Supplementary Material of
Efficient and Robust Color Consistency for Community Photo Collections

Jaesik Park
Intel Labs

Yu-Wing Tai
SenseTime

Sudipta N. Sinha
Microsoft Research

In So Kweon
KAIST

In the supplementary material we present experimental results and demonstrate applications that we had to omit from the main paper due to lack of space.

1. Additional Experiments

1.1. Validation on Synthetic Images

We first tested our approach on two sets of synthetically generated images. In the first set, we randomly changed the white balance and gamma of one image to obtain 11 images (shown in Fig. 1). In the second set, we further modified the colors in hue space, and with non-gamma monotonic tone curve, transformed some images using auto tone and auto color correction function in Photoshop CS6 and converted one image to grayscale (shown in Fig. 2). The extreme color transformations in the second set does not directly follow our model. In both cases, our color correction results were satisfactory and the second set confirmed that our method was robust to unmodeled illumination effects in the input images. Here, the method described in the main paper for matching features in non-rigid scenes was used.

1.2. Comparison with Adobe Lightroom

We present two sets of results – ST. BASIL CATHEDRAL (Figure 3) and DRESDEN FRAUENKIRCHE (Figure 4), where we compare our method with the color adjustment feature in Adobe Photoshop Lightroom[1] which is widely used for batch photo editing and color adjustment. The results using Lightroom were obtained by applying *auto white balance adjustment* and *auto tone* on the respective photo collections. Lightroom recovers images with consistent color histograms where darker images are brightened and exaggerated colors are reduced. In comparison, our method achieves higher color consistency. As described in the main paper, by default our technique utilizes the median intensity values of matched pixels as a soft regularization for the unknown albedos. This default choice causes the whitish appearance in our results (see Fig. 4). However, as shown in the paper and the supplementary video, our approach also enables consistent color transfer from any target image to the whole collection.

[1]Homepage of Adobe Lightroom software. http://www.adobe.com/products/photoshop-lightroom.html

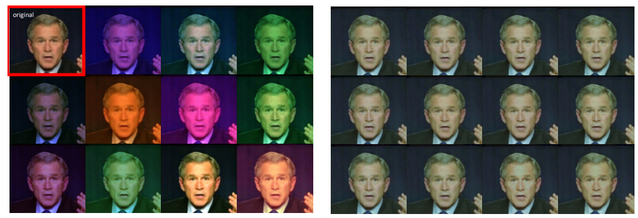


Figure 1. (Left) Synthetic images obtained by applying random white balance and gamma functions. (Right) Our result.

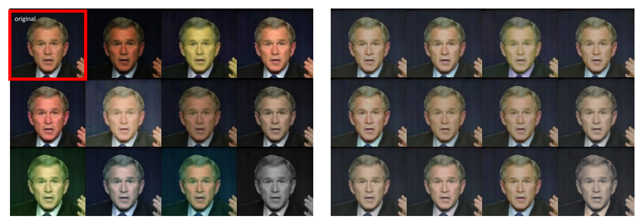


Figure 2. (Left) Various transforms that do not follow our color correction model – color, hue, non-gamma curve, auto tone, auto color and grayscale conversion was applied to generate the inputs. (Right) Our result.

2. Applications

2.1. Image Based Rendering

Hyperlapse photography, a form of motion timelapse was developed by photographers (video posted in geofftompkinson.com or vimeo.com/bzoomi) to convey camera motion in a scene. Originally, careful camera path planning and professional equipment was needed. Although tools for first person video hyperlapse [3] and mobile apps such as Instagram Hyperlapse are available nowadays, they cannot be applied to crowdsourced photos. Snavely *et al*. [5] and Kushal *et al*. [4] addressed automatic source image selection and virtual camera path optimization for rendering smooth 3D photo tours through a reconstructed scene. We show that our color correction method can im-

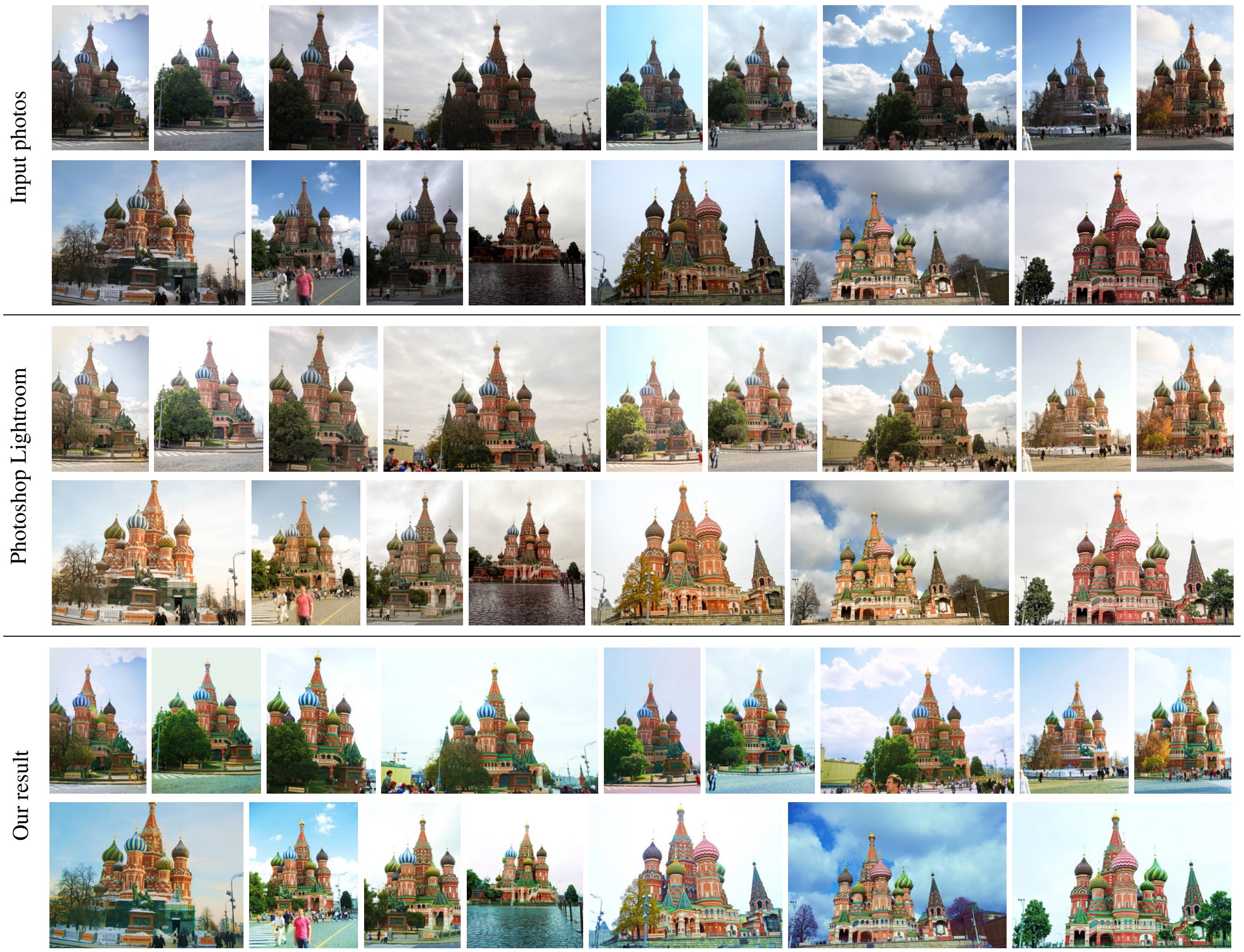


Figure 3. Batch color correction on ST. BASIL CATHEDRAL. (Top) Input photos. (Middle) Results obtained using Adobe Photoshop Lightroom CC (R2015). (Bottom) our results.

prove the appearance and color consistency of such results (see Figure 5 and the supplementary video). The color transfer examples shown in the video give the impression that the imagery was captured at a specific time. We now describe our image-based rendering system.

First, a 2D paint interface overlaid on a top view of the SfM reconstruction was used to guide the selection of the source images along a smooth path (this step can be automated [4, 5]), after which images with extreme focal lengths were removed. For TREVI FOUNTAIN and STATUE OF LIBERTY, subsets of 204 and 123 images were selected for the subsequent rendering. A smooth path for the virtual camera is estimated by fitting a smooth curve that lies near the source camera positions and a fixed look-at point is computed by robustly estimating the 3D point that lies closest to all the camera optical axes. To render video frames, the virtual camera is moved along the smooth path and for each location, the nearest left and right source cameras are used to perform two-image view dependent texture mapping. The two rendered images are blended with weights proportional to the distance to the virtual camera position. For rendering an image from a novel viewpoint near one of the source cameras, we used per-camera dense depth map proxies obtained from the sparse SfM point cloud. Specifically, the sparse depth values obtained from 3D points visible in an image were interpolated using an edge aware technique. Hole-filling on image borders was done using the pull-push method [2].

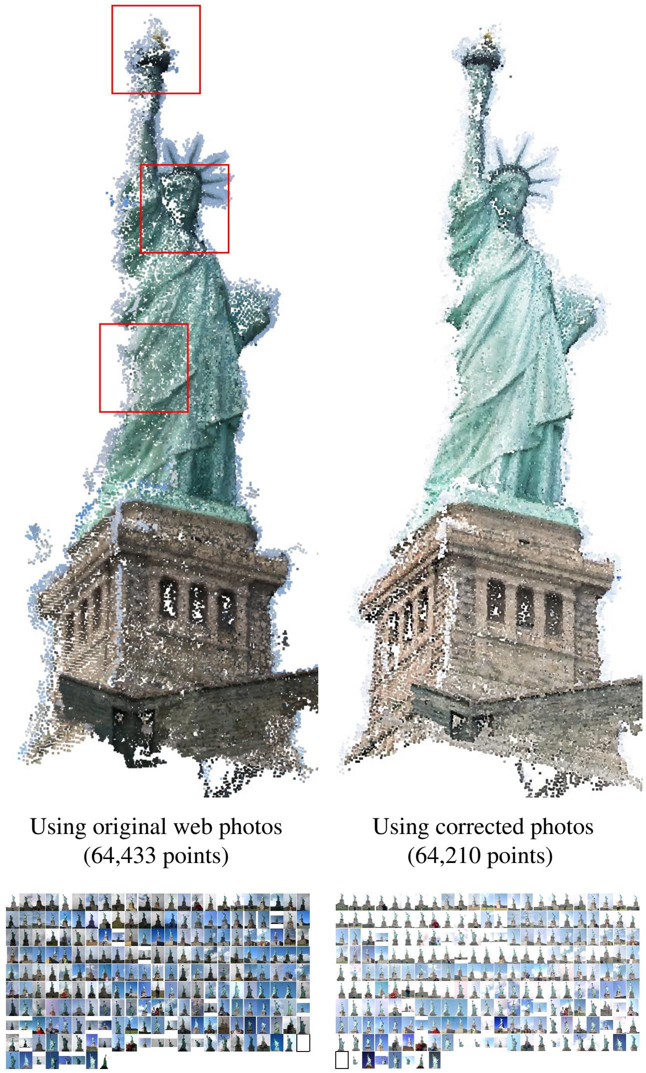
2.2. Image Stitching

Our technique can also improve image stitching on crowdsourced images. Figure 6 shows panoramas created using the Photomerge feature in Photoshop CS6 from the DRESDEN FRAUENKIRCHE and TREVI-FOUNTAIN collections. These results were stitched from images captured with different cameras from nearby but different view-

Figure 4. Batch color correction on DRESDEN FRAUENKIRCHE. (Top) Input photos. (Middle) Results obtained using Adobe Photoshop Lightroom CC (R2015). (Bottom) our results. See discussion in Sec. 1.2.

Figure 5. Image based rendering: (Top) Source images. (Middle) Novel views rendered from a stabilized camera path. (Bottom) Novel views obtained using images corrected with our technique. The supplementary video demonstrates more results.

Figure 6. Image stitching. (a) Input images. (b)-(d) Results obtained using (b) original input images, (c) images corrected using the Auto Color feature in Adobe Photoshop CS6 and (d) images corrected by our method. The same image order is applied on (e)-(h). Our corrections are with respect to the entire collections of DRESDEN FRAUENKIRCHE and TREVI-FOUNTAIN. All images were stitched using the Photomerge feature in Adobe Photoshop CS6.



Using original web photos (64,433 points)

Using corrected photos (64,210 points)

Original web photos

Corrected photos

Figure 7. 3D reconstruction of STATUE OF LIBERTY using two photo collections: original web photos (left side) and color corrected photos (right side). The red boxes highlight that the 3D reconstruction obtained from the original photos is less reliable than the result obtained from the corrected photos.

points. Figure 6(b) and (f) shows the results obtained using the original images – the first result has inconsistent brightness on different structures whereas the second result has inconsistent colors (whitish and yellowish hues) on the two sides of the same building. Figure 6(c) and (g) shows results obtained when Photoshop's Auto Color feature was used for color correction. Although the results are better, overall color consistency is still lacking. With our correction technique, both type of artifacts are significantly reduced as shown in Figure 6(d) and (h).

2.3. Multiview stereo

As described in Sec. 2.2, consistent intensities in the color corrected images help the image matching. Figure 2.1 shows another application on the multiview stereo that indicates this benefit. For the experiment, we apply the Structure from Motion (SfM) [6] and the consequent Multiview Stereo [1] pipelines using two image sets: original images and color corrected images. The 220 images of the STATUE OF LIBERTY dataset are used for the displayed result. It is notable that the 3D points reconstructed from color corrected photos show the more reliable structures although the number of recovered 3D points is slightly smaller.

References

[1] Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(8):1362–1376, Apr. 2010.

[2] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 43–54, New York, NY, USA, 1996. ACM.

[3] J. Kopf, M. Cohen, and R. Szeliski. First-person hyperlapse videos. *ACM Trans. Graphics (SIGGRAPH)*, 33(4), 2014.

[4] A. Kushal, B. Self, Y. Furukawa, D. Gallup, C. Hernandez, B. Curless, and S. M. Seitz. Photo tours. In *3DImPVT*, 2012.

[5] N. Snavely, R. Garg, S. M. Seitz, and R. Szeliski. Finding paths through the world's photos. *ACM Transactions on Graphics (TOG)*, 27(3):11–21, 2008.

[6] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3d. In *Proceedings of ACM SIGGRAPH*, pages 835–846. ACM Press, 2006.