# Image Completion with Intrinsic Reflectance Guidance

Soomin Kim
soo.kim813@gmail.com

Taeyoung Kim
retupmoc@kaist.ac.kr

Min H. Kim
minhkim@kaist.ac.kr

Sung-Eui Yoon
sungeui@kaist.edu

KAIST
School of Computing
291 Daehak-ro, Yuseong-gu
Daejeon, Korea

## Abstract

Patch-based image completion methods often fail in searching patch correspondences of similar materials due to shading caused by scene illumination, resulting in inappropriate image completion with dissimilar materials. We therefore present a novel image completion method that additionally accounts for intrinsic reflectance of scene objects, when searching patch correspondences. Our method examines both intrinsic reflectances and color image structures to avoid false correspondences of different materials so that our method can search and vote illumination-invariant patches robustly, allowing for image completion mainly with homogeneous materials. Our results validate that our reflectance-guided inpainting can produce more natural and consistent images than state-of-the-art inpainting methods even under various illumination conditions.

## 1 Introduction

Single-image inpainting has been broadly applied to complete a blank region, such as occlusion by an unexpected object, without requiring any auxiliary source image [6]. Since there is no source image available, except the input image itself, the inpainting task that fills the blank region is a severely ill-posed problem. Previous inpainting algorithms can be classified into two main categories: diffusion-based or exemplar-based methods.

First, the diffusion-based approach recovers a damaged image, where missing regions are sparse and thin [3, 5]. The solution is straightforward and robust, based on a local smoothness prior, assuming that gradients near close areas are similar. However, this approach does not work well with a large blank area. Second, the exemplar-based approach has been proposed to overcome the area limitation of the diffusion-based approach [8, 12, 16], inspired by example-based texture synthesis [10]. The exemplar-based approach separates an input image into the source and the target region to search small source patches in the source region to inpaint the target region with them.

The *similarity* of the source patch to the target region is determined as a distance in terms of colors [16], gradients [9], Laplacians [13], depth [2], and planar structures [11].

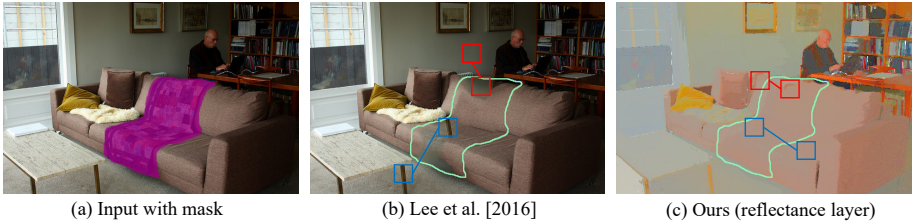|  (a) Input with mask  |  (b) Lee et al. [2016]  |  (c) Ours (reflectance layer)  |

Figure 1: Image (a) shows an input image, where the blanket is removed (masked as red). Image (b) presents correspondences of source patches using color and Laplacian properties [13]. Image (c) visualizes our correspondence search using intrinsic reflectance. Note that our method brings source patches from virtually similar materials. Complete inpainting results are available in Figure 5.

These various image properties have been considered to identify more reliable source patches for inpainting the target region. In the exemplar-based approach, the patch-similarity test using image properties often fails in searching patch correspondences of similar materials due to scene illumination. The exemplar-based approach frequently results in inappropriate inpainting output with dissimilar materials in the target region.

We are therefore motivated to examine an intrinsic property of *reflectance* of scene objects, when searching source-patch correspondences. The intrinsic reflectances of the scene objects help us identify source patches of similar materials without being affected by scene illumination. We examine both intrinsic reflectances and color image structures to avoid false correspondences of source patches from different materials (see Figure 1 for an example of the search process) so that we can search and vote illumination-invariant patches robustly, allowing for inpainting with homogeneous materials over the target region. We provide an in-depth analysis of all the factors and parameters used in the proposed inpainting algorithm and compare our results with state-of-the-art methods to validate the effectiveness of our image completion with intrinsic reflectance guidance.

## 2 Reflectance-Guided Image Completion

**Overview** We first decompose an input image into two layers of reflectance and shading using an intrinsic image decomposition. While many existing techniques can be used for intrinsic image decomposition, we choose a recent CNN-based technique [17]. Since we are motivated to devise a guidance term that is less affected by illumination, we make use of the reflectance layer only from the decomposed layers. The reflectance layer allows us to vote more similar patches to a hole pixel in terms of material colors without being affected by the illumination or image structures of patches. We use the reflectance information as an additional guidance feature term to find candidate patches. We also adopt the multiscale approach for the overall inpainting process, following other exemplar-based approaches [13, 16], and carefully design how to use the reflectance guidance term to harmonize it with other energy terms. Figure 2 provides an overview of the workflow of our algorithm.

### 2.1 Extracting Intrinsic Reflectance

Barrow and Tenenbaum [5] introduced the term of the intrinsic images that refers to the decomposition of an image $I(x, y)$ into two components: reflectance (albedo) $R(x, y)$ and
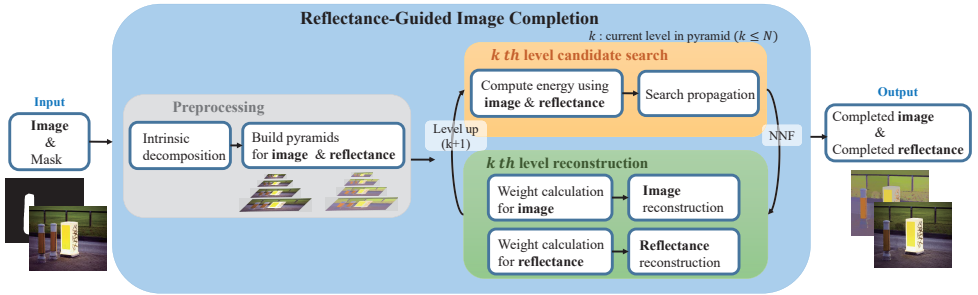
Figure 2: Overview of our algorithm. From a given image with a target mask, we first estimate an intrinsic reflectance image using a data-driven intrinsic decomposition method (Section 2.1). We then compute patch similarity through pyramids of color image structures and reflectances, respectively, to yield nearest neighbor fields of patch correspondences. For reconstructing colors, we examine not only image color structures but also intrinsic reflectances (Section 2.2).

illumination (shading) $S(x,y)$ under the assumption that the product of reflectance and shading elements becomes the original image. We then define the following simple relationship between the original and element images:

$$I(x,y) = S(x,y)R(x,y). \tag{1}$$

Splitting into two intrinsic images from a single input image is a long-lasting ill-posed problem since it is under-constrained. In our method, we adopt a recent data-driven approach [17] to obtain the intrinsic reflectance layer from the input image.

Nonetheless, the reflectance information cannot be solely used for calculating patch similarity in our inpainting method, since the reflectance layer loses high-frequency details significantly, compared to the original image [see Figure 1(c)]. To overcome the over-smoothed property of the reflectance with our similarity calculation, we utilize the original image's detail using Laplacian coefficients [13]. Since those high-frequency details in the image are less sensitive to illumination than colors are, we can utilize the original image's high-frequency information along with the reflectance's color information. See Section 2.2.1 for more details.

## 2.2 Reflectance-Guided Image Completion

Our inpainting method is based on the traditional patch-based approach [16], aiming to preserve global consistence in addition to image structures. The inpainting process starts from the coarsest level to the finest level in the image pyramid (see Figure 2). The coarsest level image is a mostly downsampled image, where the hole region is also reduced as small. We can inpaint the small hole merely using neighboring information based on a diffusion-based technique [2]. Once the coarse level image is filled, the higher-level image is reconstructed by performing upsampling. We then perform the inpainting process in the level. This process is iterated until we inpaint the finest level of the original resolution in the pyramid. Our filling-in method is based on a recent Laplacian-based method [13] with the main difference of utilizing reflectance guidance for similarity calculation and reconstruction.

Our inpainting process starts with two images: an input image and its reflectance image. To inpaint the original image with the guide of reflectance, it is necessary to fill the hole in the reflectance image as well. For this process, we also build two image pyramids of

CIELAB colors and Laplacians for not only the color image, but also the reflectance image, respectively. The total number of pyramids is four. In the multiscale approach, inappropriate candidate patches in a coarse level can propagate in a larger area in finer levels. To mitigate this issue, we utilize reflectance information to find better matching candidate patches, which might be hard to be found only in the original image due to the illumination.

### 2.2.1   Searching Candidate Patches

Measuring similarity between candidate patches and hole patches is one of the critical parts in inpainting that determines the quality of completed images. While our method utilizes two different image properties of color and intrinsic reflectance for inpainting, we accommodate them into a single energy function that contains a new reflectance guidance term. We use this function for guiding the search of candidate patches.

We define $I$ as an original image and $R$ as its reflectance layer. With the user-defined mask, we can divide the input image and the reflectance layer into the target hole area $T$ and the source area $S$. Since we utilize the multiscale approach, symbols with under subscript of $k$ refer to the $k$-th level of the multiscale pyramids, for instance, $I_k$ and $R_k$ mean the $k$-th levels of the image pyramids on the original image and the reflectance, respectively. Given a hole pixel $p$ in $k$-th level ($p \in T_k$), its patch $P_{I_k}$ and $P_{R_k}$ refer to the $w \times w$ patch centered at the pixel $p$ in each image $I_k$ and reflectance $R_k$, respectively; we use $7 \times 7$ for all the results.

Our similarity equation consists of two terms: the original image coherence term and the reflectance guidance term, as follows:

$$E(p) = \sum_{Q_{I_k}, Q_{R_k} \in S_k} \min\{E_{coherence}(P_{I_k}, Q_{I_k}) + E_{guidance}(P_{R_k}, Q_{R_k})\}. \tag{2}$$

The above equation aims to find out the best candidate patch $Q$ for a patch $P$, that minimizes the energy function. Note that $Q_{I_k}$ and $Q_{R_k}$ are extracted in the same location, but in different image spaces $I_k$ and $R_k$. Once we found the candidate patch for each hole pixel that minimizes our energy function, the location of it is stored as nearest neighbor fields (NNFs) and used for reconstructing the hole area.

**Image Coherence**   Our image coherence term $E_{coherence}$ measures the distance of both colors and Laplacian edges between target and source patches using the original image information only:

$$E_{coherence}(P_{I_k}, Q_{I_k}) = \alpha_k C(P_{I_k}, Q_{I_k}) + \gamma C(\nabla^2 P_{I_k}, \nabla^2 Q_{I_k}), \tag{3}$$

where $C$ is the sum of squared distances between two patches, and $\alpha_k$ is a weighting term to determine the contribution of the original color information in the $k$-th level. Laplacian coefficients are used to detect the edge information, and $\gamma$ is used to modulate the weight of the edge distance for the overall patch similarity. For obtaining Laplacian coefficients, we inherit the inpainting approach using the difference of Gaussians (DoG), following [13].

**Reflectance Guidance**   Since the reflectance image is more independent of the illumination effects than the input color image, we expect that the reflectance distance can lead us to find candidate patches of more similar materials, which might be considered as different patches due to illumination changes in the input image. The reflectance weight in the guidance term is critical for balancing the color and the reflectance similarity. Similarity search using the reflectance only often fails due to significant color differences between the

(a) Input image  (b) Reflectance image  (c) Laplacian of the input image  (d) Laplacian of the reflectance
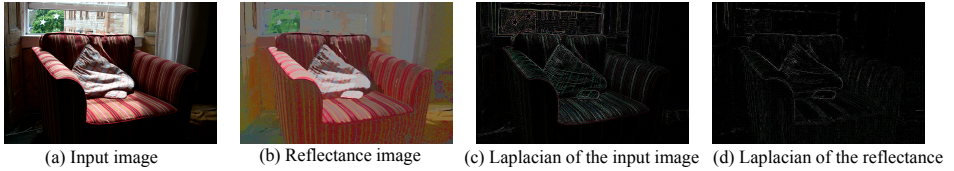
Figure 3: Comparison between the Laplacian coefficients of the input image (a) and its reflectance image (b). Even though the reflectance image is less affected by illumination than the input image, its detail is lost significantly compared to the original image. Accordingly, we use Laplacian coefficients (c) of the input image for the edge information. (d) shows Laplacian coefficients of the reflectance.

reflectance and the shaded color. We carefully modulate the reflectance weight with consideration of the level of operation.

Our reflectance guidance term $E_{guidance}$ measures the L2 color distance between the target and source patches in terms of reflectance:

$$E_{guidance}(P_{R_k}, Q_{R_k}) = \beta_k C(P_{R_k}, Q_{R_k}), \qquad (4)$$

where $\beta_k$ is the weighting term to modulate a contribution of the reflectance distance in the similarity function in the $k$-th level. For accounting for color distances, we employ the CIELAB [14] space for both original color and the reflectance image.

Note that we utilize only color information of the reflectance, not its edge information, since the reflectance image contains less details compared to the original image (Section 2.1). Figure 3 shows the Laplacian coefficients of the original and reflectance images. Even though the Laplacian of the reflectance image contains overall object edges, they lose details of the image. Instead, we compensate the detail information from Laplacian coefficients of color layers mainly. For finding the candidate patch location as NNFs, which minimizes our similarity equation [Equation (2)], we utilize the Patchmatch algorithm [4] for efficient searching.

### 2.2.2 Adaptive Reflectance Weights

The reflectance information can guide us to find appropriate candidate patches, but without original color information, the result images tend to suffer from artifacts due to small color offsets. We therefore address this issue by adaptively adjusting weights of our reflectance term in the multiscale approach.

In the multiscale approach, coarser level images contain more general scene information, and finer level images contain more detail information. As a result, it is desirable to use the reflectance layer more in coarser levels for properly guiding patch search within similar materials. It is more appropriate to weight the image color layer more in finer levels. We formulate this idea simply by adaptively adjusting weights upon the level. In other words, we assign higher reflectance weights to coarser level images and relatively smaller reflectance weights to finer level images. For this, we use level-dependent $\alpha_k$ and $\beta_k$ weights, and they are defined as the following:

$$\beta_k = \beta - \Delta \ \text{ and } \ \alpha_k = \alpha + \Delta,$$
$$\Delta = \beta \times \frac{k-1}{N-1} \times \frac{\sigma_R}{\sigma_I}, \qquad (5)$$

where $0 < k \leq N$. $N$ indicates the maximum level of our multiscale approach. Initial reflectance and image weights are given by input parameters. For all of our experiments, we

assign 0.05 to $\alpha$, 0.65 to $\beta$. Note that we aim to consider the reflectance layer more in coarser images and thus we set a higher value to $\beta$ than $\alpha$. As the level goes up, the reflectance weight then decreases, while the image weight increases as an amount of $\Delta$. The $\Delta$ value is proportional to the ratio of reflectance and image standard deviation, to equalize the reflectance's distance according to its image statistics [15]. Among parameter values, we set $\alpha_k + \beta_k + \gamma = 1$ for all the results.

Based on the weight adjustment according to the image level, we effectively guide the appropriate candidate area considering the reflectance in coarser levels, while preserving the image consistency in finer levels, which get closer to the original image size. The other weight, the edge weight, does not need to be changed, so we keep the initial edge weight ($\gamma$) the same throughout all the levels.

### 2.2.3   Reconstruction

Our reconstruction process follows the standard voting scheme used in the common exemplar-based methods [16]. The main difference for our reconstruction method over the existing methods is that our method utilizes both the original and reflectance images. Note that in order to find out NNFs, we account for three different types of image properties: image colors, image edges and reflectances (as described in Section 2.2.1), while we examine an additional property of reflectance edges, in addition to the three properties to reconstruct inpainted colors. We therefore need to recalculate patch similarity to determine weights for the voting process, using Equation (6), rather than inheriting similarity calculated from NNF search [Equation (2)] .

To reconstruct colors, we copy patches from the source to the target region with $w \times w$ sized patches. Each pixel in the hole can have several overlapped information from different patches. To reconstruct a color value with such overlapped patches, we calculate the weighted sum of the overlapped information. A weight term for a pixel $p$, $w(p)$, consists of two factors: (1) patch similarity about another pixel $q$, $\text{sim}(p,q)$, whose patch is nearest to that of $p$, and (2) the distance from the boundary information, $\Lambda(p)$ for avoiding disconnected completion around the boundary by giving a more weight to the information near the hole boundary. The weight is the product of these two values: $w(p) = \text{sim}(p,q) \cdot \Lambda(p)$, where the patch similarity $\text{sim}(p,q)$ is defined as a Gaussian function, $\exp\left(\frac{-D_{I|R}(p,q)}{2\sigma^2}\right)$, where $\sigma$ is a term to adjust similarity smoothness.

For reconstructing the target region of the original input image and its reflectance image, we use two separate similarity functions for each image. The distance functions of similarity, $\text{sim}(p,q)$, in the original image and the reflectance layer are denoted as $D_I(p,q)$ and $D_R(p,q)$, respectively. They are defined as follow:

$$D_{I|R}(p,q) = (1-\gamma)C(P_{I|R}, Q_{I|R}) + \gamma C(\nabla^2 P_{I|R}, \nabla^2 Q_{I|R}), \tag{6}$$

where $\gamma$ is the same weight term for adjusting the edge information in the patch similarity [Equation (3)], and $I|R$ indicates the original image or the reflectance layer.

## 3   Results

We evaluate the performance of the proposed method with various scenes (Figures 4 and 5), where scene illumination varies significantly, compared with state-of-the-art inpainting methods: Photoshop content-aware fill (based on [16]), Huang *et al.* [11], Darabi *et al.* [9], and

(a) Input image with hole    (b) Photoshop    (c) Darabi et al. [2012]    (d) Lee et al. [2016]    (f) Ours
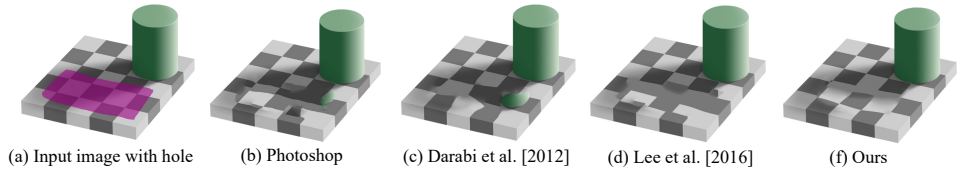
Figure 4: Our method inpaints the input image by separating intrinsic reflectance and shading so that our method can complete the image with intrinsic scene materials, rather than shadowed color values. Other state-of-the-art methods inpaint the blank region using color values directly, resulting in unrealistic artifacts.



(f) Input and mask

(e) Photoshop

(d) Darabi et al. [2012]

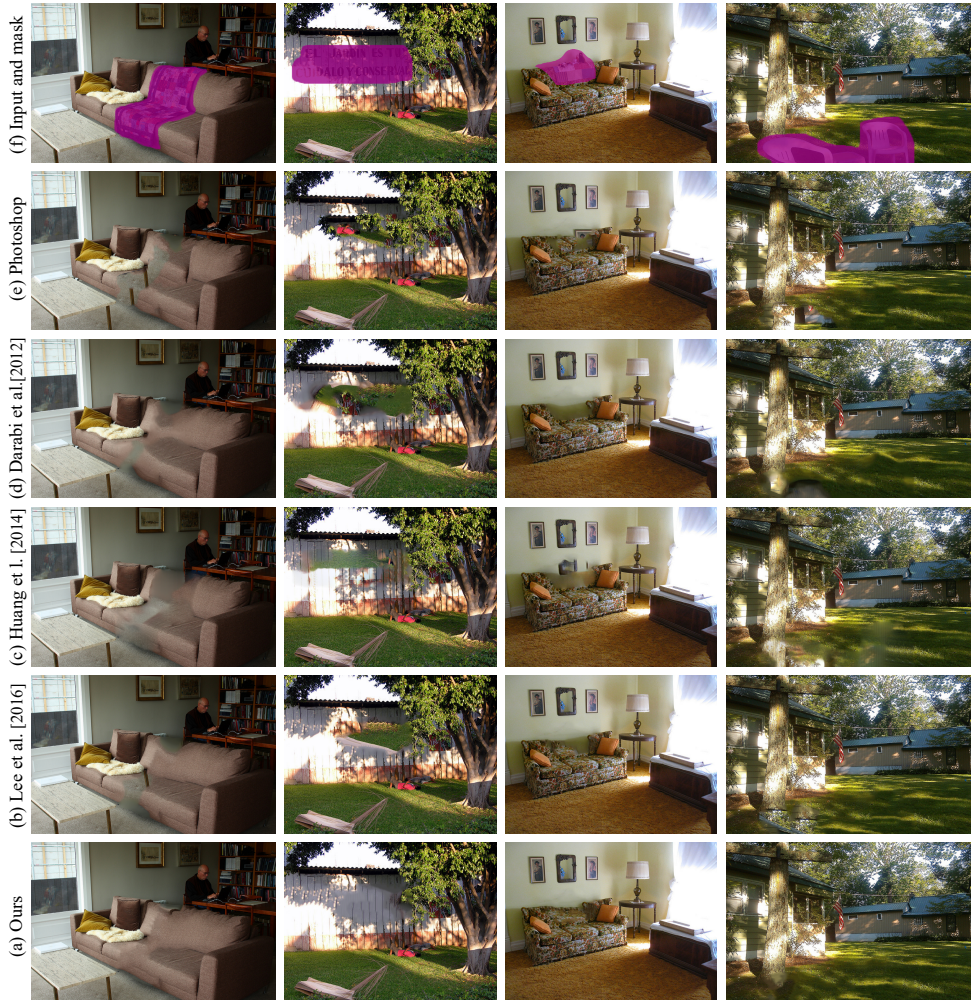(c) Huang et l. [2014]

(b) Lee et al. [2016]

(a) Ours

Figure 5: Natural scenes under various illumination are inpainted. While our method inpaints target regions with patches from similar materials of the scene objects, other methods often inpaint the region with patches from different materials. Refer to the supplemental material for more examples.
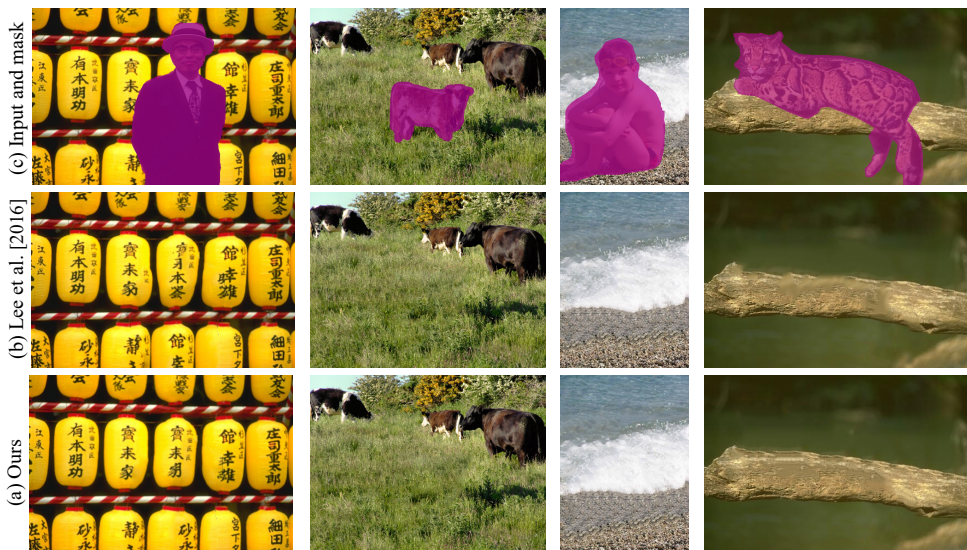
Figure 6: With common datasets used by other inpainting algorithms, our reflectance-based method is strongly competitive to a state-of-the-art method using Laplacian [13], in addition to the robust performance on natural scenes under strong illumination.

Lee *et al.* [13]. We use available codes written by authors and commercial software for comparison. In addition, we compare results on common reference images for inpainting (Figure 6) to validate that our method can perform consistently with general scenes. Refer to the supplemental material for high-resolution result images.

Figure 4 compares results of the checker shadow illusion image, rendered by Adelson [1]. Scene illumination that drops soft shadow over checkerboard patterns alternates the lightness intensities of the patterns. Since pixel color-based approaches just recognize the shaded pixel values as colors of the object materials, the original structure of the object cannot be inpainted properly due to shading. However, our method uses intrinsic reflectance as guidance so that it can reconstruct the original structure more plausibly than the state-of-the-art methods.

Figure 5 compares a series of inpainting results of real scenes, where some parts of the scene are strongly illuminated with significant variation. In these natural scenes, the searching and voting processes of source patches are affected by scene illumination in previous inpainting methods. For instance, the first column image shows very subtle differences of colors near the target blank region, since most regions of the scene belong to shadow. Our method can inpaint the target region with patches from similar materials, while other previous methods suffer from artifacts on image structures due to source patches searched from significantly different materials that might present similar image structures in the scene. The rest of the columns present results with a wider variation of scene illumination, which hinders the example search process in the previous inpainting algorithms. Our method can produce more natural inpainting results consistently by completing the target regions with patches from similar materials.

For fairness in comparison, we evaluate the performance of our method with reference image datasets, commonly used by other inpainting algorithms. Figure 6 compares the results of our method with those of a state-of-the-art method using Laplacian [13]. The per-

formance of our method is highly competitive to the Laplacian method, in addition to the enhanced performance with natural scenes under strong illumination, as shown in Figure 5. As shown in the third column, when the variation of colors in the input image is small, our method can outperform the existing method thanks to intrinsic reflectance guidance.

# 4 Discussion



(a) Chrominance guide
(synthetic scene)

(b) Reflectance guide
(synthetic scene)

(c) Chrominance guide
(real-world scene)

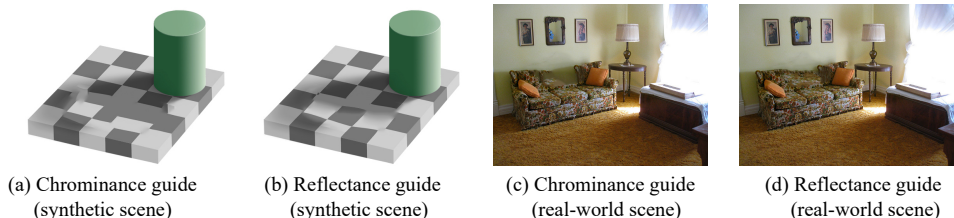(d) Reflectance guide
(real-world scene)

Figure 7: Our results using two different guidances: (a),(c) chrominance and (b),(d) reflectance. Reflectance-guided inpainting reconstructs intrinsic image structures successfully, while chrominance-based inpainting suffers from artifacts in both synthetic scene and real-world scene. We chose the use of reflectance guide.

Among various color properties, both chrominance and reflectance are luminance-invariant features in the color space. Chrominance can be obtained easily by simply decomposing the input color images into the luminance and the chrominance layers, while intrinsic reflectance can be achieved by using an intrinsic algorithm, resulting in heavier computation than using chrominance. We therefore examine if chrominance can substitute reflectance in our algorithm. We evaluate the performance difference of our method with the different choice of the image property. As shown in Figure 7, the performance of our method is more robust when it uses reflectance guidance, rather than chrominance guidance. We found that when we use chrominance as guidance of materials for inpainting, chrominance guidance disregards the brightness information of materials. In contrast, reflectance guidance accounts for both chrominance and brightness (as luminance) when examining intrinsic material properties. We therefore chose reflectance guidance over chrominance guidance despite of computational costs.

In addition, we adopt a state-of-the-art intrinsic image method using a deep neural network [17] to calculate intrinsic reflectance from an input image. However, since decomposing reflectance from a single image is also a severely ill-posed problem, the adopted method could fail sometimes in calculating accurate reflectance. It is not surprising that our method performance is also affected by the accuracy of the employed intrinsic image algorithm. Intrinsic image decomposition has been extensively developed to achieve significant enhancements in recent years. We believe an intrinsic decomposition method in near future could be an alternative to be beneficial to enhance the robustness of our method.

# 5 Conclusion and Future Work

We have presented an reflectance-guided image completion method that can handle a wide variety of images that include significant illumination variation. Our method utilizes reflectance as a guidance term when calculating patch similarity, so that we can find appropriate patches, even when we have strong illumination variations. In addition, the reflectance

term can identify different materials through the search process of source patches, reducing a chance of bringing source patches of different materials, even though they are similar in terms of color and structure. Currently, we adopt an intrinsic image decomposition method without any modification for our method. Revising the decomposition method into our framework remains as future work.

# Acknowledgements

# References

[1] E. H. Adelson. Lightness perception and lightness illusions. *The New Cognitive Neurosciences*, pages 339–351, 2000.

[2] Seung-Hwan Baek, Inchang Choi, and Min H. Kim. Multiview image completion with space structure propagation. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR 2016)*, pages 488–496, Las Vegas, USA, 2016. IEEE.

[3] Coloma Ballester, Marcelo Bertalmio, Vicent Caselles, Guillermo Sapiro, and Joan Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE transactions on image processing*, 10(8):1200–1211, 2001.

[4] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics-TOG*, 28(3):24, 2009.

[5] Harry Barrow and J Tenenbaum. Recovering intrinsic scene characteristics. *Comput. Vis. Syst., A Hanson & E. Riseman (Eds.)*, pages 3–26, 1978.

[6] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 417–424. ACM Press/Addison-Wesley Publishing Co., 2000.

[7] Marcelo Bertalmio, Andrea L Bertozzi, and Guillermo Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–355. IEEE, 2001.

[8] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on image processing*, 13(9):1200–1212, 2004.

[9] Soheil Darabi, Eli Shechtman, Connelly Barnes, Dan B Goldman, and Pradeep Sen. Image melding: Combining inconsistent images using patch-based synthesis. *ACM Trans. Graph.*, 31(4):82–1, 2012.

[10] Alexei A Efros and Thomas K Leung. Texture synthesis by non-parametric sampling. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1033–1038. IEEE, 1999.

[11] Jia-Bin Huang, Sing Bing Kang, Narendra Ahuja, and Johannes Kopf. Image completion using planar structure guidance. *ACM Transactions on Graphics (TOG)*, 33(4): 129, 2014.

[12] Vivek Kwatra, Irfan Essa, Aaron Bobick, and Nipun Kwatra. Texture optimization for example-based synthesis. *ACM Transactions on Graphics (ToG)*, 24(3):795–802, 2005.

[13] Joo Ho Lee, Inchang Choi, and Min H. Kim. Laplacian patch-based image synthesis. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR 2016)*, pages 2727–2735, Las Vegas, USA, 2016. IEEE.

[14] János Schanda. *Colorimetry: understanding the CIE system*. John Wiley & Sons, 2007.

[15] Yu-Wing Tai Taeyoung Kim and Sung-Eui Yoon. Pca based computation of illumination-invariant space for road detection. In *WACV*, 2017.

[16] Yonatan Wexler, Eli Shechtman, and Michal Irani. Space-time video completion. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–120. IEEE, 2004.

[17] Tinghui Zhou, Philipp Krahenbuhl, and Alexei A Efros. Learning data-driven reflectance priors for intrinsic image decomposition. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3469–3477, 2015.