and a nonlinear part which is modeled by radial basis functions in 1D, the equation is defined as follows:

$$\phi(u, \theta) = c_0 + c_1 x + \sum_{l=1}^{r} w_l \, \varphi(\|u - u_l\|_2) \tag{7.5}$$

Where $\theta = \{c_0, c_1, w_1, w_2, ..., w_r\}$ the parameters that control the transformation that need to be estimated, $\{u_l\}_{l=1}^r$ are the control points, $\varphi(\|u - u_l\|_2) = \|u - u_l\|_2$ is radial basis function of linear spline type, with $\| \, . \, \|_2$ denoting the Euclidean norm on $\mathbb{R}$. 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 in this case is:

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

with $\dim(u_l) = 1$, $\dim(c_0) = 1$, and $\dim(c_1) = 1$ (Eq. 7.5).

7.2.1.5 Optimization:

In order to estimate the parameters of the transformation ϕ_θ in Eq. (7.5) that minimizes the cost function in Eq. (7.1), we used Quasi-Newton method (Shanno, 1970), which is a gradient-based numerical optimization technique. We computed the analytical derivative of the cost function Eq. (7.1) and passed it to the gradient ascent algorithm to speed up the optimization (Grogan and Dahyot, 2019; Jian and Vemuri, 2011). As for the implementation specifics, we used the Matlab function (*fminunc*) that implements the Quasi-Newton minimization method, which is suitable for our problem formulation. *fminunc* function returns a vector that is a local minimizer of the scalar valued function. In our case, the scalar valued function is our cost function that returns the $\mathcal{L}_2$ distance and the returned vector that gives minimum $\mathcal{L}_2$ distance is our parameters $\theta = \{c_0, c_1, w_1, w_2, ..., w_r\}$.

7.2.2 The Pipeline

We follow INWDT (Chapter 6) in constructing high dimensional descriptors, but we replace INWDT by SL2D in the pipeline in Algorithm 5 (in step 12). For clarity, we explain the pipeline with SL2D in the following.

7.2.2.1 Patch correspondences

We use the same process explained in Section 6.2.2.1 to create a set of corresponding patches. First, giving the source colour pixels denoted by $\{x_i^c\}_{i=1}^n$ and the target colour pixels $\{y_j^c\}_{j=1}^m$, we use SIFT flow motion estimation (Liu et al., 2011) to compute motion flows (explained in Section 5.2.1) and extract the correspondences $\{(x_i^c, y_i^c)\}_{i=1}^n$ in $\mathbb{R}^N$ (for example $N = \dim(x) = 3$ for RGB colour space). Second, we

concatenate the original pixels' positions in the grid coordinate of the images with the colour channels such that $\{(x_i, y_i) = ((x_i^c, x_i^p) , (y_i^c, y_i^p))\}_{i=1}^n$ in $\mathbb{R}^N$, where $N = \dim(x) = \dim(x^c) + \dim(x^p)$. Next, we normalize the coordinates to lie between 0 and 255, and we encode overlapping neighborhoods of pixels correspondences with patch size neighborhood $d \times d$ to create high dimensional pairs of vectors $\{(x_i, y_i)\}_{i=1}^q$ in $\mathbb{R}^N$, where $N = \dim(x) \times d \times d$. We also consider patches containing only colour information of a pixel neighborhood, i.e., $\{(x_i, y_i) = (x_i^c, y_i^c)\}_{i=1}^n$.

7.2.2.2 Sliced $\mathcal{L}_2$ Distance Distribution Transfer

Our SL2D algorithm outlined in Algorithm 7 is applied to our set patch correspondences $\{(x_i, y_i)\}_{i=1}^q$ (input) to compute recoloured patches $\{x_i^{(\infty)}\}_{i=1}^q$ (output). At each iteration k, and at each projection r, following the step of projecting the correspondences $\{(x_i, y_i)\}_{i=1}^q$ originally in $\mathbb{R}^N$ space, where $N = \dim(x) \times d \times d$, to a 1D subspace creating 1D datasets $\{(u_i, v_i)\}_{i=1}^q$ in the projective space, we estimate the parameters θ without using correspondences using Eq.(7.1) with Eq.(7.3) (cf. step 14 in Algorithm 7), or by using correspondences using Eq.(7.1) with Eq.(7.4) (cf. step 20 in Algorithm 7), as explained in Section 7.2.1. Next, we apply the estimated parametric transformation on every data point in the source patch dataset $\{u_i\}_{i=1}^q$ (cf. step 15 & step 21 in Algorithm 7) to derive new estimates $\{\phi_r(u_i, \hat{\theta})\}_{i=1}^q$.

Exploring the Convergence: To explore the convergence, the standard $\mathcal{L}_2$ distance between N-dimensional probability distributions (Jian and Vemuri, 2011) 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 (at step 29 in Algorithm 7). We follow the same process in Section 6.2.2.2 to compute the standard $\mathcal{L}_2$ between N-dimensional pdfs. Figure 7.2 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.

Interpolation between solutions: We can create new transformations by interpolating 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 semi-supervised situa-

tions where correspondences are only partially available (Grogan et al., 2017; Grogan and Dahyot, 2019). For example, given the $\mathcal{L}_2$ derived solution ($\hat{g}_1$) when using $\phi^{\mathcal{L}_2}$ to transfer the source distribution (f) into the target distribution (g), and given the solution ($\hat{g}_2$) when using $\phi^{\mathcal{L}_2{}^{corr}}$ that incorporate the correspondences, we can create a linear interpolation between the two solutions as follows:

$$\forall\, \lambda \in [0,1], \quad S = (1-\lambda)\hat{g}_1 + \lambda\hat{g}_2 \tag{7.6}$$

S is the interpolation between the two solutions $\hat{g}_1$ (obtained at λ=0) and $\hat{g}_2$(obtained at λ=1). This solution is useful when we have large non-overlapped areas or occlusions, Figure 7.9 illustrates some examples of this situation.

7.2.2.3 Merge recoloured candidates.

Because the same pixel is present in overlapping patches $\{x_i^{(\infty)}\}_{i=1}^q$, the average colour value from all the candidates is selected for recolouring the source dataset $\{x_i\}_{i=1}^n$ similar to PCT_OT method (Chapter 4), OT_NW method (Chapter 5), and INWDT method (Chapter 6).
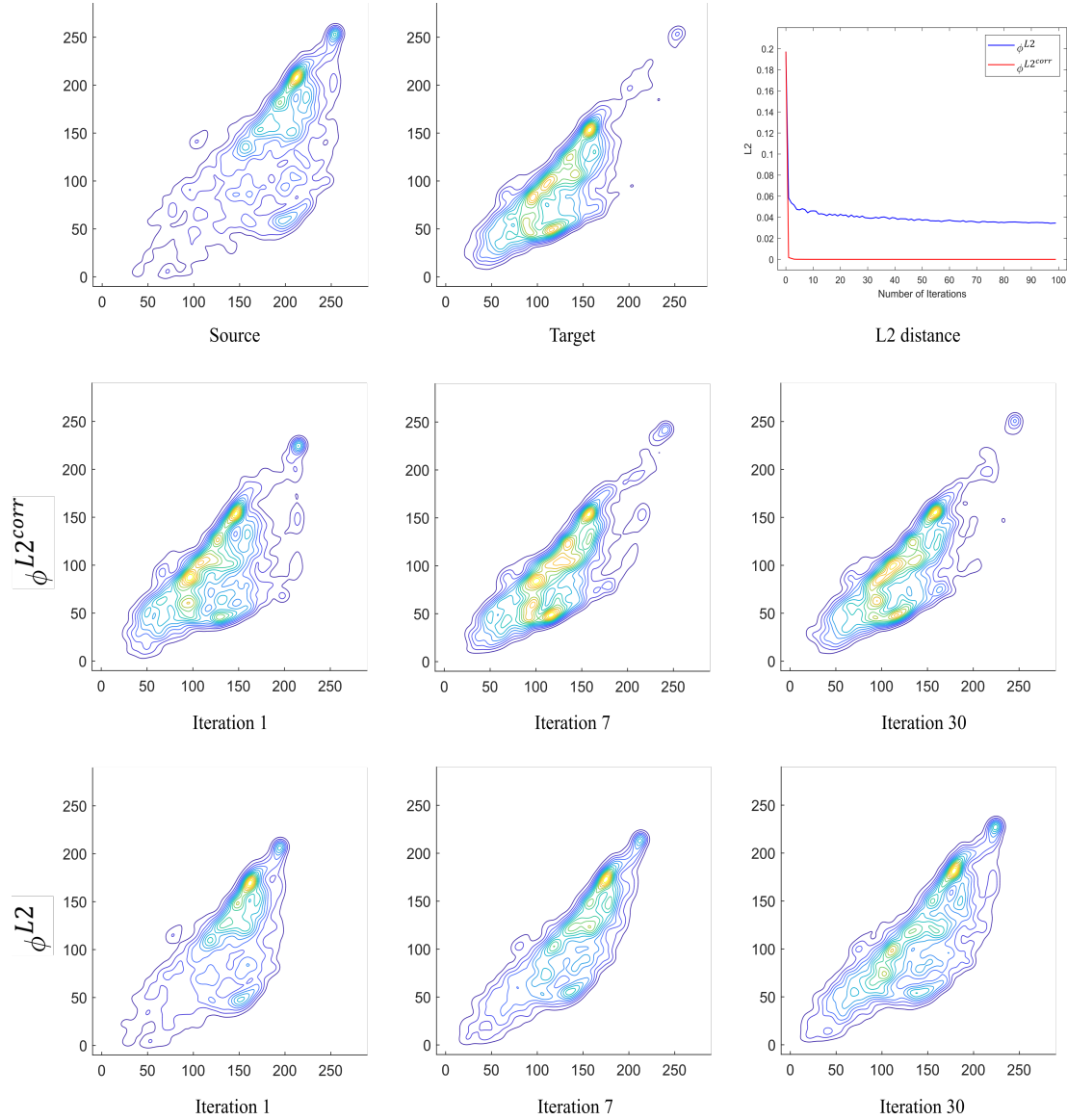
Figure 7.2: Example of pdf of the transferred source patches projected in 2D space (RG). The patch size chosen is 1×1 and only the colour information is used $N = 3$ (space RGB). The standard $\mathcal{L}_2$ distance is computed after each iteration to illustrate the convergence of the source distribution 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.

Algorithm 7: Pseudo-code for Sliced $\mathcal{L}_2$ Distance (SL2D)

1: **Input:** Source and Target datasets with correspondences $\{(x_i^s, x_i^t)\}_{i=1}^q$ if available, samples of r.v. $x \in \mathbb{R}^N$
2: **Initialisation:** $k \leftarrow 0$ and $\forall i, \ x_i^{s(0)} \leftarrow x_i^s$
3: **if** no correspondences taken into account) **then**
4:      Choose K and apply K-means to create source cluster $\{x_i^s\}_{i=1}^K$ and target cluster $\{x_j^t\}_{j=1}^K$
5: **else if** Using correspondences **then**
6:      Pass $\{(x_i^s, x_i^t)\}_{i=1}^q$ to the following steps
7: **end if**
8: **repeat**
9:      Generate N random unit vectors in $\mathbb{R}^N$ stored in matrix $\mathrm{R} = [e_1, ..., e_N]$
10:      **for** $r = 1$ to N **do**
11:          **if** no correspondences **then**
12:              Compute projections $\forall \ i, j \ \ u_i^{s(k)} = e_r^T x_i^{s(k)}$ and $u_j^t = e_r^T x_j^t$
13:              Set the means as $\{\mu_i^{s(k)} = u_i^{s(k)}\}_{i=1}^K$ and $\{\mu_j^t = u_j^t)\}_{j=1}^K$
14:              $\hat{\theta} \leftarrow \underset{\theta}{\operatorname{argmin}} \ [\ \mathcal{L}_2(f(u|\theta), g(u))]$ (Eq. (7.1) with Eq. (7.2) and Eq. (7.3))
15:              Compute 1D $\mathcal{L}_2$ derived estimates $\{\phi_r(\mu_i^{s(k)}, \hat{\theta})\}_{i=1}^K$ & source dataset $\{\phi_r(u_i^{s(k)}, \hat{\theta})\}$
16:          **else if** with correspondences **then**
17:              Compute projections $\forall \ i \ \ u_i^{s(k)} = e_r^T x_i^{s(k)}$ and $u_i^t = e_r^T x_i^t$
18:              Choose K bins and create binned correspondences $\{\bar{u}_i^s, \bar{u}_i^t\}_{i=1}^K$
19:              Set the means $\{(\mu_i^s, \mu_i^t) = (\bar{u}_i^s, \bar{u}_i^t)\}_{i=1}^K$
20:              $\hat{\theta} \leftarrow \underset{\theta}{\operatorname{argmin}} \ [\ \mathcal{L}_2(f(u|\theta), g(u))]$ (Eq. (7.1) with Eq. (7.4))
21:              Compute 1D $\mathcal{L}_2$ derived estimates $\{\phi_r(u_i^{s(k)}, \hat{\theta})\}$ for the source dataset
22:          **end if**
23:      **end for**
24:      **if** no correspondences **then**
25:          Update the means $\{x_i^{s(k+1)} \leftarrow \Phi(\mu_i^{s(k)}) = \sum_{r=1}^D (\phi_r(\mu_i^{s(k)}, \hat{\theta}) \ e_r\}_{i=1}^K$ (Eq. (3.15)
26:      **end if**
27:      Update source dataset $\{x_i^{s(k+1)} \leftarrow \Phi(x_i^{s(k)}) = \sum_{r=1}^D (\phi_r(u_i^{s(k)}, \hat{\theta})) \ e_r\}$
28:      $k \leftarrow k + 1$
29: **until** convergence $\mathcal{L}_2(f, g) \to 0$ (noted $k \equiv \infty$)
30: **Result:** The final one-to-one mapping Φ in $\mathbb{R}^N$ is given by
     $\forall i, \{ \ x_i\} \mapsto \{\Phi(x_i)\} = \{x_i^{(\infty)}\}$

7.3 Experimental Assessment

We provide extensive quantitative and qualitative evaluations to validate our four SL2D based solutions: the first without correspondences using colour patches only annotated in the results as SL2D_c, the second without correspondences using colour patches with pixel location information annotated as SL2D_cp, the third with correspondences using colour patches only annotated in the results as $\text{SL2D}_\text{c}^\text{corr}$, and the fourth with correspondences using colour patches with pixel location information annotated as $\text{SL2D}_\text{cp}^\text{corr}$. We use the evaluation tools presented in Section 4.4.4 which consist of using numerical metrics, summary descriptive statistics, statistical significance test, and visual analysis. We compare our methods with different state of the art colour transfer methods noted by B-PMLS (Hwang et al., 2019), L2 (Grogan and Dahyot, 2019), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), PMLS (Hwang et al., 2014), IDT (Pitie et al., 2007), also we compare them with PCT_OT (Chapter 4), OT_NW (Chapter 5) and INWDT (Chapter 6). In these evaluations we use image pairs with similar content from an existing dataset provided by Hwang et al. (Hwang et al., 2014). The dataset includes registered pairs of images (source and target) taken with different cameras, different in-camera settings, and different illuminations and recolouring styles.

Note that the results using PMLS and B-PMLS were provided by the authors Hwang et al. (2014); Hwang et al. (2019), respectively. We also compared them with two other recent techniques (Xia et al., 2017; Park et al., 2016) that account for correspondences into their framework, but PMLS has been shown to outperforms these two (Grogan and Dahyot, 2019), hence PMLS is the one reported here. We also compared them with (Bonneel et al., 2016; Ferradans et al., 2013) that do not account for correspondences, but IDT has been shown to be superior (Grogan and Dahyot, 2019), hence IDT is the one reported here for ease of comparison.

7.3.1 Colour space and parameters settings

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 $\text{SL2D}_\text{cp}^\text{corr}$ 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 $\texttt{SL2D}_\texttt{c}$ and $\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$, pixel position is not accounted for, and only RGB colours are used, which create patch vectors in 27 dimensions ($N = 3 \times 3 \times 3 = 27$).

7.3.2 Experimental Results

Quantitative evaluation: The numerical results of each metric with means and standard errors (SE) are shown in Tables 7.3 to 7.6, along with box plots shown in Figures 7.3 to 7.6, respectively. The purpose of the box plots is to visualize differences among methods and to show how close our method is to the state of the art algorithms. By looking at the tables (Tables 7.3 to 7.6), with all metrics we can see that Sliced $\mathcal{L}_2$ with correspondences ($\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$ and $\texttt{SL2D}_\texttt{cp}^{\texttt{corr}}$) significantly outperforms the Sliced $\mathcal{L}_2$ solutions without correspondences ($\texttt{SL2D}_\texttt{c}$ and $\texttt{SL2D}_\texttt{cp}$), and this can be clearly visualized in the box plots. Also, in all the metrics results, we find that incorporating colour and spatial information ($\texttt{SL2D}_\texttt{cp}^{\texttt{corr}}$) improves the performance over using colour space only ($\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$), this can be visualized in the box plots. Moreover, all metrics results in the tables show that the iterative projection approach with $\mathcal{L}_2$ solutions that accounts for correspondences (colour only $\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$ and combined colour and position $\texttt{SL2D}_\texttt{cp}^{\texttt{corr}}$) outperform the iterative projection approach with OT solution (IDT algorithm). Both solutions ($\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$ and $\texttt{SL2D}_\texttt{cp}^{\texttt{corr}}$) aim to extend the Grogan and Dahyot's full $\mathcal{L}_2$ solution annotated in the results by $\texttt{L2}$ (Grogan and Dahyot, 2019) to higher dimensions (note that we report here the correspondences based $\texttt{L2}$ solution as reported in (Grogan and Dahyot, 2019) to be a superior to the solution that does not account for correspondences), and when comparing their performances with $\texttt{L2}$, we find that by looking at the tables (Tables 7.3 to 7.6), all metrics show that the Sliced $\mathcal{L}_2$ approach which solves the high dimensional problem (up to 27 and 45 dimensions) outperforms $\texttt{L2}$ (which is limited to 3D space) in some cases, and provide similar performance in other cases. However, by looking at the summary statistics (the box plots) we see that both solutions $\texttt{SL2D}_\texttt{c}^{\texttt{corr}}$ and $\texttt{SL2D}_\texttt{cp}^{\texttt{corr}}$ and $\texttt{L2}$ solution (Grogan and Dahyot, 2019) and also the rest of the best performances from the sate of the art, namely, $\texttt{B-PMLS}$ (Hwang et al., 2019), $\texttt{PMLS}$ (Hwang et al., 2014), also $\texttt{OT\_NW}$ (Chapter 5) and $\texttt{INWDT}$ (Chapter 6) all greatly overlap with each other, the median and mean values (the means shown as red dots in the plots, and the medians shown as horizontal black lines) are very close in value, and this indicates similar

Metrics	P-values
PSNR	0.954
SSIM	0.8662
CID	0.379
FSIMc	0.9934

Table 7.1: This table shows the results of Kruskal-Wallis Rank Sum statistical test for each metric separately. The test tests if there are significant differences between $\text{SL2D}_{\text{c}}^{\text{corr}}$ and $\text{SL2D}_{\text{cp}}^{\text{corr}}$ results and the top performance methods, namely, B-PMLS, L2, PMLS, OT_NW and INWDT. With 95% confidence level, the P-values for all metrics are greater than the significance level $\alpha = 0.05$, which indicate that there is no significant difference between the aforementioned methods.

performances. We run the statistical significance analysis (Kruskal-Wallis Rank Sum test) to see if there is significant differences between the methods that greatly overlap in the box plots, namely, B-PMLS, L2, PMLS, OT_NW, INWDT, $\text{SL2D}_{\text{c}}^{\text{corr}}$ and $\text{SL2D}_{\text{cp}}^{\text{corr}}$. For each individual metric results, the Kruskal-Wallis Rank Sum test compares the resulting medians of the methods. The null hypothesis H_o states that all medians of the methods distributions are equal, the alternative hypothesis H_1 states that at least one median is different. The confidence level $= 95\%$, and the significance level $\alpha = 0.05$. If the resulting P-value $< \alpha$, the null hypothesis is rejected and alternative hypothesis is accepted. We found that with 95% confidence level, the P-values for all metrics are greater than the significance level $\alpha = 0.05$, which indicate that there is no significant difference between the aforementioned methods, Table 7.1.

Qualitative Evaluation: Figure 7.7 and Figure 7.8 provide qualitative results. For clarity, the results in Figure 7.8 are presented in image mosaics, created by switching between the target image and the transformed source image column wise. If the colour transfer is accurate, the resulting mosaic should look like a single image (ignoring the small motion displacement between source and target images), otherwise colour column differences appear. Figure 7.8 shows that $\text{SL2D}_{\text{c}}^{\text{corr}}$ and $\text{SL2D}_{\text{cp}}^{\text{corr}}$ that accounts for correspondences are successful in alleviating colour column differences compared with the other two versions SL2D_{c} and SL2D_{cp} that do not account for correspondences. Moreover, Figure 7.8 shows that $\text{SL2D}_{\text{c}}^{\text{corr}}$ and $\text{SL2D}_{\text{cp}}^{\text{corr}}$ are more successful in alleviating the column differences than the full L2 (cf. Figure 7.8 in examples "building" and "Tonal 4"), but OT_NW (Chapter 5) is still the best in eliminating

the colour column differences competing with `B-PMLS`.

Figure 7.7 shows more qualitative results. All sliced $\mathcal{L}_2$ distance solutions are robust to registration errors and do not introduce visual artifacts, unlike `PMLS` and `B-PMLS` which produce those artifacts although they provide the two top scores in terms of the numerical metrics (Tables 7.3 to 7.6). Moreover, all sliced $\mathcal{L}_2$ distance solutions can transfer colours between images that contain significant spatial differences and eliminate the shadow artifacts unlike `PCT_OT` (cf. Figure 7.7 with examples "building", "illum", "mart"). Also, we found that indeed the results of $\text{SL2D}_{\text{c}}^{\text{corr}}$ and $\text{SL2D}_{\text{cp}}^{\text{corr}}$ that accounts for correspondences are very similar, visually, to `INWDT` method (Chapter 6), they are all smooth functions and accounts for correspondences (Figure 7.1 shows the mappings ϕ^{NW} and $\phi^{\mathcal{L}_2^{corr}}$ in 1D). All sliced $\mathcal{L}_2$ distance solutions can create a smooth colour transfer results, even smoother than the full `L2` (Grogan and Dahyot, 2019), and can also alleviate JPEG compression artifacts and noise (cf. Figure 7.7 for comparison).

We further show experimental results in Figure 7.9, where we interpolate (with $\lambda = 0.2$) between our solutions $\text{SL2D}_{\text{cp}}^{\text{corr}}$ which accounts for correspondences and SL2D_{cp} which does not account for correspondences and we apply the colour transfer between images with severe scene changes. These images exhibit large non-overlapped areas where there are no correspondences. We compare our results with state of the art methods `B-PMLS` and also (HaCohen et al., 2011) that uses Non-Rigid Dense Correspondence (NRDC) which has shown effective results in finding dense correspondence between partially overlapped images. As we can see from the figure, incorporating the correspondences lead to reasonable results (cf. Figure 7.9 last column).

Time Complexity: Here we detail the time complexity of our algorithms for comparison. The 1D iterative projection approach involves iterative projection of dataset values originally in high dimensional space to a 1D subspaces and solve 1D transformation functions and then propagate the solution back to the original space. This operation is repeated with different directions in 1D space until convergence. The 1D transformation problems are independent of each other; hence the process is suitable for parallel computation and can be run independently for each pixel.

When we use the 1D transformation function as OT in the form of CDF-matching, the algorithm takes $\mathcal{O}(k)$ where k is the number of the algorithm's iterations. When we use the 1D transformation function as OT in the form of quantile-matching, the algo-

Optimal Transport		Nadaraya-Watson	$\mathcal{L}_2$ Inference
CDF-matching $\mathcal{O}(k)$			
		$\mathcal{O}(k)$	$\mathcal{O}(rk)$
Quantile-matching $\mathcal{O}(n\log(n)k)$			

Table 7.2: Time complexity: when we use the 1D transformation function as Optimal Transport (Chapter 4 and Chapter 5), when we use the 1D transformation function as Nadaraya-Watson estimator (Chapter 6) and when we use the 1D transformation function based on the $\mathcal{L}_2$ framework (Chapter 7). Where k is the the number of iterations of the algorithm, n is number of samples in the dataset and r is the number of iterations of the gradient ascent optimization function.

rithm takes $\mathcal{O}(n\log(n)k)$ where n is the number of the samples in the dataset. When we use the 1D transformation function as Nadaraya-Watson estimator, the Gaussian kernel can be computed independently for each pixel hence the algorithm takes $\mathcal{O}(k)$. When we use the 1D transformation function based on the $\mathcal{L}_2$ framework, the algorithm takes $\mathcal{O}(rk)$ where r is the number of the iterations of the gradient ascent optimization function, Table 7.2 shows a summary for the time complexity for each algorithm.

Generally, the actual execution time of an algorithm typically relies on the programming language, the style of implementation, and the hardware resources chosen (Cormen et al., 2009). Our approach is similar to neural networks architecture; hence we can use the same strategy of implementing the neural networks using libraries such as PyTorch (Paszke et al., 2019) and Tensorflow (Abadi et al., 2015) that use GPUs. However, we are not learning the projections (weights) as neural networks do. We, in fact, are learning more complex activation functions. Thus, here we only have discussed the time complexity of algorithms theoretically. The task of optimizing the implementation specifics using neural network libraries is left for future work.

Scene	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	PCT_OT 2019	OT_NW 2020	INWDT 2020	$SL2D_c$ 3x3	$SL2D_{cp}$ 3x3	$SL2D_c^{corr}$ 3x3	$SL2D_{cp}^{corr}$ 3x3
Gangnam1	25.354	35.725	24.048	23.936	36.165	35.358	31.479	33.565	32.241	24.336	25.293	30.571	32.671
Gangnam2	27.116	36.553	25.952	25.944	36.255	35.524	35.502	33.627	32.783	26.577	27.194	32.161	33.523
Gangnam3	22.372	35.007	21.908	21.913	36.022	33.284	26.393	28.217	29.269	21.845	22.183	27.483	29.445
Illum	19.822	20.167	19.785	19.960	21.960	19.079	20.306	20.858	21.066	18.097	19.131	20.213	20.975
Building	20.554	22.634	22.736	22.769	24.867	20.499	25.019	24.039	22.895	15.950	18.531	22.656	23.457
Playground	27.184	27.835	25.501	25.436	28.172	27.647	28.482	28.491	28.625	25.785	26.457	28.653	28.777
Flower1	24.238	26.981	23.765	23.706	27.194	26.857	25.186	27.158	27.530	24.141	24.298	27.130	27.669
Flower2	25.417	25.760	25.259	25.223	25.978	25.772	26.373	26.497	26.445	24.227	23.989	26.349	26.447
Tonal1	30.082	37.215	31.617	31.413	37.556	37.332	37.044	37.151	36.719	20.697	24.183	37.404	37.750
Tonal2	27.992	31.508	25.062	25.087	32.229	31.356	32.049	31.579	32.394	24.722	26.963	32.336	32.897
Tonal3	29.575	36.246	28.136	28.065	36.826	36.644	33.793	35.014	34.900	23.770	25.314	34.665	35.672
Tonal4	28.605	34.521	28.852	28.848	35.697	34.344	33.819	35.320	35.202	19.862	23.254	34.843	35.789
Tonal5	30.330	35.260	29.580	29.448	36.870	34.303	36.437	36.616	35.656	21.230	24.678	35.181	36.479
Mart	22.747	24.742	23.183	23.196	24.936	24.450	24.509	25.189	25.574	23.900	24.315	25.330	25.629
Sculpture	29.884	32.062	29.037	28.820	32.635	32.067	31.237	32.735	33.356	27.428	28.799	33.757	33.867
Mean	26.085	30.814	25.628	25.584	31.557	30.301	29.842	30.404	30.310	22.838	24.305	29.915	30.736
SE	0.905	1.459	0.841	0.821	1.408	1.518	1.306	1.291	1.249	0.823	0.718	1.296	1.320

PSNR ↑

Table 7.3: 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 (Salomon, 2004). Red, blue, and green indicate 1st, 2nd, and 3rd top scores respectively in the table (higher values are better, best viewed in colour and zoomed in).

Scene	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	PCT_OT 2019	OT_NW 2020	INWDT 2020	SL2D$_c$ 3x3	SL2D$_{cp}$ 3x3	SL2D$_c^{corr}$ 3x3	SL2D$_{cp}^{corr}$ 3x3
Gangnam1	0.900	0.992	0.892	0.891	0.990	0.990	0.964	0.973	0.965	0.915	0.917	0.959	0.969
Gangnam2	0.920	0.993	0.909	0.909	0.987	0.986	0.980	0.976	0.972	0.932	0.943	0.972	0.976
Gangnam3	0.859	0.991	0.873	0.864	0.987	0.980	0.930	0.959	0.955	0.871	0.856	0.947	0.961
Illum	0.641	0.649	0.643	0.646	0.733	0.648	0.673	0.687	0.678	0.640	0.656	0.665	0.676
Building	0.808	0.865	0.864	0.863	0.899	0.862	0.888	0.875	0.870	0.795	0.837	0.873	0.875
Playground	0.920	0.940	0.878	0.876	0.939	0.939	0.939	0.943	0.941	0.915	0.923	0.944	0.944
Flower1	0.909	0.967	0.913	0.912	0.968	0.966	0.926	0.959	0.963	0.935	0.940	0.962	0.965
Flower2	0.901	0.928	0.894	0.894	0.930	0.927	0.933	0.939	0.930	0.908	0.904	0.929	0.931
Tonal1	0.953	0.988	0.971	0.970	0.989	0.987	0.988	0.991	0.988	0.928	0.953	0.988	0.988
Tonal2	0.968	0.987	0.926	0.926	0.989	0.986	0.988	0.986	0.988	0.959	0.968	0.988	0.990
Tonal3	0.962	0.992	0.947	0.946	0.993	0.992	0.987	0.990	0.989	0.946	0.953	0.988	0.991
Tonal4	0.944	0.983	0.932	0.932	0.987	0.983	0.981	0.985	0.985	0.907	0.938	0.985	0.987
Tonal5	0.965	0.986	0.953	0.954	0.990	0.985	0.990	0.991	0.987	0.914	0.941	0.986	0.989
Mart	0.904	0.957	0.925	0.925	0.958	0.956	0.941	0.954	0.957	0.945	0.948	0.957	0.958
Sculpture	0.942	0.971	0.934	0.932	0.973	0.972	0.945	0.974	0.971	0.955	0.959	0.976	0.976
Mean	0.900	0.946	0.897	0.896	0.954	0.944	0.937	0.946	0.943	0.898	0.909	0.941	0.945
SE	0.022	0.023	0.020	0.020	0.017	0.023	0.020	0.020	0.020	0.021	0.020	0.021	0.021

SSIM ↑

Table 7.4: Comparing our algorithms SL2D$_c$, SL2D$_{cp}$, SL2D$_c^{corr}$ and SL2D$_{cp}^{corr}$ with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Red, blue, and green indicate 1st, 2nd, and 3rd top scores respectively in the table (higher values are better, best viewed in colour and zoomed in).
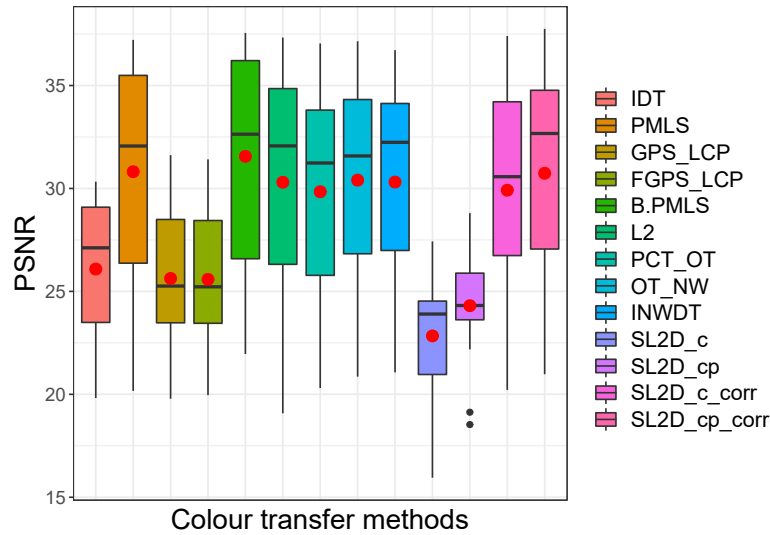
Figure 7.3: Comparing our algorithms $\mathtt{SL2D_c}$, $\mathtt{SL2D_{cp}}$, $\mathtt{SL2D_c^{corr}}$ and $\mathtt{SL2D_{cp}^{corr}}$ with the state of the art colour transfer methods using PSNR metric (Salomon, 2004). Detailed quantitative results in Table 7.3 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

Figure 7.4: Comparing $\mathtt{SL2D_c}$, $\mathtt{SL2D_{cp}}$, $\mathtt{SL2D_c^{corr}}$ and $\mathtt{SL2D_{cp}^{corr}}$ with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). Detailed quantitative results in Table 7.4 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

| Scene	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	PCT_OT 2019	OT_NW 2020	INWDT 2020	SL2D$_c$ 3x3	SL2D$_{cp}$ 3x3	SL2D$_c^{corr}$ 3x3	SL2D$_{cp}^{corr}$ 3x3
Gangnam1	0.252	0.040	0.226	0.222	0.044	0.048	0.085	0.088	0.094	0.227	0.186	0.114	0.091
Gangnam2	0.268	0.039	0.291	0.292	0.048	0.089	0.068	0.109	0.123	0.343	0.313	0.129	0.113
Gangnam3	0.496	0.108	0.472	0.487	0.140	0.193	0.261	0.267	0.297	0.465	0.483	0.316	0.263
Illum	0.386	0.390	0.395	0.396	0.315	0.397	0.377	0.376	0.373	0.446	0.406	0.400	0.372
Building	0.374	0.228	0.313	0.321	0.188	0.249	0.183	0.275	0.328	0.365	0.357	0.308	0.283
Playground	0.440	0.238	0.443	0.471	0.221	0.254	0.209	0.221	0.253	0.447	0.429	0.256	0.238
Flower1	0.389	0.163	0.396	0.400	0.164	0.174	0.285	0.194	0.233	0.320	0.318	0.260	0.215
Flower2	0.337	0.245	0.322	0.323	0.235	0.266	0.218	0.201	0.250	0.347	0.375	0.274	0.241
Tonal1	0.310	0.101	0.285	0.308	0.093	0.111	0.097	0.063	0.105	0.413	0.332	0.122	0.101
Tonal2	0.288	0.128	0.351	0.347	0.097	0.145	0.099	0.118	0.129	0.322	0.319	0.134	0.112
Tonal3	0.244	0.079	0.294	0.294	0.085	0.081	0.077	0.079	0.107	0.289	0.338	0.091	0.085
Tonal4	0.240	0.108	0.248	0.238	0.082	0.107	0.065	0.065	0.091	0.314	0.278	0.098	0.076
Tonal5	0.156	0.091	0.205	0.192	0.062	0.092	0.051	0.067	0.094	0.309	0.296	0.102	0.075
Mart	0.526	0.219	0.405	0.402	0.203	0.225	0.426	0.249	0.246	0.407	0.302	0.257	0.242
Sculpture	0.242	0.137	0.213	0.224	0.116	0.143	0.232	0.120	0.172	0.269	0.212	0.134	0.122
Mean	0.330	0.154	0.324	0.328	0.139	0.172	0.182	0.166	0.193	0.352	0.329	0.200	0.175
SE	0.027	0.024	0.022	0.023	0.020	0.024	0.031	0.025	0.025	0.018	0.020	0.026	0.024

Table 7.5: Comparing our algorithms SL2D$_c$, SL2D$_{cp}$, SL2D$_c^{corr}$ and SL2D$_{cp}^{corr}$ with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Red, blue, and green indicate 1st, 2nd, and 3rd top scores respectively in the table (lower values are better, best viewed in colour and zoomed in).

Scene	IDT 2007	PMLS 2014	GPS/LCP 2018	FGPS/LCP 2018	B-PMLS 2019	L2 2019	PCT_OT 2019	OT_NW 2020	INWDT 2020	SL2D$_c$ 3x3	SL2D$_{cp}$ 3x3	SL2D$_c^{corr}$ 3x3	SL2D$_{cp}^{corr}$ 3x3
						FSIMc ↑							
Gangnam1	0.936	0.986	0.944	0.943	0.986	0.985	0.972	0.979	0.975	0.944	0.948	0.971	0.977
Gangnam2	0.952	0.992	0.962	0.962	0.989	0.988	0.990	0.986	0.982	0.961	0.964	0.981	0.983
Gangnam3	0.946	0.992	0.962	0.961	0.993	0.990	0.987	0.982	0.980	0.959	0.954	0.978	0.981
Illum	0.800	0.819	0.824	0.824	0.859	0.818	0.839	0.844	0.828	0.815	0.820	0.821	0.827
Building	0.874	0.928	0.930	0.929	0.949	0.926	0.942	0.937	0.928	0.882	0.900	0.927	0.931
Playground	0.950	0.958	0.933	0.932	0.954	0.955	0.956	0.960	0.956	0.945	0.951	0.958	0.958
Flower1	0.954	0.975	0.968	0.967	0.976	0.976	0.971	0.977	0.975	0.967	0.966	0.975	0.976
Flower2	0.941	0.950	0.945	0.945	0.951	0.949	0.954	0.956	0.950	0.945	0.943	0.951	0.951
Tonal1	0.964	0.997	0.986	0.986	0.997	0.997	0.998	0.997	0.996	0.950	0.961	0.997	0.997
Tonal2	0.984	0.993	0.973	0.973	0.995	0.992	0.993	0.992	0.993	0.973	0.979	0.992	0.993
Tonal3	0.979	0.997	0.984	0.983	0.997	0.997	0.997	0.995	0.993	0.970	0.973	0.994	0.995
Tonal4	0.966	0.989	0.972	0.973	0.992	0.990	0.995	0.994	0.993	0.949	0.964	0.991	0.993
Tonal5	0.980	0.994	0.987	0.987	0.997	0.993	0.998	0.997	0.995	0.949	0.965	0.994	0.996
Mart	0.946	0.969	0.960	0.959	0.969	0.967	0.969	0.970	0.971	0.963	0.964	0.970	0.971
Sculpture	0.980	0.987	0.982	0.980	0.987	0.988	0.988	0.987	0.986	0.981	0.982	0.989	0.989
Mean	0.943	0.968	0.954	0.954	0.973	0.967	0.970	0.970	0.967	0.943	0.949	0.966	0.968
SE	0.012	0.012	0.010	0.010	0.009	0.012	0.010	0.010	0.011	0.011	0.010	0.011	0.011

Table 7.6: Comparing SL2D$_c$, SL2D$_{cp}$, SL2D$_c^{corr}$ and SL2D$_{cp}^{corr}$ with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Red, blue, and green indicate 1st, 2nd, and 3rd top scores respectively in the table (higher values are better, best viewed in colour and zoomed in).
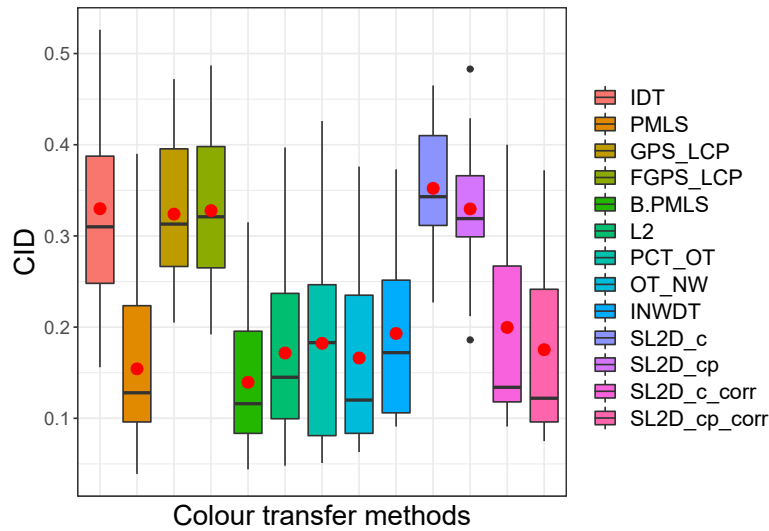
Figure 7.5: Comparing $\texttt{SL2D}_\texttt{c}$, $\texttt{SL2D}_\texttt{cp}$, $\texttt{SL2D}_\texttt{c}^\texttt{corr}$ and $\texttt{SL2D}_\texttt{cp}^\texttt{corr}$ with the state of the art colour transfer methods using CID metric (Preiss et al., 2014). Detailed quantitative results in Table 7.5 are summarized in a box plot (lower values are better, best viewed in colour and zoomed in).
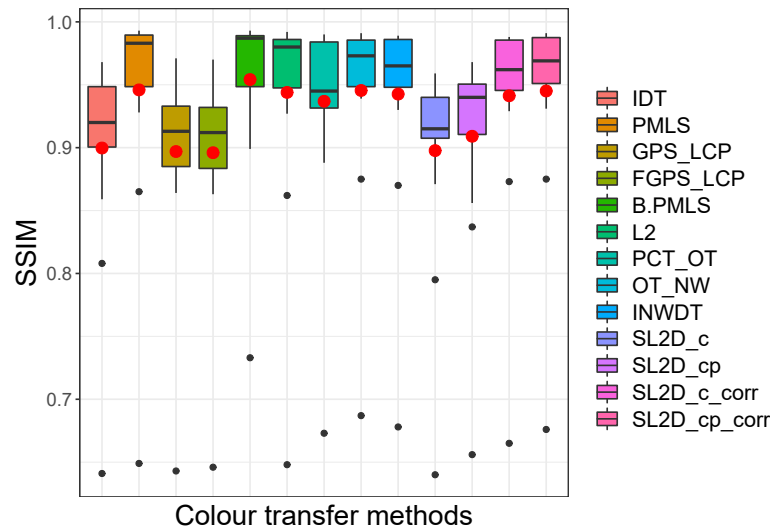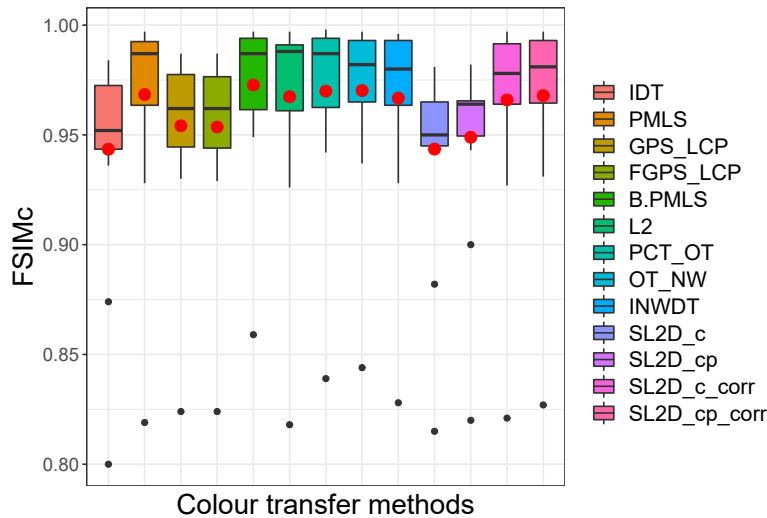

Figure 7.6: Comparing $\texttt{SL2D}_\texttt{c}$, $\texttt{SL2D}_\texttt{cp}$, $\texttt{SL2D}_\texttt{c}^\texttt{corr}$ and $\texttt{SL2D}_\texttt{cp}^\texttt{corr}$ with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). Detailed quantitative results in Table 7.6 are summarized in a box plot (higher values are better, best viewed in colour and zoomed in).

Source	Target	PMLS 2014	B-PMLS 2019	L2 2019	PCT_OT 2019

OT_NW 2020	INWDT 2020	SL2D$_c$ 3x3	SL2D$_{cp}$ 3x3	SL2D$_c^{corr}$ 3x3	SL2D$_{cp}^{corr}$ 3x3
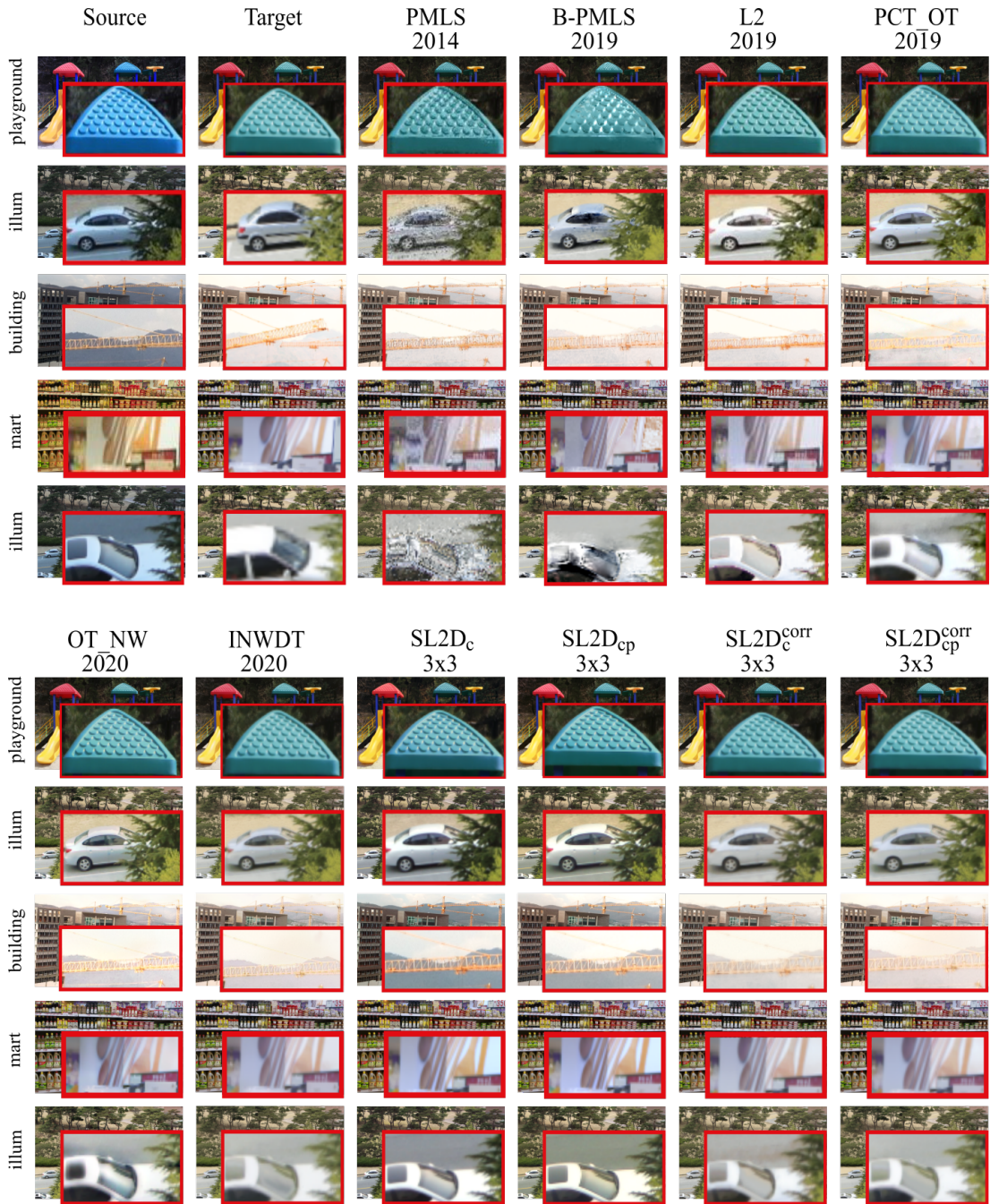
Figure 7.7: Visual analysis part 1: a close up look at some of the results generated using the PMLS (Hwang et al., 2014), B-PMLS (Hwang et al., 2019), L2 (Grogan and Dahyot, 2019), PCT_OT (Chapter 4), OT_NW (Chapter 5), INWDT (Chapter 6) and our algorithms using correspondences (SL2D$_c^{corr}$ and SL2D$_{cp}^{corr}$) and without using correspondences (SL2D$_c$ and SL2D$_{cp}$) - best viewed in colour and zoomed in.

Figure 7.8: Visual analysis part 2: a close up look at some of the results generated using the IDT (Pitie et al., 2007), PMLS (Hwang et al., 2014), GPS/LCP and FGPS/LCP (Bellavia and Colombo, 2018), B-PMLS (Hwang et al., 2019), L2 (Grogan and Dahyot, 2019), PCT_OT (Chapter 4), OT_NW (Chapter 5), INWDT (Chapter 6) 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.

| Source | Target | Flow Field | NRDC | B-PMLS | SL2D$_{cp}$ | $(1-\lambda)$ SL2D$_{cp}$ $+\lambda$ SL2D$_{cp}^{corr}$ |
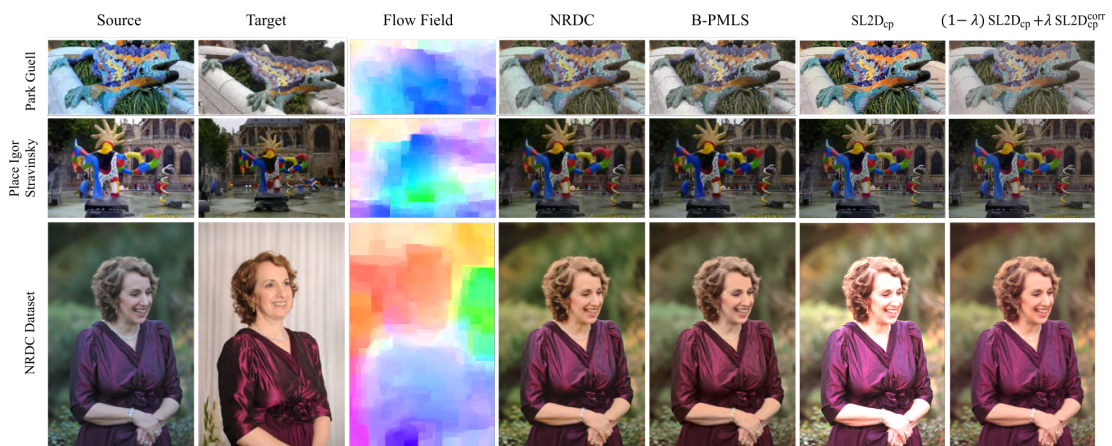
Figure 7.9: This figure shows our colour transfer results of interpolating between our solutions SL2D$_{cp}^{corr}$ and SL2D$_{cp}$ with comparisons to the colour transfer results in (Hwang et al., 2019; HaCohen et al., 2011).

7.4 Conclusion

In this chapter, several contributions to colour transfer with $\mathcal{L}_2$ distance have been made. We conducted extensive experiments and analysis that show quantitative and qualitative competitive results as compared with the leading state of the art methods of colour transfer. Using the 1D projection approach, we have introduced the Sliced $\mathcal{L}_2$ Distance approach between high dimensional probability density functions, where only 1D marginal distributions are used in the estimation process, which involves low computational costs. Unlike IDT, $\mathcal{L}_2$ with 1D projection approach allows colour correspondences that may be available to be taken into account. Also, we have shown how to combine solutions that take advantage of correspondences between the source and target images ($\phi^{\mathcal{L}_2^{corr}}$) and without correspondences ($\phi^{\mathcal{L}_2}$) to tackle semi-supervised situations where correspondences are only partially available.

SUMMARY AND CONCLUSIONS

8.1 Overview

The works undertaken in this thesis aim to contribute to the field of example-based colour transfer (photometric alignment), which modifies the colour of a source image using a colour palette of a target image. The colour transfer problem is formulated as a distribution transfer problem, where the actual probability density function (PDF) of the target colour samples is transferred to the source colour samples. We constructed high dimensional image descriptors by fusing different information such as spatial information and colour information and encoding pixel neighborhood information to guide the transfer process. We investigated solving the high-dimensional distribution transfer problem using the 1D iterative projection approach, using three statistical methods: Optimal Transport, $\mathcal{L}_2$ inference, and Nadaraya–Watson estimator. According to the 1D iterative projection approach only 1D marginal distributions are used in the estimation process, which involves low computational costs. The extensive experiments and analysis conducted in this thesis show quantitative and qualitative competitive results as compared with the leading state of the art methods of colour transfer. An overview of the work contributions achieved and the potential avenues that could be explored for future work are summarized as follows.

8.2 Summary of Contributions

We have proposed in this thesis methods for solving colour transfer problems that demonstrate three advantages: scalability to high-dimensional spaces, reduction of

computational costs, suitability for parallel optimization and providing competitive results as compared with the leading state of the art colour transfer methods. We now list the main contributions made in our work.

8.2.1 Patch-based Colour Transfer with Optimal Transport

Popular Optimal Transport solutions for colour transfer have been proposed by many researchers as described in Chapter 2. However, as the Optimal Transport process is performed in the colour space, it does not consider that coherent colours should be transferred to neighboring pixels, which can create results with blocky artifacts emphasizing JPEG compression blocks or increasing noise. To tackle this problem, we investigated the use of neighborhoods of pixels (patches) with Optimal Transport algorithm in high dimensional space in order to preserve local topology and transfer coherent colours to neighboring pixels, taking into account both pixels colour and spatial information. Using the 1D iterative projection approach, we solved the high dimensional distribution transfer problem in 1D subspaces using Optimal Transport transfer function, where only 1D marginal distributions are used in the estimation process, thus requiring low computational costs. The high dimensional construction improved the pixel descriptors in the 1D subspaces and compensated for the loss of the original dimensional space structure. This original formulation implies a new reconstruction step since each recoloured pixel benefits from the contribution of several estimated candidates, due to the overlapped patches, using an averaging solution that allows denoising and artifact removal as well as colour transfer. Our method's performance outperforms the IDT and SWD algorithms which use Optimal Transport in colour space only. Moreover, extensive experiments and analysis show quantitative and qualitative improvements over other state of the art colour transfer methods.

8.2.2 Optimal Transport with SIFT Flow

Optimal Transport solves the transfer problem with minimum cost but, by definition, does not take into account any available information about correspondences when computing the optimal solution, while the ability of taking advantage of the available correspondences would guide the transfer process and improve the results. To tackle this problem, we proposed to estimate motions flow between images using motion es-

timation (SIFT flow approach (Liu et al., 2011)) and incorporate the spatial correspondence information in the encoded overlapping neighborhoods of pixels. This new formulation makes OT implicitly take into account correspondences to guide the estimation process. Also, we introduced smoothing as part of the iterative algorithms for solving optimal transport, namely Iterative Distribution Transport (IDT) and its variant, the Sliced Wasserstein Distance (SWD). The algorithm allows denoising, shadow artifact removal, smoother colour transfer as well as improving the results over our formulation summarized in Section 8.2.1.

8.2.3 Application of Non-parametric Kernel Regression to the Colour Transfer Problem

Optimal Transport by definition does not take into account any available information about the correspondences when computing the mapping and it is not smooth. Hence, we investigated colour transfer with the non-parametric Nadaraya-Watson that explicitly accounts for input correspondences in matching high-dimensional probability density functions. The high-dimensional PDF transfer problem is tackled using the 1D iterative projection approach, where only 1D datasets are used in the Nadaraya-Watson estimation process, which therefore involves low computational costs. The algorithm significantly outperforms the Optimal Transport variant that does not account for correspondences. Extensive quantitative and qualitative experiments show competitive results as compared with the leading state of the art of colour transfer in images where spaces of dimension up to 45 have been used.

8.2.4 Sliced $\mathcal{L}_2$ Distance for Colour Transfer

The use of $\mathcal{L}_2$ divergence between source and target PDFs to align the distributions have been successfully used in 2D/3D problem spaces in the field of shape registration (Jian and Vemuri, 2011) and recently in colour transfer context (Grogan and Dahyot, 2019) but this standard $\mathcal{L}_2$ framework is not scalble to high-dimensional problems. The 1D projection approach motivated us to propose the Sliced $\mathcal{L}_2$ Distance which significantly involves lower computational requirements than the standard $\mathcal{L}_2$ distance. The high-dimensional distributions are iteratively projected to one-dimensional marginal distributions and the 1D parametric transformation function is estimated by

minimizing the $\mathcal{L}_2$ distance between 1D GMMs representing the source and target marginal distributions. We investigated how this new formulation performed when applied to high-dimensional colour transfer problem. We conducted extensive experiments and analysis that show quantitative and qualitative competitive results over the leading state of the art methods of colour transfer in images where spaces of dimension up to 45 have been used.

8.2.5 Conclusions and recommendations across the approaches

While all the approaches report a comparative performance with each other, some preferences can be taken into account when choosing a particular approach over the other. Optimal Transport solutions are popular for colour transfer applications. Extending the Optimal Transport to be a geometry-aware solution by combining colour and spatial would significantly enhance the results. If there are subtle motion changes between the source and target images, such as the case of photo retouch styles, then global image plane coordinates can be used (Chapter 4). If there are large motion changes between the images, then a motion estimation technique can be used to estimate motions between images, and the estimated spatial correspondence information can be incorporated with the colour information (Chapter 5). Optimal Transport, by definition, does not consider any available information about the correspondences; hence an alternative approach that directly incorporates correspondences in its formulation is the use of Nadaraya-Watson estimator (Chapter 6). Optimal Transport and Nadaraya-Watson have no parametric formulation of the solution; hence if a parametric formulation is needed, then the $\mathcal{L}_2$ framework can be used (Chapter 7). With parallel computing, Optimal Transport (CDF-matching) and Nadaraya-Watson share the same time complexity of $\mathcal{O}(k)$ where k is the number of the algorithm's iterations, while the $\mathcal{L}_2$ framework has the time complexity of $\mathcal{O}(rk)$, which is slower due to the r number of iterations of the gradient ascent optimization step.

8.3 Limitations and Future Work

We have proposed global methods which calculate the colour statistics by taking into account all pixels in the images. However, some inherent limitations have been identified and possible solutions could be implemented as presented in the following:

- Our experiments show that enhancing the pixel descriptors by incorporating colour and spatial information improves the performance over using colour space only. However, when colour transfer performed between images that have small overlapped areas, one issue may appear such as shadow artifact. To tackle this situation, one could extract overlapped area and estimate the colour transfer and extrapolate the solution to non-overlapped areas

- Our experiments show that using motion estimation to find correspondences can guide the transfer process when there are large motion changes (Chapter 5) and greatly improve the results. However, the existence of outlier correspondences would affect the estimated colour transfer when the position's weight w is set to a high value, and hence artifacts may appear. To tackle this situation, we experimentally found that setting the position's weight to not more than $w = 10$ produces the best visual results, but a possible alternative solution is to use the motion estimation technique to determine which pixels in the source image have moved in the target image. With pixels affected by large motion, the position's weight would be set to $w = 0$, making the algorithm account only for colour based estimation. When pixels move less, the position's weight w would be set to higher number, making the algorithm account for colour and position based estimation.

The results obtained from the work carried out in this thesis suggest to further explore the following lines of research:

Exploring new optimisation techniques: We proposed Sliced $\mathcal{L}_2$ Distance algorithm that performs the optimisation in 1D dimensional spaces which significantly reduce the computational cost when searching a high dimensional latent space using a gradient ascent technique. We can further reduce the computational cost by reducing the number of mixtures in the GMM, a smaller number of Gaussians with differing

weights and variable variances could instead be used to model the distributions. In addition, the advantage of the 1D projection approach is that the 1D transformation problems involved are independent of each other, and are thus readily adapted for parallel programming; this is therefore a path to be explored for future work.

Applying colour transfer to video sequence: In order to extend the non-parametric solutions (OT and NW) to unseen values such as recolouring multiple frames in a video sequence, a smooth Thin Plate Spline transformation could be used to fit the solution for colour transfer and estimate the colour transformation between target image and a key frame of a video sequence and encode it in a LUT (LookUp Table) to be applied to all frames in the video sequence.

Exploring different applications: Point set registration using Gaussian Mixture Models is a different application than colour transfer but with a similar problem description whereby a parametric transformation function is estimated by minimizing the distance between two GMMs which capture the source shape and the target shape. We have started investigating the registration of point clouds in $\mathbb{R}^2$ using the 1D iterative projection approach with Optimal Transport, where at each step the problem solved is shape registration of 1D datasets, and we compared the results with standard $\mathcal{L}_2$ solution for point set registration (Jian and Vemuri, 2011). Our preliminary results has been published in (Alghamdi et al., 2017). Overall our IDT based algorithms have a good performance while $\mathcal{L}_2$ remains the best. Note that IDTs solve iteratively the problem in 1D projective spaces with an unconstrained non-parametric transformation while $\mathcal{L}_2$ solves it directly in 2D considering a smooth parametric transformation (TPS). TPS does not scale well in high dimensional spaces, but IDT approach that considered 1D projective space has the potential to adapt well in higher dimensions, and it is also suitable for parallel optimization. Moreover, we extended the 2D problem space to higher dimensional spaces by encoding neighborhoods of points (called here super-points) and investigated whether this formulation preserve some topology information of the original dataset (preliminary results can be found in Appendix B). Continuing the exploration of the application of the 1D projection approach to shape registration is an interesting direction for future work to assess if a projection approach to registration can ease efficiently the computational load for registration of datasets in high dimensional spaces.

Future work with Deep learning: Our approach is similar to neural network architecture. Exploring the iterative 1D projection approach with neural network framework is an interesting direction for future work. For example, as we can see in the Figure 3.3, the 1D projection approach involves a set of operations repeated through iterations where each iteration's output is an input to the next iteration, we can re-design and describe the algorithm with a single layer of Recurrent Neural Networks (RNNs) where the output from previous step are fed as input to the current step.

Many researchers have been looking at going beyond rectified linear unit (ReLU) function and having more complex activation functions (Lauriola et al., 2020; Scardapane et al., 2019; Ramachandran et al., 2017). Our work could be seen as non-parametric activation functions modeled as the monotonically non-decreasing Optimal Transport or as the non-monotonic Nadaraya-Watson kernel regression. Future work will investigate how our Optimal Transport and Nadaraya-Watson activation functions could contribute to that space of research in deep learning that can be applied to many application domains.

Appendices

COLOUR TRANSFER

In this Appendix we provide the evaluation dataset (Section A.1) and additional colour transfer results generated using our proposed colour transfer methods (Section A.2), and results of evaluation metrics with and without the extreme outliers for comparison purposes (Section A.3).

A.1 Dataset

Figure A.1 and Figure A.2 show image pairs with a similar content from a popular dataset provided by Hwang et al. (Hwang et al., 2014). The dataset includes geometrically registered pairs of images (source and target) that exhibit different colour change sources that come from different cameras, different in-camera settings, different illuminations, and different photo retouch styles.

	Source	Target Registered	Flow Field
building			
flower1			
flower2			
gangnam1			
gangnam2			
gangnam3			
illum			
mart			
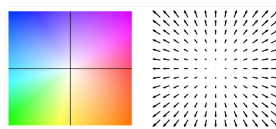
The visualization of flow fields

Figure A.1: This figure shows part 1 of the evaluation dataset that contains geometrically registered pairs of images (source and target). The flow fields show the motion changes between the target and source image. The last row shows flow field visualization based on the code in Baker et al. (2007): each pixel denotes a flow vector where the orientation and magnitude are represented by the hue and saturation of the pixel, respectively.

Figure A.2: This figure shows part 2 of the evaluation dataset that contains geometrically registered pairs of images (source and target). The flow fields show the motion changes between the target and each source image. The last row shows flow field visualization based on the code in Baker et al. (2007): each pixel denotes a flow vector where the orientation and magnitude are represented by the hue and saturation of the pixel, respectively.

A.2 Additional Colour Transfer Results

This section provides additional results generated using our proposed colour transfer methods using colour information only and combined colour and spatial information.
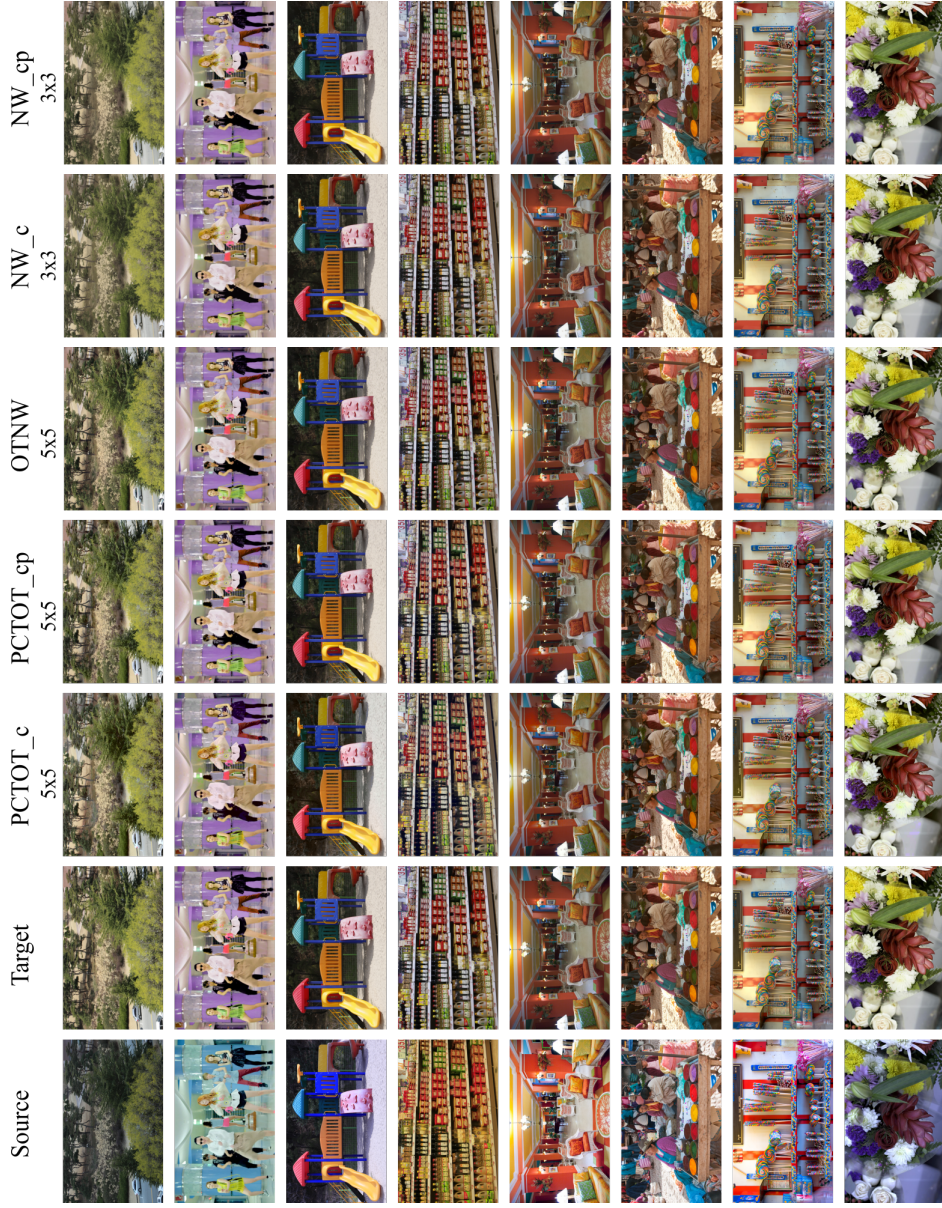
Figure A.3: Part 1: results generated using our algorithms: PCTOT (Chapter 4) with colour only (PCTOT_c) and with colour and spatial information (PCTOT_cp), OTNW (Chapter 5) , INWDT (Cpater 6) with colour only (NW_c) and with colour and spatial information (NW_cp) - best viewed in colour and zoomed in.
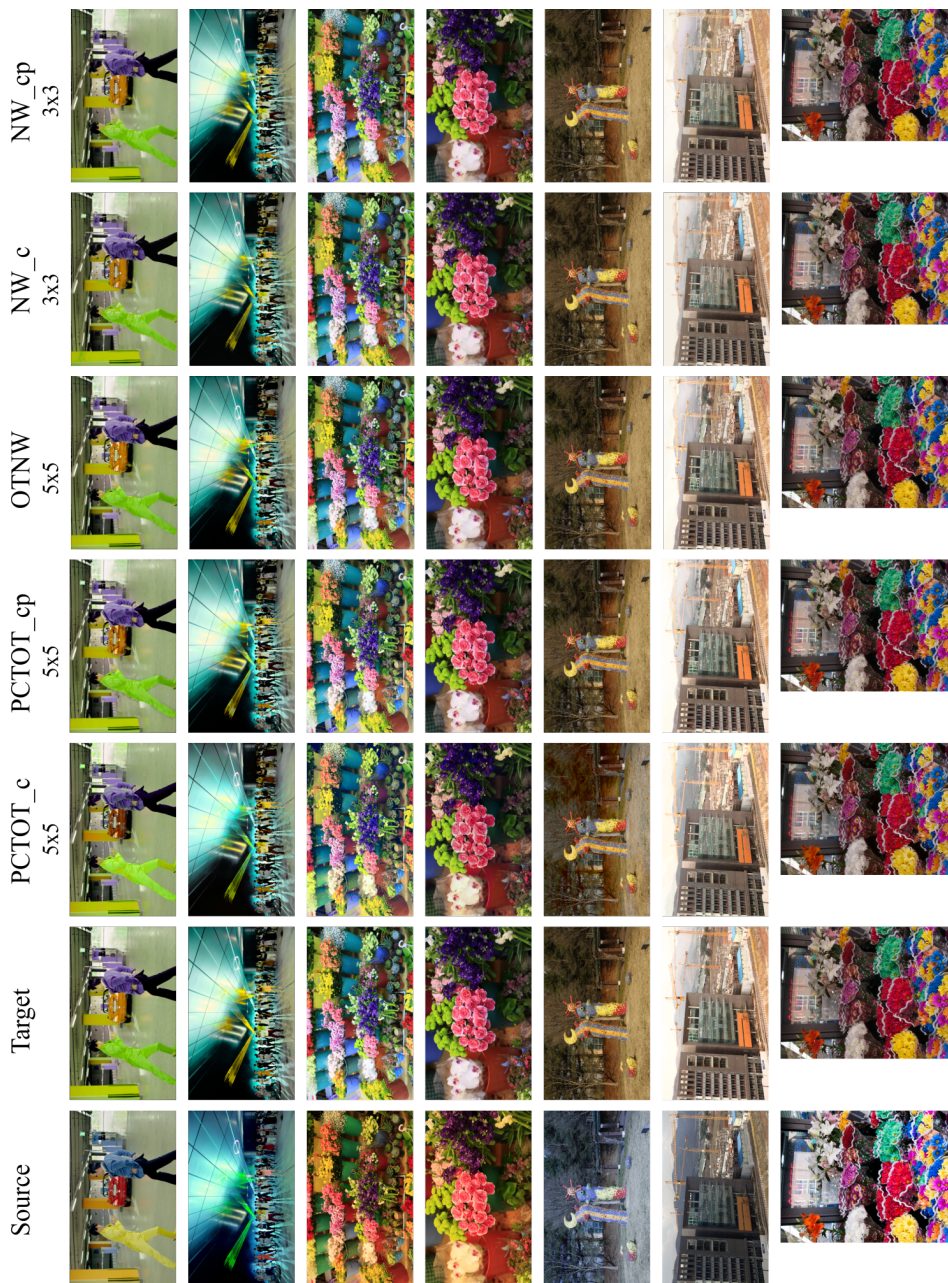
Figure A.4: Part 2: results generated using our algorithms: PCTOT (Chapter 4) with colour only (PCTOT_c) and with colour and spatial information (PCTOT_cp), OTNW (Chapter 5), INWDT (Cpater 6) with colour only (NW_c) and with colour and spatial information (NW_cp) - best viewed in colour and zoomed in.
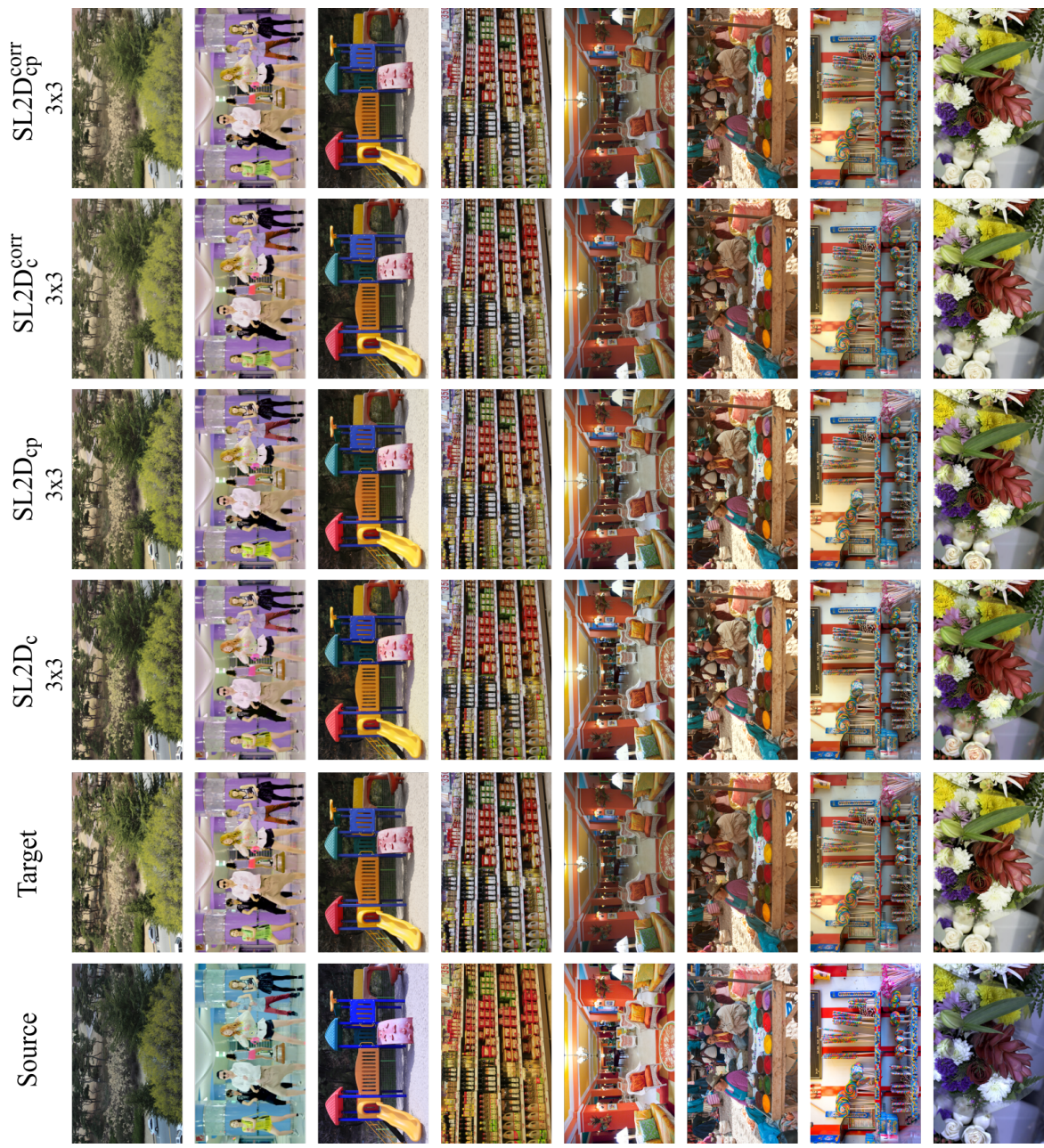
Figure A.5: Part 3: results generated using our algorithms: SL2D (Chapter 7) using correspondences ($\texttt{SL2D}_\texttt{c}^\texttt{corr}$ and $\texttt{SL2D}_\texttt{cp}^\texttt{corr}$) and without using correspondences ($\texttt{SL2D}_\texttt{c}$ and $\texttt{SL2D}_\texttt{cp}$) - best viewed in colour and zoomed in.
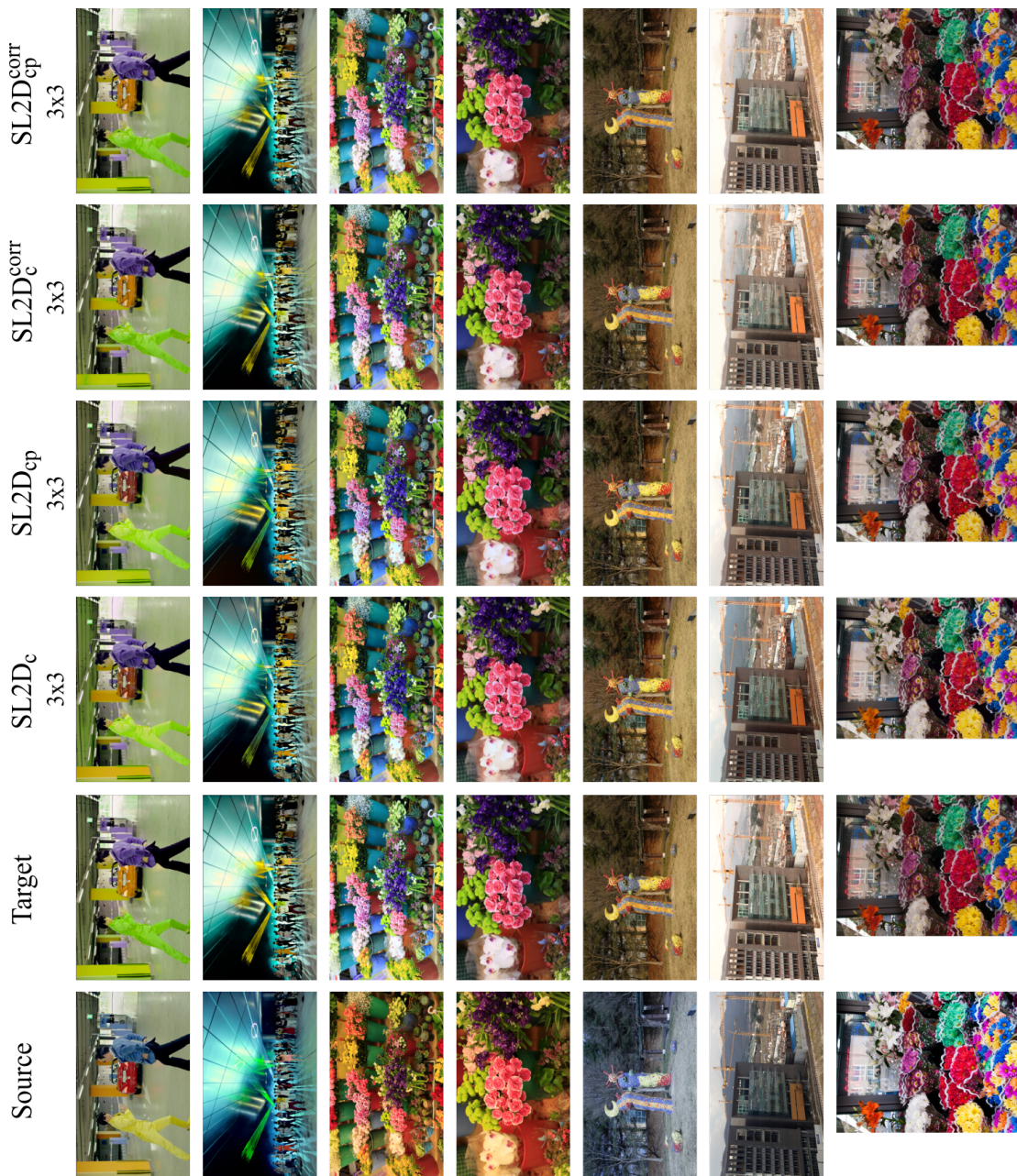
Figure A.6: Part 4: results generated using our algorithms: SL2D (Chapter 7) using correspondences ($SL2D_c^{corr}$ and $SL2D_{cp}^{corr}$) and without using correspondences ($SL2D_c$ and $SL2D_{cp}$) - best viewed in colour and zoomed in.
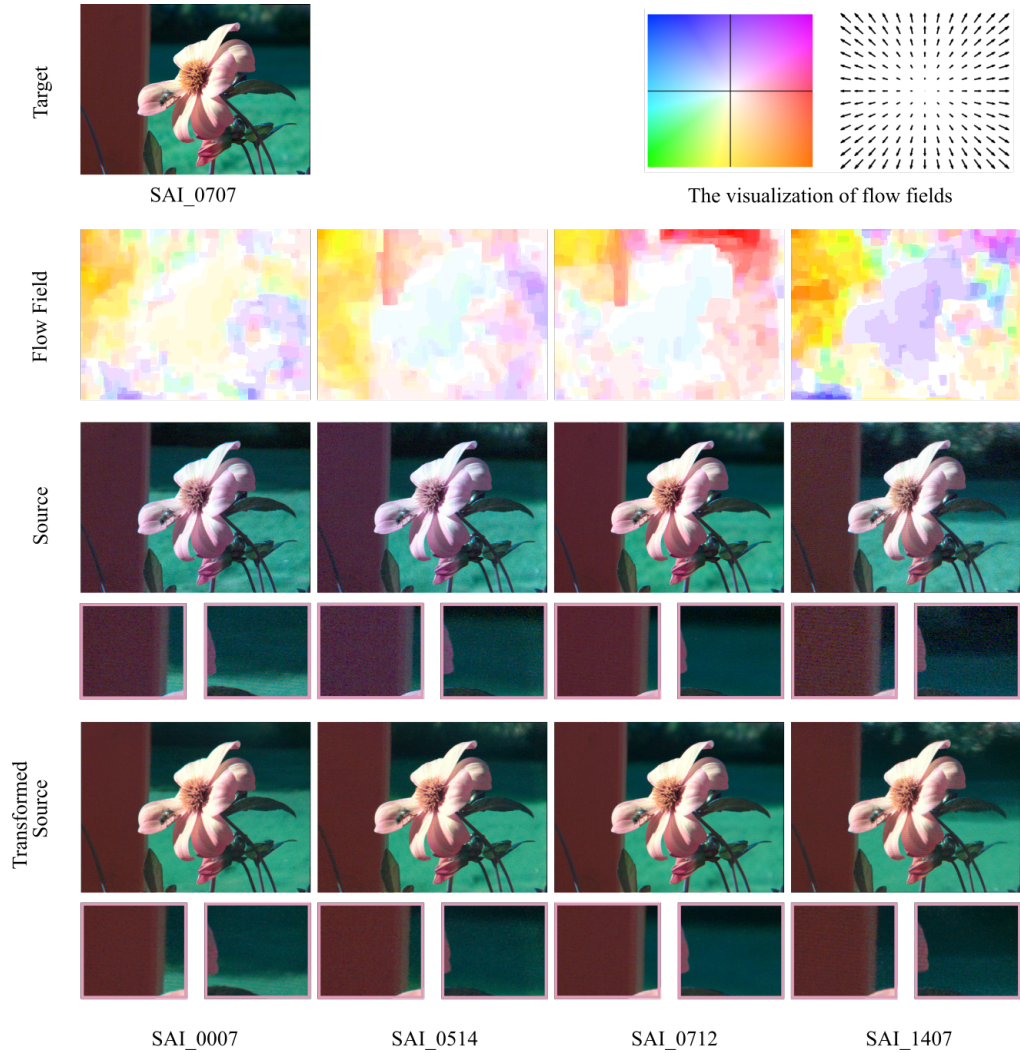
Figure A.7: Results on SAI light fields images. Our method $\mathtt{SL2D_{cp}^{corr}}$ (Chapter 7) successfully corrects artifacts such as colour inconsistencies and noise in the selected outer views images using the centre view image (SAI_0707) as the target image. Our method also produces smooth results as can be seen in the close-up patches. The flow fields show the motion changes between the target and each source image. The top row on the right shows flow field visualization based on the code in Baker et al. (2007): each pixel denotes a flow vector where the orientation and magnitude are represented by the hue and saturation of the pixel, respectively.

A.3 Metrics Comparisons with/without Extreme Outliers

With the metrics SSIM (cf. Figure A.8 and Figure A.9) and FSIMc (cf. Figure A.10 and Figure A.11), having mean values that are located at the edge of the box plots of some methods is due to having outliers in the results (shown as black dots in the plots). However, if we removed those outliers, this will not affect our conclusions regarding the performances.
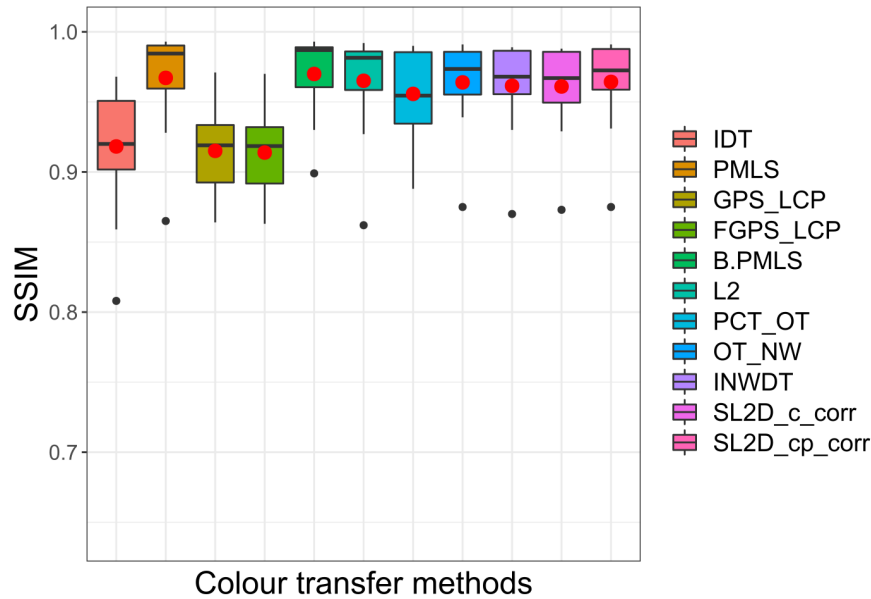
Figure A.8: Comparing our proposed methods PCT_OT (colours and positions Ours_cp in Chapter 4), OT_NW (Chapter 5), INWDT (Chapter 6), SL2D$_C^{corr}$ and SL2D$_{cp}^{corr}$ (Chapter 7 with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004). In this plot we removed the extreme outliers (higher values are better, best viewed in colour and zoomed in).
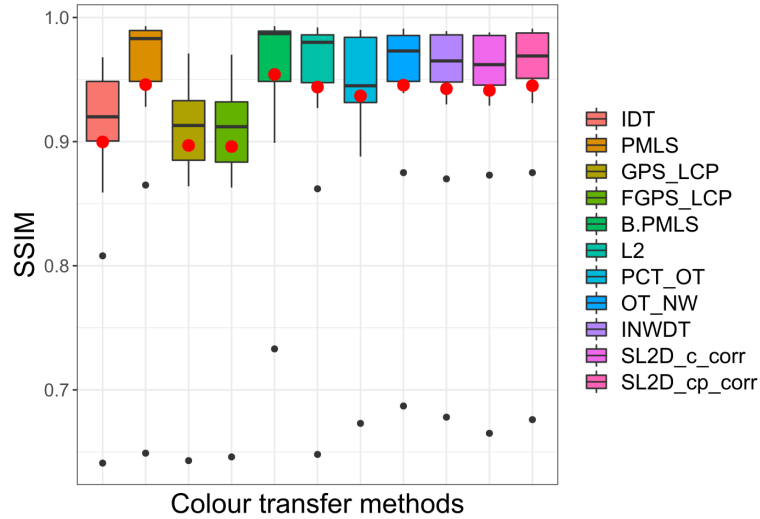
Figure A.9: Comparing our proposed methods `PCT_OT` (colours and positions `Ours_cp` in Chapter 4), `OT_NW` (Chapter 5), `INWDT` (Chapter 6), $\text{SL2D}_\text{c}^\text{corr}$ and $\text{SL2D}_\text{cp}^\text{corr}$ (Chapter 7 with the state of the art colour transfer methods using SSIM metric (Zhou Wang et al., 2004) with the present of outliers (higher values are better, best viewed in colour and zoomed in).
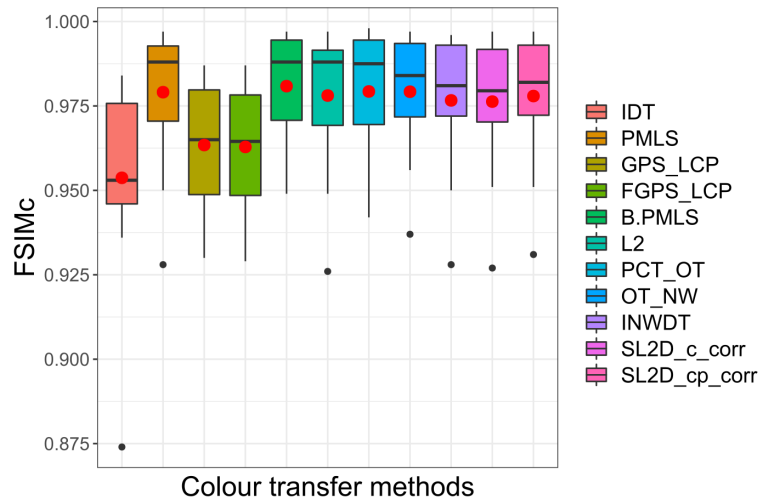
Figure A.10: Comparing our proposed methods `PCT_OT` (colours and positions `Ours_cp` in Chapter 4), `OT_NW` (Chapter 5), `INWDT` (Chapter 6), $\text{SL2D}_\text{c}^\text{corr}$ and $\text{SL2D}_\text{cp}^\text{corr}$ (Chapter 7 with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011). In this plot we removed the extreme outliers (higher values are better, best viewed in colour and zoomed in).
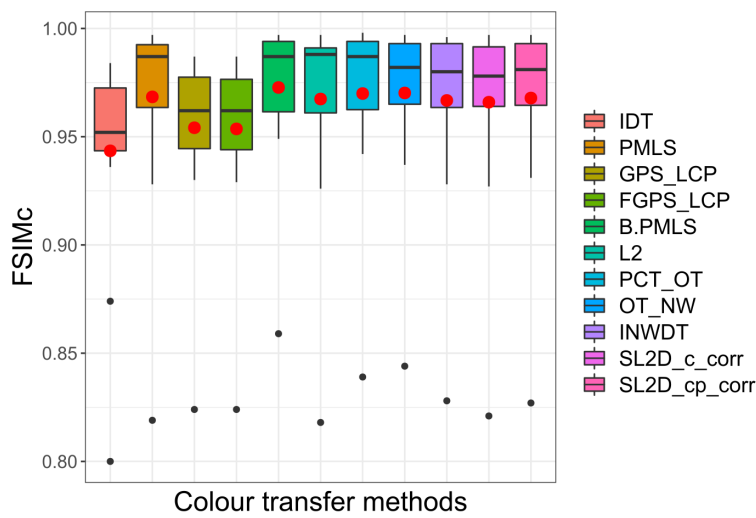
Figure A.11: Comparing our proposed methods PCT_OT (colours and positions Ours_cp in Chapter 4), OT_NW (Chapter 5), INWDT (Chapter 6), $\text{SL2D}_\text{c}^\text{corr}$ and $\text{SL2D}_\text{cp}^\text{corr}$ (Chapter 7 with the state of the art colour transfer methods using FSIMc metric (Zhang et al., 2011) with the present of outliers (higher values are better, best viewed in colour and zoomed in).

SHAPE REGISTRATION:

PRELIMINARY RESULTS

In this appendix we present preliminary results on the investigation of the application of the iterative 1D projection approach with optimal transport to shape registration. Section B.1 presents a brief introduction for our published paper "IDT vs L2 Distance for Point Set Registration" (Alghamdi et al., 2017), and Section B.2 presents results on encoding neighborhoods of points in high-dimensional spaces to preserve local structures.

B.1   Part 1: IDT vs L2 Distance for Point Set Registration

Summary: Registration techniques have many applications such as 3D scans alignment, panoramic image mosaic creation or shape matching. We focused on 2D point cloud registration using novel iterative algorithms that are inspired by the 1D Iterative projection approach implemented in the IDT algorithm, originally proposed to solve colour transfer (Pitie et al., 2007). We propose three variants to IDT algorithm that we compare with the standard $\mathcal{L}_2$ shape registration technique (Jian and Vemuri, 2011). Overall our IDT based algorithms have a good performance while standard $\mathcal{L}_2$ remains the best. Note that IDTs solve iteratively the problem in 1D projective spaces with an unconstrained non parametric transformation while $\mathcal{L}_2$ solves it directly in 2D considering a smooth parametric transformation (TPS). TPS does not scale well in high dimensional spaces, but IDT approach that considered 1D projective space has the potential to adapt well in higher dimensions, and it is also suitable for parallel

optimization. For more information, the interested reader is referred to our published work in the paper (Alghamdi et al., 2017).

B.2   Part 2: SWD with Superpoints for Point Set Registration

In this section we present preliminary results on encoding neighborhoods of points in high-dimensional spaces to preserve local structures. In the following, Section B.2.1 presents a summary and conclusion of the work, and Section B.2.2 presents the work details.

B.2.1   Summary

Context and motivation: In this work we investigate 2D point set registration. We propose to extend the Sliced Wasserstein Distance algorithm to higher dimensional spaces by encoding **non-overlapping** neighborhoods of points (called here super-points). We investigate whether superpoints reserve some structure information of the original 2D point set. One motivation to choose Sliced Wasserstein Distance to extend, is that coding superpoints will decrease the size of the number of data points dramatically and hence Sliced Wasserstein Distance approach is a proper solution for this situation as it uses Quantile-matching approach (see Section 2.2.1.3) that does not employ either histogram or cumulative distribution rather is based on the matching of two sorted arrays.

Method proposed: Extending Sliced Wasserstein Distance to higher dimensional spaces by encoding neighborhoods of points for shape registration.

Purpose: To register 2D point clouds of curves in iterative 1D projective spaces.

Dataset: 2D curves extracted from images taken from a dataset[1] provided by Lu et al. Lu et al. (2014).

Results: The proposed method is able to produce perfect alignments for all different patterns used. The superpoints technique is able to reserve the topology information

---

[1]obtained at `http://gfx.cs.princeton.edu/pubs/Lu2014DDS`

of the original 2D points. The interpolation technique used to estimate the correspondences for new points in the source points cloud fails with some results by producing noisy points. The reason for the noisy points is that while the superpoints maintain neighborhood structure within the points that constitute the superpoints but does not maintain continuity between superpoints, i.e, the new points in the source dense contour might lie in between old points used in the registration step but these old points are not adjacent and belong to separate parts of the shape.

B.2.2 Work Details

B.2.2.1 Method proposed

Creating superpoints: Given a 2D point clouds with even size (n), Assuming random starting point, we take successive non-overlapping ($\frac{n}{2}$) number of points in $\mathbb{R}^2$ space in one direction to form a superpoint in $\mathbb{R}^N$, such that $N = (\frac{n}{2}) \times 2$. The result is a data set contains two superpoints, Figure B.1. In case of odd sizes point clouds, the point cloud could be interpolated to create an even data set size.
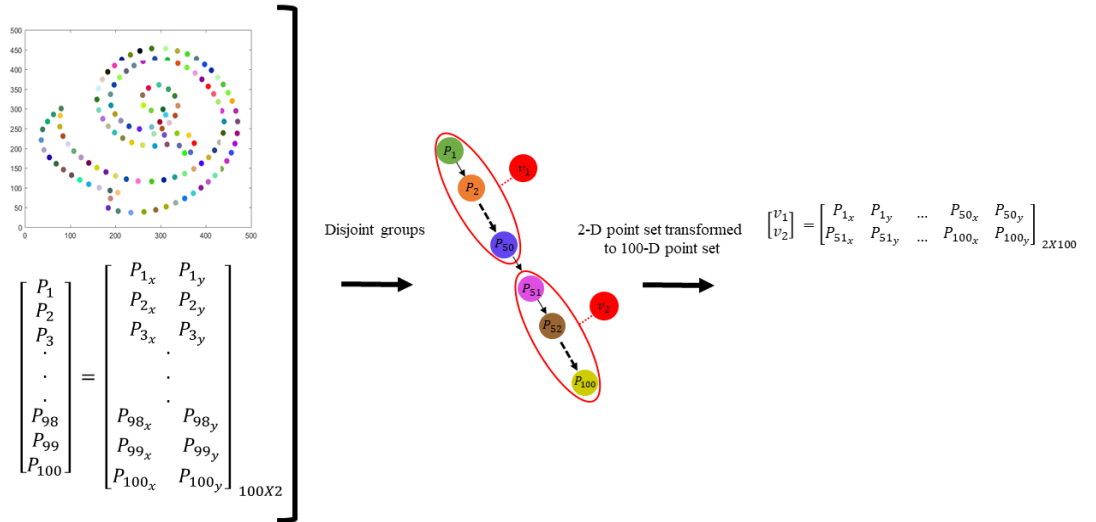
Figure B.1: Example of creating super-point set. $\{\mathbf{P}_i\}_{i=1,\cdots,100}$ are point set in $\mathbb{R}^2$ transformed to $\{\mathbf{v}_i\}_{i=1,2}$ super-point set in $\mathbb{R}^{100}$, by taking non-overlapping neighborhoods of points (best viewed in colour and zoomed in).

Generate orthogonal basis for $\mathbb{R}^N$: We generate a random subset $S = \{\mathbf{a}_r\}_{r=1,\cdots,N}$ of the vector space $\mathbb{R}^N$ as follows Muller (1959):

$$\mathbf{a}_i \sim \mathcal{N}(\mathbf{0}, \mathrm{I}_N) \tag{B.1}$$

Where $\mathcal{N}(0, \mathrm{I}_N)$ is a normal distribution in $\mathbb{R}^N$ centered on the origin in $\mathbb{R}^N$ with covariance matrix equal to the identity matrix in $\mathbb{R}^N$. The number of vectors in S equals the dimensions of the vector space $\mathbb{R}^N$. Then we generate the orthonormal basis $R = \{\mathbf{e}_r\}_{r=1,\cdots,N}$ for the subset S (using matlab function "$orth$"), that is, the basis vectors are all unit vectors and orthogonal to each other, i.e linearly independent vectors and hence these basis can be used as orthogonal basis for the vector space $\mathbb{R}^N$.

1D projections and accumulated shifts: $M = \{\mathbf{u}_i\}_{i=1,\cdots,n}$ (source) and $Q = \{\mathbf{v}_j\}_{j=1,\cdots,m}$ (target) are point clouds in $\mathbb{R}^N$ to be registered, we generate random set of orthogonal basis R_k of $\mathbb{R}^N$ for each iteration $k = 0, 1, ..., \infty$. The gradient descent step start from the source as initial state $M^{(0)}$. For some step size $\tau_\ell > 0$ each point $\mathbf{u}$ in the source is updated in each iteration as follows:

$$\mathbf{u}^{k+1} = (1 - \tau_k)\mathbf{u}^k + \tau_k \, \phi(\mathbf{u}^k) \tag{B.2}$$

Such that $\phi(\mathbf{u})$ transforms data point $\mathbf{u}$ using accumulated 1D transformations as follows:

$$\phi(\mathbf{u}) = R_k \begin{pmatrix} \phi_1(\mathbf{e_1}^T\mathbf{u}) \\ \phi_2(\mathbf{e_2}^T\mathbf{u}) \\ \vdots \\ \phi_N(\mathbf{e_N}^T\mathbf{u}) \end{pmatrix} = \sum_{r=1}^{N} \phi_r(\mathbf{e_r}^T\mathbf{u}) \, \mathbf{e_r} \tag{B.3}$$

The assumption is that the projections along the axes are independent, and the transfer functions $\{\phi_r\}_{r=1,\cdots,N}$ model independent transfer functions for projected points computed in the directions of $\{\mathbf{e}_r\}_{r=1,\cdots,N}$.

1D optimal maps ϕ_r: The transfer functions $\{\phi_r\}_{r=1,\cdots,N}$ for the source and target projected points in the directions of $\{\mathbf{e}_r\}_{r=1,\cdots,N}$ are computed using a closed form for optimal transform in 1D, more specifically Quantile-matching, see Section 2.2.1.3.

B.2.2.2 Experimental results

We present experimental results on the application of the Sliced Wasserstein distance algorithm combined with superpoints to the registration of 2D curves extracted from images taken from a dataset provided by Lu et al. Lu et al. (2014). All experiments were performed using MATLAB R2017a on a PC with 16 GB of RAM and an Intel Core i7-6700HQ (2.60 GHz) CPU. The goal of the point set registration is to align the source shape onto the target shape.

Figure B.2 Shows results of the curves registration. Seven different images were selected that represent different patterns to test the registration performance. The first column is the source cloud points and the second column is the target cloud points and the third column is the results of the registration. For a given pair of source and target shape we extract dense boundary contours (cf. Figure B.2, the black shapes in row 1,3,5,7,9,11,13) and then we extract sample points at equal interval along the boundary contours (cf. Figure B.2, row 2,4,6,8,10,12,14).

For clarity we show points that constitute each superpoints with different colour (blue and red) in the plots, Figure B.2. Please note in the plots **we show the connection between the points within each superpoint by plotting a thin line from the start point to the end point within the superpoint** so we can track the connected points after the mapping. The goal of encoding superpoints is to investigate whether this formulation reserve some topology information of the original 2D points before projecting the samples onto 1D subspaces, i.e answering questions such as do the connected points stay adjacent after the mapping or are they mapped to different places? do they stay in the same order?. Table B.1 shows the size and dimensionality of the point clouds after grouping the points for all curves used for registration.

With the resulted correspondences from the registration step, we compute the correspondence for the rest of the points in the dense contours by interpolating the points using inverse distance weighted (IDW) average Shepard (1968) of the values of **the nearest two points along the contour**, to transform the model dense contour into the target pattern (cf. Figure B.2 column 3 row 1,3,5,7,9,11,13).

From the results shown in Figure B.2, we observe that the proposed method is able to produce perfect alignments for all different patterns used. We can observe the connected points in blue and red colour where they are mapped after the registration in

	Original point set		Transformed point set	
Curve #	Number of points	$\mathbb{R}^N$	Number of points	$\mathbb{R}^N$
Curve 1	200	$\mathbb{R}^2$	2	$\mathbb{R}^{200}$
Curve 2	200	$\mathbb{R}^2$	2	$\mathbb{R}^{200}$
Curve 3	800	$\mathbb{R}^2$	2	$\mathbb{R}^{800}$
Curve 4	200	$\mathbb{R}^2$	2	$\mathbb{R}^{200}$
Curve 5	200	$\mathbb{R}^2$	2	$\mathbb{R}^{200}$
Curve 6	800	$\mathbb{R}^2$	2	$\mathbb{R}^{800}$
Curve 7	800	$\mathbb{R}^2$	2	$\mathbb{R}^{800}$

Table B.1: Sizes of the points clouds and dimensionality for all the curves used for registration.

Figure B.2, column 3 row 2,4,6,8,10,12,14. We can see that the connected points stay adjacent and in the same order after the registration, hence we can conclude that the superpoints technique reserve the topology information of the original 2D points.

The interpolation technique used to estimate the correspondences for new points in the source points cloud produces correct estimation with some patterns like curve 1 and 2 in column 3 row 1 and 3, and degrades in performance by producing noisy points with other patterns like curve 4 and 5, column 3 row 7 and 9. The reason for the noisy points shown in the curves is that the superpoints maintain neighborhood structure within the points that constitute the superpoints but does not maintain continuity between superpoints, i.e, the new points in the source dense contour might lie in between old points used in the registration step but these old points are not adjacent and belong to separate parts of the shape.

The transformations being estimated are non-parametric optimal transport. We can create new transformations by interpolating between solutions. For example, given optimal transport solution ($\hat{Q}_1$) when registering a model pattern (M) to a target pattern (Q_1), we can create a linear interpolation between the two patterns as follows:

$$\forall\, t \in [0, 1], \;\; S = (1 - t)M + t\hat{Q}_1 \tag{B.4}$$

S is the interpolation between the two patterns M (obtained at t=0) and $\hat{Q}_1$ (obtained at t=1). Given a second optimal transport solution ($\hat{Q}_2$) when registering a model
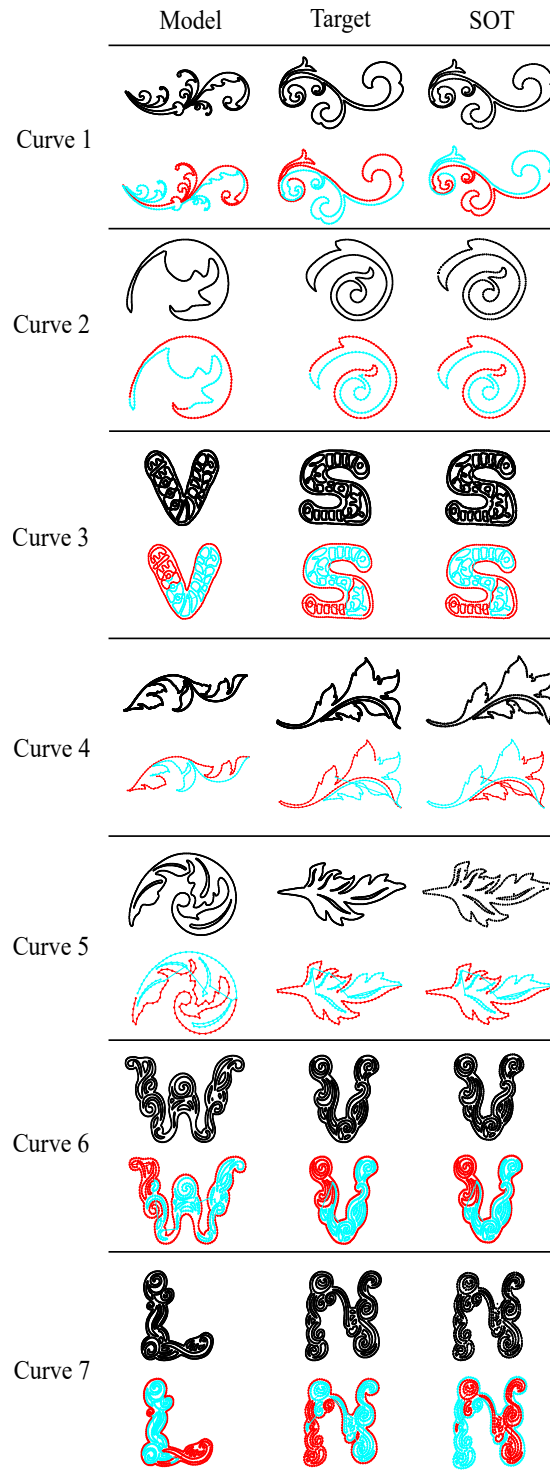
Figure B.2: Results of the curves registration using Sliced optimal transport with super-points. Row 1,3,5,7,9,11,13 represent dense boundary contours. Row 2,4,6,8,10,12,14 represent sampled point clouds. The sampled point clouds used to register the source to the target. Inverse distance weighting interpolation used to transfer the dense contour into the target using the correspondence resulted from registering step (best viewed in colour and zoomed in).
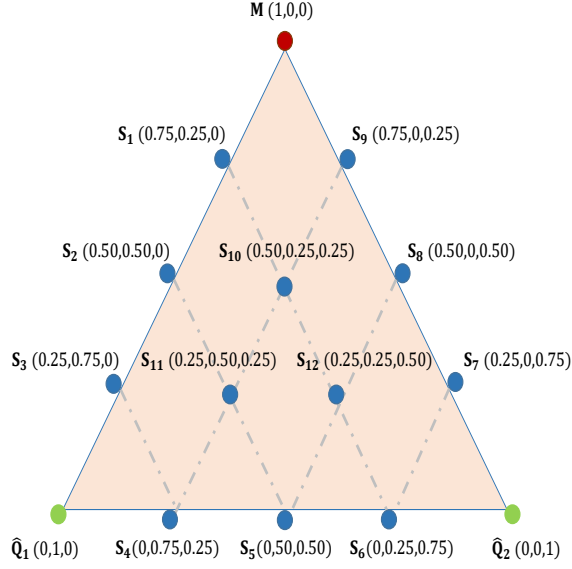
Figure B.3: Pyramid shows the weights used to interpolate curves registered using sliced optimal transport with super-points. Every circle in the figure represent the corresponding curve with the associated weights. Red circle represent the the source M and green circles represent the resulted registered curves $\hat{Q}_1$ and $\hat{Q}_2$ and blue circle S_i represent the interpolated curves between them. The associated weights (t_1, t_2, t_3) represent the weights of M, $\hat{Q}_1$ and $\hat{Q}_2$ respectively.

pattern (M) to a target pattern (Q_2), we can create a linear interpolation between the three patterns M, $\hat{Q}_1$ and $\hat{Q}_2$ as follows:

$$\forall\, t_1, t_2 \in [0,1],\; t_3 = 1 - (t_1 + t_2),\;\; S = t_1 M + t_2 \hat{Q}_1 + t_3 \hat{Q}_2 \tag{B.5}$$

Figure B.3 shows the weights used to interpolate each curve, such that (t_1, t_2, t_3) represent the weights of M, $\hat{Q}_1$ and $\hat{Q}_2$ respectively. The pyramids in Figure B.4 display samples of curves generated by interpolating between the estimated transformations.
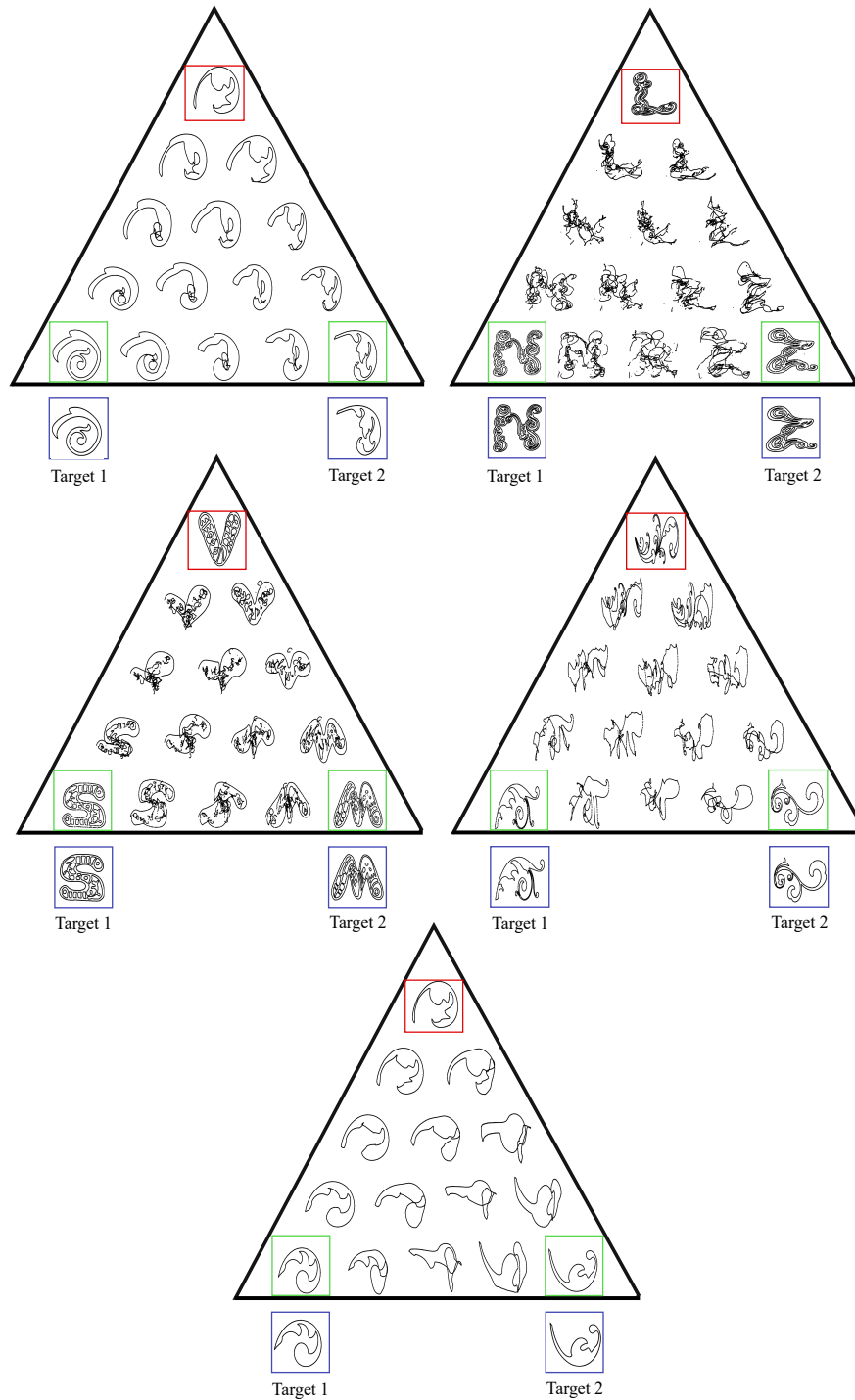
Figure B.4: Curve registration and interpolation results generated using Optimal transfer with super-points. The shape in red square is the source, the shapes in green squares are the resulted registered shapes and shapes in blue squares are the targets (best viewed in colour and zoomed in).

BIBLIOGRAPHY

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. URL: `https://www.tensorflow.org/`.

Abadpour, A., Kasaei, S., 2007. An efficient pca-based color transfer method. Journal of Visual Communication and Image Representation 18, 15 – 34.

Agarwal, S., Furukawa, Y., Snavely, N., Curless, B., Seitz, S.M., Szeliski, R., 2010. Reconstructing rome. Computer 43, 40–47.

Agustsson, E., Sage, A., Timofte, R., Gool, L.V., 2019. Optimal transport maps for distribution preserving operations on latent spaces of generative models, in: Proceedings of the 7th International Conference on Learning Representations (ICLR).

Alghamdi, H., Dahyot, R., 2020a. Iterative nadaraya-watson distribution transfer for colour grading, in: Proceedings of the IEEE 22th International Workshop on Multimedia Signal Processing (MMSP).

Alghamdi, H., Dahyot, R., 2020b. Patch based colour transfer using sift flow, in: Proceedings of the 22th Irish Machine Vision and Image Processing conference (IMVIP).

Alghamdi, H., Grogan, M., Dahyot, R., 2017. Idt vs l_2 distance for point set registration, in: Proceedings of the 19th Irish Machine Vision and Image Processing conference (IMVIP).

Alghamdi, H., Grogan, M., Dahyot, R., 2019. Patch-based colour transfer with optimal transport, in: Proceedings of the 27th European Signal Processing Conference (EUSIPCO), pp. 1–5.

Arellano, C., Dahyot, R., 2016. Robust ellipse detection with gaussian mixture models. Pattern Recognition 58, 12–26.

Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein generative adversarial networks, in: Proceedings of Machine Learning Research, pp. 214–223.

Baker, S., Roth, S., Scharstein, D., Black, M.J., Lewis, J.P., Szeliski, R., 2007. A database and evaluation methodology for optical flow, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1–8.

Basu, A., Harris, I.R., Hjort, N.L., Jones, M., 1998. Robust and efficient estimation by minimising a density power divergence. Biometrika 85, 549–559.

Bellavia, F., Colombo, C., 2018. Dissecting and reassembling color correction algorithms for image stitching. IEEE Transactions on Image Processing 27, 735–748.

Bertsekas, D.P., 1988. The auction algorithm: A distributed relaxation method for the assignment problem. Annals of operations research 14, 105–123.

Bonneel, N., Coeurjolly, D., 2019. Spot: Sliced partial optimal transport. ACM Transactions on Graphics (TOG) 38.

Bonneel, N., Peyré, G., Cuturi, M., 2016. Wasserstein barycentric coordinates: Histogram regression using optimal transport. ACM Transactions on Graphics (TOG) 35.

Bonneel, N., Rabin, J., Peyré, G., Pfister, H., 2015. Sliced and radon wasserstein barycenters of measures. Journal of Mathematical Imaging and Vision 51, 22–45.

Brown, M., Lowe, D.G., 2007. Automatic panoramic image stitching using invariant features. International Journal of Computer Vision 74, 59–73.

Brox, T., Bruhn, A., Papenberg, N., Weickert, J., 2004. High accuracy optical flow estimation based on a theory for warping, in: Proceedings of the European conference on computer vision (ECCV), pp. 25–36.

Cazelles, E., Seguy, V., Bigot, J., Cuturi, M., Papadakis, N., 2018. Geodesic pca versus log-pca of histograms in the wasserstein space. SIAM Journal on Scientific Computing 40, B429–B456.

Chen, T., Vemuri, B.C., Rangarajan, A., Eisenschenk, S.J., 2010. Group-wise point-set registration using a novel cdf-based havrda-charvát divergence. International journal of computer vision 86, 111.

Chen, Y., Georgiou, T.T., Tannenbaum, A., 2018. Optimal transport for gaussian mixture models. IEEE Access 7, 6269–6278.

Cheng, Z., Yang, Q., Sheng, B., 2015. Deep colorization, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q., 2006. Color harmonization. ACM Transactions on Graphics 25, 624–630.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2009. Introduction to algorithms. MIT press.

Courty, N., Flamary, R., Tuia, D., 2014. Domain adaptation with regularized optimal transport, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer. pp. 274–289.

Courty, N., Flamary, R., Tuia, D., Rakotomamonjy, A., 2016. Optimal transport for domain adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 1853–1865.

Dantzig, G.B., 1990. Origins of the Simplex Method. Association for Computing Machinery, New York, USA. p. 141–151.

Del Barrio, E., Gordaliza, P., Lescornel, H., Loubes, J.M., 2019. Central limit theorem and bootstrap procedure for wasserstein's variations with an application to structural relationships between distributions. Journal of Multivariate Analysis 169, 341 – 362.

Dong, W., Bao, G., Zhang, X., Paul, J.C., 2010. Fast local color transfer via dominant colors mapping, in: ACM SIGGRAPH ASIA 2010 Sketches, Association for Computing Machinery, New York, USA.

Fan, J., Gijbels, I., 1992. Variable bandwidth and local linear regression smoothers. Annals of Statistics 20, 2008–2036.

Fan, J., Marron, J.S., 1994. Fast implementations of nonparametric curve estimators. Journal of Computational and Graphical Statistics 3, 35–56.

Faridul, H.S., Pouli, T., Chamaret, C., Stauder, J., Tremeau, A., Reinhard, E., 2014. A survey of color mapping and its applications, in: Eurographics (State of the Art Reports), The Eurographics Association.

Ferradans, S., Papadakis, N., Rabin, J., Peyré, G., Aujol, J.F., 2013. Regularized discrete optimal transport, in: International Conference on Scale Space and Variational Methods in Computer Vision, Springer Berlin Heidelberg. pp. 428–439.

Freedman, D., Kisilev, P., 2010. Object-to-object color transfer: Optimal flows and smsp transformations, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 287–294.

Frigo, O., Sabater, N., Delon, J., Hellier, P., 2016. Motion driven tonal stabilization. IEEE Transactions on Image Processing 25, 5455–5468.

Frigo, O., Sabater, N., Demoulin, V., Hellier, P., 2014. Optimal transportation for example-guided color transfer, in: Proceedings of the Asian Conference on Computer Vision (ACCV), pp. 655–670.

Gasser, T., Kneip, A., Köhler, W., 1991. A flexible and fast method for automatic smoothing. Journal of the American Statistical Association 86, 643–652.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets, in: Advances in Neural Information Processing Systems, pp. 2672–2680.

Grogan, M., 2017. Colour Transfer and Shape Registration using Functional Data Representations. Ph.D. thesis. School of Computer Science and Statistics, Trinity College Dublin.

Grogan, M., Dahyot, R., 2015. l_2 registration for colour transfer in videos, in: Proceedings of the 12th European Conference on Visual Media Production (CVMP), Association for Computing Machinery, New York, USA.

Grogan, M., Dahyot, R., 2018. Shape registration with directional data. Pattern Recognition 79, 452 – 466.

Grogan, M., Dahyot, R., 2019. l_2 divergence for robust colour transfer. Computer Vision and Image Understanding 181, 39 – 49.

Grogan, M., Dahyot, R., Smolic, A., 2017. User interaction for image recolouring using l_2, in: Proceedings of the 14th European Conference on Visual Media Production (CVMP), Association for Computing Machinery, New York, USA.

Grogan, M., Prasad, M., Dahyot, R., 2015. l_2 registration for colour transfer, in: Proceedings of the 23rd European Signal Processing Conference (EUSIPCO).

Grogan, M., Smolic, A., 2019. l_2 based colour correction for light field arrays, in: In Proceedings of the 16th ACM SIGGRAPH European Conference on Visual Media Production (CVMP).

Grundland, M., Dodgson, N.A., 2005. Color histogram specification by histogram warping, in: Color Imaging X: Processing, Hardcopy, and Applications, International Society for Optics and Photonics. SPIE. pp. 610 – 621.

HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2011. Non-rigid dense correspondence with applications for image enhancement. ACM Transactions on Graphics (TOG) 30.

HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2013. Optimizing color consistency in photo collections. ACM Transactions on Graphics (TOG) 32.

Hasanbelliu, E., Giraldo, L.S., Príncipe, J.C., 2011. A robust point matching algorithm for non-rigid registration using the cauchy-schwarz divergence, in: IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6.

He, M., Chen, D., Liao, J., Sander, P.V., Yuan, L., 2018. Deep exemplar-based colorization. ACM Transactions on Graphics (TOG) 37.

He, M., Liao, J., Chen, D., Yuan, L., Sander, P.V., 2019. Progressive color transfer with dense semantic correspondences. ACM Transactions on Graphics (TOG) 38.

Hristova, H., Le Meur, O., Cozot, R., Bouatouch, K., 2015. Style-aware robust color transfer, in: Proceedings of the Workshop on Computational Aesthetics, Eurographics Association. p. 67–77.

Hwang, Y., Lee, J., Kweon, I.S., Kim, S.J., 2014. Color transfer using probabilistic moving least squares, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3342–3349.

Hwang, Y., Lee, J.Y., Kweon, I.S., Kim, S.J., 2019. Probabilistic moving least squares with spatial constraints for nonlinear color transfer between images. Computer Vision and Image Understanding 180, 1 – 12.

Jeong, K., Jaynes, C., 2008. Object matching in disjoint cameras using a color transfer approach. Machine Vision and Applications 19, 443–455.

Jian, B., Vemuri, B.C., 2011. Robust point set registration using gaussian mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence 33, 1633–1645.

Jones, M., Hjort, N.L., Harris, I.R., Basu, A., 2001. A comparison of related density-based minimum divergence estimators. Biometrika 88, 865–873.

Kantorovich, L.V., 1942. On translation of mass (in russian), in: Dokl. Akad. Nauk. USSR, pp. 199—201. English translation in J. Math. Sci. 133, 4 (2006), 1381–1382.

Kingma, D.P., Welling, M., 2014. Auto-encoding variational bayes, in: Proceedings of the 2nd International Conference on Learning Representations (ICLR).

Knothe, H., 1957. Contributions to the theory of convex bodies. Michigan Mathematical Journal 4, 39–52.

Kolouri, S., Park, S.R., Thorpe, M., Slepcev, D., Rohde, G.K., 2017. Optimal mass transport: Signal processing and machine-learning applications. IEEE Signal Processing Magazine 34, 43–59.

Kolouri, S., Rohde, G.K., Hoffmann, H., 2018. Sliced wasserstein distance for learning gaussian mixture models, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Kuhn, H.W., 1955. The hungarian method for the assignment problem. Naval Research Logistics Quarterly 2, 83–97.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Annals of Mathematical Statistics 22, 79–86.

L. Arellano Vidal, C., 2014. Inference for Shape Parameter Estimation. Ph.D. thesis. School of Computer Science and Statistics, Trinity College Dublin.

Lauriola, I., Gallicchio, C., Aiolli, F., 2020. Enhancing deep neural networks via multiple kernel learning. Pattern Recognition 101, 107194.

Li, L., Xia, M., Liu, C., Li, L., Wang, H., Yao, J., 2020. Jointly optimizing global and local color consistency for multiple image mosaicking. ISPRS Journal of Photogrammetry and Remote Sensing 170, 45 – 56.

Liao, J., Yao, Y., Yuan, L., Hua, G., Kang, S.B., 2017. Visual attribute transfer through deep image analogy. CoRR abs/1705.01088.

Lischinski, D., Farbman, Z., Uyttendaele, M., Szeliski, R., 2006. Interactive local adjustment of tonal values. ACM Transactions on Graphics (TOG) 25, 646–653.

Lissner, I., Preiss, J., Urban, P., Lichtenauer, M.S., Zolliker, P., 2013. Image-difference prediction: From grayscale to color. IEEE Transactions on Image Processing 22, 435–446.

Liu, C., Yuen, J., Torralba, A., 2011. Sift flow: Dense correspondence across scenes and its applications. IEEE Transactions on Pattern Analysis and Machine Intelligence 33, 978–994.

Liu, M., Vemuri, B.C., Amari, S.I., Nielsen, F., 2010. Total bregman divergence and its applications to shape retrieval, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3463–3468.

Lu, J., Barnes, C., Wan, C., Asente, P., Mech, R., Finkelstein, A., 2014. Decobrush: drawing structured decorative patterns by example. ACM Transactions on Graphics (TOG) 33.

Luan, F., Paris, S., Shechtman, E., Bala, K., 2017. Deep photo style transfer, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Ma, J., Qiu, W., Zhao, J., Ma, Y., Yuille, A.L., Tu, Z., 2015. Robust l_2e estimation of transformation for non-rigid registration. IEEE Transactions on Signal Processing 63, 1115–1129.

Matysiak, P., Grogan, M., Le Pendu, M., Alain, M., Zerman, E., Smolic, A., 2020. High quality light field extraction and post-processing for raw plenoptic data. IEEE Transactions on Image Processing 29, 4188–4203.

Meng, C., Ke, Y., Zhang, J., Zhang, M., Zhong, W., Ma, P., 2019. Large-scale optimal transport map estimation using projection pursuit, in: Advances in Neural Information Processing Systems, pp. 8118–8129.

Micusik, B., Kosecka, J., 2009. Piecewise planar city 3d modeling from street view panoramic sequences, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2906–2912.

Monge, G., 1781. Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris , 666–704.

Morovic, J., Shaw, J., Sun, P.L., 2002. A fast, non-iterative and exact histogram matching algorithm. Pattern Recognition Letters 23, 127–135.

Muller, M.E., 1959. A note on a method for generating points uniformly on n-dimensional spheres. Communications of the ACM 2, 19–20.

Murphy, J.K., 2014. Hidden states, hidden structures: Bayesian learning in time series models. Ph.D. thesis. University of Cambridge.

Muzellec, B., Cuturi, M., 2019. Subspace detours: Building transport plans that are optimal on subspace projections, in: Advances in Neural Information Processing Systems, pp. 6917–6928.

Nadaraya, E.A., 1965. On non-parametric estimates of density functions and regression curves. Theory of Probability & Its Applications 10, 186–190.

Neumann, L., Neumann, A., 2005. Color Style Transfer Techniques using Hue, Lightness and Saturation Histogram Matching, in: Computational Aesthetics in Graphics, Visualization and Imaging, pp. 111–122.

Oliveira, M., Sappa, A.D., Santos, V., 2015. A probabilistic approach for color correction in image mosaicking applications. IEEE Transactions on Image Processing 24, 508–523.

Panaretos, V.M., Zemel, Y., 2019. Statistical aspects of wasserstein distances. Annual Review of Statistics and Its Application 6, 405–431.

Park, J., Tai, Y., Sinha, S.N., Kweon, I.S., 2016. Efficient and robust color consistency for community photo collections, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 430–438.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison,

M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. Pytorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems, pp. 8024–8035.

Peyré, G., Cuturi, M., 2019. Computational Optimal Transport: With Applications to Data Science. Foundations and Trends in Machine Learning Series, Now Publishers.

Pitie, Kokaram, Dahyot, 2005a. Towards automated colour grading, in: Proceedings of the 2nd European Conference on Visual Media Production (CVMP), pp. 9–16.

Pitie, F., Kokaram, A., 2007. The linear monge-kantorovitch linear colour mapping for example-based colour transfer. IET Conference Proceedings , 23–23(1).

Pitie, F., Kokaram, A., Dahyot, R., 2008. Enhancement of digital photographs using color transfer techniques. Single-Sensor Imaging: Methods and Applications for Digital Cameras , 295–299.

Pitie, F., Kokaram, A.C., Dahyot, R., 2005b. N-dimensional probability density function transfer and its application to color transfer, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1434–1439.

Pitie, F., Kokaram, A.C., Dahyot, R., 2007. Automated colour grading using colour distribution transfer. Computer Vision and Image Understanding 107, 123 – 137.

Pouli, T., Reinhard, E., 2011. Progressive color transfer for images of arbitrary dynamic range. Computers and Graphics 35, 67 – 80.

Preiss, J., Fernandes, F., Urban, P., 2014. Color-image quality assessment: From prediction to optimization. IEEE Transactions on Image Processing 23, 1366–1378.

Rabin, J., Delon, J., Gousseau, Y., 2011. Removing artefacts from color and contrast modifications. IEEE Transactions on Image Processing 20, 3073–3085.

Rabin, J., Ferradans, S., Papadakis, N., 2014. Adaptive color transfer with relaxed optimal transport, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 4852–4856.

Rabin, J., Peyré, G., 2011. Wasserstein regularization of imaging problem, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 1541–1544.

Rabin, J., Peyré, G., Delon, J., Bernot, M., 2012. Wasserstein barycenter and its application to texture mixing, in: International Conference on Scale Space and Variational Methods in Computer Vision, Springer Berlin Heidelberg. pp. 435–446.

Rachev, S., Rüschendorf, L., 1998. Mass Transportation Problems: Volume I: Theory. Mass Transportation Problems, Springer.

Radford, A., Metz, L., Chintala, S., 2016. Unsupervised representation learning with deep convolutional generative adversarial networks, in: Proceedings of the 4th International Conference on Learning Representations (ICLR).

Ramachandran, P., Zoph, B., Le, Q.V., 2017. Searching for activation functions. CoRR abs/1710.05941. arXiv:1710.05941.

Redko, I., Courty, N., Flamary, R., Tuia, D., 2019. Optimal transport for multi-source domain adaptation under target shift, in: Proceedings of Machine Learning Research, pp. 849–858.

Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P., 2001. Color transfer between images. IEEE Computer Graphics and Applications 21, 34–41.

Rolland, J.P., Vo, V., Bloss, B., Abbey, C.K., 2000. Fast algorithms for histogram matching: Application to texture synthesis. Journal of Electronic Imaging 9, 39–46.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. The annals of mathematical statistics 23, 470–472.

Roy, A.S., Gopinath, A., Rangarajan, A., 2007. Deformable density matching for 3d non-rigid registration of shapes, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 942–949.

Rubner, Y., Tomasi, C., Guibas, L.J., 2000. The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision 40, 99–121.

Ruderman, D.L., Cronin, T.W., Chiao, C.C., 1998. Statistics of cone responses to natural images: implications for visual coding. Journal of the Optical Society of America A 15, 2036–2045.

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

Sanjabi, M., Ba, J., Razaviyayn, M., Lee, J.D., 2018. On the convergence and robustness of training gans with regularized optimal transport, in: Advances in Neural Information Processing Systems, pp. 7091–7101.

Santambrogio, F., 2015. Optimal transport for applied mathematicians. Birkäuser, NY 55, 94.

Scardapane, S., Van Vaerenbergh, S., Totaro, S., Uncini, A., 2019. Kafnets: Kernel-based non-parametric activation functions for neural networks. Neural Networks 110, 19–32.

Scott, D.W., 2001. Parametric statistical modeling by minimum integrated square error. Technometrics 43, 274–285.

Shanno, D.F., 1970. Conditioning of quasi-newton methods for function minimization. Mathematics of Computation 24, 647–656.

Shao, J., Tu, D., 2012. The jackknife and bootstrap. Springer Science & Business Media.

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 23rd ACM national conference, pp. 517–524.

Shu, Z., Hadap, S., Shechtman, E., Sunkavalli, K., Paris, S., Samaras, D., 2017. Portrait lighting transfer using a mass transport approach. ACM Transactions on Graphics (TOG) 36.

Shuchang Xu, Yin Zhang, Sanyuan Zhang, Xiuzi Ye, 2005. Uniform color transfer, in: Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. III–940.

Su, Z., Wang, Y., Shi, R., Zeng, W., Sun, J., Luo, F., Gu, X., 2015. Optimal mass transport for shape matching and comparison. IEEE Transactions on Pattern Analysis and Machine Intelligence 37, 2246–2259.

Sugiyama, M., 2013. Direct approximation of divergences between probability distributions, in: Empirical Inference. Springer, pp. 273–283.

Tai, Y., Jia, J., Tang, C., 2007. Soft color segmentation and its applications. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 1520–1537.

Tanaka, A., 2019. Discriminator optimal transport, in: Advances in Neural Information Processing Systems, pp. 6816–6826.

Terrell, G.R., 1990. The maximal smoothing principle in density estimation. Journal of the American Statistical Association 85, 470–477.

Terrell, G.R., Scott, D.W., 1992. Variable kernel density estimation. The Annals of Statistics 20, 1236–1265.

Turlach, B.A., 1993. Bandwidth selection in kernel density estimation: A review, in: CORE and Institut de Statistique.

Vazquez-Corral, J., Bertalmío, M., 2014. Color stabilization along time and across shots of the same scene, for one or several cameras of unknown specifications. IEEE Transactions on Image Processing 23, 4564–4575.

Villani, C., 2003. Topics in Optimal Transportation. Graduate studies in mathematics, American Mathematical Society.

Villani, C., 2008. Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg.

Wang, F., Vemuri, B., Syeda-Mahmood, T., 2009. Generalized l_2-divergence and its application to shape alignment, in: International Conference on Information Processing in Medical Imaging, Springer. pp. 227–238.

Wang, F., Vemuri, B.C., Rangarajan, A., 2006. Groupwise point pattern registration using a novel cdf-based jensen-shannon divergence, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1283–1288.

Wang, Q., Sun, X., Wang, Z., 2010. A robust algorithm for color correction between two stereo images, in: Proceedings of the Asian Conference on Computer Vision (ACCV), pp. 405–416.

Watson, G.S., 1964. Smooth regression analysis. Sankhyā: The Indian Journal of Statistics, Series A 26, 359–372.

Xia, M., Renping, J.Y., Zhang, X.M., Xiao, J., 2017. Color consistency correction based on remapping optimization for image stitching, in: Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW), pp. 2977–2984.

Xiang, Y., Zou, B., Li, H., 2009. Selective color transfer with multi-source images. Pattern Recognition Letters 30, 682 – 689.

Xiao, C., Han, C., Zhang, Z., Qin, J., Wong, T.T., Han, G., He, S., 2020. Example-based colourization via dense encoding pyramids. Computer Graphics Forum 39, 20–33.

Xiao, X., Ma, L., 2009. Gradient-preserving color transfer. Computer Graphics Forum 28, 1879–1886.

Yoo, J.D., Park, M.K., Cho, J.H., Lee, K.H., 2013. Local color transfer between images using dominant colors. Journal of Electronic Imaging 22, 1 – 11.

Zhang, L., Zhang, L., Mou, X., Zhang, D., 2011. Fsim: A feature similarity index for image quality assessment. IEEE Transactions on Image Processing 20, 2378–2386.

Zhou Wang, Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13, 600–612.