

# JH2R: Joint Homography Estimation for Highlight Removal

Sungmin Eum  
smeum@umiacs.umd.edu

Hyungtae Lee  
htlee@umiacs.umd.edu

David Doermann  
doermann@umiacs.umd.edu

University of Maryland Institute of  
Advanced Computer Studies (UMIACS)  
University of Maryland College Park  
College Park, MD, USA

---

## Abstract

This paper addresses the problem of removing highlight regions caused by the light sources reflecting off glossy surfaces in indoor environments. We devise an efficient method to detect and remove the highlights from the target scene by jointly estimating separate homographies for the target scene and the highlights. Our method is based on the observation that when given two images captured at different viewpoints, the displacement of the target scene is different from that of the highlight regions. We show the effectiveness of our method in removing the highlight reflections by comparing it with the related state-of-the-art methods. Unlike the previous methods, our method has the ability to handle saturated and relatively large highlights which completely obscure the content underneath.

## 1 Introduction

Imagine being in an art museum or any other indoor environment where there are numerous paintings, pictures, documents or posters held inside glass-frames for protection. There are pieces which you wish to capture using a camera, but you experience difficulty avoiding highlights which are generated by bright indoor lighting reflected off the glossy surfaces. Similar problems occur when trying to capture contents off of whiteboards, documents printed on glossy surfaces or objects such as books or CDs with plastic covers. Figure 1(a) illustrates typical examples.

In this paper, we address the problem of removing unwanted highlight regions in images generated by reflections of light sources on glossy surfaces. Although there have been efforts made to synthetically fill in the missing regions using the neighboring patterns by applying methods like inpainting [1, 2], it is impossible to recover the actual missing information in completely saturated regions. Therefore, it is prudent to consider using multiple images where corresponding regions are not covered by the saturated highlights.

We make the following observations in devising our approach:

- The distance between the camera and the virtual location of the light source is typically larger than the distance between the camera and the target content. (Figure



Figure 1: (a) Examples of highlights shown on the glossy surfaces obscuring the desired content and degrading visual quality (b) Result (right) obtained using our algorithm to remove the highlights using two images (left and middle) captured at different viewpoints

- 2). Thus, it is reasonable to use two separate homographies in distinguishing the objects at difference distances. [7]

  - When two images are captured with a change of view point, the displacement of the desired content is different from the displacement of the highlight regions. This is referred to as ‘motion parallax’.

Our method works with two images with slightly different viewpoints and applies a novel algorithm called, Joint Homography Estimation for Highlight Removal (JH2R) which performs a fast joint estimation of the two homographies, foreground and highlight, and provides a visually pleasing output with the highlights removed. (Figure 1(b))

To the best of our knowledge, no previous work has addressed an approach which can successfully handle relatively large and saturated highlight regions obscuring the content underneath. We show the effectiveness of our approach by comparing it with closely related state-of-the-art methods.

## 2 Related Work

Several methods have been suggested to explicitly address highlight issues based on the dichromatic reflection model [23]. Tan et al. [25] uses a user-assisted inpainting and show that highlight pixels contain useful information for highlight removal. Similarly, [26] asserts that the color texture data lying outside the highlights can assist in filling in the missing diffuse surface colors inside the highlights. Yang et al. [28, 29] introduced a method which propagates the diffuse color information into the highlight regions using an iterative bilateral filter. Tan et al. [7] proposed a local operation based method which does not require explicit color segmentation. They strongly assume that surface color is chromatic and ignores cases with saturated regions.

Solutions based on reflection removal or layer separation can also be taken into consideration. Some suggest that it is possible to solve this ill-posed problem using a single image supported by additional priors. Levin et al. [8] showed that layer decomposition can be performed by minimizing the total number of edges and corners. In [10], the prior information for layer separation is strengthened by bringing the user into the loop for manual gradient labeling. Li and Brown [12] recently suggested an approach which assumes that one layer is smoother than the other. Since all of these methods use only one image as input, it is virtually impossible to recover the content obscured by the highly-saturated or large highlights unless the region is homogeneous and smooth.

Numerous approaches to exploit multiple images have also been explored. Some approaches have used the polarizing effect on specularities [6, 11, 20, 21] while others have

used focus [22] or flash [10] as priors. However, using polarizers, different focuses or flash may require use of additional hardware which is not always feasible or convenient for typical users.

Techniques using multiple images with different viewpoints have also proven effective. Szeliski et al. [24] showed that relative motion between the layers in multiple images can be used effectively. In [13], gradients across the aligned image set are used to distinguish pixels in different layers. Lin et al. [15, 16] integrated color analysis and multi-baseline stereo. This, however, requires large set ( $>50$ ) of images captured by moving the camera along a linear path with constant velocity. The approach also suffers when images contain color saturations. Recently, Guo et al. [8] showed that by harnessing correlation, sparsity, and the independence prior, reflection separation can be performed.

These methods share a similar perspective with our approach in that they use multiple viewpoints and incorporate the relative motion difference in different layers. However, our method does not employ any sophisticated optimization which usually requires significant processing time [8, 13], nor does it require any user intervention [8]. Most importantly, unlike others, our method uses the relationship between the highlight regions resulting in more robust removal of saturated highlights. A detailed comparison is presented in the experiments section where our method is shown to outperform the representative state-of-the-arts.

## 3 Our Method

### 3.1 Overview

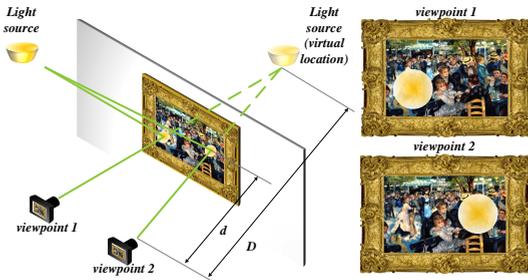


Figure 2: The illustration depicts the overhead view of the camera, the desired content, and the light source.

Our method was motivated by a widely acknowledged physical phenomenon known as ‘motion parallax’. Motion parallax states that as the viewer moves, the movement of the objects in the vicinity is greater across the field of view than those in the distance. A driver can easily observe that the objects close to the window (e.g., roadside traffic signs) pass by quickly while those in distance (e.g., clouds) remain in one’s field of view longer.

Without loss of generality, we can similarly view the relationship between the desired content (e.g., a painting) and the highlights as shown in Figure 2. Since the highlights caused by the light source are the result of the reflection on the glossy surface before they reach the camera, the light source can be modeled to virtually exist on the other side of the content. Note that the distances of the two sources (target content and light) from the camera are different. Unless the light source is attached on the same wall as the painting, in which case

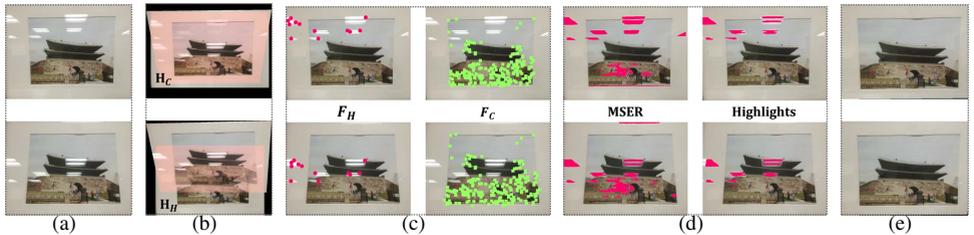


Figure 3: Schematic overview of our method (a) Input images (b) Joint homography estimation (c) Feature-level labeling (d) Pixel-level labeling (e) Final results

no reflection would exist, the distances can never be the same. In fact, the distance from the light source is always larger than the distance from the content ( $D > d$ , in Figure 2).

In order to distinguish the movements of the highlights, we need at least two images captured in different views. We detect where the highlights are by searching for the two separate homography matrices: one for the content ( $\mathbf{H}_C$ ) and the other for the highlights ( $\mathbf{H}_H$ ). Applying two different homographies for scenes at different distances proved to be effective by Gao et al. in [14]. We exploit the fact that the homography ( $\mathbf{H}_C$ ) which can properly overlay the desired contents in the two images will generate an erroneous overlap between the corresponding highlight regions. Similarly, the desired contents will display incorrect overlap when  $\mathbf{H}_H$  is employed. This is shown in the second step of Figure 3(b).

Unlike the intrinsic layer separation problem, removing the saturated highlights from images requires another image which can provide the corresponding non-highlight pixels. To perform such “pixel-transfer”, it is necessary to have the pixel-level detection results of the highlights. In our approach, we first detect the highlight regions at the feature level by jointly estimating the two homographies using the proposed JH2R algorithm. Then  $\mathbf{H}_H$  is used to estimate the highlight regions at the pixel-level. Finally, we remove the highlights in both of the images by transferring the corresponding pixels from the complementary image using Poisson blending [19]. Figure 3 shows the schematic overview of our method. Details on each steps of the algorithm are explained in the following subsections.

### 3.2 Joint homography estimation and highlight feature labeling

In our approach, we attempt to estimate the two different homographies. We devise a novel, yet efficient algorithm which only requires feature correspondences between the two images along with Maximally Stable Extremal Region (MSER) [18] features for those images as input. Although we have utilized the SIFT [17] features in our implementation, any type of feature extractor and descriptor can be used as long as the features can be stably matched throughout the image including the highlight regions. Before triggering our algorithm, a set of all the feature correspondences ( $F$ ) is acquired by thresholding the Euclidean distances between tentative feature pairs as described in [17]. Our algorithm is shown in Algorithm 1. Note that  $F$  and  $M$  represent the set of all feature correspondences and the set of all MSER features, respectively. The framework of our algorithm was inspired by the Random Sample Consensus algorithm [9].

Our algorithm begins by estimating the homography for the content ( $\mathbf{H}_C$ ) using four randomly selected feature correspondences from  $F$ . Using  $\mathbf{H}_C$ , we temporarily label all the feature correspondences in  $F$  as either the content feature  $F_C$  or the outlier feature  $F_O$  by

thresholding ( $T$ ) their symmetric transfer errors [9]. The threshold  $T$  is empirically acquired. For estimating the symmetric transfer error of a feature correspondence  $F_i$ , we consider both the forward ( $\mathbf{H}_C$ ) and backward ( $\mathbf{H}_C^{-1}$ ) transformations and use them to compute the sum of geometric errors as follows:

$$e(F_i, \mathbf{H}_C) = d(x_i, \mathbf{H}^{-1}x'_i)^2 + d(x'_i, \mathbf{H}x_i)^2, \quad (1)$$

where  $x_i$  and  $x'_i$  are the corresponding feature points in  $F_i$ , while  $d(p, q)$  represents the Euclidean distance between the inhomogeneous points  $p$  and  $q$ .

At this point, we assume that the set of outlier correspondences,  $F_O$ , should include the highlight feature correspondences since they do not follow the homography for the desired content ( $\mathbf{H}_C$ ). Based on that, a second random sampling from set  $F_O$  is carried out to compute the homography for the highlights ( $\mathbf{H}_H$ ). The results for the joint homography estimation is depicted in Figure 3(b).

Once both of  $\mathbf{H}_C$  and  $\mathbf{H}_H$  are estimated, all the feature correspondences are relabeled into three different mutually exclusive sets:  $F_C$ ,  $F_H$  and  $F_O$ . Figure 3(c) depicts a sample result of the feature-level labeling step. If a feature correspondence  $F_i$  is not labeled as either desired content or highlight, it is labeled as an outlier. In order for a correspondence  $F_i$  to be categorized into the desired content correspondence set ( $F_C$ ), the symmetric transfer error using  $\mathbf{H}_C$  (i.e.,  $e(F_i, \mathbf{H}_C)$ ) should be smaller than the threshold  $T$ . At the same time, the error using  $\mathbf{H}_C$  has to be smaller than the error using  $\mathbf{H}_H$ , which indicates that  $F_i$  favors  $\mathbf{H}_C$  over  $\mathbf{H}_H$ . If  $F_i$  does not get categorized into  $F_C$ , the algorithm checks if it can be categorized as one of the highlights by evaluating the symmetric transfer error using the highlight homography ( $\mathbf{H}_H$ ) in a similar manner.

One additional criterion is employed for  $F_i$  to be categorized into  $F_H$ . It constrains the features in  $F_i$  to be present on the “bright-on-dark” MSERs [18]. The “bright-on-dark” MSER regions indicate the MSER regions which are brighter than the vicinity. As the intensity values in highlight regions tend to be stable and lighter than the neighboring regions, MSER is a reasonable choice for obtaining potential highlight regions. Yet, MSER also detects some other non-highlight regions as shown in Figure 3(d) which will be eliminated by the pixel-level labeling and the blending scheme in Section 3.3.

Having obtained the labeling for all the feature correspondences along with the two homographies, the cost  $J$  for the current iteration is computed as

$$J = E(F_C, \mathbf{H}_C) + E(F_H, \mathbf{H}_H) - \gamma \left( \frac{n(F_C) + n(F_H)}{n(F)} \right). \quad (2)$$

The first and the second term incorporate the symmetric transfer error while the third incorporates the number of inlier (desired content and highlights) feature correspondences.  $\gamma$  is a parameter which balances the three terms.  $n(F_C)$  and  $n(F_H)$  indicates the number of feature correspondences in each of the sets, respectively, while  $n_{tot}$  represents the total including the outliers. The first term which measures the average symmetric transfer error for the set ( $F_C$ ) is computed using Equation 3. The second term is computed in the same manner.

$$E(F_C, \mathbf{H}_C) = \sum_{F_i \in F_C} \frac{e(F_i, \mathbf{H}_C)}{n(F_C)} \quad (3)$$

**Algorithm 1:** Joint homography estimation for highlight removal (JH2R)

---

```

Input :  $F, M$ 
Output:  $\mathbf{H}_C, \mathbf{H}_H, F_C, F_H$ 
1  $k \leftarrow 1$  /* iteration index */
2 repeat
3   Randomly select 4 correspondences  $\in F$ , Compute  $\mathbf{H}_C$ 
4   for  $\forall F_i \in F$  do
5     if  $e(F_i, \mathbf{H}_C) > T$  then
6        $F_O \leftarrow F_O \cup F_i$ 
7     end
8   end
9   Randomly select 4 correspondences  $\in F_O$ , Compute  $\mathbf{H}_H$ 
10  for  $\forall F_i \in F$  do
11    if  $e(F_i, \mathbf{H}_C) \leq T$  &  $e(F_i, \mathbf{H}_C) \leq e(F_i, \mathbf{H}_H)$  then
12       $F_C \leftarrow F_C \cup F_i$ 
13    end
14    if  $e(F_i, \mathbf{H}_H) \leq T$  &  $e(F_i, \mathbf{H}_C) > e(F_i, \mathbf{H}_H)$  &  $F_i \in M$  then
15       $F_H \leftarrow F_H \cup F_i$ 
16    end
17  end
18  Compute  $J_{curr}$  (Eqn. 2)
19  if  $J_{curr} < J$  then
20     $J \leftarrow J_{curr}$  and update  $\mathbf{H}_C, \mathbf{H}_H, F_C, F_H$ 
21  end
22   $k \leftarrow k + 1$ 
23  compute and update  $N$  (Eqn. 5)
24 until  $k < N$ 

```

---

If the cost for the current iteration is smaller than the best previous case, the two homographies along with the two feature correspondence sets are updated. This process is repeated until the termination criteria are met.

**Termination criteria** We determine a maximum iteration number  $N$  adaptively after every iteration. We define  $w_C$  as the probability that any correspondence randomly selected from  $F$  is included in  $F_C$ . We assume that  $w_H$  is the probability that any correspondence randomly selected from  $F - F_C$  is included in  $F_H$ . These probabilities can be iteratively updated at the end of each iteration as  $w_C = n(F_C)/n(F)$  and  $w_H = n(F_H)/(n(F) - n(F_C))$ . In Equation 4,  $p$  is defined as the probability that 4 randomly selected samples are from  $F_C$  in the first selection and 4 randomly selected correspondences are from  $F_H$  in the second selection within  $N$  iterations, at least once.

$$p = 1 - ((1 - w_C^4) + w_C^4(1 - w_H^4))^N = 1 - (1 - w_C^4 w_H^4)^N \quad (4)$$

Here,  $(1 - w_C^4)$  is the probability that all 4 correspondences in the first selection are not from  $F_C$ .  $w_C^4(1 - w_H^4)$  indicates the probability that the 4 correspondences from the first selection are from  $F_C$  but at least one sample from the second selection are from the outlier

set. Therefore, the adaptive maximum iteration number  $N$  can be derived from equation 4 as

$$N = \frac{\log(1-p)}{\log(1-w_C^4 w_H^4)}. \quad (5)$$

### 3.3 Pixel-level highlight detection and blending

Using the JH2R algorithm, the two homographies (Figure 3(b)) along with the two feature correspondence sets (Figure 3(c)) for the desired content and the highlight regions can be acquired. However, the feature-level detection of the highlight regions is insufficient to properly eliminate the highlights. Instead, it needs to be extended up to the pixel-level so that the non-highlight pixels can be transferred complementarily to recover the obscured contents.

We make use of two previously acquired results which make this step computationally efficient: the estimated homographies ( $\mathbf{H}_C$ ,  $\mathbf{H}_H$ ) and the MSER detection. Pink regions in the left column of Figure 3(d) depict the MSER detection result. Then the homography  $\mathbf{H}_H$  is used to warp the two MSER images onto a common plane. This overlays the highlight regions on one image onto the corresponding highlight regions on the other. Thus, the intersection between the two MSER images, when projected onto the same plane using  $\mathbf{H}_H$ , should be the estimated region for the highlights in pixel-level sense. The right column of Figure 3(d) shows the final highlight detection result. Note that we are assuming that the two images both contain the highlights which we wish to eliminate.

Given the pixel-level highlight regions in both of the images,  $\mathbf{H}_C$  is used to project the two images onto a common plane so that the desired contents are overlaid properly while the highlight regions do not overlap. In other words, highlight regions in one image are overlaid by the non-highlight regions in the other image. This enables us to easily recover missing information for all the highlight regions in both of the images. Lastly, Poisson blending [14] is applied to assist the pixel transfers at the highlight regions with smooth boundaries. Figure 3(e) shows the sample result with all the highlights eliminated with visually pleasing quality.

## 4 Experimental Results

Our method is implemented in Matlab and run on Intel Core i5 PC (2.6GHz CPU, 4GB RAM). All the data used in the experiments are captured in real world scenes under different indoor lighting conditions. Each input image set contains two images with two different viewpoints.

**Comparison with state-of-the-art** We have compared our method with four state-of-the-art algorithms [8, 13, 14, 19]. They are chosen to represent three different approaches to solve the given problem : 1) highlight removal, 2) single image-based reflection removal, and 3) multiple image-based reflection removal. We have used the implementations provided by the authors using author-recommended parameters. Since [14] and [19] only use a single image, we have used only one of the two images per set as input.

Figure 4 shows five sample results of real world images. As can be observed in Figure 4(c) and 4(d), both [14] and [19] are incapable of removing the highlights due to the lack of information within the regions. Li et al. [14] fails to obtain sufficient amount of gradient information which they use to separate the reflection layer. Yang et al. [19] also suffers since the saturated highlights are void of diffuse color information which is supposed to change smoothly from outside the highlights to the inside.

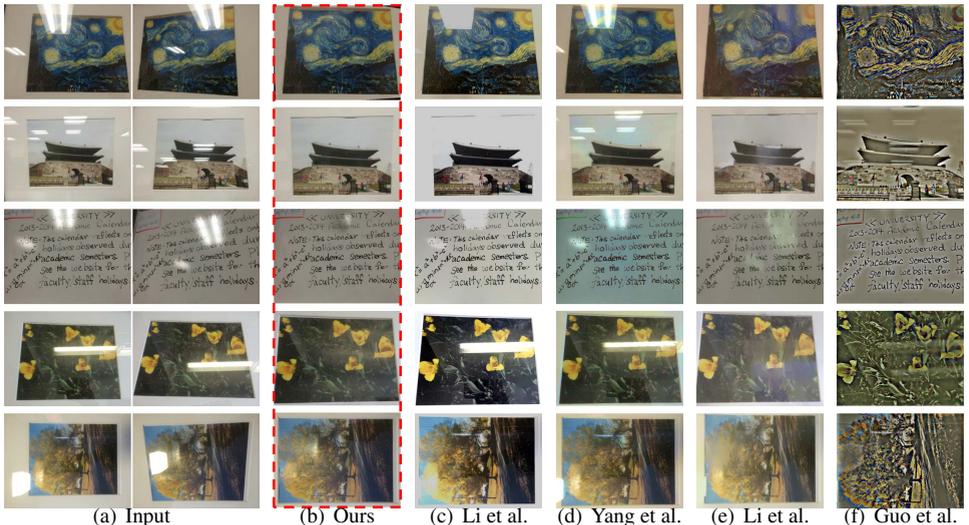


Figure 4: Five examples of highlight removal results using (b) our method compared with those produced by (c) Li et al. [14], (d) Yang et al. [29], (e) Li et al. [3], (f) Guo et al. [8]

Multiple-image based approaches by [8, 3] produce results where the highlights are only partially removed. In [3], gradients with variation across the aligned images are assumed to belong to the reflected scenes while constant gradients are assumed to belong to the desired scene. Thus, when the gradients on the highlights are too weak to be distinguished from the underlying smooth texture, this approach may suffer as shown in Figure 4(e). While [8] uses several priors including the independence between the desired content and the reflection to separate the two layers, none of the priors explain the inherent characteristics of highlights. Thus, in most cases (Figure 4(f)), color components were falsely categorized into the reflection layers, generating unnaturally colored results.

Our method, unlike others, specifically uses the relationship between the highlight regions resulting in more precise detection and removal. One may observe from Figure 4 that our method can also handle dim highlights as there still exist geometrical distinction between desired contents and dim highlights in terms of homography. In overall, our method produces the most visually pleasing results.

**Homography estimation evaluation** In Figure 5, we show the efficacy of JH2R by comparing the warped images using the estimated homographies with those using the groundtruth. The estimated  $\mathbf{H}_C$  for the desired content are very accurate. Although the estimated  $\mathbf{H}_H$  may not be equivalent to the groundtruth as illustrated in the third example, notice that the highlight regions are still well aligned. As long as the highlights overlap properly, pixel-level labeling can be performed. The groundtruth homographies are computed using manually labeled correspondences for content and highlights, separately.

**Processing time** Our method spends 25.3 seconds on average which is much faster than Li et al. [3] and Guo et al. [8] by almost the order of magnitude as shown in Table 1. Although Li et al. [24] and Yang et al. [29] both spend less processing time compared to ours, their performance in removing the highlights are unsatisfactory. We have used a single image ([24, 29]) or a pair of images (ours, [8, 3]) according to each methodology. The size of the images used in the experiments is 640 x 480.



Method	Num of Imgs	Processing Time
Ours	2	25.3 s
Li et al. [4]	1	24.5 s
Yang et al. [9]	1	< 1s
Li et al. [3]	2	221.7s
Guo et al. [8]	2	260.2s

Table 1: Quantitative processing time comparison with previous methods

view methods [9]. It will be worthwhile to further investigate an automatic capture scheme which can smartly overcome the challenging scenarios.

## Acknowledgement

The partial support of this research by the US Government through NSF Awards IIS-0812111 and IIS-1262122, the Xerox Foundation is gratefully acknowledged.

## References

- [1] A. Agrawal, R. Raskar, S.K. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Transactions on Graphics*, 24(3):828–835, Jul 2005.
- [2] A. Artusi, F. Banterle, and D. Chetverikov. A survey of specular removal methods. *Computer Graphics Forum*, 30(8):2208–2230, 2011.
- [3] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *ACM Transactions on Graphics*, pages 417–424, 2000.
- [4] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *Image Processing, IEEE Transactions on*, 13(9):1200–1212, Sep 2004.
- [5] H. Farid and E.H. Adelson. Separating reflections and lighting using independent components analysis. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, page 267, 1999.
- [6] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, Jun 1981.
- [7] J. Gao, S.J. Kim, and M.S. Brown. Constructing image panoramas using dual-homography warping. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 49–56, 2011.
- [8] X. Guo, X. Cao, and Y. Ma. Robust separation of reflection from multiple images. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 2195–2202, 2014.

- [9] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [10] N. Kong, Y-W. Tai, and J.S. Shin. A physically-based approach to reflection separation: From physical modeling to constrained optimization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(2):209–221, Feb 2014.
- [11] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(9):1647–1654, Sept 2007.
- [12] A. Levin, A. Zommet, and Y. Weiss. Separating reflections from a single image using local features. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 306–313, 2004.
- [13] Y. Li and M.S. Brown. Exploiting reflection change for automatic reflection removal. In *Computer Vision (ICCV), IEEE International Conference on*, pages 2432–2439, 2013.
- [14] Y. Li and M.S. Brown. Single image separation using relative smoothness. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 2752–2759, 2014.
- [15] S. Lin and H.Y. Shum. Separation of diffuse and specular reflection in color images. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 341–346, 2001.
- [16] S. Lin, Y. Li, S.B. Kang, X. Tong, and H.Y. Shum. Diffuse-specular separation and depth recovery from image sequences. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 210–224, 2002.
- [17] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [18] J. Matas, O. Chum, M. Urban, and T. Padjla. Robust wide baseline stereo from maximally stable extremal regions. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 36.1–36.10, 2002.
- [19] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Transactions on Graphics*, 22(3):313–318, Jul 2003.
- [20] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 328–341, 2004.
- [21] Y.Y. Schechner, J. Shamir, and N. Kiryati. Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. In *Computer Vision (ICCV), IEEE International Conference on*, pages 814–819, 1999.
- [22] Y.Y. Schechner, N. Kiryati, and R. Basri. Separation of transparent layers using focus. *International Journal of Computer Vision*, 39(1):25–39, Aug 2000.
- [23] S.A. Shafer. Using color to separate reflection components. *Color Research and Application*, 10(4):210–218, 1985.

- [24] R. Szeliski, S. Avidan, and P. Anandan. Layer separation from multiple images containing reflections and transparency. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 246–253, 2000.
- [25] P. Tan, S. Lin, L. Quan, and H-Y. Shum. Highlight removal by illumination-constrained inpainting. In *Computer Vision (ICCV), IEEE International Conference on*, pages 164–169, 2003.
- [26] P. Tan, S. Lin, and L. Quan. Separation of highlight reflections on textured surfaces. In *Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.*, pages 1855–1860, 2006.
- [27] R. T. Tan and K. Ikeuchi. Separating reflection components of textured surfaces using a single image. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(2):178–193, Feb 2005.
- [28] Q. Yang, S. Wang, and N. Ahuja. Real-time specular highlight removal using bilateral filtering. In *Computer Vision (ICCV), IEEE International Conference on*, pages 87–100, 2010.
- [29] Q. Yang, J. Tang, and N. Ahuja. Efficient and robust specular highlight removal. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1(1):1, 1 2014.