

L2 REGISTRATION FOR COLOUR TRANSFER

Mairéad Grogan, Mukta Prasad & Rozenn Dahyot

School of Computer Science and Statistics
Trinity College Dublin Ireland

ABSTRACT

This paper proposes to perform colour transfer by minimising a divergence (the L2 distance) between two colour distributions. We propose to model each dataset by a compact Gaussian mixture which is designed for the specific purpose of colour transfer between images which have different scene content. A non rigid transformation is estimated by minimising the Euclidean distance (L2) between these two distributions, and the estimated transformation is used for transferring colour statistics from one image to another. Experimental results show that this is a very promising approach for transferring colour and it performs very well against an alternative reference approach.

Index Terms— Colour transfer, registration, L2, Gaussian Mixtures

1. INTRODUCTION

Adjusting colour statistics from one image to another is a common task in image processing. Applications range from image and video restoration, artistic and aesthetic after effects, and image corrections to facilitate further processing amongst others. Inspired by recent advances in non-rigid shape registration, we propose a new technique for performing colour transfer. The problem is formulated explicitly as estimating a transformation to register two colour distributions from two images (the target image and the palette image). The estimated transformation is non rigid and this transfer function is then applied to the colour values of each pixel in the target image to recolor it. Section 2 reviews both the topics of colour transfer and shape registration. Our method is presented in detail in section 3 and it is shown to compare very well against a leading method for colour transfer (section 4). Possible future directions of research to extend this approach are discussed in section 5.

This work has been supported by a Ussher scholarship from Trinity College Dublin (Ireland), and partially supported by EU FP7-PEOPLE-2013-IAPP GRAISearch grant (612334).

2. STATE OF THE ART

We first start with a review on both the topics of colour transfer (section 2.1) and shape registration (section 2.2) to highlight the connections between the two fields. We point out in section 2.3 the differences between our method and the current state of the art.

2.1. Colour transfer

To change the colour of a target image by mapping the colour palette of another image, Reinhard [3] proposes to first convert RGB values into another uncorrelated colour space, allowing the three axes to be treated independently. Then the means and variances along each of the three axes are registered such that the target image colour values are transformed to have the same mean and standard deviation as the source image.

Pitié et al. [1, 2, 4] proposed to perform non-rigid transformation of the target colour cumulative distribution onto the palette. Their algorithm computes successive solutions in one-dimensional spaces computed as a projection of the 3 dimensional colour space. This iterative process is shown to decrease the Kullback-Leiber divergence between the two colour density functions [1]. To avoid potential artifacts arising from the colour transfer process, constraints are then added to the colour transfer process to maintain the gradient field of the coloured image [1, 4].

In [5], the problem considered is to correct the color between two images which have the same content. Correspondences between pixels in the two images are assumed to be known and each colour channel is treated independently. While Oliveira et al. use pixel based correspondences between images which have the same content, Wu et al. [6] consider images with different content and compute high level semantic scene content correspondences in both images to constraint colour transfer.

2.2. Shape Registration

Jian et al. [7] propose to register two point sets (i.e. sets of vertices in 2D or 3D spaces), denoted $\{x_t^{(i)}\}_{i=1, \dots, n_t}$ and $\{x_p^{(i)}\}_{i=1, \dots, n_p}$, by first fitting a kernel density estimate to

each point set using the Gaussian kernel. The target distribution is formulated as [8]:

$$p_t(x|\theta) = \frac{1}{n_t} \sum_{i=1}^{n_t} \mathcal{N}(x; \mu_t(x_t^{(i)}, \theta), h^2\mathbf{I}) \quad (1)$$

with notation $\mathcal{N}(x; \mu, \Sigma)$ indicating a Normal distribution for random vector x , with mean μ and covariance Σ . The function $\mu_t(x_t^{(i)}, \theta)$ moves vertex $x_t^{(i)}$ and this displacement is controlled by a latent vector θ to be estimated. The transfer functions μ_t considered in this paper are two non-rigid transformations: affine and Thin Plate Splines (TPS) [7, 8]. The affine transformation has a small number of parameters to be estimated. The TPS model is more complex and can be decomposed into a linear part which is modelled by an affine motion, and a nonlinear part which is controlled by TPS warping coefficients [7, 8].

The second kernel density estimate is defined using the second point set as:

$$p_p(x) = \frac{1}{n_p} \sum_{i=1}^{n_p} \mathcal{N}(x; x_p^{(i)}, h^2\mathbf{I}). \quad (2)$$

Isotropic covariance matrices controlled by a bandwidth h (with \mathbf{I} a 3×3 identity matrix) are used for both distributions. The latent parameters θ for mapping the two Gaussian mixtures are estimated by minimising their L_2 distance [7, 9]:

$$\hat{\theta} = \arg \min_{\theta} \{L_2(\theta) = \|p_t - p_p\|^2\} \quad (3)$$

with notation $\|p_t - p_p\|^2 = \int (p_t(x|\theta) - p_p(x))^2 dx$. The advantage of computing L_2 over the Kullback Leiber divergence between probability density functions is that it can be computed explicitly with Gaussian mixtures and it is also more robust to outliers [7, 9]. Additive regularisation terms can be added to L_2 to constraint its estimation [7, 10].

2.3. Remarks

Like in Pitié et al. [2], we consider target and palette images that have different content and no constraint is imposed in our framework when minimising L_2 to estimate the transformation parameters θ . For colour transfer, a transfer function mapping the 3D colour space onto itself needs to be inferred. Many methods have been proposed to consider dimensions of the colour space separately, sometimes in an iterative manner. We propose to infer this transformation directly in the multidimensional colour space by formulating the problem as finding the transformation minimising a divergence between two probability density functions. The L_2 distance computed with Gaussian mixtures is chosen over Kullback Leiber to make the approach computationally tractable and more robust to any mismatch that can occur when comparing colour distributions from different images. A key novelty in our approach is in the choice of weights in the Gaussian mixtures to represent the colour distribution (see step 3 section 3).

3. PROPOSED APPROACH

Our method consists of the following steps:

1. **Downsampling.** To reduce the computation time involved in step 2, both images, target and palette, are downsampled to a lower resolution of 300×350 . This creates two very large ($300 \times 350 = 105000$) point clouds of colour RGB values in the RGB space to represent both the target and palette images.
2. **Clustering.** For each image, the low resolution image's RGB pixel values are clustered using the k-means algorithm in the 3 dimensional RGB space. We denote the computed centers as $\{c_t^{(i)}\}_{i=1, \dots, k}$ for the target image and $\{c_p^{(i)}\}_{i=1, \dots, k}$ for the palette image. The number k of clusters can be set by the user according to the quality required for the output of the colour transfer process. Note also that k controls the computational complexity $2k^2$ of the cost function L_2 (eq. (5)).
3. **Gaussian Mixture Models for Colour Transfer.** We propose to use the following Gaussian mixture models (GMMs):

$$\begin{cases} p_t(x|\theta) = \frac{1}{k} \sum_{i=1}^k \mathcal{N}(x; \mu_t(c_t^{(i)}, \theta), h^2\mathbf{I}) \\ p_p(x) = \frac{1}{k} \sum_{i=1}^k \mathcal{N}(x; c_p^{(i)}, h^2\mathbf{I}) \end{cases} \quad (4)$$

Here $x \in \mathbb{R}^3$ is a 3 dimensional random vector in the RGB space, \mathbf{I} is the 3×3 identity matrix and h is scalar value controlling the isotropic covariance. Note that in both GMMs all cluster centers are given equal weights $\frac{1}{k}$: the target and palette images have different content and therefore it is not expected that the number of pixels associated with any cluster has any correlation between the two images.

4. **Non-Rigid Registration.** The parameters θ of the transfer function μ_t (affine or TPS) are estimated by minimising L_2 between p_t and p_p (eq. (4)) using the following results [9]:

$$\begin{aligned} L_2(\theta) &= \|p_t\|^2 + \|p_p\|^2 - 2\langle p_t | p_p \rangle \text{ with} \\ \|p_t\|^2 &= \frac{1}{k^2} \sum_{i,j} \mathcal{N}(0; \mu_t(c_t^{(i)}, \theta) - \mu_t(c_t^{(j)}, \theta), 2h^2\mathbf{I}), \\ \langle p_t | p_p \rangle &= \frac{1}{k^2} \sum_{i,j} \mathcal{N}(0; \mu_t(c_t^{(i)}, \theta) - c_p^{(j)}, 2h^2\mathbf{I}). \end{aligned} \quad (5)$$

Note that the term $\|p_p\|^2$ does not depend on θ and can be ignored in the estimation [9].

5. **Re-colouring.** The estimated transfer function $\mu_t(x, \hat{\theta})$ is then applied to the RGB values x of the pixels in the original (full resolution) target image.

4. EXPERIMENTAL RESULTS

We compare our technique with Pitié et al. [2]. Figure 2 shows the target and palette images used in our experiments.

Figure 3 and 4 show the results of our implementation in comparison to Pitié et al. [2]. We present the results of our method with both the affine and TPS transfer mapping functions. We have set the number of clusters to $k = 150$ in our implementation to capture the variety in colour content of the images well, while keeping the registration computationally efficient. In some cases, the final transfer results are very stable for different values of k as can be seen in Figure 1. For all results

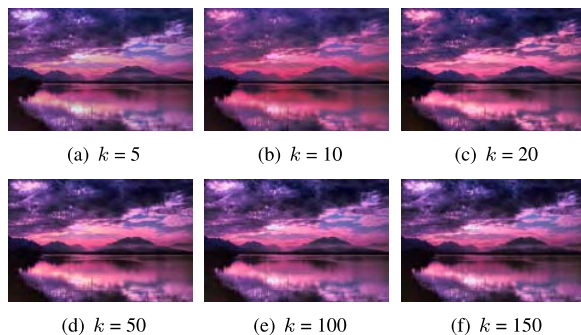


Fig. 1: Colour transfer results with our approach (TPS) for images 2(d) (palette) and 2(c) (target) with varying number of clusters k .

shown, there has been no constraint imposed to restrict transformed pixel values to lie in the interval $[0, 255]$. If a pixel has an R,G or B colour channel value greater than 255 we set it to 255. Similarly, if a pixel has an R,G or B value less than 0, we set it to 0. We do this for both our and Pitié et al.'s result images.

From Figure 3 and 4 we can see that the results generated using the TPS transformation model outperform those generated using the affine transformation. In general, our TPS colour transfer function provides smoother results with richer contrasts. In Figure 4, rows 2,4 and 6 show that our affine model fails to correctly transfer the colours from the palette image to the target. This occurs when some of the cluster centres chosen by the k means algorithm have RGB values very close to 0 or 255 (near the limits of the RGB cube in \mathbb{R}^3). All pixels that have a value close to the original cluster centre will be transformed in a similar way and may also be transformed to values lying outside of the interval $[0, 255]$. If this happens to a lot of pixels in the target image, the artifact seen in Figure 4 (second column, rows 2, 4 and 6) will arise. Having more flexibility in its modelling, our TPS transfer model performs better in these cases. In Figure 3 and 4 we can also see that in most cases the affine transformation model performs better than Pitié et al. However the artifacts seen in Figure 4 (rows 2 and 4) arise in some cases. These artifacts do not occur when the TPS model is used and this model outperforms Pitié et al in all cases. In Figure 3, rows 1 and 2, we can see that Pitié et al's algorithm suffers from grainy artifacts in some areas of the image such as the sky regions. The authors proposed a grain artifact reducer step to help overcome this

problem [1,4]. In comparison, our approach gives a smoother colour transfer result without needing any further processing.

5. CONCLUSIONS

We have shown that colour transfer between images with different content can be formulated as finding the non-rigid transformation minimising the Euclidean distance between two tailor made Gaussian mixtures in the Colour space. Future efforts will consider adding spatial constraint to simultaneously register images of the same scene, and simultaneously correct unwanted colour differences.

REFERENCES

- [1] F. Pitié, A. C. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer," *Computer Vision and Image Understanding Journal (Special Issue on Color Image Processing)*, 2007.
- [2] F. Pitié, A. C. Kokaram, and R. Dahyot, "N-dimensional probability density function transfer and its application to color transfer," in *Tenth IEEE International Conference on Computer Vision (ICCV 2005)*, 17-21 October 2005, vol. 2, pp. 1434 – 1439.
- [3] M. Ashikhmin, E. Reinhard and B. Gooch, "Color transfer between images," *Applied Perception*, September/October 2001.
- [4] F. Pitié, A. C. Kokaram, and R. Dahyot, "Towards automated colour grading," in *2nd European Conference on Visual Media Production (IEE CVMP 2005)*, London, November 2005.
- [5] M. Oliveira, A. D. Sappa, and V. Santos, "A probabilistic approach for color correction in image mosaicking applications," *IEEE Transactions on image Processing*, vol. 24, no. 2, February 2015.
- [6] F. Wu, W. Dong, Y. Kong, X. Mei, J.C. Paul, and X. Zhang, "Content-based colour transfer," *Computer Graphics Forum*, vol. 32, no. 1, pp. 190–203, 2013.
- [7] B. Jian and B. Vemuri, "Robust point set registration using gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633 – 1645, 2011.
- [8] B. Jian and B. C. Vemuri, "A robust algorithm for point set registration using mixture of gaussians," in *International Conference on Computer Vision (2005)*, 2005.
- [9] D. W. Scott, "Parametric statistical modeling by minimum integrated square error," *Technometrics*, vol. 43, no. 3, pp. pp. 274–285, 2001.
- [10] C. Arellano and R. Dahyot, "Shape model fitting algorithm without point correspondence," in *20th European Signal Processing Conference (Eusipco)*, Bucharest, Romania, August, 27-31 2012, pp. 934–938.



Fig. 2: These images were used as the target and palette images in our experiments. (Images sourced from: hqwallbase.com; wallpapercraft.com; afremov.com; restorationpath.org; helene-brennan.com; pl.forwallpaper.com)

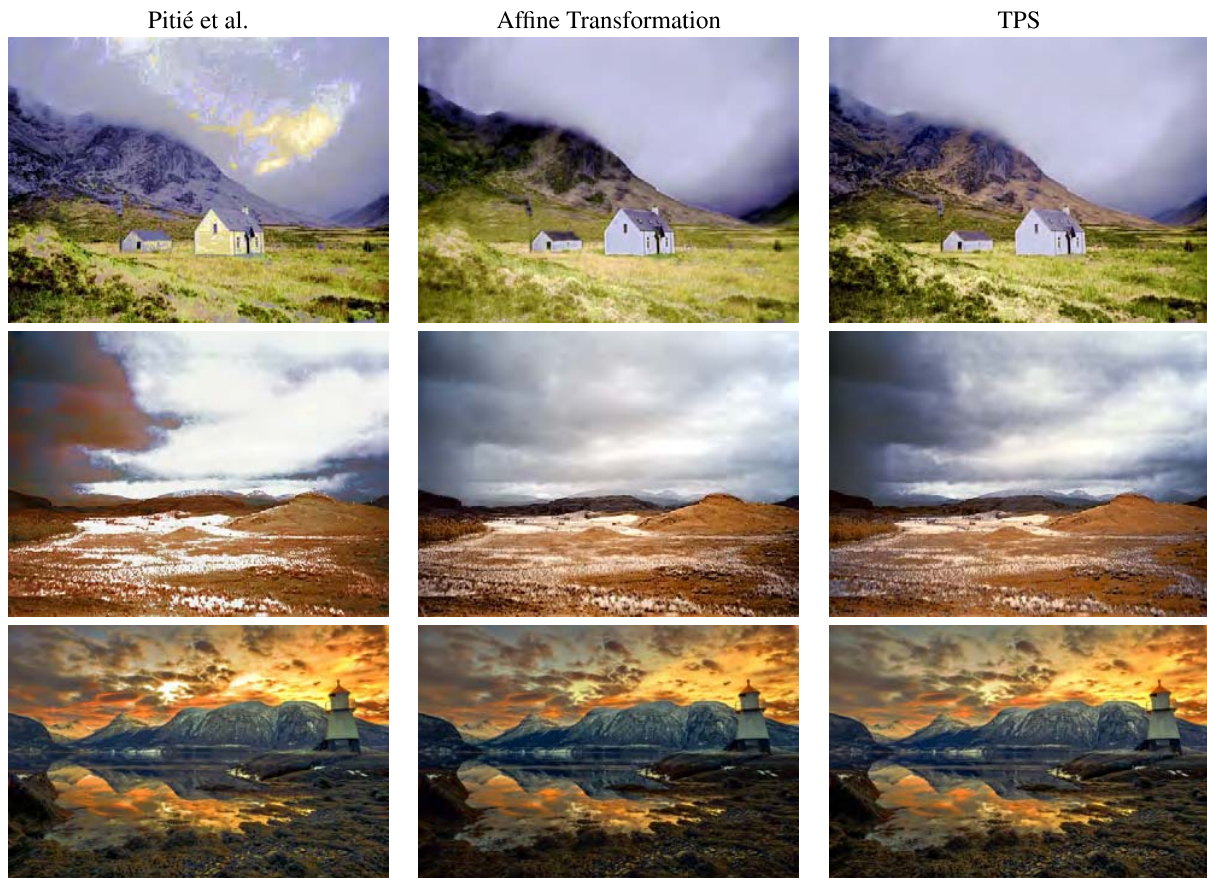


Fig. 3: The first column shows Pitié et al's results, the second column shows our results with the affine transformation model and the third column shows our results with the TPS model. From top to bottom, target images used are 2(a)(b)(d) and palette images used are 2(b)(a)(c). (More results can be found at www.scss.tcd.ie/Rozenn.Dahyot/ColourTransfer/).

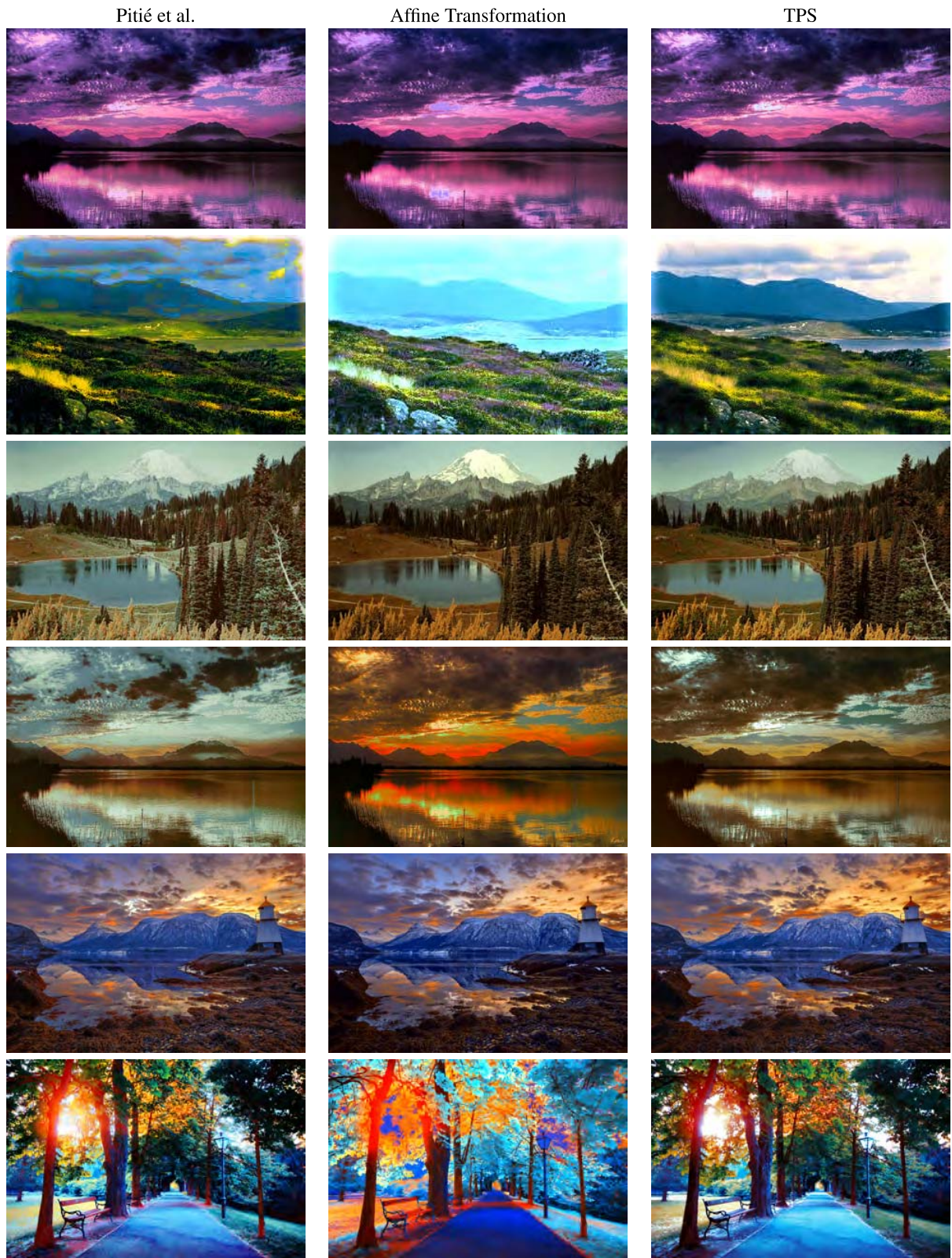


Fig. 4: The first column shows Pitié et al's results, the second column shows our results with the affine transformation model and the third column shows our results with the TPS model. From top to bottom, target images used are 2(c)(f)(e)(c)(d)(h) and palette images used are 2(d)(e)(f)(f)(g)(i). (More results can be found at www.secs.tcd.ie/Rozenn.Dahyot/ColourTransfer/).