Lighting Transfer across Multiple Views through Local Color Transforms

Qian Zhang Nanyang Technological University Pierre-Yves Laffont ETH Zurich

Terence Sim National University of Singapore



Figure 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.

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 a colorization 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 is able to transfer strong shadows across images. We show applications of our image relighting method, for browsing collections of photographs with harmonized lighting and for generating synthetic timelapses.

Keywords: relighting, photo collection, time-lapse, image editing

Concepts: •Computing methodologies \rightarrow Image manipulation; Computational photography;

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 be present at the right place at the perfect time of day. A

SA '16 Technical Briefs, December 05 - 08, 2016, Macao

DOI: http://dx.doi.org/10.1145/3005358.3005360

ISBN: 978-1-4503-4541-5/16/12

majority of photographs taken by casual users are captured in the middle of the day, when lighting is not ideal. While photo editing software such as 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, along with talented artists.

In this paper, we propose an automatic technique for changing the lighting in a photograph, given a photo collection depicting the same scene under varying viewpoint and illumination. In order to deal with the large variability of appearance changes in outdoor landmarks, we use local color transforms to model the color variations for different parts of the scene. We cast lighting transfer as a colorization 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. Our image relighting method facilitates browsing collections of photographs with harmonized lighting and for generating synthetic timelapses.

Our main contributions are as follows:

- We cast lighting transfer as a colorization problem, learning local color transforms from correspondences and propagating the transforms in an edge-aware manner, using pixel intensities in the source image as a guide.
- We introduce a confidence map to indicate the reliability of propagated transforms, which helps to preserve the color of objects that should not be modified, e.g. spurious people or objects.

2 Related work

Color transfer. The objective of color transfer is to map the colors of one example image to a given image. Global color transfer methods such as [Reinhard et al. 2001; Pouli and Reinhard 2011; Pitie et al. 2005] apply a global color mapping to reshape the color distribution of the input image such that it approaches the color statistical properties of the reference. They work well in style transfer

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. © 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM.

when the input and reference images depict semantically similar scenes, but do not account for spatial layout of the scene. In comparison, our method uses 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. Alternatively, some methods find correspondences across images and transfer a source image using the learned color changes. [Shih et al. 2013] successfully hallucinate different time-of-day images by learning color transformations from time-lapse videos. A similar method by [Laffont et al. 2014] enables drastic appearance transfer by observing color changes in the 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. Our system targets a more general case that does not need image pairs from a static viewpoint. It relies on the vastly available images of the same scene from various online photo communities. [Martin-Brualla et al. 2015] use a simple but effective new temporal filtering approach to stabilize appearance, but it depends on the computed depthmap. [Laffont et al. 2012] 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.

Colorization. Colorization is a computer-assisted process of adding color to a monochrome image or movie. [Levin et al. 2004] use manually specified scribbles and propagate colors based on pixel intensities. The image-guided propagation is based on a simple premise that two neighboring pixels should have similar colors if their intensities are similar. A related method is also used in [An and Pellacini 2008], which propagates rough user edits for spatially-varying image editing. [Liu et al. 2008] decompose image into illumination and reflectance, and transfer color to grayscale reflectance image using corresponding features. Inspired by these approaches, we use image-guided propagation to propagate local color transforms learned at sparse image pixels. We show a comparison of our transform propagation approach and the direct color propagation method for the purpose of lighting transfer in Fig 3.

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 views and under different lighting conditions. We first learn local transforms from sparse correspondences obtained from photo collections. Then, we propagate these transforms with an image-guided method inspired from image colorization. In order to detect potentially inaccurate transforms, we introduce a confidence map to indicate regions where colors are different from the correspondences.

Here is an overview of the pipeline of our approach, which consists of four main parts:

- 1. Sparse correspondences from photo collection. (Section 3.1)
- 2. Learning local color transforms. (Section 3.2)
- 3. Propagation of local color transforms. (Section 3.3)
- 4. Detecting transform outliers. (Section 3.4)

We extend our method to enable lighting transfer from multiple target images in Section 3.5.



Figure 2: Given a pair or source and target images (a-b), our method uses sparse correspondences (c-d) to learn local color transforms, which are then propagated in an image-guided manner (e) in regions with no correspondences.

3.1 Sparse correspondences from photo collections

Our method transfers lighting changes based on sparse correspondences between images. We utilize photo collections of famous landmark, which consist of images of the same scene of different viewpoints and lighting conditions. There are two reasons why we use photo collections. First, these collections with lots of lighting variations provide good examples for our lighting transfer. Besides, we can easily find sparse correspondences across pictures. We first apply structure from motion [Wu et al. 2011] to estimate the parameters of cameras and then use patch-based multi-view stereo [Furukawa and Ponce 2010] 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 correspondences.

3.2 Learning local color transforms

We learn color transforms from the correspondences between images. These color transfers model local color variations across a pair of pictures of the same scene under varying lighting, and depend on the scene geometry and incident lighting.

Given the sparse correspondences between the source image S and target image T, we estimate local transforms [Shih et al. 2013] which represent the color changes for the corresponding pixels in a local neighborhood. We use a linear model [Laffont et al. 2014] to represent the mappings in RGB color space:

$$\underset{A_{k}}{\arg\min} \|v_{k}(T) - A_{k}v_{k}(S)\|_{F}^{2} + \gamma \|A_{k} - G\|_{F}^{2}$$
(1)

$$A_{k} = (v_{k}(T)v_{k}(S)^{T} + \gamma G)(v_{k}(S)v_{k}(S)^{T} + \gamma Id_{3})^{-1}$$
(2)

Here, k corresponds to a specific correspondence. We denote by $v_k(S)$ the patch centered on pixel in the source image and by $v_k(T)$ the corresponding patch in the target image. Both are represented as $3 \times N$ matrices in the RGB color space. G is a global linear matrix estimated on the entire image ($\gamma = 0.01$), used for regularization. The obtained linear transform is represented by a 3×3 matrix A_k .

3.3 Propagation of local color transforms

We then propagate the transforms learned from correspondences to other regions of the source image (Fig. 2). We use the imageguided propagation algorithm introduced by [Levin et al. 2004], which was originally designed to propagate colors to gray images.



Figure 3: We show comparisons of our method to image warping by homography and naive propagation of color.

Here the transforms learned from correspondences are serving as color scribbles in the colorization problem. Instead of propagating the RGB pixel values, we propagate the color transforms estimated in Section 3.2.

We wish to impose the constraint that in a very small neighborhood, two pixels p_j and p_k are more likely to have similar transforms if their color intensities are similar. We formalize this using a set of weights for a pair of pixels p_j and p_k :

$$w_{jk} \propto e^{-\|rgb(j) - rgb(k)\|^2 / 2\sigma_j^2}$$
 (3)

where w_{jk} is a weighting function that sums to one, large when the RGB values of pixel p_j is similar to that of pixel p_k , and small when the two RGB intensities are different. Given transforms A_k at a sparse set of pixels p_k (computed from Section 3.2), the set of local transforms A_j for all pixels p_j in regions with no correspondences can be estimated with a least-squares minimization:

$$\operatorname*{arg\,min}_{A_j} \left(A_j - \sum_k w_{jk} A_k \right)^2 \tag{4}$$

This global optimization problem yields a large, sparse system of linear equations, which can be solved by standard methods. This allows us to propagate sparse transforms to all pixels without correspondences in image. All the A_j are optimized simultaneously. We use the backslash operator in Matlab.

3.4 Detecting transform outliers

The propagated transforms might be inaccurate for regions in the source image with very different texture from correspondences, e.g., the spurious people and green leaves. 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, then its obtained transform is less reliable, as their transforms are propagated from correspondences with a different color.

$$C(p) \propto -ln\left(\sum_{minimumK} \|rgb(p) - rgb(q)\|^2\right)$$
(5)

For each pixel p in the source image, we calculate its color differences with all correspondences q in the image. To avoid occasionality, we sum up the minimum K differences and use the minus natural logarithm of the sum as a confidence factor C(p). All factors are then normalized to [0, 1]. We find that a pixel only needs a few neighboring constraints to get an appropriate transform. We use K = 10 and set a threshold to detect possibly wrong transforms. When applying color transforms, the detected regions remain unchanged as their color in the source image.

3.5 Extending to multiple targets

We extend our method and enable lighting transfer with multiple target images. More target images of the same lighting condition provide more correspondences from different view points. We combine the transforms of the same corresponding pixels in the source image, and propagate using the same method in Section 3.3. We make a comparison on a synthetic dataset between the single-targetimage method and multiple one. It shows that the multiple-target method produces an output image more similar to the ground truth.

4 Results and comparisons

We apply our method on two types of data. First we show results of our method for photo collections of famous landmarks. We then apply our method to synthetic data which allows a comparison to ground truth.

Internet photo collections. We utilize the datasets from [Laffont et al. 2012]. When applying transforms directly to the input image, the noise existing in the input may be magnified. We use bilateral filtering to decompose the input 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 [Shih et al. 2013].

Our method enables dramatic lighting changes between images. Fig. 3 illustrates our results on several scenes, namely St.Basil, Ma-



Figure 4: We illustrate that how our method can be used for harmonizing image collections of multiple views (top row) and hallucinating image time-lapse (bottom row). The insets represent the source images.

narola and RizziHaus. Our method successfully relight the input images, where the image warping based on homography and direct propagation of pixel color fail. Homography is a projective mapping between any two planes with the same center of projection. We estimate the homography based on pixel correspondences. The propagation of pixel color uses code from [Levin et al. 2004]. In Fig. 4 we present that our method can be used for harmonized multi-view image browsing and time-lapse hallucination of a single view scenery.

Synthetic scene. We evaluate the effectiveness of our method on the synthetic dataset of St.Basil [Laffont and Bazin 2015], which contains rendered images from 3 different view points and under 30 lighting conditions. Comparing the result of our lighting transfer, to the ground truth rendering from the same viewpoint with the same lighting condition, quantitative evaluation shows that method using multiple target images produce a more plausible result.

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 propagation of transforms.

5 Conclusion

In this paper we propose method to for transferring lighting across photographs of a static scene. We take as input a photo collection of a famous landmark from different viewpoints and under varying lighting conditions. We use multi-view stereo to reconstruct 3D points, and learn local color transforms from pixel correspondences. The transforms are then propagated to other image regions in an image-guided manner inspired by image colorization techniques. We illustrate that our method can be used for harmonizing image collections of multiple views and hallucinating image timelapse.

Acknowledgements

This research, carried out in part at the BeingThere Centre, is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

References

- AN, X., AND PELLACINI, F. 2008. Appprop: all-pairs appearancespace edit propagation. In ACM Transactions on Graphics (TOG).
- FURUKAWA, Y., AND PONCE, J. 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*
- LAFFONT, P.-Y., AND BAZIN, J.-C. 2015. Intrinsic decomposition of image sequences from local temporal variations. In *ICCV*.
- LAFFONT, P.-Y., BOUSSEAU, A., PARIS, S., DURAND, F., AND DRETTAKIS, G. 2012. Coherent intrinsic images from photo collections. *ACM Transactions on Graphics (TOG)*.
- LAFFONT, P.-Y., REN, Z., TAO, X., QIAN, C., AND HAYS, J. 2014. Transient attributes for high-level understanding and editing of outdoor scenes. ACM Transactions on Graphics (TOG).
- LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2004. Colorization using optimization. In ACM Transactions on Graphics (TOG).
- LIU, X., WAN, L., QU, Y., WONG, T.-T., LIN, S., LEUNG, C.-S., AND HENG, P.-A. 2008. Intrinsic colorization. *ACM Transactions on Graphics (TOG)*.
- MARTIN-BRUALLA, R., GALLUP, D., AND SEITZ, S. M. 2015. Time-lapse mining from internet photos. *ACM Transactions on Graphics (TOG)*.
- PITIE, F., KOKARAM, A. C., AND DAHYOT, R. 2005. Ndimensional probability density function transfer and its application to color transfer. In *ICCV*.
- POULI, T., AND REINHARD, E. 2011. Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics*.
- REINHARD, E., ASHIKHMIN, M., GOOCH, B., AND SHIRLEY, P. 2001. Color transfer between images. *IEEE Computer Graphics and Applications*.
- SHIH, Y., PARIS, S., DURAND, F., AND FREEMAN, W. T. 2013. Data-driven hallucination of different times of day from a single outdoor photo. ACM Transactions on Graphics (TOG).
- WU, C., AGARWAL, S., CURLESS, B., AND SEITZ, S. M. 2011. Multicore bundle adjustment. In *CVPR*, CVPR '11.