Computer Vision and Image Understanding 181 (2019) 39–49

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

$\mathcal{L}_2$ Divergence for robust colour transfer

Mairéad Grogan *, Rozenn Dahyot

School of Computer Science & Statistics, Trinity College Dublin, Ireland

ARTICLE INFO

Communicated by N. Paragios

MSC:
41A05
41A10
65D05
65D17

Keywords:
Colour Transfer
L2 Registration
Re-colouring
Colour Grading

ABSTRACT

Optimal Transport (OT) is a very popular framework for performing colour transfer in images and videos. We have proposed an alternative framework 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 (Grogan and Dahyot, 2015). In this paper, we show that our approach combines many advantages of state of the art techniques and outperforms many recent algorithms as measured quantitatively with standard quality metrics, and qualitatively using perceptual studies (Grogan and Dahyot, 2017). Mathematically, our formulation is presented in contrast to the OT cost function that shares similarities with our cost function. Our formulation, however, is more flexible as it allows colour correspondences that may be available to be taken into account and performs well despite potential occurrences of correspondence outlier pairs. Our algorithm is shown to be fast, robust and it easily allows for user interaction providing freedom for artists to fine tune the recoloured images and videos (Grogan et al., 2017).

1. Introduction

Colour transfer refers to a set of techniques that aim to modify the colour feel of a target image or video using an exemplar colour palette provided by another image or video. Most techniques are based on the idea of warping some colour statistics from the target image colour distribution to match those of the palette image colour distribution. The transfer (or warping) function ϕ, once estimated, is then used to recolour a colour pixel value x to $\phi(x)$.

Many factors can be considered when assessing a colour transfer technique, with the most important ones summarised in Fig. 1. The first consideration is the type of data processed by the system. While most applications process images, they can require considerable effort to ensure that the colour transfer function can be applied to video data without creating temporal artifacts (Bonneel et al., 2013). Many techniques also either consider target and palette images with overlapping content (Xia et al., 2017; Park et al., 2016; Oliveira et al., 2015; Hwang et al., 2014; HaCohen et al., 2011), or those with completely different content (Pitié et al., 2007; Bonneel et al., 2015). For the former, pixel correspondences can be used to enhance the colour transfer estimation, while purely statistical properties of the images are considered in the latter case. The quality of the colour transfer results generated is also an important factor, and can be assessed by considering whether the application successfully transfers the colour feel from the palette to the target image, whether the results are good even in the case of bad pixel correspondences (if used), and whether the results generated are repeatable and consistently successful. As many techniques aim to work in real time, speed is also an important aspect. With the increased use of parallel architectures, the ability to parallelise some or all of the computation is also advantageous. The final factor to consider is the type of user interface that will be provided. Methods typically fall into one of two categories — automatic methods and those that allow for user interaction. While automatic methods generate result images easily, they greatly reduce the artistic freedom of the user. This is the main drawback of automatic methods, including many of those based on neural networks, which have become popular in recent years (He et al., 2017; Luan et al., 2017; Liao et al., 2017).

In this paper we explore a colour transfer technique that tries to satisfy all of the requirements shown in Fig. 1. Our colour transfer approach can be applied to target and palette images both with and without pixel correspondences, and can be easily applied to video content (Grogan and Dahyot, 2015). Recolouring using our parametric transfer function is highly parallelisable, and once estimated the transfer function can be easily stored and reused. Our $\mathcal{L}_2 E$ based cost function is robust to correspondence outliers and our colour transfer results outperform other state of the art techniques. Finally, this paper presents colour transfer results automatically generated given a target and palette image, with many and without any pixel correspondences between the two input images. We have also recently shown that our colour transfer algorithm can be efficiently tuned further by a user (or artist), specifying a small number of correspondences to minimise interaction (Grogan et al., 2017).

In the next section, we present relevant past work in the area of colour transfer and many of these past ideas are integrated as part of

* Corresponding author.
E-mail address: mgrogan@tcd.ie (M. Grogan).

https://doi.org/10.1016/j.cviu.2019.02.002
Received 9 April 2018; Received in revised form 6 February 2019; Accepted 8 February 2019
Available online 15 February 2019

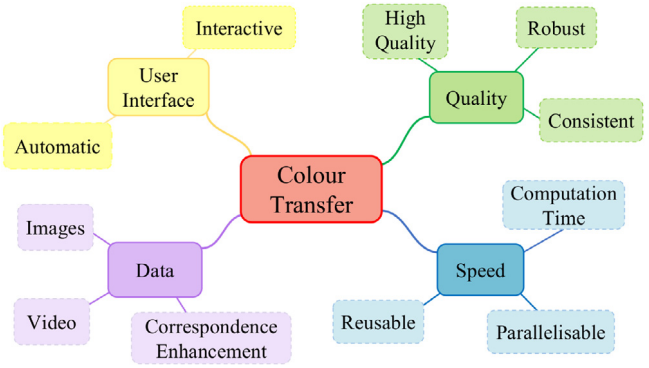

Fig. 1. Factors to consider when assessing a colour transfer technique.

our robust modelling with the $\mathcal{L}_2 E$ divergence (Section 3). Performance of our approach is then compared qualitatively and quantitatively with state of the art techniques in Section 4.

2. State of the art

The body of work in recolouring is very large and the reader is referred to this recent exhaustive review by Faridul et al. (2014). Early work on colour transfer started with registering statistical moments of colour distributions (Section 2.1) using a parametric affine formulation of the transfer function ϕ. This methodology soon shifted to using the Optimal Transport (OT) framework (Section 2.2). OT techniques use non-parametric estimates of colour distributions, and the resulting algorithms for colour corrections do not provide an explicit expression of ϕ but instead an estimated correspondence $(x, \phi(x))$ for every colour pixel x. This can be memory consuming when capturing ϕ on a 256^3 discrete RGB colour space for instance. Alternative methods (Section 2.3) instead propose to capture colour distributions with Gaussian Mixture models and use some correspondences (Section 2.4) between Gaussian components of the palette and target distributions.

2.1. Registration of colour statistical moments

The pioneering work of Reinhard et al. (2001) proposed to use a warping function ϕ with parametric form:

$$\phi(x, \theta = \{\mathrm{G}, o\}) = \mathrm{G}\ x + o \tag{1}$$

with the vector o representing an offset and the 3×3 diagonal matrix G representing the gains for each colour channel. The estimation of the parameter θ is performed by registering the probability density functions of the colours in the palette and target images, denoted p_p and p_t, represented as simple multivariate Gaussians ($p_p \equiv \mathcal{N}(x; \mu_p, \Sigma_p)$ and $p_t \equiv \mathcal{N}(x; \mu_t, \Sigma_t)$) with diagonal covariance matrices Σ_p and Σ_t. Since Normal distributions are fully described by their first two statistical moments, means and covariance matrices, the optimal mapping ϕ is specified by the solution $\hat{\theta}$ that maps the empirical estimates of (μ_t, Σ_t) computed using the pixel values $\{x_t^{(i)}\}_{i=1,\dots,n_t}$ in the target image, to the empirical estimates of (μ_p, Σ_p) computed using the pixel values $\{x_p^{(i)}\}_{i=1,\dots,n_p}$ in the palette image. Pitié and Kokaram (2007) proposed an OT solution to this affine transfer function modelling p_p and p_t as simple multivariate Gaussians likewise. However these two methods use affine transformations which are limited in the type of results that they can achieve, therefore recent methods have considered more complex transformations.

2.2. Optimal transport

Monge's formulation (Villani, 2009) sets the deterministic decoupling $y = \phi(x)$ linking two random variables $y \sim p_t(y)$ and $x \sim p_p(x)$ imposing the solution ϕ that verifies:

$$p_p(x) = p_t(\phi(x)) \times |\det \nabla\phi(x)| \tag{2}$$

In practice, finding ϕ such that Eq. (2) is true is difficult when considering multidimensional space (Villani, 2003). A solution when $x \in \mathbb{R}$ and $y \in \mathbb{R}$ can however be defined with the cumulative distribution of colours in the target and palette images P_t and P_p:

$$\phi(x) = P_t^{-1} \circ P_p(x) \tag{3}$$

when P_t is strictly increasing (i.e. $p_t(y) > 0, \forall y \in \mathbb{R}$). P_t and P_p can be approximated with cumulative histograms for instance. Such a process for finding the warping function ϕ is powerful since no strong hypotheses are made about the distributions (as opposed to the Gaussian assumption in Section 2.1). Moreover, no parametric form is imposed on the warping function ϕ.

A solution to Eq. (2) becomes non trivial in multidimensional colour spaces. Of particular interest is the pioneering work of Pitié et al. (2005) who proposed an iterative algorithm that first projects the colour pixels $\{x^{(k)}\}$ on a 1D Euclidean space, and then estimates $\hat{\phi}$ using Eq. (3) and applies it to move all values $\{x^{(k)}\}$ along the direction of the 1D space. This operation is repeated with different directions in 1D space until convergence. The intuitive explanation for this approach comes from the Radon Transform (Pitié et al., 2007): if Eq. (3) is verified in all 1D projective spaces, then Eq. (2) will be verified in the multidimensional space. However, this approach alters the colours of the target image so they match the colour distribution of the palette image exactly, which is not ideal when the colour distributions of the images are very different, and causes grainy artifacts to appear in the gradient of the result image. The subsequent extension by Pitié et al. (2007) proposed a post processing step to correct this artifact and ensure that the gradient field of the recoloured target image is as close as possible to the original target image. Similar to Pitié et al. (2007) and Bonneel et al. (2015) recently proposed to use a Radon Transform inspired strategy for colour transfer in their OT method. Their approach is a generalisation of the method proposed by Pitié et al. which uses 1-D Wasserstein distances to compute the barycentre of a number of input measures. The OT problem is formulated with a cost function c:

$$\hat{\pi} = \inf_{\pi} \iint c(x, y)\ \pi(x, y)\ dx\ dy \tag{4}$$

where the marginals of the joint density function π are p_t and p_p. Bonneel et al. (2015) used the $\mathcal{L}_2$-Wasserstein distance with the Euclidean distance $c(x, y) = \|x - y\|^2$: note that this $\mathcal{L}_2$ Wasserstein distance between random vectors x and y is different to that used in our framework, where our $\mathcal{L}_2$ refers to the Euclidean distance between marginals p_t and p_p (see Section 3). When this method is used for colour transfer between target and palette images, it matches only the chrominance channels of the target and palette images, leaving the target image's luminance channel unchanged after recolouring. As well as colour transfer, this OT approach can also be used to find the barycentre of three or more weighted image palettes.

Histograms are often employed to approximate the colour distributions of images (Neumann and Neumann, 2005; Papadakis et al., 2011; Pouli and Reinhard, 2011) and used in OT methods (Freedman and Kisilev, 2010; Ferradans et al., 2013; Pouli and Reinhard, 2011). Similar to the method proposed by Pitié et al. (2005), these discrete methods have a tendency to introduce grainy artifacts in the gradient of the result image, and like (Pitié et al., 2007), recent methods have also proposed adding a step to impose that the resulting spatial gradient of the recoloured image remains similar to the target image (Xiao and Ma, 2009; Papadakis et al., 2011; Ferradans et al., 2013). Alternatively, other methods have proposed to relax the constraint that enforces the distributions of the recoloured target image and palette image to

match exactly (Freedman and Kisilev, 2010; Pouli and Reinhard, 2011; Ferradans et al., 2013). Bonneel et al. (2015) propose using a gradient smoothing technique to reduce any quantisation errors that appear in their results (Rabin et al., 2011). Frigo et al. (2015b) propose to remove artifacts by first estimating an OT solution and using it to compute a smooth Thin Plate Spline (TPS) transformation to ensure that a smooth parametric warping function is used for recolouring, allowing them to apply their method to video content easily.

2.3. Using Gaussian Mixture models

Gaussian Mixture models are often used to model the colour distribution of images in colour transfer applications. Jeong and Jaynes (2007) use colour transfer techniques to harmonise the colour distributions of non overlapping images of the same object for tracking purposes in a multiple camera setting. The colour chrominance (2D) distribution is modelled using Gaussian Mixture models, and the transfer function is parametric with an affine form estimated by minimising the Kullback–Leibler divergence between Gaussian components using a robust procedure to tackle outlier pairs. Xiang et al. (2009) model the colour distribution of the target image with a Gaussian Mixture estimated by an EM algorithm. Each Gaussian component in the mixture defines a local region in the target image, and each segmented local region is recoloured independently using multiple palette images, likewise segmented in regions using Gaussian Mixture model fitting of their colour distributions. Colour transfer is performed by associating the best Gaussian component from the palettes to the Gaussian target region using the Gaussian affine transfer function proposed by Reinhard et al. (2001). Segmentation using alternative approaches such as meanshift (Comaniciu and Meer, 2002) and K-means (Xu et al., 2005) is also tested to define the colour Gaussian Mixtures. This approach relies on homogeneous colour regions each captured with one multivariate normal in the 3D colour space. Localised colour transfer using Gaussian Mixture models between overlapping colour images have also been proposed for colour corrections, motion deblurring, denoising, grey scale colouring (Tai et al., 2005).

2.4. Using correspondences

Other methods propose taking pixel correspondences into account when finding the colour transformation (Oliveira et al., 2015; HaCohen et al., 2013; Hwang et al., 2014; Park et al., 2016; Vazquez-Corral and Bertalmo, 2014). Oliveira et al. (2015) proposed to find the mapping of 1-dimensional truncated Gaussian Mixture model representations computed for each colour channel of the target and palette images. For colour correction of images capturing a scene taken under the same illumination, Vazquez-Corral and Bertalmo (2014) use colour correspondences to compute a colour warping function consisting of both a linear 3×3 matrix and gamma correction power terms. They show it performs well when images have been captured using different cameras, with different camera settings, and with different white balance and exposure. Frigo et al. (2015a) tackle the problem of tonal fluctuation in video by combining a 6 parameter colour warping function ϕ, similar to that proposed by Vazquez-Corral and Bertalmo (2014), with motion estimation to perform a fast, temporally weighted colour correction of the video frames. Park et al. (2016) propose 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 less correspondences than similar techniques (HaCohen et al., 2013). Xiao et al. (2006) also tackle colour inconsistencies across images due to imaging conditions, with an emphasis on maintaining colour consistency, but also ensuring individual images maintain high definition gradients and wide dynamic range. But all of these algorithms estimate warping functions controlled by a small number of parameters and these cannot successfully tackle more complex colour mapping needed in unconstrained scenarios of applications of colour transfer.

Hwang et al. (2014) on the other hand, use moving least squares to estimate an affine transformation for each pixel in the target image with the least squares algorithm. This technique can correct a larger variety of colour differences between images, include large non linear colour differences. To tackle sensitivity to outlier correspondences, only subsets of the correspondences are used as control points and a probabilistic framework is deployed to remove correspondences that are likely to be incorrect. However all of these methods are only applicable when the correspondences between target pixels and palette pixels are available.

3. Robust colour transfer with $\mathcal{L}_2 E$

We propose to solve colour transfer by estimating the warping function ϕ that minimises a divergence between two probability density functions. Unlike the OT formulation, which depends on two random vectors x and y (see Eq. (4)), only one random vector y appears explicitly in our modelling, and both the target pdf p_t and palette pdf p_p are candidate pdfs for random vector y defined in the colour space $\mathbb{R}^3$. Our approach to solving colour transfer is to propose a parametric model $p_t(y|\theta)$ for the target distribution such that the parameter θ controls the mapping function ϕ. We propose to formulate the colour transfer problem as a parameter estimation one:

$$\begin{aligned}\hat{\theta} &= \arg\min_{\theta} \left\{ \mathcal{L}_2 = \|p_t - p_p\|^2 = \int (p_t(y|\theta) - p_p(y))^2 dy \right\} \\ &= \arg\min_{\theta} \left\{ \mathcal{L}_2 E = \|p_t\|^2 - 2\,\langle p_t | p_p \rangle \right\}\end{aligned} \tag{5}$$

While many divergences exist, the $\mathcal{L}_2$ divergence (or $\mathcal{L}_2 E$) is chosen for its robustness and its ease of computation when using Gaussian Mixtures (Scott, 2001; Jian and Vemuri, 2011; Ma et al., 2015).

In Eq. (5) the term $\|p\|^2$ is related to the quadratic Renyi entropy of p noted $\mathcal{R}(p)$ (for both target and palette distributions):

$$\|p\|^2 = \exp\left(-\mathcal{R}(p)\right) \tag{6}$$

while the scalar product term $\langle p_t | p_p \rangle$ is related similarly to the Renyi cross entropy between p_t and p_p (Escolano et al., 2009). The entropy of the palette $\|p_p\|^2$ is not needed for solving this estimation: we therefore use $\mathcal{L}_2 E$ as our cost function, maximising the entropy of the target model $\mathcal{R}(p_t)$, and maximising the overlap between the target and palette pdfs. This framework was first investigated in Grogan et al. (2015), Grogan and Dahyot (2015), and it is extended further here to take advantage of correspondences that may be available between pixels of the target image to be recoloured, and the pixels of the palette image used as an exemplar (see Section 3.2.2).

Comparing OT and $\mathcal{L}_2 E$. Interestingly, recent advances in OT also include an entropy based regularisation term extending Eq. (4) (Bonneel et al., 2016; Courty et al., 2017) such that:

$$\hat{\pi} = \arg\min_{\pi} \{\langle c \mid \pi \rangle + \lambda\, \mathcal{H}(\pi)\} \tag{7}$$

Parameter λ weights the contribution of the negentropy $\mathcal{H}$ in the regularisation (Bonneel et al., 2016). The regularised OT cost (Eq. (7)) has a similar form to our $\mathcal{L}_2 E$ cost function (Eq. (5)) where $\|p_t\|^2$ acts as the Renyi entropy based regularisation (albeit not adjusted with a tuning parameter λ) added to the cross product $\langle p_t \mid p_p \rangle$ (mimicking the term $-\langle c \mid \pi \rangle$). Our $\mathcal{L}_2 E$ formulation considers the modelling of a single random variable (or vector) y for which two pdfs are defined (p_t and p_p), hence integration $\langle \mid \rangle$ is w.r.t. y. In contrast, integration in the OT formulation is w.r.t. two variables x and y whose interaction is modelled with the joint density function $\pi(x, y)$.

Table 1
Road-map for experiments. Our framework is able to take advantage of correspondences (Corr) between Target (T) and Palette (P) images when available, and without correspondences (No Corr). While correspondences are easily available when palette and target images capture the same visual content ($P \simeq T$), they are not available when using images of different content ($P \neq T$).

Method	Reference	Correspondences	No correspondences	Open source	Image	Video	Test $P \simeq T$ (Section 4.1)	Test $P \neq T$ (Section 4.2)
Ours		✓	✓	✓	✓	✓	✓	✓
Bonneel	Bonneel et al. (2015)		✓	✓	✓		✓	✓
PMLS	Hwang et al. (2014)	✓			✓	✓	✓	
Ferradans	Ferradans et al. (2013)		✓	✓	✓		✓	✓
Pitié	Pitié et al. (2007)		✓	✓	✓		✓	✓
Park	Park et al. (2016)	✓		✓	✓			✓
Xia	Xia et al. (2017)	✓		✓	✓			✓

3.1. Parametric pdf p_t and pdf p_p

The parametric pdf p_t representing the distribution of colours in the target image is modelled as a Gaussian Mixture model with parameterised means $\{\phi_\theta(\mu_t^{(k)})\}_{k=1,\dots,K}$, where $\{\mu_t^{(k)}\}_{k=1,\dots,K}$ are selected by applying the K-means algorithm to the colours in the target image. The K-means algorithm is equivalent to using an EM algorithm enforcing identical isotropic covariance matrices (Wu et al., 2007). The faster Minimum Variance Quantisation (MVQ) algorithm can be used in practice (see Section 4). Similarly, the pdf p_p representing the distribution of colours in the palette image is modelled as a Gaussian Mixture model with means $\{(\mu_p^{(k)})\}_{k=1,\dots,K}$, again where $\{\mu_p^{(k)}\}_{k=1,\dots,K}$ are selected by the K-means algorithm. Both target and palette Gaussian Mixture models use isotropic identical covariance matrices controlled by a bandwidth h:

$$\Sigma = h^2 \mathrm{I} \tag{8}$$

with the identity matrix I.

The transfer function $\phi : x \in \mathbb{R}^d \rightarrow \phi_\theta(x) \in \mathbb{R}^d$ can be conveniently defined as a rigid, affine or Thin Plate Splines parametric function (Escolano et al., 2009; Jian and Vemuri, 2011) which can be written as:

$$\phi_\theta(x) = \underbrace{\mathrm{A}\ x + o}_{\text{Affine}} + \underbrace{\sum_{j=1}^{m} w_j\ \psi(\|x - c_j\|)}_{\text{nonlinear}} = \mathrm{A}\ x + o + \mathrm{W}\ \Psi(x) \tag{9}$$

Discarding the non linear part of ϕ provides an affine warping function which can be further limited to a rigid transformation by choosing A as a rotation matrix. While earlier colour transfer approaches estimated affine transfer functions (Reinhard et al., 2001; Pitié and Kokaram, 2007), they are limited in the type of results that they can achieve (Hwang et al., 2014). Recent methods have therefore considered more complex transformations (Hwang et al., 2014; Bonneel et al., 2013). Similarly, we decided to use the non-linear TPS transformation for our colour transfer function. For the TPS transformation, the latent parameter of interest, θ, corresponds to the set of coefficients $\{w_j \in \mathbb{R}^d\}$, the $d \times d$ matrix A and the d-dimensional vector offset o. The non linear part can be rewritten in matrix form with a $d \times m$ matrix W gathering coefficients $\{w_j \in \mathbb{R}^d\}$, and an m dimensional vector $\Psi(x)$ collating all values $\{\psi(\|x - c_j\|) = -\|x - c_j\|\}$. While many algorithms using TPS chose the set of control points $\{c_j\}_{j=1,\dots,m}$ as a subset of the observations $\{x^{(i)}\}$ (Chui and Rangarajan, 2003; Myronenko and Song, 2010), we choose the m control points $\{c_j\}_{j=1,\dots,m}$ on a regular grid spanning the 3D colour space such that the 3D grid has $m = 5 \times 5 \times 5 = 125$ control points in the colour space. The dimension of our latent space to explore for the estimation of θ is:

$$\dim(\theta) = 125 \times 3 + 9 + 3 = 387$$

with $\dim(c_j) = 3,\ \forall j$, $\dim(A) = 3 \times 3 = 9$ and $\dim(o) = 3$ (cf. Eq. (9)). Our choice of control points allows us to combine and easily create new warping functions using interpolation between parameter estimates (see Section 3.4).

3.2. Computing $\mathcal{L}_2E$

In $\mathcal{L}_2E$, the term $\|p_t\|^2$ can be computed as follows:

$$\|p_t\|^2 = \sum_{i=1}^{K}\sum_{k=1}^{K} \mathcal{N}(0; \phi_\theta(\mu_t^{(k)}) - \phi_\theta(\mu_t^{(i)}), 2h^2\mathrm{I})\ \pi_t^{(k)}\pi_t^{(i)} \tag{10}$$

where $\mathcal{N}(0; \phi_\theta(\mu_t^{(k)}) - \phi_\theta(\mu_t^{(i)}), h^2\mathrm{I})$ is the Normal distribution computed at 0 with mean $\phi_\theta(\mu_t^{(k)}) - \phi_\theta(\mu_t^{(i)})$ with covariance $2h^2\mathrm{I}$. The weights $\{\pi_t^{(i)}\}_{i=1,\dots,K}$ are chosen equiprobable with $\pi_t^{(i)} = \frac{1}{K}$.

3.2.1. Computing $\langle p_t|p_p\rangle$ without correspondences

The cross product can likewise be computed as:

$$\langle p_t|p_p\rangle = \sum_{i=1}^{K}\sum_{k=1}^{K} \mathcal{N}(0; \phi_\theta(\mu_t^{(k)}) - \mu_p^{(i)}, 2h^2\mathrm{I})\ \pi_t^{(k)}\pi_p^{(i)} \tag{11}$$

with the weights $\{\pi_p^{(i)}\}_{i=1,\dots,K}$ chosen equiprobable $\pi_p^{(i)} = \frac{1}{K}$ likewise as for p_t.

3.2.2. Computing $\langle p_t|p_p\rangle$ using correspondences

Pixel pairs between target and palette images, denoted $\{(x_t^{(k)}, x_p^{(k)})\}_{k=1,\dots,n}$, can be computed when considering target and palette images capturing the same scene. Colour transfer techniques are indeed often used in this context to harmonise colours across a video sequence and/or across multiple view images such as in the bullet time visual effect. In this case, the means of the Gaussian Mixtures are set such that $\{(\mu_t^{(k)}, \mu_p^{(k)}) = (x_t^{(k)}, x_p^{(k)})\}_{k=1,\dots,n}$ imposing $K_t = K_p = n$ when defining the distributions p_t and p_p. Moreover, the scalar product $\langle p_t|p_p\rangle$ in our cost function is then simplified as follows:

$$\langle p_t|p_p\rangle = \sum_{k=1}^{n} \mathcal{N}(0; \phi_\theta(\mu_t^{(k)}) - \mu_p^{(k)}, 2h^2\mathrm{I})\ \pi_t^{(k)}\ \pi_p^{(k)} \tag{12}$$

The computational complexity of this term is then $n = K_t = K_p$ when using n correspondences, and $K_t \times K_p$ without correspondences. Performance of our approach with correspondences is assessed in Section 4 using palette and target images with similar content. Pixel correspondences are not used when considering palette and target images with different content and MVQ is used instead to define the K centroids for each target and model pdfs.

3.3. Regularisation of $\mathcal{L}_2E$ and estimation of θ

Regularisation terms, priors on θ or ϕ_θ, can also be added to $\mathcal{L}_2E$ to constrain the estimation of θ (Jian and Vemuri, 2011; Arellano and Dahyot, 2012). To enforce that a smooth solution ϕ is estimated, the $\mathcal{L}_2E$ divergence is minimised with a roughness penalty on the transfer function ϕ (Escolano et al., 2009) and the estimation is performed as:

$$\hat{\theta} = \arg\min_\theta \left\{ C(\theta) = \|p_t\|^2 - 2\langle p_t|p_p\rangle + \lambda \int \|D^2\phi(x,\theta)\|^2 dx \right\} \tag{13}$$

with $\|D^2\phi(x,\theta)\|^2 = \sum_{i,j\in\{1,\dots,d\}} \left(\frac{\partial^2\phi}{\partial x_i \partial x_j}\right)^2$ with $d = 3$ for the colour spaces. Computation of $\hat{\theta}$ is performed by standard gradient descent algorithm with simulated annealing where the bandwidth h is used as a temperature (Grogan and Dahyot, 2017).

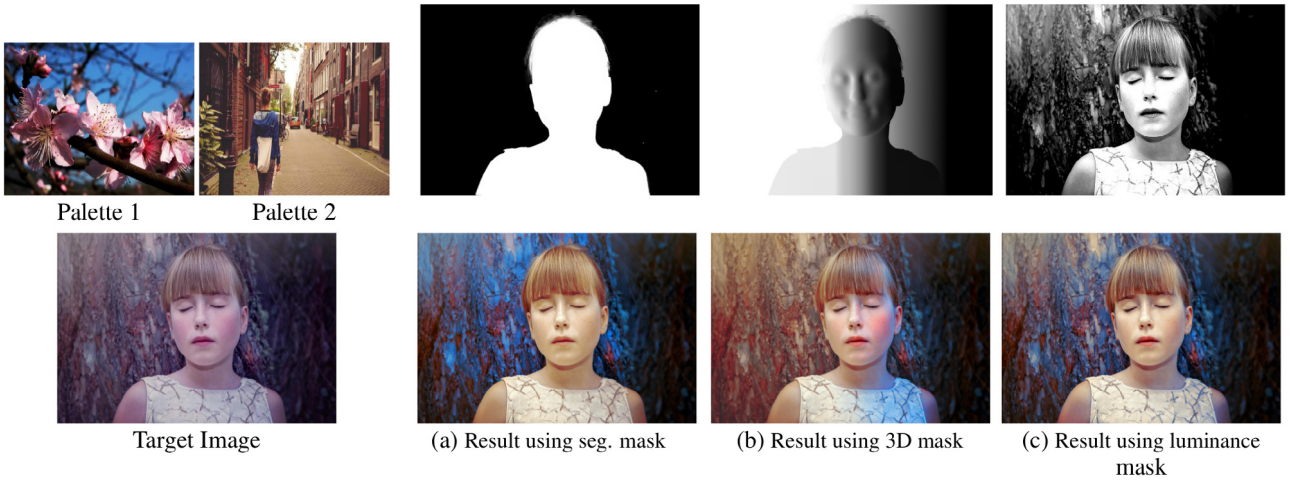

Fig. 2. The effect different masks γ(p) can have when recolouring the target image (bottom left) using two palettes, Palette 1 and 2 (top left). Result (a) was generated using an alpha mask which segments foreground from background, (b) was generated using a mask that was created taking into account the 3D face information, and (c) was generated using a mask that varies depending on the luminance of the original image.
Source: Images are sourced from pixabay.com

3.4. Parallel computing of recolouring step

One of the main advantages of this method is the fact that our transformation is controlled by a parameter θ and any pixel x can be recoloured by computing the new colour value $\phi_\theta(x)$. This computation can be done in parallel by distributing the pixels to be recoloured to the multiple processors that are available. Moreover, with one target image and N palette images of choice, the transformations $\{\hat{\theta}_1, \ldots, \hat{\theta}_N\}$ can also be estimated in parallel, and any interpolated new value θ can be used for recolouring:

$$\theta_{new} = \sum_{j=1}^{N} \gamma_j \, \hat{\theta}_j, \quad \text{with} \sum_{j=1}^{N} \gamma_j = 1, \text{and } \gamma_j \geq 0 \; \forall j \tag{14}$$

Interpolating in the θ space to create a new warping function is made easy thanks to the fact that we chose control points on a regular grid and not as sub samples of palette images. Of particular interest is the creation of smooth temporal transitions between the identity warping function and a colouring warping function for instance.

Once estimated, the parametric transformation can be easily applied to video content (Grogan and Dahyot, 2015; Frigo et al., 2015b). Interpolating between two parametric transformations $\hat{\theta}_1$ and $\hat{\theta}_2$ can also be used to create interesting effects in images.

The interpolating weight can also vary over time $\gamma(t) \in [0; 1]$ to create dissolve effects between colour palettes in videos. In this case, for each colour pixel x at pixel spatial location $p = (p_1, p_2)$ in the sequence at time t, the recolouring is computed as (Grogan and Dahyot, 2015):

$$x(p,t) \rightarrow \phi\left(x, \gamma(t)\,\hat{\theta}_1 + (1-\gamma(t))\,\hat{\theta}_2\right) \tag{15}$$

We extend this idea further using interpolation weights that vary spatially as well as in the temporal domain where pixels are recoloured as follows:

$$x(p,t) \rightarrow \phi\left(x, \gamma(p,t)\,\hat{\theta}_1 + (1-\gamma(p,t))\,\hat{\theta}_2\right) \tag{16}$$

$\gamma(p,t)$ is a dynamic grey scale mask with values between 0 and 1.

Fig. 2 presents some examples of results on images.[1] In this figure, pixels in the target image whose corresponding pixel in the grey scale mask is black have been recoloured using palette 1, and those whose corresponding pixel is white have been recoloured using palette 2. The remaining pixels have been recoloured using an interpolation between the two transformations, $\hat{\theta}_1$ and $\hat{\theta}_2$. For each pixel this interpolation is determined by the value of its corresponding pixel in the grey scale mask. Fig. 2 highlights the different effects that can be easily created using the same palettes as input, but by changing the mask used for interpolation. The mask in Fig. 2(a) is an alpha matte that segments background from foreground. In (b), the mask was generated by first fitting a 3D morphable model to the face in the image (Bas et al., 2016; Ramanan, 2012), and as the direction of the normal vectors of the 3D face change, the colours in the mask change from white to black. The background of this mask also has a fade effect from left to right. The mask in (c) was generated by first converting the target image to grey scale and then editing the brightness and contrast of this image.

These effects can be extended to mixing more than 2 transformations (or palettes). A simpler interpolation between the identity transformation $\theta^{(Id)}$ and an estimated transformation $\hat{\theta}$ with a selected colour palette can also be created:

$$x(p,t) \rightarrow \phi\left(x, \gamma(t)\,\hat{\theta}^{(Id)} + (1-\gamma(t))\,\hat{\theta}\right) \tag{17}$$

This gives the user the flexibility to transition smoothly from one colour mood to another.

Having a quick recolouring step is essential to give the user instant feedback about the new effects being applied to the image or video. Our transformation can be applied to each pixel independently and it is therefore highly parallelisable. A CPU or GPU parallel implementation would ensure that the target image is recoloured almost instantly. For our implementation we parallelised the recolouring step on the CPU using OpenMP, and performance is assessed in Section 4.4.

4. Experimental results

Table 1 summarises some state of the art techniques compared to our method in different scenarios. In this section we present further qualitative and quantitative comparisons between our technique and these methods. To quantitatively assess recolouring results, three metrics – peak signal to noise ratio (PSNR), structural similarity index (SSIM) and colour image difference (CID) – are often used when considering palette and target images of the same content for which correspondences are easily available (Lissner et al., 2013; Oliveira et al., 2015; Hwang et al., 2014; Bellavia and Colombo, 2018). Alternatively user studies have also been used to assess the perceptual visual quality of the recolouring (Hristova et al., 2015). We evaluate our proposed algorithms and show that they are comparable to current state of the art colour transfer

[1] Supplementary results for videos are provided by the authors.

Fig. 3. Results on images with similar content on the 'playground', 'mart', 'illum', 'tonal4' and 'gangnam1' images. On close inspection grainy artifacts can be seen appearing in some PMLS results (Row 9) - for example, around the car in Column 3 or in the top right corner of the image in Column 2. In comparison, the results generated by TPS^{Corr}_{rgb} remain smooth (Row 10). For zoom see Fig. 4.

algorithms in terms of the perceptual quality of the results (Sections 4.1–4.3), and superior in terms of computational speed (Section 4.4). We use colour values represented in the RGB space, but further comparison between LAB and RGB colour spaces can be found in Grogan and Dahyot (2017). We use TPS^{KM}_{rgb} to notate the version of our colour transfer algorithm that uses K-Means to estimate the Gaussian Mixture model

means (Section 3.2.1), and TPS_{rgb}^{Corr} to notate the version that takes into account pixel correspondences instead (Section 3.2.2). The code we used to implement our method has also been made available online.[2]

4.1. Images with similar content $P \simeq T$

One important application of colour transfer is in harmonising the colour palette of several images or videos capturing the same scene. To evaluate our algorithm applied to images with similar content, we use the 15 images in the dataset provided by Hwang et al. (2014)[3] which includes images with many different types of colour changes including different illuminations, different camera settings and different colour touch up styles. This dataset provides palette images which have been aligned to match the target image. To define correspondences, pixels at the same location in the target and aligned palette images are selected together to form a pair. Note that results using PMLS were provided by the authors (Hwang et al., 2014).

Table 2 provides a quantitative evaluation of our proposed methods in comparison to leading state of the art colour transfer methods, and Fig. 3 presents results for visual comparison (c.f. notations explained in Table 1). Here, we compare algorithms that account for correspondences (in green) and those that do not (in blue). We found that as the warping functions proposed by Park and Xia are controlled by a small number of parameters, these methods fail to successfully correct more non-linear colour differences between images (see Columns 4 and 5 of Fig. 3), which is also reflected in their low quantitative scores in Table 2. Optimal transport techniques such as those proposed by Pitié, Bonneel[4] and Ferradans, do not account for colour correspondences between images. When the content of the target and palette images are practically identical, OT methods perform reasonably well as they try to match the colour distributions of the target and palette images exactly, but similar to our K-means approach TPS_{rgb}^{KM}, they cannot compete with correspondence based techniques such as PMLS or our TPS_{rgb}^{Corr} approach, which perform significantly better in terms of all three metrics (Table 2).

While PMLS performs best in terms of the PSNR, SSIM and CID metrics, the accuracy of these metrics relies on the fact that the input images are registered correctly. If this is not the case, as in Fig. 4, these metrics may not accurately capture all artifacts. In fact, on visual inspection, it is evident that the PMLS technique introduces visual artifacts when registration errors are present, as can be seen in Fig. 4. Their pixel by pixel least squares recolouring approach causes neighbouring pixels to differ a lot after recolouring when outliers are present. While PMLS is not robust to registration errors, our algorithm indeed is thanks to the robust $\mathcal{L}_2$ distance. Our approach allows us to maintain the structure of the original image and to create a smooth colour transfer result (cf. Fig. 4 for comparison). So to summarise, while PMLS and our TPS_{rgb}^{Corr} algorithm provide similar quantitative performances as measured by PSNR, SSIM and CID, TPS_{rgb}^{Corr} in fact provides better qualitative visual results. (Additional qualitative and quantitative results on all 15 images can be found in the supplementary material.)

4.2. Images with different content $P \neq T$

We compare our algorithm with other optimal transport based colour transfer techniques (Bonneel et al., 2015; Ferradans et al., 2013; Pitié et al., 2007) applied to images of different content and without correspondences. In the case of Ferradans et al. (2013), all images were generated using the parameters $\lambda_X = \lambda_Y = 10^{-3}$ and $\kappa = (0.1, 1, 0.1, 1)$. Fig. 5 shows that Bonneel and Ferradans methods can create blocky artifacts in the result image gradient (Row 2, 6, 7). On the other hand, the gradient smoothing step added to Pitié's algorithm ensures that these errors do not appear in their results, creating images that are more visually pleasing. However, matching the exact colour distributions of the images can still create unwanted colour changes in Pitié's results when the distributions are very different (Rows 9–11). Ferradans technique, which relaxes the OT constraint forcing the colour distributions to match exactly after recolouring, suffers less from these unexpected colour changes (row 7, 10), although the colours in their results can be less similar to those in the palette image (Column 4, row 8, 10, 11). The colours appearing in Bonneel's results can also differ from those in the palette (Row 10–11) as this method only alters the chrominance values of the target image, leaving the luminance values unchanged. On the other hand, our algorithm produces results that match the colours in the palette image well, and the thin plate spline transformation ensures that the image gradient remains smooth, even when the colour distributions of the target and palette images are very different (Rows 10–11).

Table 2
Comparison of our algorithms TPS_{rgb}^{Corr} and TPS_{rgb}^{KM} against state of the art techniques. Mean and standard error results for the SSIM, PSNR and CID metrics computed for recoloured images with similar content: Highest PSNR and SSIM values indicate the best results, while lowest CID values indicate the best results. The best techniques are highlighted in red.

	μ_{PSNR}	SE_{PSNR}	μ_{SSIM}	SE_{SSIM}	μ_{CID}	SE_{CID}
Ferradans	22.13	0.8	0.877	0.02	0.369	0.024
Bonneel	20.07	1.1	0.838	0.03	0.377	0.024
Pitié	25.85	0.9	0.899	0.02	0.330	0.023
TPS_{rgb}^{KM}	24.20	0.8	0.908	0.02	0.325	0.024
Park	23.71	0.7	0.901	0.02	0.367	0.023
Xia	22.13	0.8	0.876	0.02	0.471	0.048
PMLS	30.82	1.5	0.946	0.02	0.154	0.024
TPS_{rgb}^{Corr}	30.30	1.5	0.944	0.024	0.172	0.02

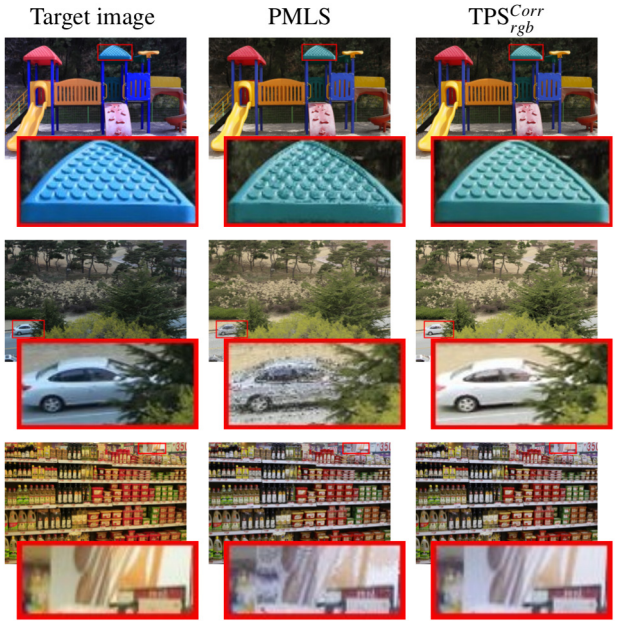

Fig. 4. A close up look at some of the errors generated using the PMLS algorithm (Hwang et al., 2014) in comparison to our smooth result with TPS_{rgb}^{Corr}. The structure of the original target image (left) is altered after recolouring using PMLS but not TPS_{rgb}^{Corr}.

[2] https://github.com/groganma/gmm-colour-transfer.
[3] https://sites.google.com/site/unimono/pmls.
[4] Results for Bonneel were generated using their code and exclude a gradient smoothing postprocessing step.

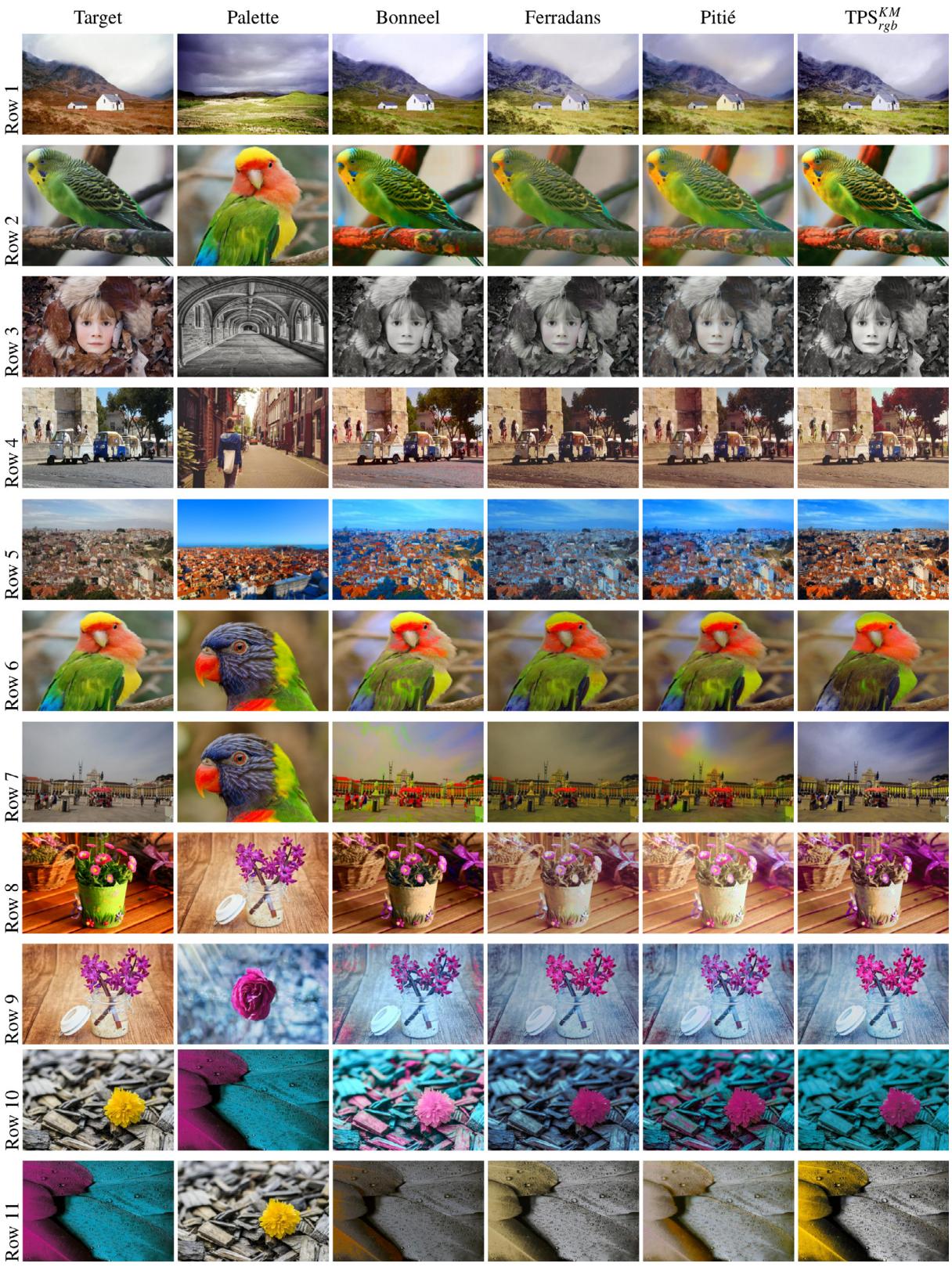

Fig. 5. Comparison with state of the art techniques with images ($P \neq T$).

4.3. Qualitative assessment

Colour transfer methods have also been evaluated using two subjective user studies, each with 20 participants evaluating 53 sets of images. In each experiment the participants were asked to choose the colour transfer result that they thought was the most successful. Out of these 53 sets of images, 38 of them had a palette and target image with different content (no correspondences available), and 15 of them contained a palette and target image with the same content (correspondences available and used). These 15 images were taken from the dataset provided

by Hwang et al. (2014) (cf. Section 4.1). For both user studies, each participant evaluated the results individually. The display properties and indoor conditions were kept constant for all participants. The order in which the image sets appeared was randomised for each participant and a short trial run took place before each user study to ensure the users adapted to the task at hand.

4.3.1. User study I

In our first experiment, each participant was presented with six images at a time - a target image, a palette image, and four result images generated using different colour transfer techniques. They then had 20 s to view the images (presented simultaneously side by side) and decide which result image was the best.

For target and palette images with similar content the four methods compared were TPS^{Corr}_{rgb}, PMLS, Pitié and Bonneel (For ease of comparison, we chose only four methods, including TPS^{Corr}_{rgb}, PMLS and Pitié, the three top methods according to the quantitative results (see Table 2). The total number of times each method was chosen can be seen in Table 3. Converting these totals to proportions, we used the z-test to compute the confidence interval on each proportion and compared methods by determining if their confidence intervals overlap. We found that TPS^{Corr}_{rgb} was the best method in terms of votes however the overlapping confidence intervals indicate that TPS^{Corr}_{rgb} and PMLS are not statistically significantly different. Hence these two methods, PMLS and ours, can be thought of as performing equally well and both being superior to Pitié and Bonneel methods.

Similarly for images with different content we compared TPS^{KM}_{rgb}, Ferradans, Pitié and Bonneel. The total number of times each method was chosen can be seen in Table 3. Again, by comparing the confidence interval on the proportions we determined that TPS^{KM}_{rgb} performs better than both Ferradans and Bonneel, however there is no statistically significant difference between TPS^{KM}_{rgb} and Pitié.

4.3.2. User study II

For our second experiment we used a 2 alternate forced choice comparison. In this case each participant was presented with two images at a time - a target image, a palette image, and two result images generated using different colour transfer techniques. They then had 15 s to view the images and decide which result image was the best. For target and palette images with similar content we compared the two methods that performed best in the previous user study — TPS^{Corr}_{rgb} and PMLS, and the total number of times each method was chosen can be seen in Table 4. Again we converted these totals to proportions and computed the confidence interval on each, and determined that there was no significant difference between the two. We then used the Thurstone Case V analysis to compute preference scores and confidences intervals for each method, as seen in Table 4, and in this case found that the confidence intervals do not overlap, indicating a significant difference between the two methods.

Similarly for images with different content we compared the two techniques that performed best in the previous user study, TPS^{KM}_{rgb} and Pitié, the results of which can be seen in Table 4. In this case both methods were chosen an equal number of times, indicating again that there is no perceivable difference between the results of TPS^{KM}_{rgb} and Pitié.

4.4. Computation time

Our algorithm is split into three parts: the clustering step, the estimation step of θ and the recolouring step $x \rightarrow \phi(x, \theta)$. As K-means can be quite time consuming, we investigated some fast quantisation methods including those provided by Matlab and the GNU Image Manipulation Program (GIMP). We found that using Matlab's minimum variance quantisation method (MVQ) provided almost identical results to K-means as well as being much faster and is used as an alternative to K-means to speed up the clustering step (Table 5).

Table 3
Number of votes given to each method by participants in our first perceptual study, their corresponding proportion and confidence interval. According to votes, the best method is indicated in red, second best in green, third best in blue.

User Study I

Similar Content $P \simeq T$			
Method	Votes	Prop	CI on Prop
Bonneel	38	0.127	(0.089, 0.164)
Pitié	63	0.210	(0.164, 0.256)
PMLS	98	0.327	(0.274, 0.380)
TPS^{Corr}_{rgb}	101	0.337	(0.283, 0.390)
Total	300	1	
Different Content $P \neq T$			
Method	Votes	Prop	CI on Prop
Bonneel	163	0.215	(0.185, 0.244)
Pitié	211	0.278	(0.246, 0.310)
Ferradans	152	0.20	(0.172, 0.228)
TPS^{KM}_{rgb}	234	0.308	(0.275, 0.341)
Total	760	1	

Table 4
Number of votes given to each method by participants in our second perceptual study, corresponding proportion and confidence interval, Thurstone score and associated confidence interval.

User Study II

Similar Content $P \simeq T$					
Method	Votes	Prop	CI on Prop	Score	CI on score
PMLS	134	0.447	(0.390, 0.503)	-0.19	(-0.30, -0.08)
TPS^{Corr}_{rgb}	166	0.553	(0.497, 0.610)	0.19	(0.08, 0.30)
Total	300	1			
Different Content $P \neq T$					
Method	Votes	Prop	CI on Prop	Score	CI on score
Pitié	380	0.5	(0.443, 0.557)	0	(-0.11, 0.11)
TPS^{KM}_{rgb}	380	0.5	(0.443, 0.557)	0	(-0.11, 0.11)
Total	760	1			

Table 5
Computation times in seconds for each step of our algorithm for a HD (1080×1920) image with over 2 million pixels. For the clustering step, the images were first downsampled to 300×350 pixels to reduce computation time.

Clustering		**Estimating $\hat{\theta}$**		**Recolouring**	
MVQ	0.005s	K = 50	6s	TPS	1.04s

The recolouring step on the other hand is highly parallelisable and can be applied independently to each pixel. Our implementation uses OpenMP within a Matlab mex file to parallelise this step on 8 CPU threads. All times are computed on a 2.93 GHz Intel CPU with 3GB of RAM with 4 cores and 8 logical processors. In comparison, the GPU implementation of PMLS takes 4.5 s to recolour a 1 million pixel image using an nVIDIA Quadro 4000 as reported in Hwang et al. (2014), which

is 9 times slower than our implementation with TPS. Similarly, Bonneel et al. (2013) report a time of approximately 3 min to recolour 108 frames of a HD segmented video on an 8 core machine with their algorithm. In comparison, our algorithm would take less than 2 min using TPS in the same situation.

5. $\mathcal{L}_2$ vs. OT for colour transfer

In this paper we have shown that for colour transfer applications, one of the main advantages of using the proposed $\mathcal{L}_2$ framework over OT techniques is its flexibility and its ability to use correspondences. Indeed OT cannot take correspondences into account when computing its solution, whereas our $\mathcal{L}_2$ framework can seamlessly take advantage of such correspondences when they are available. Such correspondences occur naturally in practice in a wide range of applications (e.g. photo stitching, video recolouring) when target and source images capture the same scene (i.e. images with similar content) and colour transfer offers a natural solution for harmonising their colours. We have also recently demonstrated that this $\mathcal{L}_2$ framework caters for user defined correspondences (i.e. for images of same or different content, or using a restricted colour palette) allowing user interaction during the recolouring process (Grogan et al., 2017).

When images have very different content, we have shown that by matching the true colour distributions of the target and palette images exactly, OT methods cause unexpected colour changes to appear. Our chosen GMM modelling framework prevents this from happening since, rather than capturing the true distribution of the transformed image, it approximates it. With equals weights and covariance matrices, our transformed GMM gives less emphasis to the spread of colours associated with each Gaussian component, and models only how the TPS transformation effects the colours captured by the Gaussian means. In this way, our $\mathcal{L}_2$ framework matches the colour content of the images, rather than the true colour distribution, giving improved results even in comparison to other OT methods that also propose approximating the colour distributions (Ferradans et al., 2013). Note that kernels other than Gaussian have been successfully used with $\mathcal{L}_2$ for shape registration (Grogan and Dahyot, 2018) and our solution for colour transfer with $\mathcal{L}_2$ is in principle not limited to using GMMs. Finally, constraining GMMs to lie within boundary values of the colour space (e.g. 0 and 255) can easily be enforced in the proposed $\mathcal{L}_2$ based framework (Grogan et al., 2017) preventing saturation from occurring.

Frigo et al. (2015b) have suggested fitting a TPS function to the estimated OT solution as a post process, which would create a smoother transfer function for OT and also removes the OT generated blocky artifacts. This step however may require robust fitting of TPS to deal with discontinuous mapping in the OT solution. The strength of the TPS regularisation term that controls how close or far the TPS solution is from the original OT solution, would also need to be carefully chosen. In contrast, $\mathcal{L}_2$ is naturally robust and our GMM formulation tackles class imbalance for images of different content, both are contributing factors in our approach for fitting a good TPS warping function.

In terms of computation time (see Table 6), our TPS_{rgb}^{KM} technique is slightly faster than Pitié when applied to a 1 million pixel image, while Bonneel and Ferradans are more computationally intensive. The time taken for our TPS_{rgb}^{Corr} approach increases linearly as the number of correspondences increases, and so depends on the accuracy required by the user — more correspondences will create a more accurate colour correction result, but will take longer to compute.

6. Conclusion

We have presented a new framework for colour transfer that is able to take into account correspondences when these are available. Our algorithms compete very well with current state of the art approaches. We have shown that TPS_{rgb}^{Corr}, our correspondence based approach, performs similarly in terms of the SSIM, PSNR and CID metrics when compared with the current top techniques, and visual inspection of our results shows that our algorithms are more immune to artifacts that can be created in the gradient field of the recoloured image by other methods (Hwang et al., 2014; Ferradans et al., 2013; Bonneel et al., 2015). When correspondences are not available, our $\mathcal{L}_2$ formulation creates good results even when the colour distribution of the target and palette images are very different, unlike other state of the art OT techniques (Pitié et al., 2007; Bonneel et al., 2015). Our TPS formulation of the transfer function also ensures that the gradient of the recoloured images stay smooth, removing the need to add an additional smoothing step typically required by OT techniques (Pitié et al., 2005; Bonneel et al., 2015; Papadakis et al., 2011). Our parametric transfer function also allows for fast recolouring of images and videos. Moreover transfer functions can be stored and easily combined and interpolated for creating visual effects. This $\mathcal{L}_2E$ framework has been extended further to let users enforce correspondences and interactively refine the recolouring (Grogan et al., 2017).

Table 6
Time taken for each method to recolour a 1 million pixel image using their Matlab implementations on a 4.0 GHz Intel CPU with 16 GB of RAM with 4 cores and 8 logical processors. TPS_{rgb}^{Corr} is applied with 10 000 pixel correspondences.

Ferradans	Bonneel	Pitié	TPS_{rgb}^{KM}	TPS_{rgb}^{Corr}
98s	170s	12s	7s	16s

Acknowledgements

This publication has emanated from research supported by an Ussher Scholarship from Trinity College Dublin, Ireland and by a research grant from Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776.

Appendix A. Supplementary data

Supplementary material related to this article, including additional quantitative and qualitative comparisons between our methods and other state of the art techniques, can be found online at http://dx.doi.org/10.1016/j.cviu.2019.02.002.

References

Arellano, C., Dahyot, R., 2012. Shape model fitting algorithm without point correspondence. In: 20th European Signal Processing Conference (Eusipco). Bucharest, Romania, pp. 934–938.

Bas, A., Smith, W.A.P., Bolkart, T., Wuhrer, S., 2016. Fitting a 3D Morphable Model to Edges: A Comparison Between Hard and Soft Correspondences, ArXiv e-prints abs/1602.01125.

Bellavia, F., Colombo, C., 2018. Dissecting and reassembling color correction algorithms for image stitching. IEEE Trans. Image Process. 27 (2), 735–748. http://dx.doi.org/10.1109/TIP.2017.2757262.

Bonneel, N., Peyré, G., Cuturi, M., 2016. Wasserstein Barycentric coordinates: histogram regression using optimal transport. ACM Trans. Graph. 35 (4), 71:1–71:10. http://dx.doi.org/10.1145/2897824.2925918, URL http://doi.acm.org/10.1145/2897824.2925918.

Bonneel, N., Rabin, J., Peyr, G., Pfister, H., 2015. Sliced and radon wasserstein barycenters of measures. J. Math. Imag. Vis. 51 (1), 22–45. http://dx.doi.org/10.1007/s10851-014-0506-3, URL http://dx.doi.org/10.1007/s10851-014-0506-3.

Bonneel, N., Sunkavalli, K., Paris, S., Pfister, H., 2013. Example-based video color grading. ACM Trans. Graph. 32 (4), 39:1–39:12. http://dx.doi.org/10.1145/2461912.2461939, URL http://doi.acm.org/10.1145/2461912.2461939.

Chui, H., Rangarajan, A., 2003. A new point matching algorithm for non-rigid registration. Comput. Vis. Image Underst. 89 (2), 114 – 141. http://dx.doi.org/10.1016/S1077-3142(03)00009-2, URL http://www.sciencedirect.com/science/article/pii/S1077314203000092.

Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Machine Intell. 24 (5), 603–619. http://dx.doi.org/10.1109/34.1000236.

Courty, N., Flamary, R., Tuia, D., Rakotomamonjy, A., 2017. Optimal transport for domain adaptation. IEEE Trans. Pattern Anal. Machine Intell. 39 (9), 1853–1865. http://dx.doi.org/10.1109/TPAMI.2016.2615921.

Escolano, F., Suau, P., Bonev, B., 2009. Information theory in Computer Vision and Pattern Recognition. Springer.

Faridul, H.S., Pouli, T., Chamaret, C., Stauder, J., Tremeau, A., Reinhard, E., 2014. A survey of color mapping and its applications. In: Lefebvre, S., Spagnuolo, M. (Eds.), Eurographics 2014 - State of the Art Reports. The Eurographics Association, http://dx.doi.org/10.2312/egst.20141035.

Ferradans, S., Papadakis, N., Rabin, J., Peyr, G., Aujol, J.F., 2013. Regularized discrete optimal transport. In: Kuijper, A., Bredies, K., Pock, T., Bischof, H. (Eds.), Scale Space and Variational Methods in Computer Vision. In: Lecture Notes in Computer Science, vol. 7893, Springer Berlin Heidelberg, pp. 428–439. http://dx.doi.org/10.1007/978-3-642-38267-3_36.

Freedman, D., Kisilev, P., 2010. Object-to-object color transfer: optimal flows and smsp transformations. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. pp. 287–294. http://dx.doi.org/10.1109/CVPR.2010.5540201.

Frigo, O., Sabater, N., Delon, J., Hellier, P., 2015a. Motion driven tonal stabilization. In: 2015 IEEE International Conference on Image Processing (ICIP). pp. 3372–3376. http://dx.doi.org/10.1109/ICIP.2015.7351429.

Frigo, O., Sabater, N., Demoulin, V., Hellier, P., 2015b. Optimal transportation for example-guided color transfer. In: Cremers, D., Reid, I., Saito, H., Yang, M.H. (Eds.), Computer Vision – ACCV 2014. In: Lecture Notes in Computer Science, vol. 9005, Springer International Publishing, pp. 655–670. http://dx.doi.org/10.1007/978-3-319-16811-1{_}43, URL http://dx.doi.org/10.1007/978-3-319-16811-1{_}43.

Grogan, M., Dahyot, R., 2015. L2 registration for colour transfer in videos. In: Proceedings of the 12th European Conference on Visual Media Production. ACM, New York, NY, USA, http://dx.doi.org/10.1145/2824840.2824862, 16:1–16:1.

Grogan, M., Dahyot, R., 2017. Robust Registration of Gaussian Mixtures for Colour Transfer, ArXiv e-prints URL http://arxiv.org/abs/1705.06091.

Grogan, M., Dahyot, R., 2018. Shape Registration with Directional Data. Pattern Recognition, URL https://www.sciencedirect.com/science/article/pii/S003132031830075X, https://doi.org/10.1016/j.patcog.2018.02.021.

Grogan, M., Dahyot, R., Smolic, A., 2017. User interaction for image recolouring using L2. In: Proceedings of the 14th European Conference on Visual Media Production (CVMP 2017). ACM, New York, NY, USA, pp. 6:1–6:10. http://dx.doi.org/10.1145/3150165.3150171, URL http://doi.acm.org/10.1145/3150165.3150171.

Grogan, M., Prasad, M., Dahyot, R., 2015. L2 registration for colour transfer. In: Proceedings of the 23rd European Signal Processing Conference (EUSIPCO). Nice, France, pp. 2366–2370.

HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2011. Non-rigid dense correspondence with applications for image enhancement. ACM Trans. Graph. 30, 70:1–70:10. http://dx.doi.org/10.1145/2010324.1964965, URL http://doi.acm.org/10.1145/2010324.1964965.

HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2013. Optimizing color consistency in photo collections. ACM Trans. Graph. 32, 38:1–38:10. http://dx.doi.org/10.1145/2461912.2461997, URL http://doi.acm.org/10.1145/2461912.2461997.

He, M., Liao, J., Yuan, L., Sander, P.V., 2017. Neural Color Transfer between Images, ArXiv e-prints abs/1710.00756 URL http://arxiv.org/abs/1710.00756.

Hristova, H., Le Meur, O., Cozot, R., Bouatouch, K., 2015. Style-aware robust color transfer. In: Proceedings of the Workshop on Computational Aesthetics. In: CAE '15, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 67–77, URL http://dl.acm.org/citation.cfm?id=2811239.2811244.

Hwang, Y., Lee, J.Y., Kweon, I.S., Kim, S.J., 2014. Color transfer using probabilistic moving least squares. In: Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. pp. 3342–3349. http://dx.doi.org/10.1109/CVPR.2014.427.

Jeong, K., Jaynes, C., 2007. Object matching in disjoint cameras using a color transfer approach. Machine Vis. Appl. 19, 443–455. http://dx.doi.org/10.1007/s00138-007-0079-x, URL http://dx.doi.org/10.1007/s00138-007-0079-x.

Jian, B., Vemuri, B., 2011. Robust point set registration using gaussian mixture models. IEEE Trans. Pattern Anal. Machine Intell. 33, 1633–1645. http://dx.doi.org/10.1109/TPAMI.2010.223.

Liao, J., Yao, Y., Yuan, L., Hua, G., Kang, S.B., 2017. Visual attribute transfer through deep image analogy. ACM Trans. Graph. 36, 120:1–120:15. http://dx.doi.org/10.1145/3072959.3073683, URL http://doi.acm.org/10.1145/3072959.3073683.

Lissner, I., Preiss, J., Urban, P., Lichtenauer, M.S., Zolliker, P., 2013. Image-difference prediction: from grayscale to color. IEEE Trans. Image Process. 22, 435–446. http://dx.doi.org/10.1109/TIP.2012.2216279.

Luan, F., Paris, S., Shechtman, E., Bala, K., 2017. Deep Photo Style Transfer, ArXiv e-prints abs/1703.07511, URL http://arxiv.org/abs/1703.07511.

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 Trans. Signal Process. 63, 1115–1129. http://dx.doi.org/10.1109/TSP.2014.2388434.

Myronenko, A., Song, X., 2010. Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Machine Intell. 32, 2262–2275. http://dx.doi.org/10.1109/TPAMI.2010.46.

Neumann, L., Neumann, A., 2005. Color style transfer techniques using hue, lightness and saturation histogram matching. In: Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 111–122. http://dx.doi.org/10.2312/COMPAESTH/COMPAESTH05/111-122, URL http://dx.doi.org/10.2312/COMPAESTH/COMPAESTH05/111-122.

Oliveira, M., Sappa, A., Santos, V., 2015. A probabilistic approach for color correction in image mosaicking applications. IEEE Trans. mage Process. 24, 508–523. http://dx.doi.org/10.1109/TIP.2014.2375642.

Papadakis, N., Provenzi, E., Caselles, V., 2011. A variational model for histogram transfer of color images. IEEE Trans. Image Process. 20, 1682–1695. http://dx.doi.org/10.1109/TIP.2010.2095869.

Park, J., Tai, Y., Sinha, S.N., Kweon, I.S., 2016. Efficient and robust color consistency for community photo collections. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 430–438. http://dx.doi.org/10.1109/CVPR.2016.53.

Pitié, F., Kokaram, A., 2007. The linear monge-kantorovitch colour mapping for example-based colour transfer. In: IEEE European Conference on Visual Media Production (CVMP'07). London, UK.

Pitié, F., Kokaram, A., Dahyot, R., 2005. N-dimensional probability density function transfer and its application to color transfer. In: Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, Vol. 2. pp. 1434–1439. http://dx.doi.org/10.1109/ICCV.2005.166.

Pitié, F., Kokaram, A.C., Dahyot, R., 2007. Automated colour grading using colour distribution transfer. Comput. Vis. Image Underst. 107, 123–137. http://dx.doi.org/10.1016/j.cviu.2006.11.011, URL http://dx.doi.org/10.1016/j.cviu.2006.11.011.

Pouli, T., Reinhard, E., 2011. Progressive color transfer for images of arbitrary dynamic range. Comput. & Graph. 35, 67 – 80. http://dx.doi.org/10.1016/j.cag.2010.11.003, URL http://www.sciencedirect.com/science/article/pii/S009784931000172X, Extended Papers from Non-Photorealistic Animation and Rendering (NPAR) 2010.

Rabin, J., Delon, J., Gousseau, Y., 2011. Removing artefacts from color and contrast modifications. IEEE Trans. Image Process. 20, 3073–3085. http://dx.doi.org/10.1109/TIP.2011.2142318.

Ramanan, D., 2012. Face detection, pose estimation, and landmark localization in the wild. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, Washington, DC, USA, pp. 2879–2886, URL http://dl.acm.org/citation.cfm?id=2354409.2355119.

Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P., 2001. Color transfer between images. IEEE Comput. Graph. Appl. 21, 34–41. http://dx.doi.org/10.1109/38.946629.

Scott, D.W., 2001. Parametric statistical modeling by minimum integrated square error. Technometrics 43, 274–285, URL http://www.jstor.org/stable/1271214.

Tai, Y.W., Jia, J., Tang, C.K., 2005. Local color transfer via probabilistic segmentation by expectation-maximization. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1. pp. 747–754. http://dx.doi.org/10.1109/CVPR.2005.215.

Vazquez-Corral, J., Bertalmo, M., 2014. Color stabilization along time and across shots of the same scene, for one or several cameras of unknown specifications. IEEE Trans. Image Process. 23, 4564–4575. http://dx.doi.org/10.1109/TIP.2014.2344312.

Villani, C., 2003. Topics in Optimal Transportation, Vol. 58. American Mathematical Society.

Villani, C., 2009. Optimal transport : old and new. In: Grundlehren der mathematischen Wissenschaften. Springer, Berlin.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Yu, P.S., Zhou, Z.H., Steinbach, M., Hand, D.J., Steinberg, D., 2007. Top 10 algorithms in data mining. Knowl. Inf. Syst. 14, 1–37. http://dx.doi.org/10.1007/s10115-007-0114-2, URL http://dx.doi.org/10.1007/s10115-007-0114-2.

Xia, M., Renping, J.Y., Zhang, X.M., Xiao, J., 2017. Color consistency correction based on remapping optimization for image stitching. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). pp. 2977–2984. http://dx.doi.org/10.1109/ICCVW.2017.351.

Xiang, Y., Zou, B., Li, H., 2009. Selective color transfer with multi-source images. Pattern Recognit. Lett. 30, 682 – 689. http://dx.doi.org/10.1016/j.patrec.2009.01.004, URL http://www.sciencedirect.com/science/article/pii/S0167865509000191.

Xiao, X., Ma, L., 2009. Gradient-preserving color transfer. Comput. Graph. Forum 28, 1879–1886. http://dx.doi.org/10.1111/j.1467-8659.2009.01566.x, URL http://dx.doi.org/10.1111/j.1467-8659.2009.01566.x.

Xiao, X., Ma, L., Kunze, M., 2006. Object-based image recoloring using alpha matte and color histogram specification. In: Zha, H., Pan, Z., Thwaites, H., Addison, A., Forte, M. (Eds.), Interactive Technologies and Sociotechnical Systems. In: Lecture Notes in Computer Science, vol. 4270, Springer Berlin Heidelberg, pp. 194–203. http://dx.doi.org/10.1007/11890881{_}22, URL http://dx.doi.org/10.1007/11890881{_}22.

Xu, S., Zhang, Y., Zhang, S., Ye, X., 2005. Uniform color transfer. In: IEEE International Conference on Image Processing 2005. pp. III–940–3. http://dx.doi.org/10.1109/ICIP.2005.1530548.