

# Lighting transfer across multiple views through local color transforms

**Journal Article****Author(s):**

Zhang, Qian; Laffont, Pierre-Yves; Sim, Terence

**Publication date:**

2017-12

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000281693>

**Rights / license:**

[Creative Commons Attribution 4.0 International](#)

**Originally published in:**

Computational Visual Media 3(4), <https://doi.org/10.1007/s41095-017-0085-5>

# Lighting transfer across multiple views through local color transforms

Qian Zhang<sup>1</sup> (✉), Pierre-Yves Laffont<sup>2</sup>, and Terence Sim<sup>3</sup>

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

**Abstract** We present a method for transferring lighting between photographs of a static scene. Our method takes as input a photo collection depicting a scene with 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 to 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 between images. We show applications of our image relighting method for enhancing photographs, browsing photo collections with harmonized lighting, and generating synthetic time-lapse sequences.

**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 softwares 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, as shown in Fig. 1. There are millions of photographs of famous landmarks on 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:

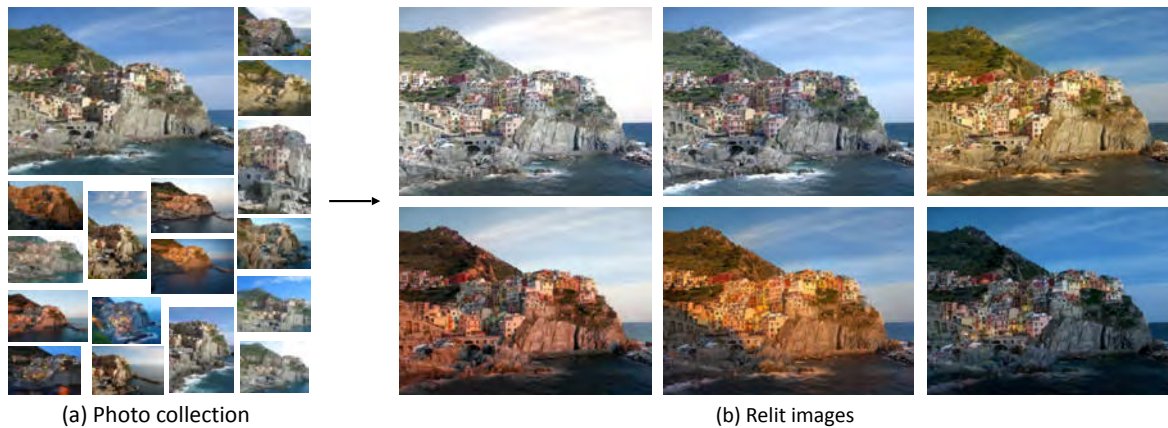
- We cast lighting transfer as an edit propagation problem, learning local color transforms from sparse geometric correspondences and propagating the transforms in an edge-aware manner.
- We introduce a confidence map to indicate the reliability of propagated transforms, which helps to preserve the color of pixels with transform outliers.

1 Nanyang Technological University, 639798, Singapore.  
E-mail: zhangqian@ntu.edu.sg (✉).

2 ETH Zurich, 8092 Zurich, Switzerland.

3 National University of Singapore, 119077, Singapore.

Manuscript received: 2017-03-31; accepted: 2017-05-26



**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.

- We extend our method to transfer lighting based on multiple target images, exploiting the information from different viewpoints.

We have run our method on 6 scenes, including 5 Internet photo collections and a synthetic benchmark with ground truth images, allowing us to give a quantitative evaluation. We also show comparisons with baselines and previous approaches. Our image relighting method enables enhancement of photographs, photo collection browsing with harmonized lighting, and synthetic time-lapse generation.

## 2 Related work

### 2.1 Color transfer and correction

Lighting transfer mainly concerns color. Approaches for color transfer manipulate color distributions. Example-based transfer methods such as those in Refs. [1–3] reshape the color distribution of the input image so that it approaches the statistical color properties of the example image. Huang et al. [4] recolor a photo by learning from database correlations between color property distributions and geometric features of regions. Li et al. [5] recolor images using geodesic distance based on harmonization. More recently, Luan et al. [6] propose a deep learning approach for photographic style transfer. These methods produce visually pleasing recolored images but cannot change local lighting. Color transfer methods can also be used for tone adjustment and correction. Park et al. [7] recover sparse pixel correspondences and compute color correction parameters with a

low-rank matrix factorization technique. Spatial-temporal correspondences are also used in Refs. [8, 9] 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, so can transfer strong shadows.

### 2.2 Image relighting

A number of image relighting methods have been proposed by various researchers over the years, such as those in Refs. [10–13]. These sophisticated systems make use of detailed geometric models and require registration or non-linear fitting. Laffont et al. [14] 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. [15] compute a parametric color model based on dense correspondences, but do not take into account local color changes. Shih et al. [16] successfully synthesize different-time-of-day images by learning color transformations from time-lapse videos. A similar approach by Laffont et al. [17] enables appearance transfer of time-of-day, weather, or season by observing color changes in a 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. [18] use a simple but effective new temporal filtering approach to stabilize appearance. In work developed concurrently, Shen et al. [19] propose regional foremost matching for image morphing and time-lapse sequence generation. In our system, we target a more general case that does not need highly accurate geometry, time-lapse sequences from a static viewpoint, or densely computed correspondences. Our method relies on the vast numbers of available images of the same scene in various online photo communities and sparse geometric correspondences.

### 2.3 Edit propagation

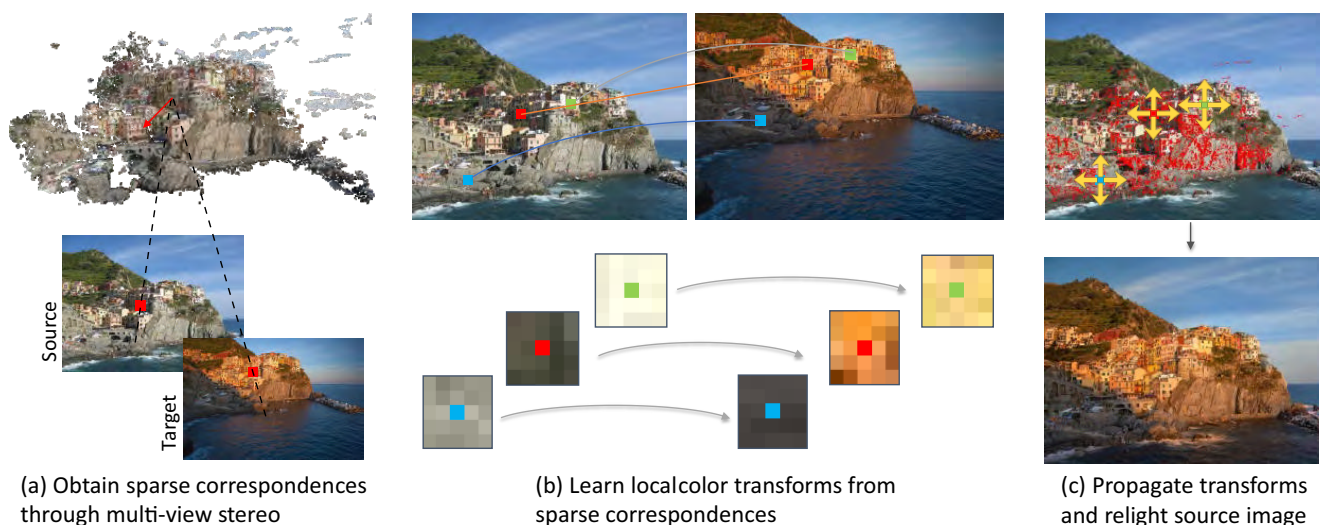
Also related are edit propagation methods, which propagate user specified edits under the guidance of image gradients. Levin et al. [20] first introduce a framework for colorization, a computer-assisted process for adding color to a monochrome image or movie. They use manually specified color scribbles and propagate the colors in an edge-aware manner. Liu et al. [21] decompose images into illumination and reflectance layers, and transfer color to grayscale reflectance images using a similar color propagation scheme. Lischinski et al. [22] extend the framework for image tone manipulation, propagating user constraints with edge-preserving optimization. A similar method is used in Ref. [23], which propagates coarse user edits for spatially-varying image editing.

Chen et al. [24] 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. Edge-aware propagation originates at 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 between photographs of a static scene. Our method takes as input a landmark scene photo collection, which includes images from 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 in different parts of the scene. The transforms are learned from paired sparse correspondences between 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.

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



**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 to regions with no correspondences, generating a relit image (c).



- (1) Extracting sparse correspondences from a photo collection (see Section 3.1).
- (2) Learning local color transforms from paired sparse correspondences (see Section 3.2).
- (3) Propagating local color transforms and relighting the source image (see 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. Further 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. Moreover, we can reconstruct a sparse point cloud from multi-view photos and find correspondences between images, which allows local analysis of lighting changes. We use off-the-shelf VisualSfM [25]: we first apply structure from motion [26] to estimate the parameters of the cameras and then use patch-based multi-view stereo [27] to generate a 3D point cloud of the scene. For each point, the algorithm also estimates a list of images in which 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 [16] model color variations between a pair of images under varying lighting. Let  $k$  denote a correspondence in the source image. We express the transform  $\mathbf{A}_k$  for  $k$  as a linear matrix that maps the color of a pixel in the source image  $S$  to another

pixel in the target  $T$ . We learn the local transforms as linear models [17] in RGB color space. The local color transforms are modeled as the solutions to an optimization problem:

$$\arg \min_{\mathbf{A}_k} \|\mathbf{v}_k(T) - \mathbf{A}_k \mathbf{v}_k(S)\|_{\mathbb{F}}^2 + \gamma \|\mathbf{A}_k - \mathbf{G}\|_{\mathbb{F}}^2 \quad (1)$$

which can be solved in closed form as

$$\mathbf{A}_k = (\mathbf{v}_k(T)\mathbf{v}_k(S)^{\top} + \gamma \mathbf{G})(\mathbf{v}_k(S)\mathbf{v}_k(S)^{\top} + \gamma \mathbf{I}_{d_3})^{-1} \quad (2)$$

The obtained linear transform  $\mathbf{A}_k$  is represented by a  $3 \times 3$  matrix. We denote by  $\mathbf{v}_k(S)$  the patch centered on the pixel in the source image and by  $\mathbf{v}_k(T)$  the corresponding patch in the target image. Both are represented as  $3 \times P$  matrices in RGB color space, where  $P = 5 \times 5$  is the number of pixels in the patch.  $\mathbf{G}$  is a global linear matrix estimated on the entire image ( $\gamma = 0.01$ ), used for regularization.  $\mathbf{I}_{d_3}$  is a  $3 \times 3$  identity matrix. Trying different color spaces, e.g., HSV, CIELAB, and RGB, a visual comparison of results in Fig. S1 in the Electronic Supplementary Material (ESM) shows that local transforms work slightly better in RGB space.

### 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. [20] and other edit propagation methods, we use an image-guided propagation algorithm. Instead of propagating 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 likely to have similar transforms. We sample every pixel  $i$  in the source image, and assign a weight  $w$  for each pixel  $j$  in the  $3 \times 3$  sampling window. We wish to minimize the difference between the transform at pixel  $i$  and the weighted average of transforms at neighboring pixels. We assign  $w = 1$  for the center pixel  $i$ . If  $i$  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 pixels  $j$  and  $i$  are similar, and small when they are different. We express the weighting function in the equation below. For each  $j$  in the sampling window  $D_i$ :

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

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

$$\arg \min_{\mathbf{A}_j} \left( \mathbf{A}_i - \sum_{j \in D_i} w_{ij} \mathbf{A}_j \right)^2 \quad (4)$$

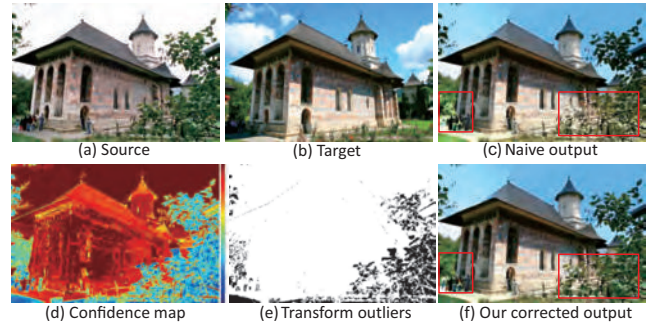
where  $\mathbf{A}_i = \mathbf{A}_k$  for all  $i \in \{k\}$ . We can rewrite Eq. (4) in the form of a matrix product, and formalize it as a global optimization problem:

$$\mathbf{W}\mathbf{A} = \mathbf{C} \quad (5)$$

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

### 3.4 Detecting transform outliers

For pixels without correspondences, the color transforms obtained by propagation may not be accurate, especially when these pixel have colors very different from those 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 the green leaves. The propagated transforms in these regions are thus inaccurate, and will transfer the source image wrongly as indicated by the 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 computed transform of that pixel is less reliable,



**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 leaving the associated source pixels' colors unchanged, we have an output with correct colors (f), e.g., people and leaves (in the red rectangle).

as the propagation of transforms is based on color similarities.

$$C(p) \propto -\ln \left( \sum_m \|\text{rgb}(p) - \text{rgb}(q)\|^2 \right) \quad (6)$$

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  $[0, 1]$ , with a small value when the transforms are unreliable. We use  $m = 10$  and set a threshold to detect possibly wrong transforms. Such transforms are removed when applying color transformations, and the associated pixels retain the same colors 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).

### 3.5 Extension to multiple targets

If the viewpoints of source and target images are drastically different, there are fewer correspondences. This makes 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 with similar lighting. We learn the local color transforms from correspondences using the same

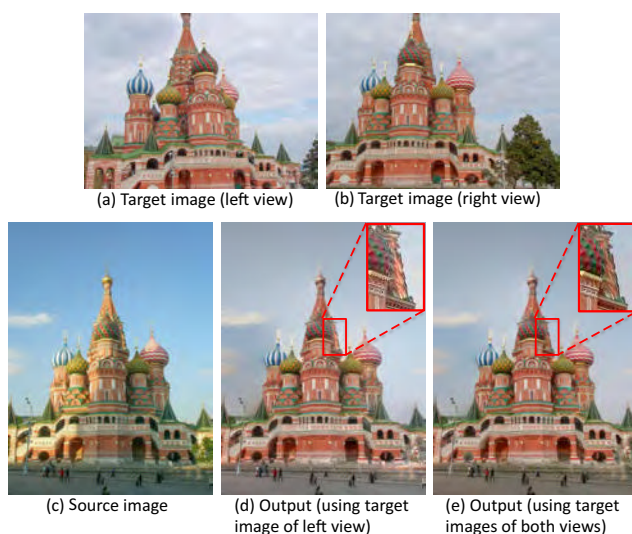
method described in the previous sections, but combine the transforms before propagation. For pixels in the source image that have correspondences in both target images, the learned transforms are combined by calculating their arithmetic mean. Figure 4 shows that with the help of target images from different viewpoints, appropriate local lighting is transferred to the source image (see the regions highlighted by red rectangles). 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 to two types of data. First we show results of our method for photo collections from online photo-sharing websites. We also apply our method to a synthetic dataset which allows a comparison to ground truth.

### 4.1 Internet photo collections

We utilize the datasets in Ref. [14]. When applying transforms directly to the source image, noise in the image may be magnified. We use bilateral filtering [28] to decompose 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



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

add back the detail layer to obtain the final result. A similar method is used in Ref. [16].

Our method enables dramatic lighting transfer between images. Figure 5 illustrates our results for several scenes, namely St. Basil, Manarola, and Rizzi Haus. We compare to two baselines: an image warping method based on homography, and direct propagation of pixel colors. While the image warping method distorts the image, and propagating pixel colors blurs image details, our method successfully relights the source images. We estimate the homography based on pixel correspondences, using linear least squares in MATLAB. The propagation of pixel colors uses code from Ref. [20]. Propagating colors produce blurred results, especially for regions with no correspondences and thus no guidance from “color scribbles”.

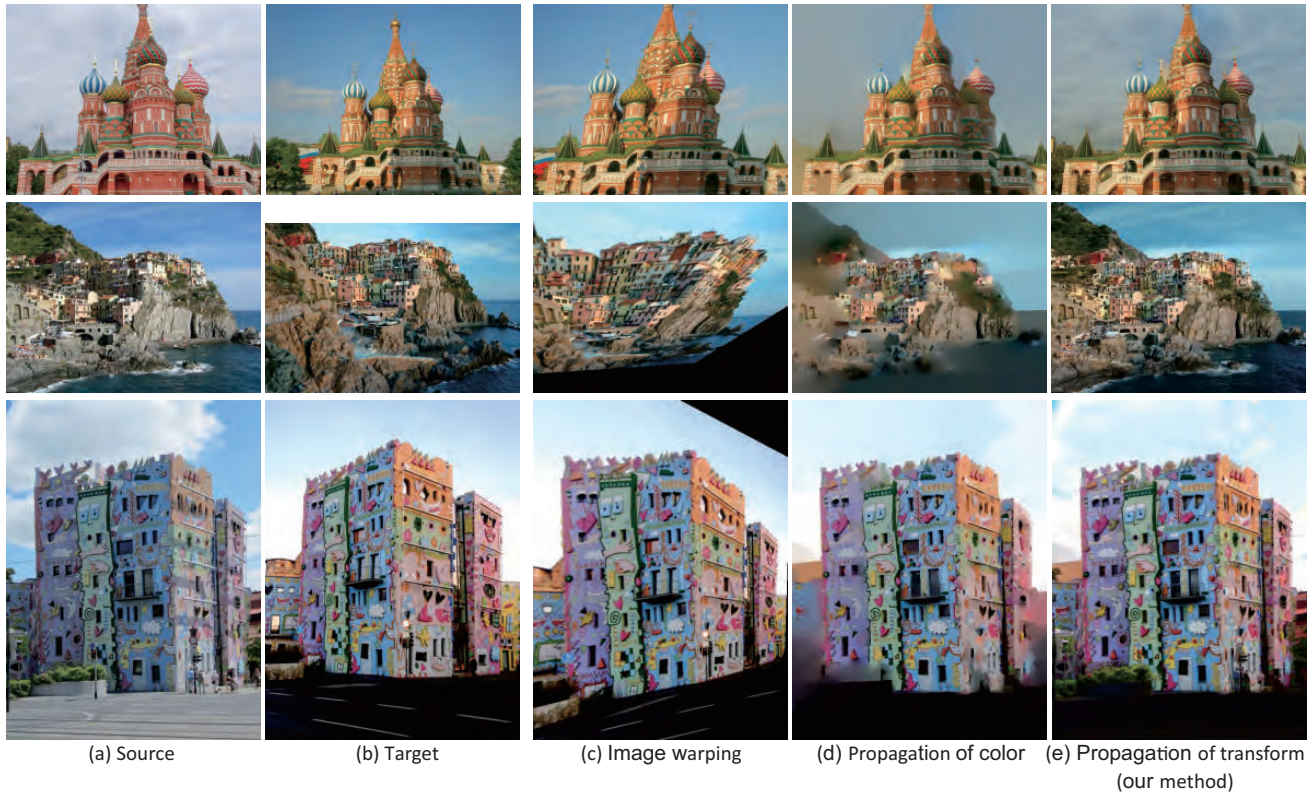
### 4.2 Synthetic scene

We evaluate the effectiveness of our method on the synthetic dataset of St. Basil [29], 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 conditions. Quantitative evaluation of absolute differences between relit images and ground truth in Fig. 6 shows that the method using multiple target images produces a more plausible result.

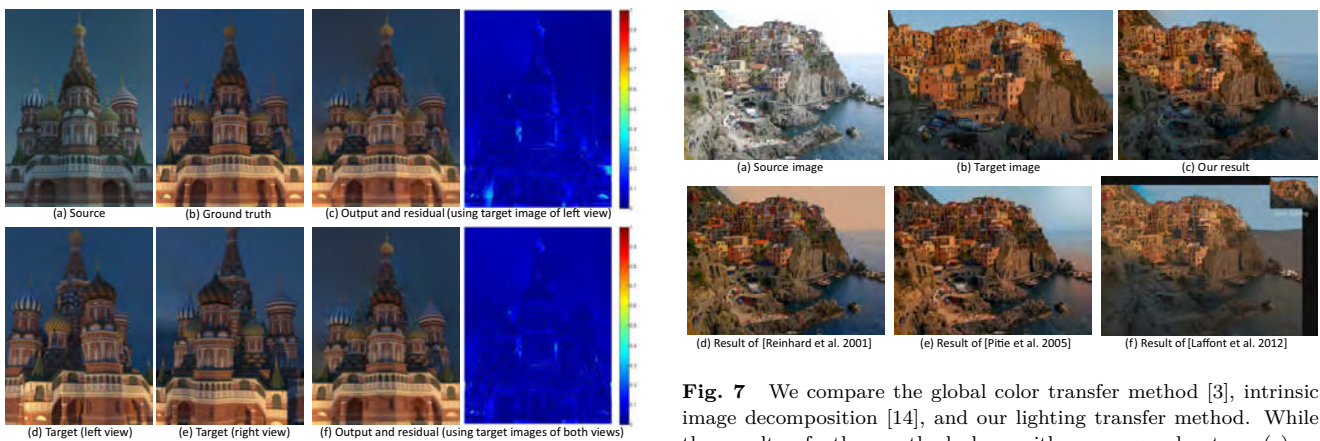
### 4.3 Comparisons

In order to further evaluate our lighting transfer method, we show a comparison with previous approaches in Fig. 7. Reinhard et al.’s method [3], which computes a global color mapping, gives the overall image a warm tone. Pitie et al.’s method [2] produces a more similar tone to the target, but both of them do not properly transfer local lighting. In the result of Laffont et al.’s method [14], 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 lighting similar to that of the target image, and people and objects in shadow still have their colors. We show more comparison results with Shih et al. [16] and deep photo style transfer [6] in Fig. S4 and Fig. S5 in the ESM.





**Fig. 5** We compare 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 image details, our method (e) successfully relights the source images.



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

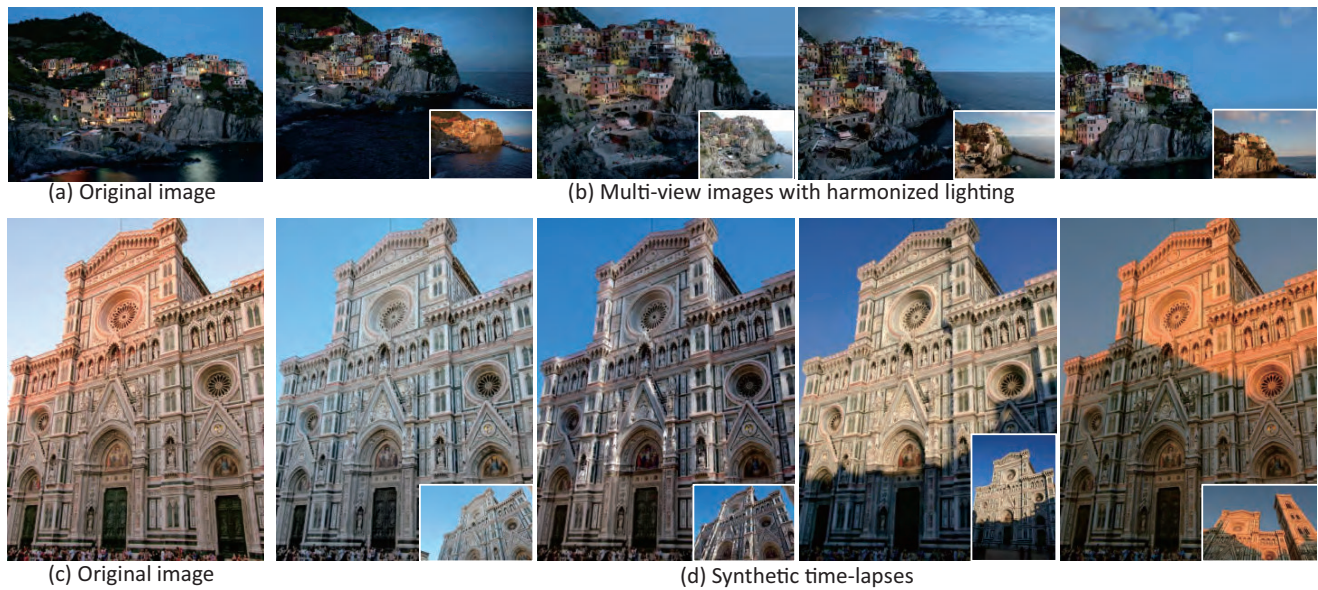
**Fig. 7** We compare the global color transfer method [3], intrinsic image decomposition [14], and our lighting transfer method. While the results of other methods have either wrong color tone (c) or artifacts (d), our result has appropriate lighting similar to the target image (b).

### 4.4 Applications

In Fig. 8, we show that our method can be used for harmonized multi-view image collection browsing and time-lapse hallucination of single view scenery. We refer to the supplementary video in

the ESM for results for these two applications. We show image-based view transitions [30] 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 time-lapse sequences synthesized by transferring all illumination conditions to a single

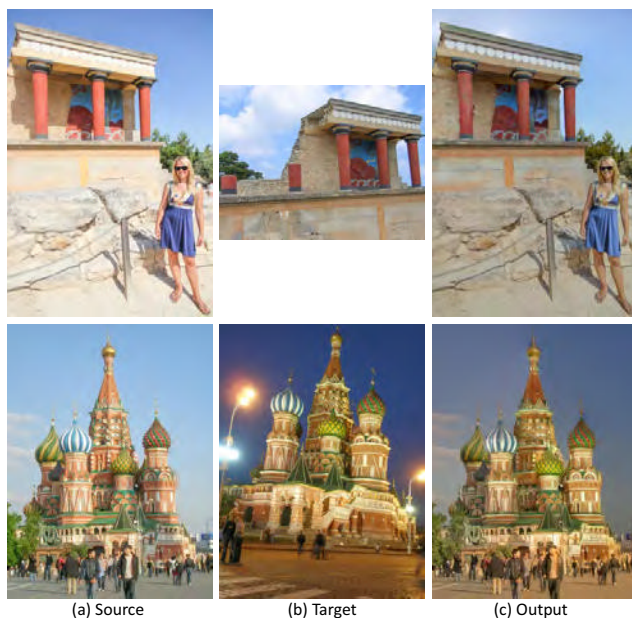




**Fig. 8** Our method can be used for harmonizing a photo collection with multi-view images (b) and hallucinating time-lapses (d). The insets represent source images in (b) and target images with desired lighting in (d). Additional results are available in the supplementary video in the ESM.

viewpoint. In addition, we show a side-by-side comparison with the results of Laffont et al. [14].

We also include relighting results where a person is present in the landmark photos and occupies 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.



**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 to the scene.

The background scene has proper local lighting transferred from the target, and people have a similar color to the scene.

#### 4.5 Performance

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

#### 4.6 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 between images. Picking a 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 a different depth. The propagation of transforms guided by the source image would be the same, and thus the generated output would not be as desired. For high-quality results in a small region, an RGB-D camera in the scene may greatly increase the number of correspondences and allow more accurate analysis of the spatially-varying lighting.

## 5 Conclusions

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, time-lapse videos from static viewpoints, or densely computed correspondences. The color transforms model the large variability of local lighting changes between images in different parts of the scene. We demonstrate that our method can be used for enhancing photographs, harmonizing image collections of multiple viewpoints, and hallucinating time-lapse sequences.

## Acknowledgements

We would like to thank all reviewers for their comments and suggestions. The first author carried out the earlier phase of the research at the 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 the 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 is administered by the Interactive Digital Media Programme Office.

**Electronic Supplementary Material** Supplementary materials (a supplementary document and video with further results) are available in the online version of this article at <https://doi.org/10.1007/s41095-017-0085-5>.

## References

- [1] Pouli, T.; Reinhard, E. Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics* Vol. 35, No. 1, 67–80, 2011.
- [2] Pitie, F.; Kokaram, A. C.; Dahyot, R. N-dimensional probability density function transfer and its application to color transfer. In: Proceedings of the 10th IEEE International Conference on Computer Vision, Vol. 2, 1434–1439, 2005.
- [3] Reinhard, E.; Ashikhmin, M.; Gooch, B.; Shirley, P. Color transfer between images. *IEEE Computer Graphics and Applications* Vol. 21, No. 5, 34–41, 2001.
- [4] Huang, H.-Z.; Zhang, S.-H.; Martin, R. R.; Hu, S.-M. Learning natural colors for image recoloring. *Computer Graphics Forum* Vol. 33, No. 7, 299–308, 2014.
- [5] Li, X.; Zhao, H.; Nie, G.; Huang, H. Image recoloring using geodesic distance based color harmonization. *Computational Visual Media* Vol. 1, No. 2, 143–155, 2015.
- [6] Luan, F.; Paris, S.; Shechtman, E.; Bala, K. Deep photo style transfer. *arXiv preprint arXiv:1703.07511*, 2017.
- [7] Park, J.; Tai, Y.-W.; Sinha, S. N.; Kweon, I. S. Efficient and robust color consistency for community photo collections. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 430–438, 2016.
- [8] Ye, S.; Lu, S.-P.; Munteanu, A. Color correction for large-baseline multiview video. *Signal Processing: Image Communication* Vol. 53, 40–50, 2017.
- [9] Lu, S.-P.; Ceulemans, B.; Munteanu, A.; Schelkens, P. Spatio-temporally consistent color and structure optimization for multiview video color correction. *IEEE Transactions on Multimedia* Vol. 17, No. 5, 577–590, 2015.
- [10] Yu, Y.; Debevec, P.; Malik, J.; Hawkins, T. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, 215–224, 1999.
- [11] Debevec, P.; Tchou, C.; Gardner, A.; Hawkins, T.; Poullis, C.; Stumpfel, J.; Jones, A.; Yun, N.; Einarsson, P.; Lundgren, T.; Fajardo, M.; Martinez, P. Estimating surface reflectance properties of a complex scene under captured natural illumination. USC ICT Technical Report ICT-TR-06. 2004.
- [12] Kopf, J.; Neubert, B.; Chen, B.; Cohen, M.; Cohen-Or, D.; Deussen, O.; Uyttendaele, M.; Lischinski, D. Deep photo: Model-based photograph enhancement and viewing. *ACM Transactions on Graphics* Vol. 27, No. 5, Article No. 116, 2008.
- [13] Yu, Y.; Malik, J. Recovering photometric properties of architectural scenes from photographs. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, 207–217, 1998.
- [14] Laffont, P.-Y.; Bousseau, A.; Paris, S.; Durand, F.; Drettakis, G. Coherent intrinsic images from photo collections. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 202, 2012.
- [15] HaCohen, Y.; Shechtman, E.; Goldman, D. B.; Lischinski, D. Non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics* Vol. 30, No.4, Article No. 70, 2011.

- [16] Shih, Y.; Paris, S.; Durand, F.; Freema, W. T. Data-driven hallucination of different times of day from a single outdoor photo. *ACM Transactions on Graphics* Vol. 32, No. 6, Article No. 200, 2013.
- [17] Laffont, P.-Y.; Ren, Z.; Tao, X.; Qian, C.; Hays, J. Transient attributes for high-level understanding and editing of outdoor scenes. *ACM Transactions on Graphics* Vol. 33, No. 4, Article No. 145, 2014.
- [18] Martin-Brualla, R.; Gallup, D.; Seitz, S. M. Time-lapse mining from internet photos. *ACM Transactions on Graphics* Vol. 34, No. 4, Article No. 62, 2015.
- [19] Shen, X.; Tao, X.; Zhou, C.; Gao, H.; Jia, J. Regional foremost matching for internet scene images. *ACM Transactions on Graphics* Vol. 35, No. 6, Article No. 178, 2016.
- [20] Levin, A.; Lischinski, D.; Weiss, Y. Colorization using optimization. *ACM Transactions on Graphics* Vol. 23, No. 3, 689–694, 2004.
- [21] Liu, X.; Wan, L.; Qu, Y.; Wong, T.-T.; Lin, S.; Leung, C.-S.; Heng, P.-A. Intrinsic colorization. *ACM Transactions on Graphics* Vol. 27, No. 5, Article No. 152, 2008.
- [22] Lischinski, D.; Farbman, Z.; Uyttendaele, M.; Szeliski, R. Interactive local adjustment of tonal values. *ACM Transactions on Graphics* Vol. 25, No. 3. 646–653, 2006.
- [23] An, X.; Pellacini, F. AppProp: All-pairs appearance-space edit propagation. *ACM Transactions on Graphics* Vol. 27, No. 3, Article No. 40, 2008.
- [24] Chen, X.; Zou, D.; Zhao, Q.; Tan, P. Manifold preserving edit propagation. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 132, 2012.
- [25] Wu, C. VisualSFM: A visual structure from motion system. 2011. Available at <http://ccwu.me/vsfm/>.
- [26] Wu, C.; Agarwal, S.; Curless, B.; Seitz, S. M. Multicore bundle adjustment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3057–3064, 2011.
- [27] Furukawa, Y.; Ponce, J. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 32, No. 8, 1362–1376, 2010.
- [28] Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In: Proceedings of the 6th International Conference on Computer Vision, 839–846, 1998.
- [29] Laffont, P.-Y.; Bazin, J.-C. Intrinsic decomposition of image sequences from local temporal variations. In: Proceedings of the IEEE International Conference on Computer Vision, 433–441, 2015.
- [30] Roberts, D. A. Pixelstruct, an opensource tool for visualizing 3D scenes reconstructed from photographs. 2009. Available at <https://github.com/davidar/pixelstruct>.



of Science and Technology, China.

**Qian Zhang** is a research assistant at Nanyang Technological University, Singapore. Her research interests include image processing, computational photography, and image-based rendering. Qian Zhang has her B.S. degree in electronics and information engineering from Huazhong University



appearance transfer, and image-based rendering and relighting. He has his Ph.D. degree in computer science from Inria Sophia-Antipolis.

**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



in computer vision problems in general, such as shape-from-shading, photometric stereo, and 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).

**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 the School. For research, Dr. Sim works primarily in the areas of facial image analysis, biometrics, and computational photography. He is also interested

**Open Access** The articles published in this journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Other papers from this open access journal are available free of charge from <http://www.springer.com/journal/41095>. To submit a manuscript, please go to <https://www.editorialmanager.com/cvmj>.