Patch based Colour Transfer using SIFT Flow

Hana Alghamdi & Rozenn Dahyot

School of Computer Science & Statistics
Trinity College Dublin, Ireland
alghamdh@tcd.ie, rozenn.dahyot@tcd.ie

Abstract

We propose a new colour transfer method with Optimal Transport (OT) 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 OT does not take into account any available information about correspondences when computing the optimal solution. To tackle this problem we propose to encode overlapping neighborhoods of pixels using both their colour and spatial correspondences estimated using motion estimation. We solve the high dimensional problem in 1D space using an iterative projection approach. We further 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 previous state of the art colour transfer methods.

Keywords: Optimal Transport, Nadaraya-Watson estimator, Iterative Distribution Transfer, Sliced Wasserstein Distance, Colour Transfer

1 Introduction

Colour variations between photographs often happen due to illumination changes, using different cameras, different in-camera settings or due to tonal adjustments of 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. Colour transfer has many applications in image processing problems, ranging from generating colour consistent image mosaicing and stitching [1] to colour enhancement and style manipulation [2].

When computing the transfer function, considering colour information only 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. To tackle this problem, Alghamdi et al. [3] proposed the Patch based Colour Transfer (PCT_OT) approach that encodes overlapping neighborhoods of pixels, taking into account both their colour and pixel positions. 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. In this paper we propose to improve PCT_OT by first improving the data preparation step for defining patches thanks to SIFT flow [4]. We estimate motions between images using the SIFT flow approach and incorporate the spatial correspondence information in the encoded overlapping neighborhoods of pixels. This formulation makes OT implicitly take 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).

2 PCT_OT with SIFT Flow

2.1 Combine colour and spatial information

The spatial information for the target image is calculated using the SIFT flow method which estimates dense spatial correspondences by robustly aligning complex scene pairs containing significant spatial differences
[4], while in PCT_OT [3] 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 to be computed, and let \( p = (a, b) \) be the 2D grid coordinate of the target image and \( w(p) = (u(p), v(p)) \) be the flow vector at \( p \) computed using the SIFT flow method, then \( y^p = p + w(p) = (a + u(p), b + v(p)) \) is the new pixel position in the target image that matches a pixel position in the source image. The pixel’s colour \( y^c \) and its 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. \( x^p = p \) and similarly to the target image the pixel’s colour \( x^c \) and its pixel position \( x^p \) are concatenated into a vector \( x = (x^c, x^p)^T \) such that \( \dim(x) = \dim(x^c) + \dim(x^p) \).

### 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 \( \mathbb{R}^d \) 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 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 and hence the more overlapped areas we have the higher \( w \) value we can set.

### 2.3 Create patch vectors

In a similar way to PCT_OT [3] 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 \( k \times k \) to extract overlapping patches. From each individual patch we create a high dimensional vector in \( \mathbb{R}^{d \times k \times k} \). We apply this process to the source and target images to create patch vector sets \( \{x_i\} \) and \( \{y_j\} \) for each respectively.

### 3 Smoothed solution for 1D Optimal Transport

The OT problem consists of estimating the minimum cost (referred to as the Wasserstein Distance [5] or as the Earth Mover’s Distance [6]) of transferring a source distribution to a target distribution. As a byproduct of OT distance estimation, the mapping \( \phi \) itself between the two distributions is also provided. Monge’s formulation of OT [5] 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:

\[
\arg\min_{\phi} \int \| x - \phi(x) \|^2 f(x) \, dx \quad \text{such that: } \quad f(x) = g(\phi(x)) \quad |\det \nabla \phi(x)|
\]

where \( f \) is the probability density function (pdf) of \( x \) and \( g \) is the pdf of \( y \). The solution for \( \phi \) can be found using existing algorithms such as linear programming, and the Hungarian and Auction algorithms [7]. However, in practice it is difficult to find a solution for colour images when \( \dim(x) = \dim(y) = d > 1 \) as the computational complexity of these solvers increases in multidimensional spaces [8]. But for \( d = 1 \), with \( x, y \in \mathbb{R} \), a solution for \( \phi \) is straightforward and can be defined using the increasing rearrangement [5]:

\[
\phi^{OT} = G^{-1} \circ F
\]

where \( F \) and \( G \) are the cumulative distributions of the colour values in the source and target images respectively.

### 3.1 Iterative Distribution Transfer (IDT)

The 1D solution \( \phi^{OT} \) Eq. (2) has been used to tackle problems in multidimensional colour spaces and of particular interest is the Iterative Distribution Transfer (IDT) algorithm for colour transfer proposed by Pitié et al. [9]. They proposed to iteratively project colour values \( \{x_i\}_{i=1}^n \) and \( \{y_j\}_{j=1}^m \) originally in \( \mathbb{R}^d \) to a 1D subspace.
and solve the OT using $\phi^{OT}$ Eq. (2) in this 1D subspace and then propagate the solution back to $\mathbb{R}^d$ space. This operation is repeated with different directions in 1D space until convergence. This strategy was inspired by the idea of the Radon Transform [9] 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 $\mathbb{R}^d$ 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 solution [9].

3.2 Sliced-Wasserstein Distance (SWD)

The Sliced Wasserstein Distance (SWD) algorithm follows from the iterative projection approach of IDT but computes the 1D solution $\phi^{OT}$ with quantile matching instead of cumulative histogram matching [10, 11]. 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 $O(n \log(n))$ operations using a fast sorting algorithm [10]. When a small number of observations are available, using SWD is best but with a large number of observations, histogram matching with IDT is more efficient.

3.3 Smoothing $\phi^{OT}$ with Nadaraya Watson Estimator

Giving the correspondences $\{(x_i, y_i)\}_{i=1,\ldots,n}$, the Nadaraya Watson (NW) estimator is defined as follows:

$$E[y|x] = \int y \ p(y|x) \ dy = \int y \ \frac{p(y,x)}{p(x)} \ dy = \frac{n^{-1} \sum_{i=1}^{n} y_i \ K_h(x-x_i)}{n^{-1} \sum_{i=1}^{n} K_h(x-x_i)} = \phi^{NW}_h(x) \quad (3)$$

With this form NW can be seen as locally weighted average of $\{y_i\}_{i=1,\ldots,n}$, 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 $\phi^{NW}_h$ gets smoothed. We propose to smooth $\phi^{OT}$ computed in IDT or SWD by using non-parametric Nadaraya Watson estimator. At each iteration $t$, following the step of calculating the optimal map $\phi^{OT}$, we feed the OT estimated correspondences $\{(x_i, \phi^{OT}_t(x_i))\}_{i=1}^{n}$ to the NW estimator to compute a smoother OT solution, denoted as $\phi^{OT}_h$, defined as follows:

$$\phi^{OT}_h(x) = \sum_{i=1}^{n} \phi^{OT}_t(x_i) \ K_h(x-x_i) \sum_{i=1}^{n} K_h(x-x_i) \quad (4)$$

Figure 1: Results shows the smoothed Optimal Transport solution using non-parametric Nadaraya-Watson ($\phi^{OT}_h$) with different bandwidth values $h = \{3, 10, 20\}$. Nadaraya-Watson significantly reduces the grainy artifacts produced by the original Optimal Transport function ($\phi^{OT}$), mapping the source patch projections to the target patch projections, the bigger $h$ value the more smoothed mapping. The results processed without post processing step. Note that the graph is a zoom in on 0-255 range pixel value.

Figure 1 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 [9]. Optimal
Transport solutions are suitable in situations where the function that we need to estimate must satisfy important side conditions, such as being strictly increasing, and 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 explicit expression controlled by parameters on the regression function which makes them directly employable. In the following sections we are applying OT and NW smoothing in the relevant context of colour transfer where the function that we need to estimate must satisfy the condition of being an increasing function.

4 Experimental Assessment

We provide here quantitative and qualitative evaluations of our approach noted OT\_NW with comparisons to different state of the art colour transfer methods noted IDT [9], PMLS [2], GPS/LCP and FGPS/LCP [12], L2 [13] and PCT\_OT [3]. In these evaluations we use image pairs with similar content from an existing dataset provided by Hwang et al [2]. The dataset includes registered pairs of images (source and target) taken with different cameras and settings, and different illuminations and recolouring styles.

4.1 Colour space and parameters settings

We use the RGB colour space where each pixel is represented by its 3D RGB colour values and its 2D spatial position. 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 component (2D). We found that a patch size of $5 \times 5$ captures enough of a pixel's neighbourhood. We stretch the hypercube space in $\mathbb{R}^d$ in the direction of the spatial coordinates by a factor $w = 10$ to make it harder to move the pixels in the spatial domain than in the colour domain. We experimented with different bandwidth values and we found a fixed value of $h = 10$ gives best results.

4.2 Evaluation metrics

To quantitatively assess the recolouring results, four metrics are used: peak signal to noise ratio (PSNR) [14], structural similarity index (SSIM) [15], colour image difference (CID) [16] and feature similarity index (FSIMc) [17]. These metrics are often used when considering source and target images of the same content [18, 19, 2, 12]. Note that the results using PMLS were provided by the authors [2]. It has already been shown in [13] that PMLS performs better than two other more recent techniques using correspondences [20, 21], so PMLS is the one reported here with [3, 13] as algorithms that account for correspondences.

4.3 Experimental Results

Figures 2–5 show detailed tables of quantitative results for each metric along with box plots carrying a lot of statistical details. 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. Figure 2 (b) and Figure 5 (b) show PSNR and FSIMc metrics results respectively, by examining the box plots in both figures we see that the four methods PMLS, L2, PCT\_OT and OT\_NW are greatly overlapped with each other, the median and mean values (the mean shown as red dots in the plots) are the highest among all algorithms and are very close in value and the whiskers length almost similar indicating similar data variation and consistency. Figure 3 (b) shows the SSIM box plot where we can see that OT\_NW performs similarly to PMLS and L2’s highest scoring values while here the median line of PCT\_OT box lies outside the three top algorithms scoring the lowest value among them. With CID metric in Figure 4, OT\_NW performs similarly to PMLS, L2 and PCT\_OT. In conclusion, the quantitative metrics show that our algorithm with Nadaraya Watson OT\_NW performs similarly with top methods PMLS, L2 and PCT\_OT and outperforms the rest of the state of the art algorithms.

Figure 7 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 (Figure 7, 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. As can be noted, our approach OT\_NW with Nadaraya Watson step is visually the best at removing the column differences.

While PMLS and PCT\_OT provide equivalent results to our method in terms of metrics measures, PMLS introduces visual artifacts if the input images are not registered correctly (Figure 6), while our method is robust.
Several contributions to colour transfer with OT have been made in this paper, showing quantitative and qualitative improvements over state of the art methods. In particular, first, correspondences information as well as colour content of pixels are both encoded in the high dimensional feature vectors, 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 smooth colour transfer between images that may contain large motion changes.

5 Conclusion

Several contributions to colour transfer with OT have been made in this paper, showing quantitative and qualitative improvements over state of the art methods. In particular, first, correspondences information as well as colour content of pixels are both encoded in the high dimensional feature vectors, 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 smooth colour transfer between images that may contain large motion changes.
Figure 4: Metric comparison, using CID [16]. (a) Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (lower values are better), (b) visualized in box plot (best viewed in colour and zoomed in).

Figure 5: Metric comparison, using FSIMc [17]. (a) Red, blue, and green indicate 1st, 2nd, and 3rd best performance respectively in the table (higher values are better), (b) visualized in box plot (best viewed in colour and zoomed in).

Figure 6: A close up look at some of the results generated using the PMLS [2], L2 [13], PCT_OT [3] and our algorithm OT_NW (best viewed in colour and zoomed in).
Figure 7: A close up look at some of the results generated using the IDT [9], PMLS [2], GPS/LCP and FGPS/LCP [12], L2 [13], PCT_OT [3] and our algorithm OT_NW. 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 (best viewed in colour and zoomed in).
Acknowledgments

This work is partly funded by a scholarship from Umm Al-Qura University, Saudi Arabia, and in part by a research grant from Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776, and the ADAPT Centre for Digital Content Technology (www.adaptcentre.ie) that is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

References


