Computational Visual Media
DOI 10.1007/s41095-xxx-xxxx-x Vol. x, No. x, month year, xx–xx

Research Article

Lighting Transfer across Multiple Views through Local Color Transforms

Qian Zhang[1] (✉), **Pierre-Yves Laffont**[2], **and Terence Sim**[3]

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract We present a method for transferring lighting across photographs of a static scene. Our method takes as input a photo collection depicting a scene under varying viewpoints and lighting conditions. We cast lighting transfer as an edit propagation problem, where the transfer of local illumination across images is guided by sparse correspondences obtained through multi-view stereo. Instead of directly propagating color, we learn local color transforms from corresponding patches in pairs of images and propagate these transforms in an edge-aware manner in regions with no correspondences. Our color transforms model the large variability of appearance changes in local regions of the scene, and are robust to missing or inaccurate correspondences. The method is fully automatic and can transfer strong shadows across images. We show applications of our image relighting method for enhancing photographs, browsing photo collections with harmonized lighting and generating synthetic timelapses.

Keywords relighting; photo collection; time-lapse; image editing.

1 Introduction

If there is one thing that can make or break a photograph, it is lighting. This is especially true for outdoor photography, as the appearance of a scene changes dramatically with the time of day. In order to capture the short, transient moments of interest, photographers have to wait at the right place for the perfect time of day. A majority of photographs taken by casual users are captured in the middle of the day, when lighting is not ideal. While photo retouching software such as Adobe Photoshop and Lightroom enable after-the-fact editing to some extent, achieving convincing manipulations such as drastic changes in lighting requires significant time and effort even for talented artists.

In this paper, we propose an automatic technique for transferring lighting across photographs, given a photo collection depicting the same scene under varying viewpoint and illumination. There are millions of photographs of famous landmarks at online photo-sharing websites, providing rich information for lighting transfer. For a pair of source and target images chosen by a user from the photo collection, our method modifies the source image by transferring the desired lighting from the target image. We model the large variability of appearance changes for different parts of the scene with local color transforms. The transforms are learned from sparse geometric correspondences, which we obtain from the photo collection through multi-view stereo. For regions without correspondences, we propagate the transforms in an edge-aware manner. Compared to direct color propagation, our propagation technique is robust to missing or inaccurate correspondences.

Our main contributions are as follows:

- We cast lighting transfer as an edit propagation problem, learning local color transforms from sparse geometric correspondences and propagating the transforms in an edge-aware manner.
- We introduce a confidence map to indicate the reliability of propagated transforms, which helps to preserve the color of pixels with transform outliers.
- We extend our method to transfer lighting based on multiple target images, exploiting the information from different viewpoints.

We run our method on 6 scenes, including 5 Internet

1 Nanyang Technological University, 639798, Singapore.
2 ETH Zurich, 8092 Zurich, Switzerland.
3 National University of Singapore, 119077, Singapore.
Manuscript received: 2017-03-31; accepted: 2017-05-29.

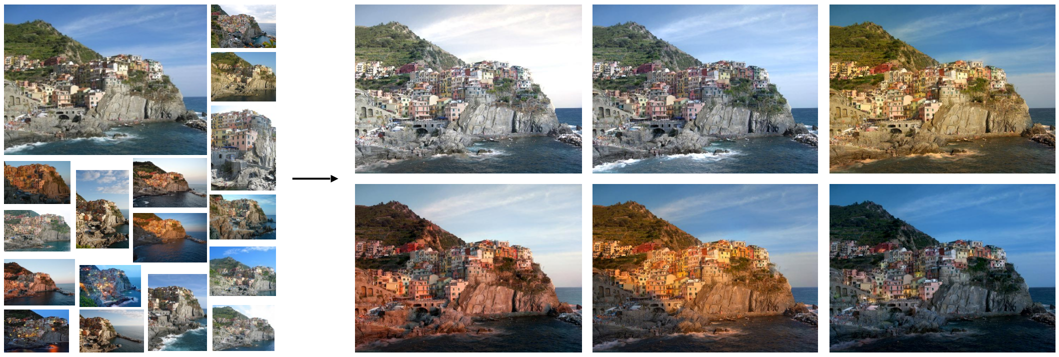


Fig. 1 Given a photo collection of a landmark scene under varying lighting, our method transfers the illumination between images from different viewpoints, synthesizing images with new combinations of viewpoint and time of day.

photo collections and a synthetic benchmark with ground truth images, which allows for quantitative evaluations. We also show comparisons with baselines and previous approaches. Our image relighting method enables enhancement of photographs, photo collections browsing with harmonized lighting and synthetic timelapse generation.

2 Related work

Color transfer and correction. Lighting transfer is mainly about color. Approaches for color transfer manipulate the color distributions. Example-based transfer methods such as [19–21] reshape the color distribution of the input image such that it approaches the color statistical properties of the example image. Huang et al. [6] recolor a photo by learning correlations between color property distributions and geometric features of regions from a database. Li et al. [12] recolor images using geodesic distance based on harmonization. More recently, Luan et al. [16] propose deep learning approach for photographic style transfer. These methods produce visually pleasing recolored images but cannot change the local lighting. Color transfer methods can also be used for tone adjustment and correction. Park et al. [18] recover sparse pixel correspondences and compute color correction parameters with a low-rank matrix factorization technique. Spatial-temporal correspondences are also used in [15, 28] for multi-view color correction. These methods work well for optimizing color consistency for image collections or videos, but they are not intended to transfer spatially-varying lighting. In our case, we use local transforms to model the large variability of appearance changes in local regions of the scene, which is able to transfer strong shadows.

Image relighting. A number of image relighting methods have been proposed by various researchers over the years, such as [3, 7, 29, 30]. These sophisticated systems make use of detailed geometric models and require registration or non-linear fitting. Laffont et al. [9] show that intrinsic image decomposition can be used for illumination transfer, but the extraction of consistent reflectance and illumination layers is a challenging and computationally expensive problem. Alternatively, some methods transfer an image by learning color changes from correspondences of image pairs. HaCohen et al. [5] compute a parametric color model based on dense correspondences, but do take into account local color changes. Shih et al. [24] successfully synthesize different time-of-day images by learning color transformations from time-lapse videos. A similar approach by Laffont et al. [10] enables appearance transfer of time of day, weather or season by observing color changes in webcam database. However, both methods rely on the availability of images of different appearance from the same webcam. While these image pairs may be available for some scenes with a static camera, this data does not exist in many cases. More recently, Martin-Brualla et al. [17] use a simple but effective new temporal filtering approach to stabilize appearance. In work developed concurrently, Shen et al. [23] propose regional foremost matching for image morphing and timelapse generation. In our system, we target a more general case that does not need highly accurate geometry, timelapses from a static viewpoint or densely computed correspondences. Our method relies on the vastly available images of the same scene from various online photo communities and the sparse geometric correspondences.

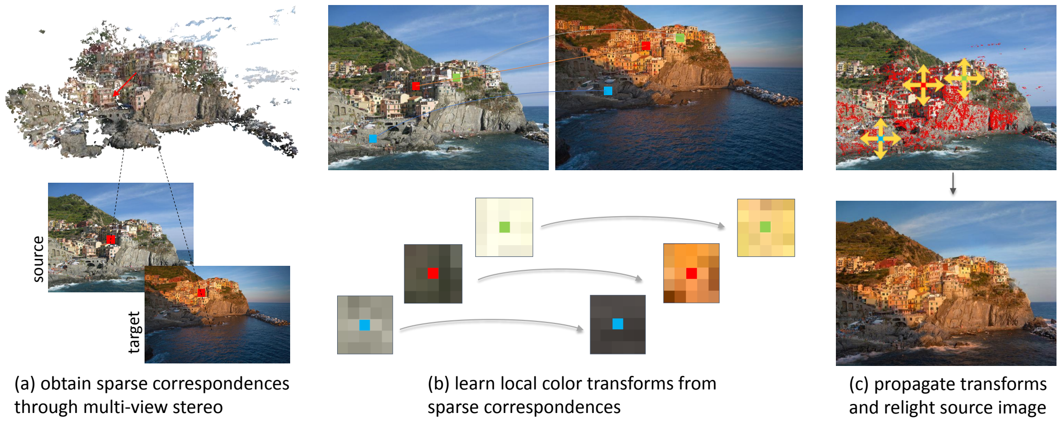


Fig. 2 Given a pair of source and target images from a photo collection, our method uses sparse correspondences (a) to learn local color transforms (b), which are then propagated in an image-guided manner in regions with no correspondences, and generates a relit image (c).

Edit propagation. Also related are the edit propagation methods, which propagate user specified edits with guidance of image gradients. Levin et al. [11] first introduce the framework for colorization, a computer-assisted process for adding color to a monochrome image or movie. They use manually specified color scribbles and propagate the colors in an edge-ware manner. Liu et al. [14] decompose image into illumination and reflectance layers, and transfer color to grayscale reflectance image using similar color propagation scheme. Lischinski et al. [13] extend the framework for image tonal manipulation, propagating user constraints with an edge-preserving optimization. A similar method is used in [1], which propagates rough user edits for spatially-varying image editing. Chen et al. [2] propose a manifold preserving edit propagation algorithm for video object recoloring and grayscale image colorization. Inspired by these approaches, we propagate local color transforms for lighting transfer. The edge-aware propagation originates from the sparse correspondences obtained from a pair of images. A key difference between our method and previous approaches is that we propagate transforms rather than simply color, which allows us to preserve texture in the source image.

3 Method

We propose a method for transferring lighting across photographs of a static scene. Our method takes as input a landmark scene photo collection, which includes images of multiple viewpoints and under different lighting conditions. The user chooses from the photo collection a source image to be edited, and a target image with desired lighting condition. We cast lighting transfer as an edit propagation problem. We use local color transforms to model the large variability of lighting changes for different parts of the scene. The transforms are learned from paired sparse correspondences across source and target images. Then, we propagate these transforms to relight the source image in an image-guided manner, and output a result image. The process is fully automatic.

Fig. 2 shows an overview of the pipeline of our approach, which consists of three main steps:

1. Extracting sparse correspondences from a photo collection. (Section 3.1)
2. Learning local color transforms from paired sparse correspondences. (Section 3.2)
3. Propagating local color transforms and relighting the source image. (Section 3.3)

To be robust to missing or inaccurate correspondences, we introduce a confidence map to detect potentially unreliable transforms in Section 3.4. We further extend our method for relighting based on multi-view target images in Section 3.5. More results and comparisons are presented in Section 4.

3.1 Sparse correspondences from a photo collection

We take as input a photo collection, consisting of images of the same scene with different viewpoints and lighting conditions. There are two reasons why we utilize photo collections. Photo-sharing websites contain millions of photographs of famous landmarks, and these collections of scenes with varying illumination provide rich information for lighting transfer. Besides, we can reconstruct a sparse point cloud from multi-view photos and find correspondences across images, which allows local analysis of lighting changes. We use off-the-shelf VisualSfM [26]: we first apply structure from motion [27] to estimate the parameters of cameras and then use patch-based multi-view stereo [4] to generate a 3D point cloud of the scene. For each point, the algorithm also estimates a list of images where it appears. The visible 3D points are projected to each image to obtain paired correspondences.

3.2 Learning local color transforms

We learn the lighting changes from sparse correspondences between the source image S and target image T. These correspondences can be represented by three-dimensional points in a given color space. We estimate transformations for corresponding pixel pairs to represent the color changes in a local neighborhood. The local color transforms [24] model color variations across a pair of images under varying lighting. Let $\boldsymbol{k}$ denote a correspondence in the source image. We express transform A_k around $\boldsymbol{k}$ as a linear matrix that maps the color of a pixel in the source image S to another pixel in target T. We learn the local transforms by applying linear models [10] in RGB color space. The local color transforms are modeled as the solutions to an optimization problem:

$$\arg\min_{A_k} \|v_k(T) - A_k v_k(S)\|_F^2 + \gamma \|A_k - G\|_F^2 \tag{1}$$

which can be solved in closed form as

$$A_k = (v_k(T)v_k(S)^T + \gamma G)(v_k(S)v_k(S)^T + \gamma I_{d_3})^{-1} \tag{2}$$

The obtained linear transform is represented by a 3×3 matrix A_k. Here, $\boldsymbol{k}$ corresponds to a specific correspondence. We denote by $v_k(S)$ the patch centered on the pixel in source image and by $v_k(T)$ the corresponding patch in target image. Both are represented as $3 \times P$ matrices in the RGB color space, where $P = 5 \times 5$ is the number of pixels in the patch. G is a global linear matrix estimated on the entire image ($\gamma = 0.01$), used for regularization. I_{d_3} is a 3×3 identity matrix. Among different color spaces, e.g. *HSV*, *CIELAB* and *RGB*, we find that local transforms work slightly better in the RGB space according to visual comparisons of results.

3.3 Propagation of local color transforms

We then propagate the transforms learned from correspondences to other regions of the source image. Inspired by the work of Levin et al. [11] and other edit propagation methods, we use an image-guided propagation algorithm. Instead of propagating the RGB pixel values, we propagate the color transforms estimated in the previous section.

Our propagation algorithm builds on the assumption that in a very small neighborhood, two pixels with similar colors are more likely to have similar transforms. We sample every pixel $\boldsymbol{i}$ in the source image, and assign a weight for each pixel $\boldsymbol{j}$ in the 3×3 sampling window. We wish to minimize the difference between the transform at pixel $\boldsymbol{i}$ and the weighted average of transforms at neighboring pixels. We assign $w = 1$ for the center pixel $\boldsymbol{i}$. If $\boldsymbol{i}$ has a correspondence in the target image and thus a precomputed transform, we set the weights of its neighbors to zero. Otherwise, the weights are calculated from Euclidean distances of colors. The weight is large when the colors of pixel $\boldsymbol{j}$ and $\boldsymbol{i}$ are similar, and small when they are different. We express the weighting function in the equation below. For each $\boldsymbol{j}$ in the sampling window D_i:

$$w_{ij} = \begin{cases} 1, & \boldsymbol{j} = \boldsymbol{i} \\ 0, & \boldsymbol{j} \neq \boldsymbol{i}, \boldsymbol{i} \in \{\boldsymbol{k}\} \\ -\frac{e^{-\|rgb(j)-rgb(i)\|^2/2\sigma_i^2}}{\sum_j e^{-\|rgb(j)-rgb(i)\|^2/2\sigma_i^2}}, & \boldsymbol{j} \neq \boldsymbol{i}, \boldsymbol{i} \notin \{\boldsymbol{k}\} \end{cases} \tag{3}$$

where σ_i^2 is the variance of the colors in the sampling window. Those weights are then used as constraints and guidance when propagating transforms. Given a sparse set of pixels $\boldsymbol{k}$ with precomputed transforms A_k (from Equation 2), the set of local transforms for all pixels in regions with no correspondences can be obtained by solving:

$$\arg\min_{A_j} \left(A_i - \sum_{\boldsymbol{j} \in D_i} w_{ij}\ A_j \right)^2 \tag{4}$$

where $A_i = A_k$ for all $\boldsymbol{i} \in \{\boldsymbol{k}\}$. We can rewrite Equation 4 in the form of matrix product, and formalize it as a global optimization problem:

$$W_{ij} * A = C \tag{5}$$

where W_{ij} is a $N \times N$ sparse matrix, whose (i,j)th entry is w_{ij}. $N = \mathit{width} \times \mathit{height}$ is the number of pixels in the source image. C is a constraint matrix with $C_i = A_k$ for $\boldsymbol{i} \in \{\boldsymbol{k}\}$, and $C_i = \mathbf{0}$ for $\boldsymbol{i} \notin \{\boldsymbol{k}\}$. A is the transform matrix to be solved. This large, sparse system of linear equations can be solved by standard methods. We use the backslash operator in Matlab. All the transforms A are optimized simultaneously. This allows us to propagate the learned sparse transforms to all the pixels without correspondences.

3.4 Detecting transform outliers

For the pixels without correspondences, the color transforms obtained via propagation might not be accurate, especially when these pixel colors are very

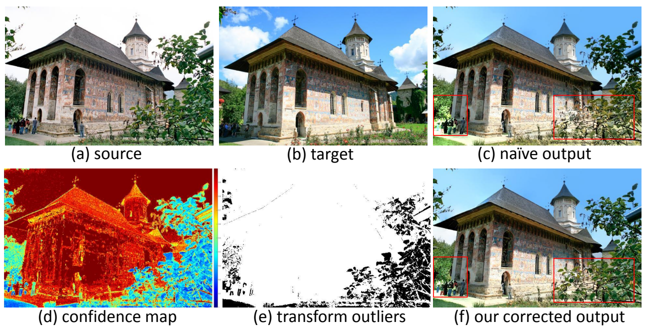


Fig. 3 Unreliable transforms result in distorted colors in the naive output (c). We compute a confidence map (d) to detect transform outliers after propagation. By removing the transforms with low confidence values (e) and remaining the associated source pixels' color unchanged, we have an output with correct colors (f), e.g. people and leaves (in red rectangle).

different from that of correspondences. We show an example of this situation in Fig. 3. The paired pixel correspondences between source image (a) and target image (b) are on the building, where the pixel colors are different from the colors of people's clothes and green leaves. The propagated transforms in these regions are thus inaccurate, and will transfer the source image wrongly as indicated with red rectangle in the naive output (c). To detect regions where transforms are potentially less reliable, we introduce a confidence map. The idea is if a source pixel's color is not similar to any of the correspondences in the source image, the obtained transform of that pixel is less reliable, as the propagation of transforms are based on color similarities.

$$C(p) \propto -ln\left(\sum_{m} \|rgb(p) - rgb(q)\|^2\right) \qquad (6)$$

For each pixel p in the source image, we calculate its color differences with all correspondences q in this image. A pixel only needs a few neighboring constraints to get an appropriate transform, so we sum up the smallest m differences and use the negative natural logarithm of the sum as a confidence factor $C(p)$. All factors are then normalized to $[0, 1]$, small when the transforms are not reliable. We use $m = 10$ and set a threshold to detect possibly wrong transforms. Those transforms are removed when applying color transformations, and their associated pixels remain their color as in the source image. As shown in Fig. 3, while there are color artifacts in the naive result (c), the leaves remain green and people's clothes seem more natural in the corrected output image (f).

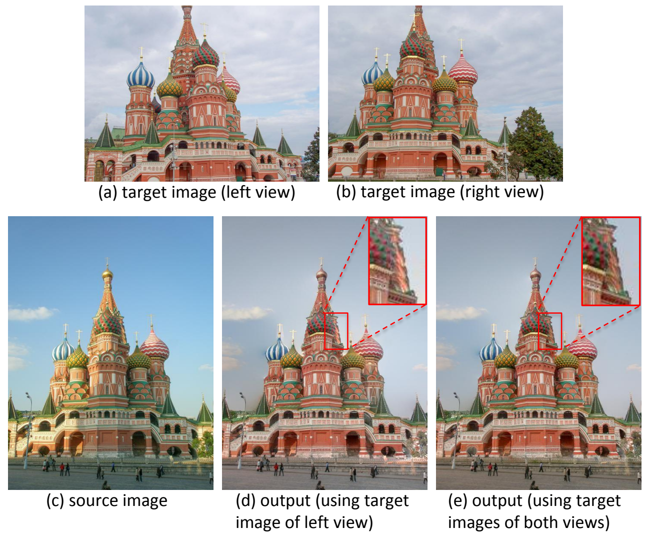


Fig. 4 We extend our method and transfer lighting from multiple target images. While the output using left-view target image (d) has inconsistent lighting in some regions, the output using target images of both views (e) has appropriate local lighting.

3.5 Extending to multiple targets

If the viewpoints of source and target images are drastically different, there are fewer correspondences. This would make it difficult to transfer lighting properly. To alleviate this issue, we extend our method by combining multiple target images with similar illumination conditions for the relighting of a source image.

Multiple target images provide more correspondences from different viewpoints. Here, we demonstrate the method using two target images of similar lighting. We learn the local color transforms from correspondences using the same method described in the previous sections, but combine the transforms before propagation. For pixels in the source image that have correspondences in both target images, the learned transforms are combined by calculating arithmetic mean. Fig. 4 shows that with the help of target images from different viewpoints, appropriate local lighting is transferred to the source image (see the regions highlighted in red rectangle). We further evaluate our method on a synthetic dataset, and make a comparison between the single-target-image method and the extended multiple-target-image one. The results are shown in the next section.

4 Results and Comparisons

We apply our method on two types of data. First we show results of our method for photo collections

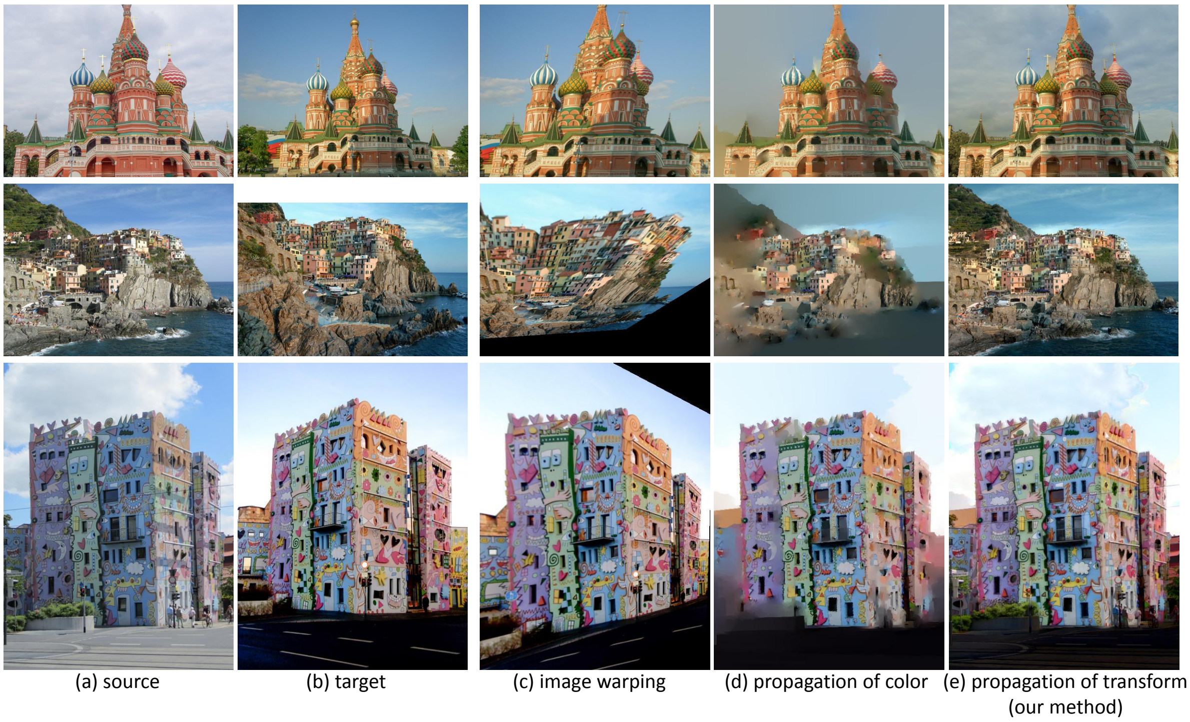


Fig. 5 We show comparisons of our method to image warping by homography and naive propagation of color. While the image warping method based on homography (c) distorts the image, and direct propagation of pixel colors (d) blurs out the image details, our method (e) successfully relights the source images.

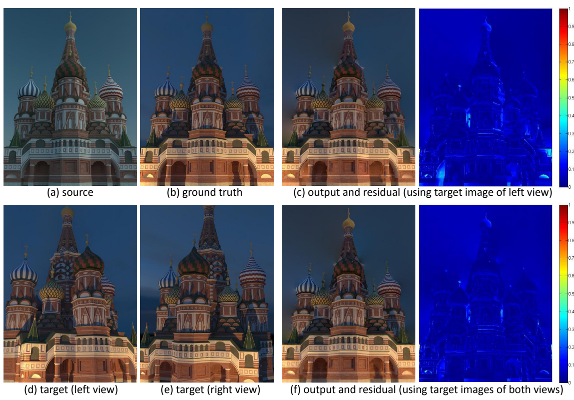


Fig. 6 We test the our lighting transfer methods on a synthetic dataset, and show results of quantitative evaluation. Compared with output using only the left-view target image (c), the output image produced with target images of both views (f) looks more similar to the ground truth (b) and has smaller residuals.

from online photo-sharing websites. We also apply our method to a synthetic dataset which allows a comparison to ground truth.

Internet photo collections. We utilize the datasets in [9]. When applying transforms directly to the source image, the noise existing in the image may be magnified. We use bilateral filtering [25] to decompose

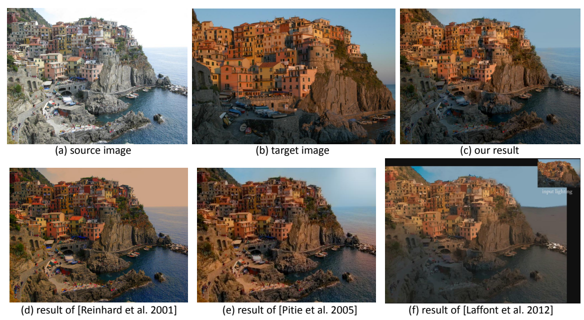


Fig. 7 We show a comparison between the global color transfer method [21], intrinsic image decomposition [9] and our lighting transfer method. While the results of other methods have either wrong color tone (c) or artifacts (d), our result has an appropriate lighting similar to the target image (e).

the source image into a detail layer and a base layer, and learn and propagate the transforms based on the base layer. We then apply the linear transforms to the base layer and add back the detail layer to obtain the final result. Similar method is used in [24].

Our method enables dramatic lighting transfer between images. Fig. 5 illustrates our results of several scenes, namely St.Basil, Manarola and RizziHaus. We compare to two baselines: image warping method based on homography and direct propagation of pixel colors. While the image warping method distorts

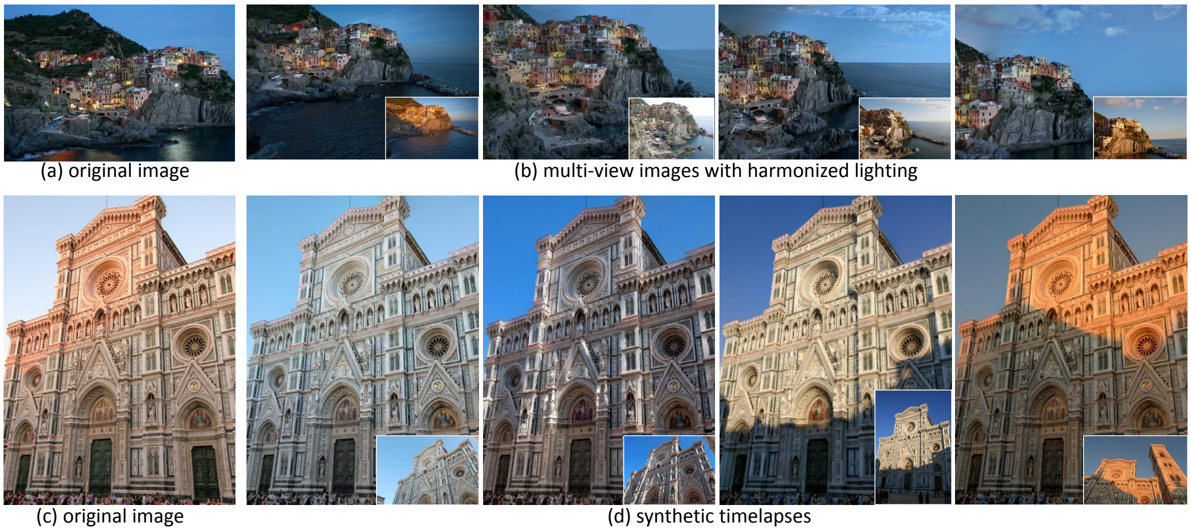


Fig. 8 We illustrate that our method can be used for harmonizing a photo collection with multi-view images (b) and hallucinating timelapses (d). The insets represent source images in (b) and target images with desired lighting in (d). Additional results are available in the supplementary video.

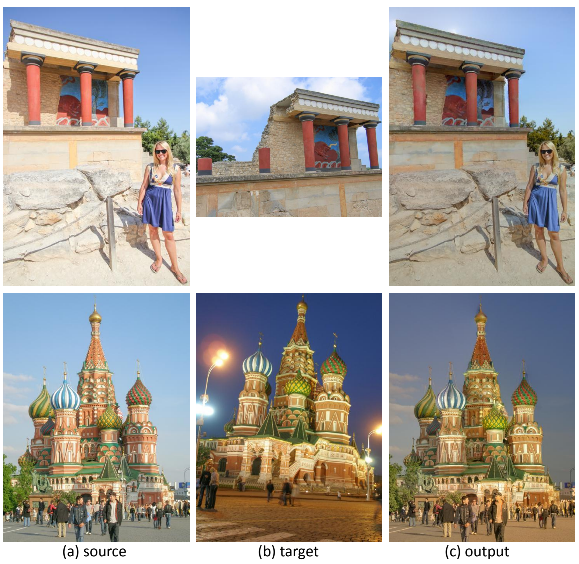


Fig. 9 Image relighting with people present in the landmark photos. Our method produces plausible results for scenes with strong local lighting. The background scene has proper local lighting transferred from the target, and people have a similar color with the scene.

the image, and propagating pixel colors blurs out the image details, our method successfully relights the source images. Homography is a projective mapping between any two images of the same planar surface in space. We estimate the homography based on pixel correspondences, using linear least squares in Matlab. The propagation of pixel colors uses code from [11]. Propagating colors produce blurry results, especially for regions with no correspondences and thus no guidance from "color scribbles".

Synthetic scene. We evaluate the effectiveness of our method on the synthetic dataset of St.Basil [8], which contains rendered images from 3 different viewpoints and under 30 lighting conditions. We compare the result of our lighting transfer to the ground truth rendering from the same viewpoint with the same lighting condition. Quantitative evaluation of absolute difference between relit images and ground truth in Fig. 6 shows that the method using multiple target images produces a more plausible result.

Comparisons. In order to further evaluate our lighting transfer method, we show a comparison with previous approaches in Fig. 7. Reinhard et al.'s method [21], which computes a global color mapping, turns the overall image into a warm tone. Pitie et al.'s method [19] produces a more similar tone to the target, but both of them do not transfer the local lighting properly. For the result of Laffont et al.'s method [9], which uses intrinsic image decomposition, the regions in shadow are washed-out, and there are artifacts around the boundaries of the sky and buildings. In contrast, our result has the lighting similar to the target image, and people and objects in the shadow still have their color. We show more comparison results with Shih et al [24] and deep photo style transfer [16] in Fig. S4, Fig. S5, in the electronic supplementary material (ESM).

Applications. In Fig. 8 we present that our method can be used for harmonized multi-view image collection browsing and time-lapse hallucination of a single view scenery. We refer to the supplementary video for more results, including the two applications of our image relighting method. We show image-based view transitions [22] with harmonized photographs. Our method produces stable transitions between views, and can transfer or remove strong shadows in the original images that could not be handled by simple color compensation. We also show timelapse sequences synthesized by transferring all illumination conditions to a single viewpoint. In addition, we show a side-by-side comparison with the results of Laffont et al. [9].

We also include relighting results where a person is present in the landmark photos and takes a significant part of the scene. Though people in the scene do not have any correspondences with the target images, Fig. 9 shows that our transform propagation method can produce a plausible result. The background scene has proper local lighting transferred from the target, and people have a similar color with the scene.

Performance. We use a 3.6 GHz Intel Core i7 CPU in this paper. All images are resized to the width of 640 pixels. Our Matlab implementation takes approximately 7s for learning and applying color transforms and 23s for propagating the transforms.

Limitations. Like all example-based techniques, our method has limitations. Processing images from varying viewpoints and under dramatically different illumination conditions can be challenging, as the multi-view stereo method may not find sufficient correspondences across images. Picking the target image with more correspondences or several targets with similar illumination may help produce better results. Another challenging case is a scene region with similar texture but distinct target lighting at different depth. The propagation of transforms guided by the source image would be the same, and thus the generated output is not desired. For high-quality result of a small region, a RGB-D camera in the scene may greatly increase the number of correspondences and allow more accurate analysis of the spatially-varying lighting.

5 Conclusion

The novelty of this paper is that we cast lighting transfer as an edit propagation problem. We learn local color transforms from sparse correspondences reconstructed from multi-view stereo, and propagate in an image-guided manner. Compared to previous image relighting methods, our approach does not rely on highly accurate geometry, timelapse videos from static viewpoints or densely computed correspondences. The color transforms model the large variability of local lighting changes for different parts of the scene across images. We demonstrate that our method can be used for enhancing photographs, harmonizing image collections of multiple viewpoints and hallucinating timelapses.

Acknowledgements

We would like to thank all the reviewers for their comments and suggestions. The first author carried out the earlier phase of the research at National University of Singapore with support from the School of Computing. This research is supported by the BeingThere Centre, a collaboration between Nanyang Technological University Singapore, Eidgenössische Technische Hochschule Zürich, and University of North Carolina at Chapel Hill. The BeingThere Centre is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the Interactive Digital Media Programme Office.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Electronic Supplementary Material (ESM) Supplementary document and video with more results are available in the online version of this article.

References

[1] X. An and F. Pellacini. Appprop: all-pairs appearance-space edit propagation. *ACM Transactions on Graphics (TOG)*, 2008.

[2] X. Chen, D. Zou, Q. Zhao, and P. Tan. Manifold preserving edit propagation. *ACM Transactions on Graphics (TOG)*, 2012.

[3] P. Debevec and et al. Estimating surface reflectance properties of a complex scene under captured natural illumination. Technical report, USC Institute for Creative Technologies, 2004.

[4] Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

[5] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with applications for image enhancement. *ACM transactions on graphics (TOG)*, 2011.

[6] H.-Z. Huang, S.-H. Zhang, R. R. Martin, and S.-M. Hu. Learning natural colors for image recoloring. In *Computer Graphics Forum*, 2014.

[7] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski. Deep photo: Model-based photograph enhancement and viewing. *ACM Transactions on Graphics (TOG)*, 2008.

[8] P.-Y. Laffont and J.-C. Bazin. Intrinsic decomposition of image sequences from local temporal variations. In *ICCV*, 2015.

[9] P.-Y. Laffont, A. Bousseau, S. Paris, F. Durand, and G. Drettakis. Coherent intrinsic images from photo collections. *ACM Transactions on Graphics (TOG)*, 2012.

[10] P.-Y. Laffont, Z. Ren, X. Tao, C. Qian, and J. Hays. Transient attributes for high-level understanding and editing of outdoor scenes. *ACM Transactions on Graphics (TOG)*, 2014.

[11] A. Levin, D. Lischinski, and Y. Weiss. Colorization using optimization. *ACM Transactions on Graphics (TOG)*, 2004.

[12] X. Li, H. Zhao, G. Nie, and H. Huang. Image recoloring using geodesic distance based color harmonization. *Computational Visual Media*, 2015.

[13] D. Lischinski, Z. Farbman, M. Uyttendaele, and R. Szeliski. Interactive local adjustment of tonal values. *ACM Transactions on Graphics (TOG)*, 2006.

[14] X. Liu, L. Wan, Y. Qu, T.-T. Wong, S. Lin, C.-S. Leung, and P.-A. Heng. Intrinsic colorization. *ACM Transactions on Graphics (TOG)*, 2008.

[15] S.-P. Lu, B. Ceulemans, A. Munteanu, and P. Schelkens. Spatio-temporally consistent color and structure optimization for multiview video color correction. *IEEE Transactions on Multimedia*, 2015.

[16] F. Luan, S. Paris, E. Shechtman, and K. Bala. Deep photo style transfer. *arXiv preprint arXiv:1703.07511*, 2017.

[17] R. Martin-Brualla, D. Gallup, and S. M. Seitz. Time-lapse mining from internet photos. *ACM Transactions on Graphics (TOG)*, 2015.

[18] J. Park, Y.-W. Tai, S. N. Sinha, and I. So Kweon. Efficient and robust color consistency for community photo collections. In *CVPR*, 2016.

[19] F. Pitie, A. C. Kokaram, and R. Dahyot. N-dimensional probability density function transfer and its application to color transfer. In *ICCV*, 2005.

[20] T. Pouli and E. Reinhard. Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics*, 2011.

[21] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, 2001.

[22] D. A. Roberts. Pixelstruct, an opensource tool for visualizing 3d scenes reconstructed from photographs, 2009. https://github.com/davidar/pixelstruct.

[23] X. Shen, X. Tao, C. Zhou, H. Gao, and J. Jia. Regional foremost matching for internet scene images. *ACM Transactions on Graphics (TOG)*, 2016.

[24] Y. Shih, S. Paris, F. Durand, and W. T. Freeman. Data-driven hallucination of different times of day from a single outdoor photo. *ACM Transactions on Graphics (TOG)*, 2013.

[25] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, 1998.

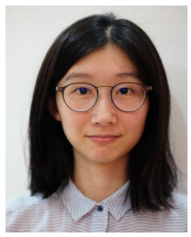
[26] C. Wu. Visualsfm: A visual structure from motion system, 2011. http://ccwu.me/vsfm/.

[27] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz. Multicore bundle adjustment. In *CVPR*, 2011.

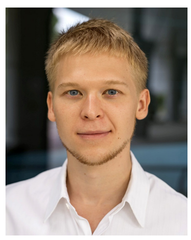
[28] S. Ye, S.-P. Lu, and A. Munteanu. Color correction for large-baseline multiview video. *Signal Processing: Image Communication*, 2017.

[29] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs. SIGGRAPH, 1999.

[30] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. SIGGRAPH, 1998.



Qian Zhang is a research assistant at Nanyang Technological University Singapore. Her research interests include image processing, computational photography, and image-based rendering. Zhang has a BS in electronics and information engineering from Huazhong University of Science and Technology.



Pierre-Yves Laffont is the CEO and co-founder of Lemnis Technologies. During this research, he was a postdoctoral researcher at ETH Zurich and a visiting researcher at Nanyang Technological University. His research interests include intrinsic image decomposition, example-based appearance transfer, and image-based rendering/relighting. Laffont has a PhD in Computer Science from Inria Sophia-Antipolis.

Terence Sim is an Associate Professor at the School of Computing, National University of Singapore. He is also an Assistant Dean of Corporate Relations at our School. For research, Dr. Sim works primarily in these areas: facial image analysis, biometrics, and computational photography. He is also interested in computer vision problems in general, such as shape-from-shading, photometric stereo, object recognition.

From 2014 to 2016, Dr. Sim served as President of the Pattern Recognition and Machine Intelligence Association (PREMIA), a national professional body for pattern recognition, affiliated with the International Association for Pattern Recognition (IAPR).