Object Image Relighting through Patch Match Warping and Color Transfer

Xin Jin[1,] *, Yulu Tian[1], Ningning Liu[3], Chaochen Ye[1], Jingying Chi[2], Xiaodong Li[1,] *, Geng Zhao[1]

[1]*Beijing Electronic Science and Technology Institute, Beijing 100070, China*

[2]*Beijing University of Chemical and Technology, Beijing 100029, China*

[3]*University of International Business and Economics, School of Information technology and Management, Beijing, 100029, .China*

[*]Corresponding Authors: {jinxin, lxd}@besti.edu.cn

***Abstract*—In this paper, we propose a novel method for image based object relighting through patch match warping and color transfer. According to our method, a pair of reference object images (A and B), without any knowledge of the 3D geometry or material information of the input object image, are needed. The input image has similar illumination condition as that of the reference image A. The reference image B has the target illumination condition. We first modify the color of the reference object images according to the color of the input object image. Then the modified reference image A is warped to the input image using the image patch match. The reference B has the same warping to that of A. At last, the illumination of the reference B is transferred to the input image through a local and global transfer model learned from the warped reference B and A. We test our method on the ALOI dataset (Amsterdam Library of Object Images) and acquire convincing relit result on numerous input and reference object images with different illumination conditions.**

Keywords—Object relighting; Image based relighting; Patch match; Color transfer; Illumination transfer

I. Introduction

Image-based photo-realistic virtual relighting without 3D model has been extensively studied in the visual computing community and widely used in visual design, digital entertainment, film production, etc.

Image-based virtual relighting can be coarsely classified into face relighting [1][2][3][4][5], scene relighting [6] and object relighting [7][9]. Besides the 3D reconstruction methods [5], another research line is to use reference images and is more practical for real world applications. In this line, the face relighting is the most extensively studied. The references are from a collection of face images [8], 2 reference face images [4] to a single reference face image [1][2]. The scene relighting [6] and the object relighting [9] still need a collection of reference images.

For more convenient use and wider application. This work is to put forward the object relighting task to only requiring 2 reference object images. Our objective is to generate photo-realistic relighting result of an object image taken under one illumination condition, so as to make the result as similar as possible in illumination effects to those of a reference object image (Ref. B in short) under another illumination. We also need a reference object image (Ref. A in short) taken under similar illumination condition to that of the input image, as shown in Fig. 1.

Human faces have similar geometry and skin reflectance. According to current image based face relighting methods, faces are first aligned with pre-defined landmarks manually or automatically with manually refine [1][2][3][4][5]. Then the reference face(s) is/are warped according to the landmarks of the input faces. After that the illumination of the reference face(s) is transferred to the input face using quotient image [4][8], man-made illumination templates [3], or edge preserving filters [1][2] etc.

Unlike human faces, general objects such as cars, cups, teapots have various geometry shapes and materials. Current face alignment algorithm for landmark detection cannot be directly used for general objects. Besides, the assumption of similar skin reflectance is no longer suitable for general objects.

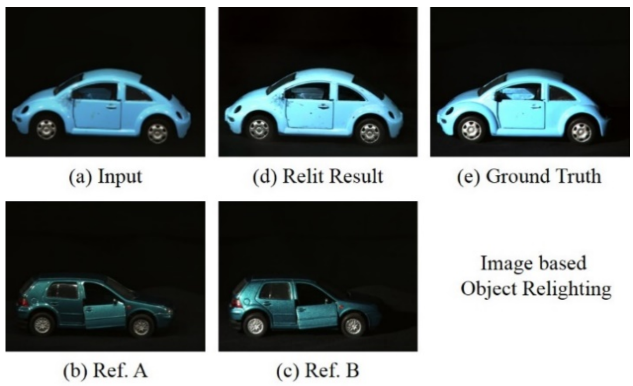

Fig. 1. The objective. (a) is the input object image. (b) and (c) are the Ref. A and the Ref. B. (d) is the relit result by our method. (e) is the ground truth of the input image taken under the illumination condition of Ref. B.

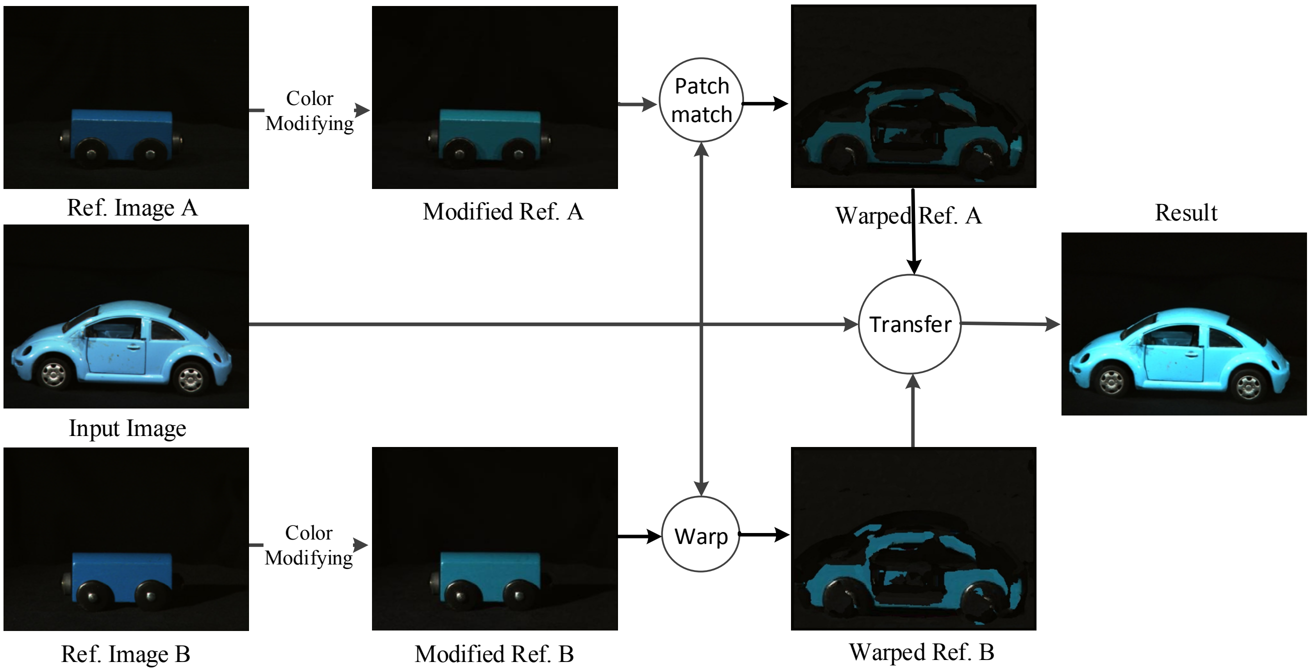

Fig. 2. The workflow of the proposed method. The Ref. image A and the input image are taken under similar illumination conditions. First, the colors of both Ref. A and B are modified according to that of the input image using the color transfer algorithm [10]. Then, the modified Ref. A is warped using the patch match method according to the input image. The modified Ref. B is warped using the same warping as the Ref. A use. At last, the warped. Ref. A and the warped Ref. B are used to transfer the illumination of the warped Ref. A to the input image through the local and global transfer model.

Shih et al. [6] propose a scene relighting method. They need a time-lapse video as the reference. They formulate the warping problem as a Markov Random Field (MRF) using a data term and pairwise term. However, obtaining the time-lapse video of the same scene is a non-trivial task.

Inspired by face relighting methods, we first use color transfer [10] to modify the color of the 2 reference object images according to the color of the input image.Thus the color difference between the reference images and the input image becomes small. The face alignment algorithms cannot be used for general objects. We need an alignment free method to warp the reference object images according to the shape of the input object image. The MRF based warping method requires no alignment algorithms, however it needs a time-lapse video. Thus, we introduce a patch match [11] based warping method to warp the reference object images. At last, the illumination of the reference B is transferred to the input image through a local and global transfer model learned from the warped reference B and A. We test our method on the ALOI dataset (Amsterdam Library of Object Images) and acquire convincing relit result on numerous input and reference object images with different illumination conditions.

Our main contributions include: (1) The first work for image based object relighting that only need a pair of another reference object images, (2) a local and global transfer model based on CIELAB color space for illumination transfer.

The following paper is organized as follows. In section II, we describe the proposed method for image based object relighting through patch match warping and color transfer. In section III, we will show our experimental results on the ALOI dataset. Section IV gives the conclusion and discussion.

II. OBJECT IMAGE RELIGHTING

The workflow of our method is illustrated in Fig. 2. The color modifying, the patch match warping, and the local and global transfer model will be described in this section.

A. *Color Modifying*

Human faces always have similar skin reflectance in the same ethnicity. Thus the image based face relighting methods often make such an assumption and the illumination transfer only happen in the illumination component. However, the general objects have various colors. The above assumption does not work anymore. Inspired by the face relighting methods, we first modify the colors of the Ref. A and Ref. B to that of the input image. In this work we use the color transfer method in [10] to modify the colors of the Ref. A and B.

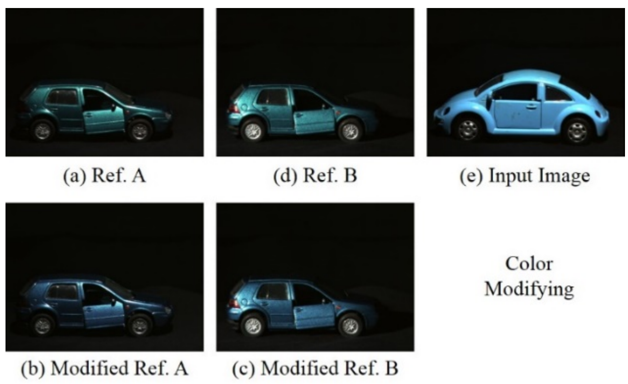

Fig. 3. The color modifying. (a) and (d) are the original reference images. (e) is the input image. (b) and (c) are color modified version of (a) and (d).

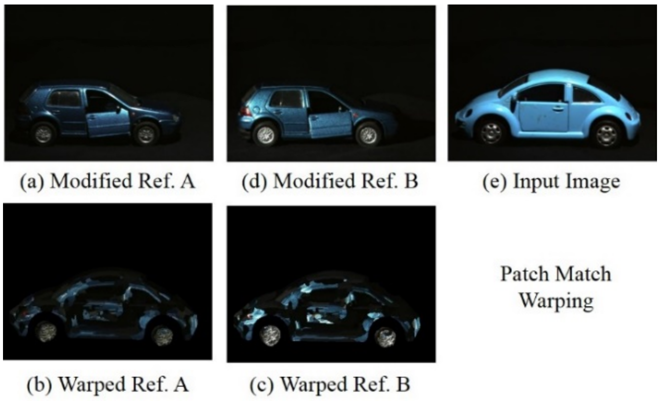

Fig. 4. The patch match warping. (a) and (d) are the modified Ref. A and Ref. B. (e) is the input image. (b) and (c) are the warped version of (a) and (d) according to the input image (e). Note that the pixels outside the object in (b) and (c) do not influence the relit result and can be ignored.

As shown in Fig. 3, we first compute the color themes (TA, TB and TI) of the 2 reference images and the input image. Then, the color theme TA and TB are mapped to TI using the manifold preserving propagation proposed by Chen et al. [10].

B. *Patch Match Warping*

Unlike human faces, who have similar geometry shapes, general objects often have various shapes even with the same sematic label. We seek to pair each pixel in the input image I with a pixel in the color modified Ref. A. However, such a dense corresponds is with high time consuming. Thus, for each patch in I, we seek a patch in A that looks similar to it. L_2 norm is used over square patches of side length $2r+1$. For pixels $p \in I$ and the corresponding pixel $q \in A$, we minimize the following term:

$$\sum_{i=-r}^{r} \sum_{j=-r}^{r} \left\| I(x_p+i, y_p+j) + A(x_p+i, y_p+j) \right\|^2 \tag{1}$$

where (x_p, y_p) is the coordinate of a pixel in the patch. We solve this using the generalized patch match correspondence algorithm [11].

As shown in Fig. 4, the color modified Ref. A is warped using the above patch match warping method according to the input image I. Then the color modified Ref. B is warped using the same correspondence derived from the warping of the color modified Ref. A. Such a setting is because that the modified Ref. A is taken under similar illumination condition to that of the input image. While, the modified Ref. B is taken under another different illumination from that of the input image. The input image and the modified Ref. A have more similar patches than those in the input image and the modified Ref. B.

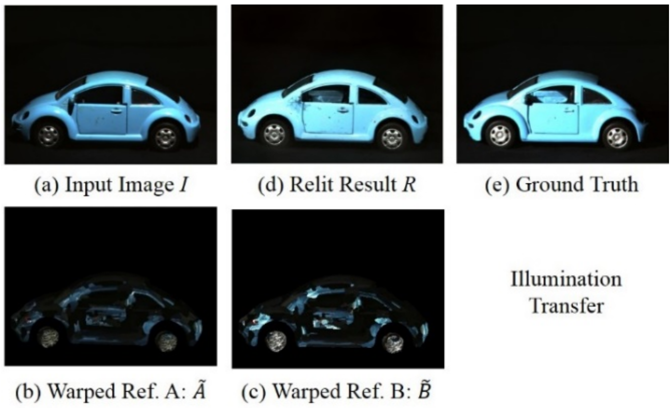

Fig. 5. The illumination transfer. (a) is the input image I. (b) and (c) are the warped Ref. A (denoted as $\tilde{A}$) and the warped Ref. B (denoted as $\tilde{B}$). (d) is the relit result R. (e) is the ground truth image of the input object really taken under the illumination of the Ref. B. Note that the pixels outside the object in (b) and (c) do not influence the relit result and can be ignored.

C. *The Local and Global Transfer*

We propose an example based local and global transfer model for illumination transfer. The transfer model is learned from the warped Ref. A (denoted as $\tilde{A}$) and the warped Ref. B (denoted as $\tilde{B}$) and is applied to the input image I . The output of this step is the relit result (denoted as *R*), (see Fig. 5).

We decouple the image into lightness and color, and the illumination effects are considered mainly retained on lightness. We choose CIE 1976 (L*, a*, b*) color space, as it could separate color image to lightness and color well, L* channel contains lightness information (similar to human perception lightness), and a* channel and b* channel contain color information. All the operations on the above image I, $\tilde{A}$, $\tilde{B}$, R are computed in the L* channel.

$\alpha + \beta = \chi$. (1) (1)

We use $P_k(\cdot)$ to denote the k[th] patch of an image given in argument. For a patch containing N pixels, $P_k(\cdot)$ is a $1\times$ N vector. The local transformation is represented as T_k for each patch. The first term in the energy function models the transformation from $\tilde{A}$ to $\tilde{B}$. We use the L_2 norm to represent this transformation as:

$$\sum_k \left\| P_k(\tilde{B}) - T_k(\tilde{A}) \right\|^2 \tag{2}$$

We use the same transformation T_k to convert the input image I to the output relit result R as:

$$\sum_k \left\| P_k(R) - T_k(I) \right\|^2 \tag{3}$$

Now, we need a regularization term to avoid overfitting. We choose the global transformation G from the entire image of $\tilde{A}$ to $\tilde{B}$. Thus, the whole energy function of our local and global transfer model is defined as:

$$R = \arg\min_{R,\{T_k\}} \sum_k \left\| P_k(\tilde{B}) - T_k(\tilde{A}) \right\|^2$$
$$+a \sum_k \left\| P_k(\tilde{B}) - T_k(\tilde{A}) \right\|^2$$
$$+b \sum_k \left\| T_k - G \right\|^2$$

Where a and b are the relative importance of each term. The above minimization can be solved as a standard local liner regression [6]. For all results in this paper, we use a = 0:01, b = 1 (pixel value 2 [0; 255]), N = 25 (5×5 patch).

III. The Experimental Results

We test our method on the ALOI dataset [7]. ALOI is a color image collection of one-thousand small objects, recorded for scientific purposes. In order to capture the sensory variation in object recordings, they systematically varied viewing angle, illumination angle, and illumination color for each object, and additionally captured wide-baseline stereo images. We recorded over a hundred images of each object, yielding a total of 110,250 images for the collection.

For each object, they have illumination direction varied in 24 configurations [7]. Each object is recorded with only one out of five lights turned on, yielding five different illumination angles (conditions l1-l5). By switching the camera, and turning the stage towards that camera, the illumination bow is virtually turned by 15 (camera c2) and 30 degrees (camera c3), respectively. Hence, the aspect of the objects viewed by each camera is identical, but light direction has shifted by 15 and 30 degrees in azimuth. In total, this results in 15 different illumination angles.

Furthermore, combinations of lights were used to illuminate the object. Turning on two lights at the sides of the object yielded an oblique illumination from right (condition l6) and left (condition l7). Turning on all lights (condition l8) yields a sort of hemispherical illumination, although restricted to a more narrow illumination sector than true hemisphere. In this way, a total of 24 different illumination conditions were generated, conditions l[1..8]c[1..3](1…24).

Thus we choose the input object and the reference one with similar sematic label and shape for relighting. The input image and the reference image A have the same illumination condition. The reference image B has another illumination condition. Each result has a ground truth to compare, which has the same illumination condition as that of the reference image B.

A. Color Modifying Results

As shown in Fig. 6, we show color modifying results according to the input image descripted in Section 2.1. The 6 reference images are taken from 6 different illumination conditions.

We compare the relit results with and without the color modifying descripted in Section 2.1. As shown in Fig. 7, the color modifying obviously obtain more convincing relit results than directly relighting without color modifying. The color of the reference has directly influence on the relit results. Some color of the reference is preserved in the relit results, which makes the results with many obvious artifacts.

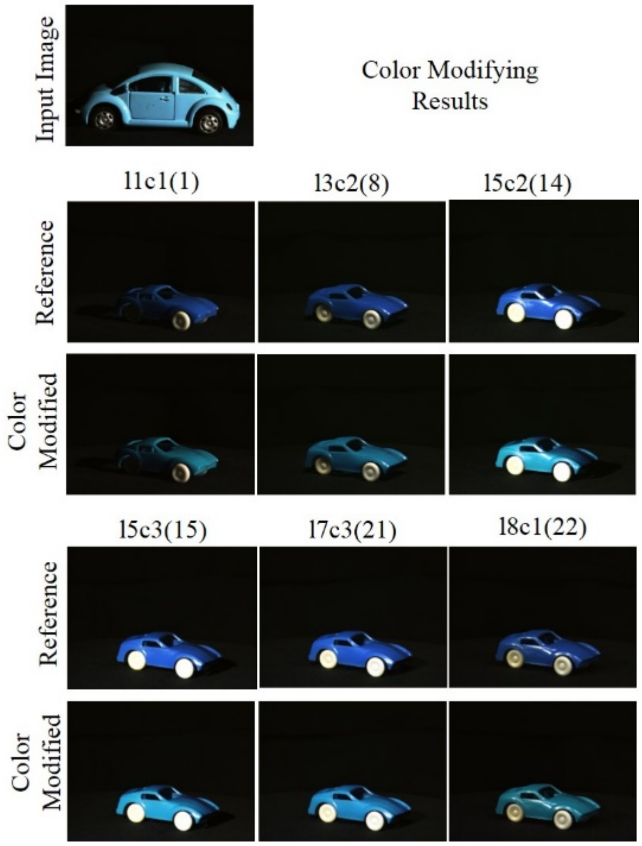

Fig. 6. The color modifying results.

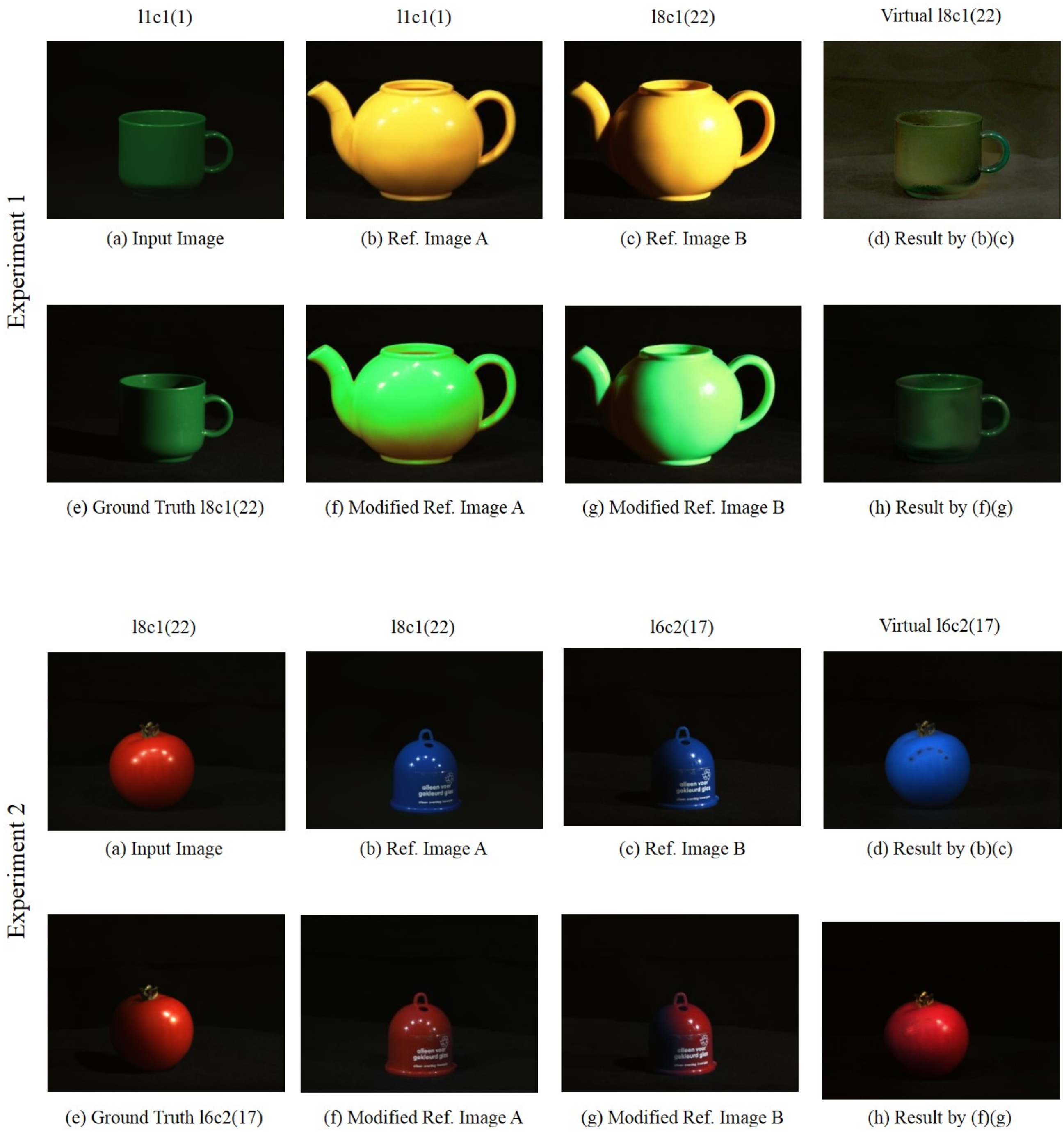

Fig. 7. Relighting with and without color modifying. (a) is the input image. (b) and (c) are the original reference image A and B. (d) is the relit result directly using (b) and (c) without color modifying. (f) and (g) are the color modified version of (b) and (c). (h) is the relit result using (f) and (g). (e) is the ground truth.

B. One Input, Multiple Reference Objects

We test our method on one input with multiple reference objects. The reference images B of multiple reference objects are taken under the same illumination condition. As shown in Fig. 8, to obtain novel illumination effects, we can use different reference objects with different sematic label and similar geometry and material to relight the input image. The relit results by different objects are similar with slightly different in color, which is caused by the color modifying algorithm.

C. One Input, Multiple Reference Illumination

We also test our method on one input with multiple illumination of the same reference object. This is a strong ability that allow one shot an object once and automatically generate a collection of images of that object in various illumination condition virtually.

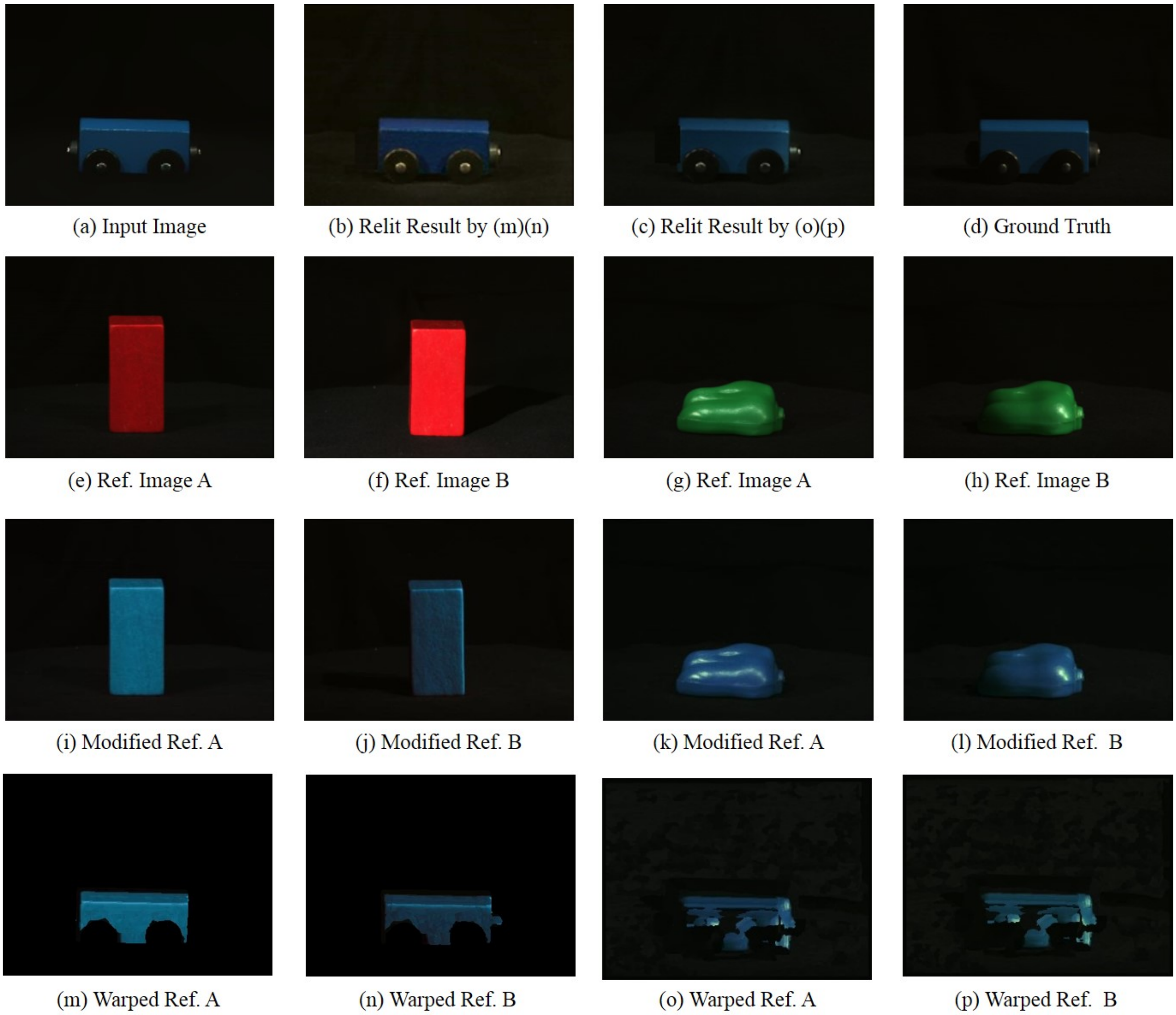

Fig. 8. One input, multiple reference objects. (a) is the input image "wood car"; (e) and (f) / (g) and (h) are the original reference image A and B of 2 different objects "cylinder" and "pepper"; (i) and (j) / (k) and (l) are the color modified version of (e) and (f) / (g) and (h); (m) and (n) / (o) and (p) are the warped version of (i) and (j) / (k) and (l); (b) is the relit result using (m) and (n); (c) is the relit result using (o) and (p). (d) is the ground truth.

As shown in Fig. 9, for one input image of one object, we choose another object as its reference with multiple illuminations. We can virtually generate multiple results of the input object under multiple illumination conditions.

IV. Conclusion and Discussion

In this work, we propose an image based object virtual relighting method. According to our method, only an image pair of another reference object taken under different illumination conditions is needed. Our method contains color modifying, patch match warping and illumination transfer. In addition, a local and global transfer model based on CIELAB color space for illumination transfer is proposed. We test our method in the ALOI dataset. We acquire convincing relit results on multiple objects.

In this work, the object matching is not taken into the consideration, which is done manually. In the future work, we will first add this feature into our framework to automatically find appropriate object for the input object relighting. Then we will optimize all the steps in our framework together instead of optimizing them independently in current version.

Acknowledgment

This work is partially supported by the National Natural Science Foundation of China (Grant No.61402021, No.61402023), the Science and Technology Project of the State Archives Administrator. (Grant No.2015-B-10), the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No. BUAA-VR-16KF-09), the Fundamental Research Funds for the Central Universities (No. 2014GCYY04, 2014GCYY02).

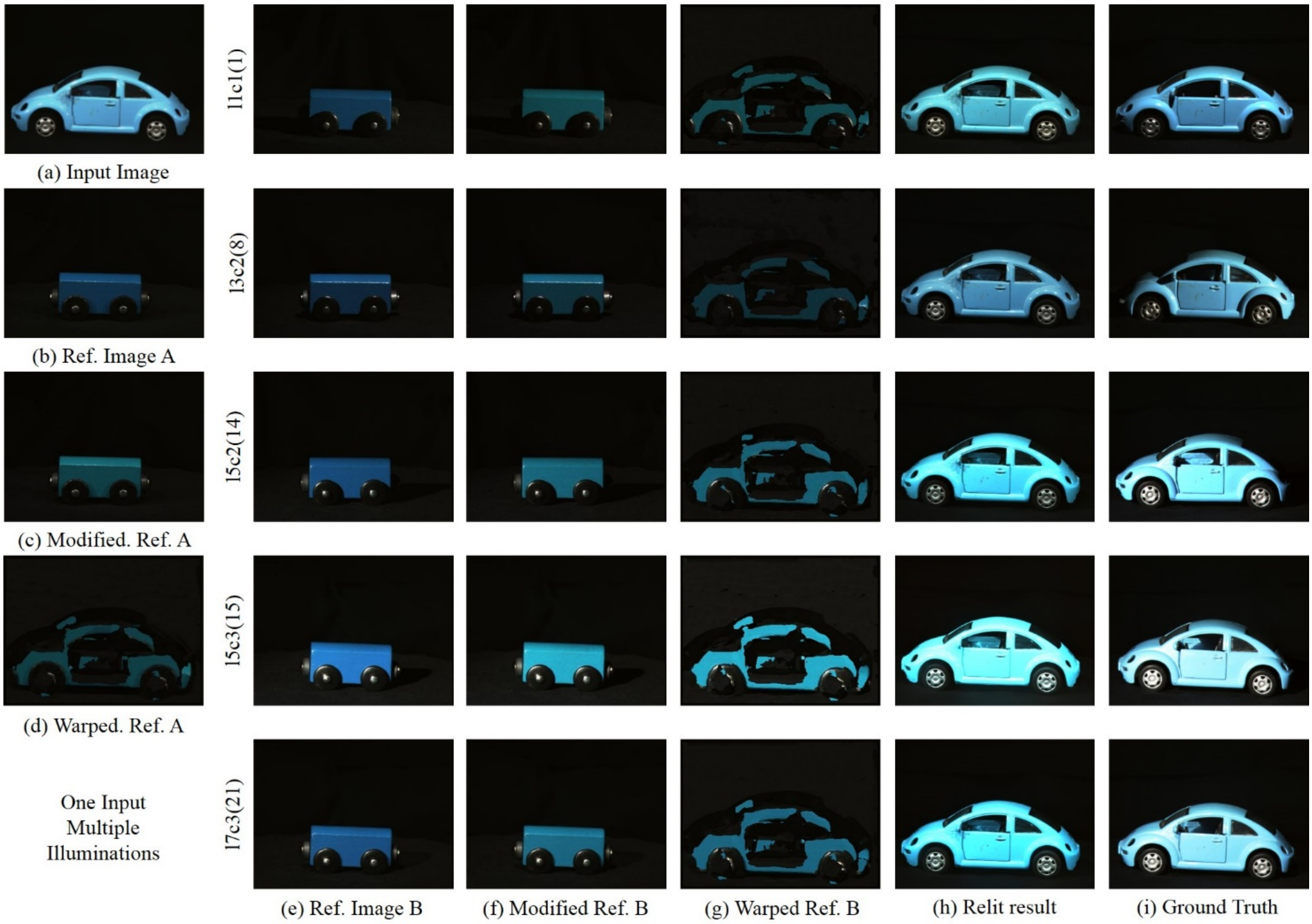

Fig. 9. One input, multiple reference illuminations. (a) is the input image "car"; (b) is the reference image A; (c) is the color modified reference A; (d) is the warped reference A. In this experiment, we choose 5 reference images Bs, as shown in the (e) column. The color modified reference Bs are shown in the (f) column. The warped reference Bs are shown in the (g) column. Our relit results are shown in the (h) column. The (g) column shows the corresponding ground truth images.

REFERENCES

[1] Xiaowu Chen, Hongyu Wu, Xin Jin and Qinping Zhao. Face Illumination Manipulation using a Single Reference Image by Adaptive Layer Decomposition. IEEE Transactions on Image Processing (TIP), Vol. 22, No. 11, pp. 4249 - 4259, Nov. 2013.

[2] Xiaowu Chen, Mengmeng Chen, Xin Jin and Qinping Zhao. Face Illumination Transfer through Edge-preserving Filters. Computer Vision and Pattern Recognition (CVPR), Colorado Springs, June 20-25, pp.281-287, 2011.

[3] Xiaowu Chen, Xin Jin, Qinping Zhao and Hongyu Wu. Artistic Illumination Transfer for Portraits. Computer Graphics Forum (CGF), Vol. 31, No. 4, pp.1425-1434, 2012.

[4] Jiansheng Chen, Guangda Su, Jinping He, Shenglan Ben. Face Image Relighting using Locally Constrained Global Optimization. Computer Vision – ECCV 2010 Volume 6314 of the series Lecture Notes in Computer Science pp 44-57.

[5] Y. Wang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras. Face re-lighting from a single image under harsh lighting conditions. In Proc. CVPR, 2007.

[6] Yichang Shih, Sylvain Paris, Frédo Durand, and William T. Freeman. 2013. Data-driven hallucination of different times of day from a single outdoor photo. ACM Trans. Graph. 32, 6, Article 200 (November 2013), 11 pages.

[7] Jan-Mark Geusebroek, Gertjan J. Burghouts, and Arnold W. M. Smeulders. 2005. The Amsterdam Library of Object Images. Int. J. Comput. Vision 61, 1 (January 2005), 103-112.

[8] P. Peers, N. Tamura, W. Matusik, and P. E. Debevec. Postproduction facial performance relighting using reflectance transfer. In Proc. ACM SIGGRAPH, 2007.

[9] Haber T, Fuchs C, Bekaer P, et al. Relighting objects from image collections. Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 627-634.

[10] Chen X, Zou D, Zhao Q, et al. Manifold preserving edit propagation[J]. ACM Transactions on Graphics (TOG), 2012, 31(6): 132.

Barnes, C., Shechtman, E., Goldman, D., and Finkelstein, A. 2010. The generalized patchmatch correspondence algorithm. Computer Vision–ECCV 2010, 29–43.