Single Image Reflection Suppression

Nikolaos Arvanitopoulos, Radhakrishna Achanta and Sabine Süsstrunk
School of Computer and Communication Sciences (IC)
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{nick.arvanitopoulos,radhakrishna.achanta,sabine.susstrunk}@epfl.ch

Abstract

Reflections are a common artifact in images taken through glass windows. Automatically removing the reflection artifacts after the picture is taken is an ill-posed problem. Attempts to solve this problem using optimization schemes therefore rely on various prior assumptions from the physical world. Instead of removing reflections from a single image, which has met with limited success so far, we propose a novel approach to suppress reflections. It is based on a Laplacian data fidelity term and an l_0 gradient sparsity term imposed on the output. With experiments on artificial and real-world images we show that our reflection suppression method performs better than the state-of-the-art reflection removal techniques.

1. Introduction

Images taken through glass or windows often contain undesirable reflections. It is possible to minimize their effect by using special hardware, multiple images, or manual post-processing. For example, professional photographers use polarizing filters to mitigate, if not eliminate, reflection artifacts. Similarly, using multiple photographs taken with small displacements makes the problem easier to tackle [5, 6, 11, 14, 18, 20, 21, 24]. However, given the skill and resources of an average user, none of these methods are feasible. For everyday photography done with consumer grade cameras, we need single image reflection removal.

Barrow and Tenenbaum [2] first presented a linear model assumption for an image $\mathbf{Y}$ that contains reflections as a sum of two other images (or layers) as follows:

$$\mathbf{Y} = \mathbf{T} + \mathbf{R}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{n \times m}$ is the observed image and $\mathbf{T}, \mathbf{R}$ are the transmission and reflection layers, respectively. Since we have one equation for two unknowns, this problem is highly ill-posed. Methods to solve the problem have to impose constraints based on assumptions from the physical world and prior knowledge.

(a) Original Image   (b) Our result

Figure 1. Our reflection suppression method applied on a real-world image taken through a glass window. Notice how we succeed in suppressing the reflections and yet preserve the details of the original image.

Existing approaches for reflection removal from a single image rely on natural image priors to formulate objective functions. Depending on the nature of these priors, all methods have their specific limitations. We discuss these limitations with experiments conducted on synthetic images.

Instead of trying to separate transmission and reflection from the original image, we develop an algorithm whose output is an image where reflections are *suppressed* (see Fig. 1). This makes the problem more tractable. We propose a novel and efficient optimization function, which is based on an l_0 prior on the gradient sparsity of the transmission layer. Additionally, we propose a data fidelity term that penalizes the difference between the Laplacian of the input image and the desired transmission layer.

We show that our formulation provides better reflection suppression and color reproduction results than the state-of-the-art in single image reflection removal on both real and synthetic images. Our main contributions are:

- We propose a novel and computationally tractable single image reflection suppression algorithm based on an l_0 gradient sparsity prior and on a Laplacian data fidelity term.

- We conduct quantitative experiments on synthetic images that we create based on reflection model assump-

tions. We show that we are better than the state-of-the-art by a significant margin.

- We show better qualitative results in real-world images with respect to previous approaches in single-image reflection removal.

The paper is organized as follows. In Section 2 we give a detailed description of the related work in reflection removal. In Section 3 we present our method and provide justifications for our choices. In Section 4 we provide comparisons with state-of-the-art methods, and finally, in Section 5 we conclude the paper.

2. Background

We separate the previous research in reflection removal into two categories. We first present the methods that use multiple images as input, then we present the methods that use a single image to remove reflections.

2.1. Multiple-image Reflection Removal

A substantial amount of work on reflection removal relies on multiple captured images. The works of Kong et al. [8] and Schechner et al. [16] build upon a physical reflection model and use several images taken with different polarization angles to estimate the reflection layer. Similarly, Farid and Adelson [4] use Independent Component Analysis to estimate the mixing matrix of two images taken with different polarization angles.

Agrawal et al. [1] rely on two photos of the same scene, taken with and without a flash. They use a gradient projection scheme based on a gradient coherence model that allows removal of reflections and highlights from flash images. Sirinukulwattana et al. [19] exploit the fact that the reflections vary in multiple images captured from slightly different viewpoints. They impose a constraint on the disparity map which smooths specific areas of the reflection layer while simultaneously preserving the sharpness of the transmission.

Some approaches rely on video sequences to decorrelate the motion between the transmission and reflection layers [14, 18, 21]. Xue et al. [24] utilize the motion differences to decompose the input image to an initial transmission and reflection layer. From the initial layers they extract motion fields. They repeat the process of updating the transmission and reflection layers, and estimating the motion fields, until convergence. Gai et al. [5] simultaneously estimate layer motions and linear mixing coefficients with a sparse blind separation algorithm. Guo et al. [6] use rank constraints and structural priors to exploit the correlation of the transmission layers from multiple images. In [11, 20] the authors use SIFT-flow to calculate the motion from photos taken from different view-points. Using a motion score they

classify edges as belonging to either transmission or reflection layers, which helps solve an optimization scheme to separate the layers.

2.2. Single-image Reflection Removal

Single image reflection removal, which is the focus of this paper, is of practical importance because in most situations the user will not have access to multiple images. However, as stated above, it is a highly ill-posed problem. Existing works therefore rely on sparse gradient priors to distinguish between transmission and reflection edges [9, 10]. Levin and Weiss [9] solve a constrained optimization problem by imposing a Laplacian mixture prior over the image gradients. In their work though, user-intervention is required to label image gradients as belonging to either the transmission or reflection layer.

The work of Li and Brown [12] automatically extracts the two layers by optimizing an objective function which imposes a smooth gradient prior over the reflection layer and a sparse gradient prior over the transmission. This gradient prior is based on the observation that reflections are often less in focus, i.e., have weaker gradients, than the transmitted image. Similarly, Wan et al. [22] compute the Depth of Field per pixel with the use of Kullback-Leibler divergence to build reflection and transmission edge maps. With these maps they use the method of [9] to extract the two layers. The work of Shih et al. [17] tries to exploit ghosting artifacts that are typical of images captured through a window in order to solve a deblurring-based optimization problem.

The performance of the methods that depend on only one image is, in general, limited in real-world scenarios. This is to be expected, due to the highly ill-posed nature of the problem. In addition, the methods proposed so far are often computationally inefficient. In this paper, we propose a method that suppresses reflections from a single input image and propose a solution that is superior to the state-of-the-art.

3. Our Algorithm

We rely on two main observations in our method for reflection suppression. The first observation is that, compared to transmission edges, reflection edges are of smaller magnitude and they are less in focus. This is often true in real-life scenarios. The camera focuses on the background objects, whose distance to the camera is usually different from the reflection components. Formally, we can express our first assumption with the following image formation model [12, 15]:

$$\mathbf{Y} = \mathbf{W} \circ \mathbf{T} + (\mathbf{1} - \mathbf{W}) \circ (\mathbf{k} * \mathbf{R}), \qquad (2)$$

where $\circ$ denotes element-wise multiplication, $\mathbf{k}$ is the blurring kernel and $*$ denotes the convolution operation. $\mathbf{W}$ is a

matrix that weighs the contribution of the transmission layer at each pixel. It is important to note here that for real images, $\mathbf{W}$ is not usually constant, but depends on the lighting conditions and on the position of the camera relative to the image plane (see [8] for a detailed discussion).

In the rest of the paper we make the simplifying assumption that $W_{i,j} = w, \; \forall i, j$. Even though this constant blending factor assumption can be incorrect in real-world scenarios, it is a reasonable approximation that makes the problem tractable, considering that we only have one image at our disposal from which to suppress reflection artifacts.

Our second observation is that in most cases, humans have an uncanny ability to tell apart reflections, likely because we rely on several visual cues, including, as we remark, Gestalt continuity of structures in images [3]. As illustrated in Fig. 1a, the human visual system discounts the intensity modulations due to the reflection in the upper-right quadrant of the image. While harder to formalize, this observation helps us choose the prior term.

To account for these two observations, we build upon the successful image smoothing approach of Xu et al. [23]. In their work, the authors smooth the image by imposing a constraint on the number of non-zero gradients on the output. Their approach globally eliminates a substantial amount of gradients of small magnitudes while simultaneously retaining large magnitude edges. The optimization problem they solve has the form:

$$\mathbf{T}^* = \arg\min_{\mathbf{T}} \|\mathbf{T} - \mathbf{Y}\|_2^2 + \lambda C(\mathbf{T}), \qquad (3)$$

where

$$C(\mathbf{T}) = \#\{(i,j) \mid |\nabla_x T_{i,j}| + |\nabla_y T_{i,j}| \neq 0\}. \qquad (4)$$

The combination of the data fidelity term with the l_0 prior on the image gradients ensures that the algorithm removes edges in increasing order of magnitudes. The larger the regularization parameter λ is, the more gradients are removed.

The prior term $C(\mathbf{T})$ encourages smoothing of the image while maintaining the continuity of large structures. However, its combination with the data-fidelity term $\|\mathbf{T} - \mathbf{Y}\|_2^2$ eliminates most of the high frequency details from the image, which is desirable for smoothing, but not for reflection suppression. We want to not only preserve the continuity of large structures but also retain as much of the transmission layer details. We thus revisit the data fidelity term.

To avoid losing important high frequency details from the image, we propose a Laplacian-based data fidelity term to modify the objective function of Eq. (3). The Laplacian of an image is defined as

$$\mathcal{L}(\mathbf{Y}) = \nabla_{xx}\mathbf{Y} + \nabla_{yy}\mathbf{Y}, \qquad (5)$$

which is equivalent to a convolution with the 3×3 kernel $\mathbf{k}_{\mathcal{L}} = [0, 1, 0; 1, -4, 1; 0, 1, 0]$. A fidelity term based on the

Laplacian better enforces consistency in structures of fine details. Our proposed optimization problem thus takes the following form:

$$\mathbf{T}^* = \arg\min_{\mathbf{T}} \|\mathcal{L}(\mathbf{T}) - \mathcal{L}(\mathbf{Y})\|_2^2 + \lambda C(\mathbf{T}), \qquad (6)$$

where $C(\mathbf{T})$ is the same as in Eq. (4).

We demonstrate the effect of our proposed method in Fig. 2. We create a synthetic blend of the letters 'T' (Fig. 2a, transmission) and 'R' (Fig. 2b, reflection) with different types of background texture. We use the model of Eq. (2) with $w = 0.7$ and $\mathbf{k}$ a Gaussian blur with $\sigma = 2$. The resulting blend is shown in Fig. 2c. We compare our method (Fig. 2f) with the approach of Xu et al. [23] (Fig. 2e). We also provide results for standard Total Variation (TV) smoothing [13] (Fig. 2d). In Fig. 2g we show the superposed middle scan lines of Figures 2d, 2e, and 2f for a more detailed visualization of the outputs.

While conventional TV (l_1-based) priors smooth edges, the l_0-based sparsity prior has the effect of flattening a signal in order to affect signal smoothness. We observe that our proposed method with the Laplacian data fidelity term more faithfully represents strong edges and detailed structures of the signal. At the same time, it retains more texture from the transmission layer. The original approach of Xu et al. [23] drastically smooths image details, even with very small regularization parameter values. This behavior results in severe loss of the transmission layer's original texture. The TV approach, on the other hand, uniformly reduces the gradient magnitudes because of the soft thresholding operation. Even large magnitude edges are smoothed out by increasing the regularization parameter. This is an unwanted effect in reflection removal, where the goal is to maintain strong edges.

We solve the optimization problem in Eq. (6) with half-quadratic splitting as done by Xu et al. [23]. We introduce auxiliary variables $\mathbf{H}$ and $\mathbf{V}$, corresponding to $\nabla_x\mathbf{T}$ and $\nabla_y\mathbf{T}$ respectively. By defining $\mathbf{G} = [\mathbf{H}; \mathbf{V}]$ and $\nabla = [\nabla_x; \nabla_y]$, the objective function can be compactly written as

$$\mathbf{T}^*, \mathbf{G}^* = \arg\min_{\mathbf{T}, \mathbf{G}} \|\mathcal{L}(\mathbf{T}) - \mathcal{L}(\mathbf{Y})\|_2^2 + \lambda C(\mathbf{G}) +$$
$$+ \beta\|\nabla\mathbf{T} - \mathbf{G}\|_2^2, \qquad (7)$$

where

$$C(\mathbf{G}) = C(\mathbf{H}, \mathbf{V}) = \#\{(i,j) \mid |H_{i,j}| + |V_{i,j}| \neq 0\}. \qquad (8)$$

The problem in Eq. (7) is solved by alternately minimizing over either $\mathbf{G}$ or $\mathbf{T}$, while keeping the other fixed. In the following, we provide the details on how to efficiently optimize each sub-problem.

(a) Transmission layer   (b) Reflection layer   (c) Synthetic blend, $w = 0.7, \sigma = 2$

(d) $l1, \lambda = 0.5$ (e) Xu et al. [23] $\lambda = 0.05$ (f) Proposed, $\lambda = 0.05$

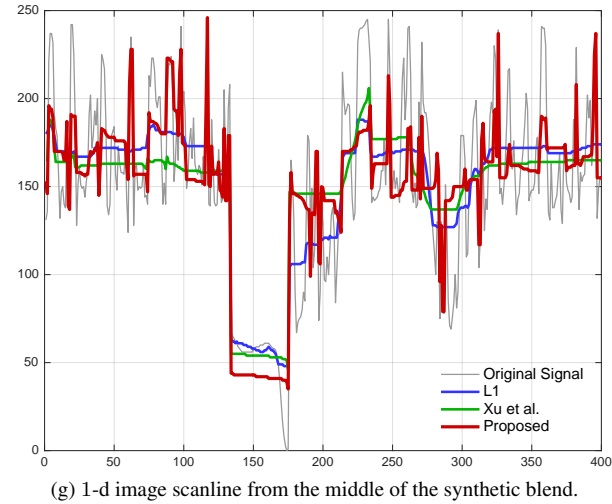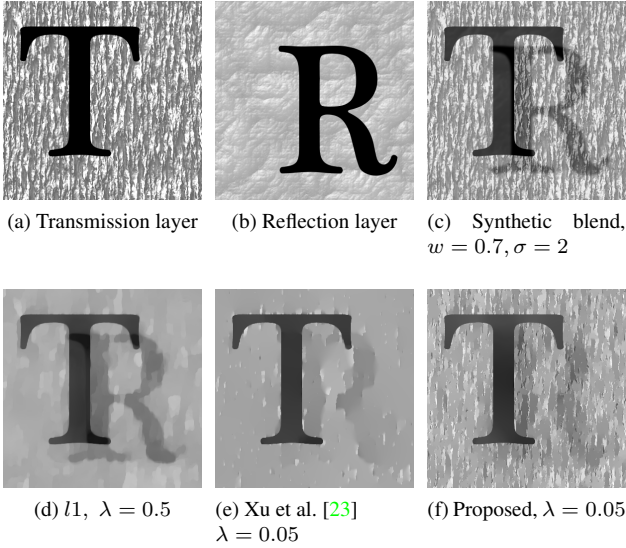(g) 1-d image scanline from the middle of the synthetic blend.

Figure 2. Comparison of different smoothing techniques on a 2-d toy example. Fig. 2a and Fig. 2b are combined to obtain Fig. 2c. Fig. 2d shows the result of l_1 smoothing [13], Fig. 2e shows the result of Xu et al. [23], and Fig. 2f the result of our proposed method. Our method is better able to retain the original texture content from the transmission layer. Fig. 2g shows a single scanline from Fig. 2d, Fig. 2e, and Fig. 2f overlapped one on top of the other. Note how our proposed method can suppress the reflection component as well as or better than the method of Xu et al. [23].

3.1. Sub-problem 1

The optimization problem over $\mathbf{T}$, keeping $\mathbf{G}$ fixed, is given by

$$\mathbf{T}^* = \arg\min_{\mathbf{T}} \|\mathcal{L}(\mathbf{T}) - \mathcal{L}(\mathbf{Y})\|_2^2 + \beta\|\nabla\mathbf{T} - \mathbf{G}\|_2^2. \quad (9)$$

The function is quadratic and can be solved analytically using the Fast Fourier Transform (FFT). The FFT diagonalizes the Laplacian and gradient operators and the linear system is solved in the Fourier domain with point-wise divisions.

However, a spatial shift can occur using this approach (also reported by Li and Brown [12]), because the Laplacian data fidelity term is insensitive to any global shift on the pixel values. In other words, for any constant signal $\mathbf{C} \in \mathbb{R}^{n \times m}$, it holds $\mathcal{L}(\mathbf{T}+\mathbf{C}) = \mathcal{L}(\mathbf{T})$. Li and Brown [12] try to compensate for this shift by re-normalizing the output to fall within a range. However, due to the large dimensionality of the problem and numerical inconsistencies, a solution based on a constant global shift is not suitable. This may explain the color shifts we observe in the results of Li and Brown [12] shown in the second column of Fig. 6.

To overcome this problem, we instead use gradient descent, applying Adam [7], an accelerated first-order gradient descent method, to minimize Eq. (9). The gradient of the objective function in Eq. (9) is given by

$$\nabla_{\mathbf{T}} = 2\mathcal{L}^T(\mathcal{L}(\mathbf{T}) - \mathcal{L}(\mathbf{Y})) + 2\beta\nabla^T(\nabla\mathbf{T} - \mathbf{G}), \quad (10)$$

where $\mathcal{L}^T$ and ∇^T denote the transposed Laplacian and gradient operators, respectively.

3.2. Sub-problem 2

This sub-problem is equivalent to the second sub-problem in [23]. We state it here for the sake of completeness. The objective function over $\mathbf{G}$ is given by

$$\mathbf{G}^* = \arg\min_{G} \|\nabla\mathbf{T} - \mathbf{G}\|_2^2 + \frac{\lambda}{\beta}C(\mathbf{G}). \quad (11)$$

The objective of Eq. (11) can be spatially decomposed over the elements of $\mathbf{H}$ and $\mathbf{V}$ and efficiently optimized. For each pixel (i, j) the analytic solution of Eq. (11) is given by

$$(H_{i,j}, V_{i,j}) = \begin{cases} (0,0), & (\nabla_x T_{i,j})^2 + (\nabla_y T_{i,j})^2 \leq \lambda/\beta \\ (\nabla_x T_{i,j}, \nabla_y T_{i,j}), & \text{otherwise} \end{cases}. \quad (12)$$

A proof of Eq. (12) is given by Xu et al. [23]. We summarize our proposed optimization method in Alg. 1.

4. Experiments

As discussed in Section 2.2, there are only a few algorithms that attempt single-image reflection removal. We compare our method against the state-of-the-art methods of Li and Brown [12] and Wan et al. [22] for a collection of synthetic and real-world images. Of the other methods, the method of Levin and Weiss [9] requires user labeling, while the method of Shih et al. [17] is computationally intractable for the image sizes (greater than 0.5 million pixels) we use in our experiments. Moreover, it assumes that the reflection

Algorithm 1: Reflection Suppression with Laplacian l_0 Minimization

---

Input: $\mathbf{Y}, \lambda, \beta_0, \beta_{\max}, \kappa$
Initialize: $\mathbf{T} \leftarrow \mathbf{Y}, \ \beta \leftarrow \beta_0$
while $\beta \leq \beta_{\max}$ **do**
    Update $\mathbf{G}$ using Eq. (12) ;
    Update $\mathbf{T}$ using Adam and Eq. (10) ;
    $\beta \leftarrow \kappa\beta$;
end
Output: Final image $\mathbf{T}$

---

image contains strong ghosting, which does not cover all the real-world scenarios.

For the method of Li and Brown [12] we fix the regularization parameters to $\lambda_{lb} = 100$. This value is a good compromise between retaining transmission components and removing reflections (see discussion in [12]). The results of Wan et al. [22] were provided by the authors using the default parameters described in their paper. For the Adam optimizer of our method we use the default parameters as described by Kingma and Ba [7].

For our algorithm, we fix the regularization parameter to $\lambda_{ours} = 0.002$. We observe empirically that a range of values in the interval $[0.001, 0.005]$ is a good starting point for images with reflections. We show the effect of the regularization parameter on the output of our method in Fig. 3. The larger the parameter, the more reflection components are removed. But simultaneously, more details from the transmission layer are also lost. We fix the parameters of Alg. 1 to $\beta_0 = 2\lambda_{ours}$, $\beta_{\max} = 10^5$ and $\kappa = 2$ as in Xu et al. [23]. Source code is available at http://ivrl.epfl.ch/research/reflection_removal.

4.1. Synthetic Images

We create synthetic images to simulate reflections using the model of Eq. (2) with constant $\mathbf{W} = w$. We fix the kernel $\mathbf{k}$ to a Gaussian with $\sigma = 2$ and we use two blending weights $w = \{0.7, 0.5\}$. In Fig. 4 we show the images we use as transmission (Figs. 4a, 4c) and reflection layers (Figs. 4b, 4d) for the synthetic experiments. Our method better suppresses the unwanted effect of the reflection components compared to the competing methods. This is shown quantitatively in Table 1 in terms of PSNR and SSIM, where we are better than the competing methods by a significant margin.

The method of Li and Brown [12] has the tendency to produce dark outputs and false colors (see Figs. 5b, 5f, 5j, 5n), a fact that strongly affects its evaluation on the objective metrics. The method of Wan et al. [22] better reproduces color, however it creates unwanted artifacts and excessive smoothing in the transmission layer (see Figs. 5c, 5g, 5k, 5o).

(a) Input   (b) Proposed, $\lambda = 0.001$

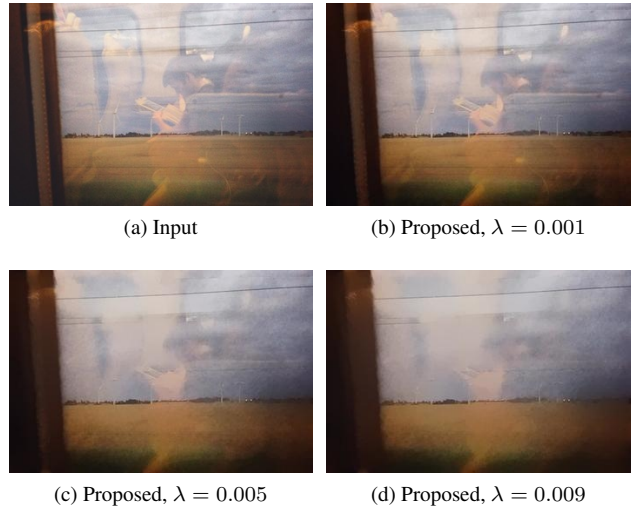(c) Proposed, $\lambda = 0.005$ (d) Proposed, $\lambda = 0.009$

Figure 3. Effect of the regularization parameter on our reflection suppression method. The larger the parameter, the more reflection components are removed. However, more details from the transmission layer are also lost. Best viewed on screen.

(a) $\mathbf{T}_1$ (b) $\mathbf{R}_1$

(c) $\mathbf{T}_2$ (d) $\mathbf{R}_2$

Figure 4. Images used as transmission (left column) and reflection layers (right column) for the synthetic experiments.

Even in the limit case of w approaching 0.5, when all methods perform poorly, our method still performs better than the state-of-the-art (see the third and fifth row of

(a) $\mathbf{T}_1 + \mathbf{R}_1, w = 0.7, \sigma = 2$ (b) [12]   (c) [22]   (d) Proposed

(e) $\mathbf{T}_1 + \mathbf{R}_1, w = 0.5, \sigma = 2$ (f) [12]   (g) [22]   (h) Proposed

(i) $\mathbf{T}_2 + \mathbf{R}_2, w = 0.7, \sigma = 2$ (j) [12]   (k) [22]   (l) Proposed

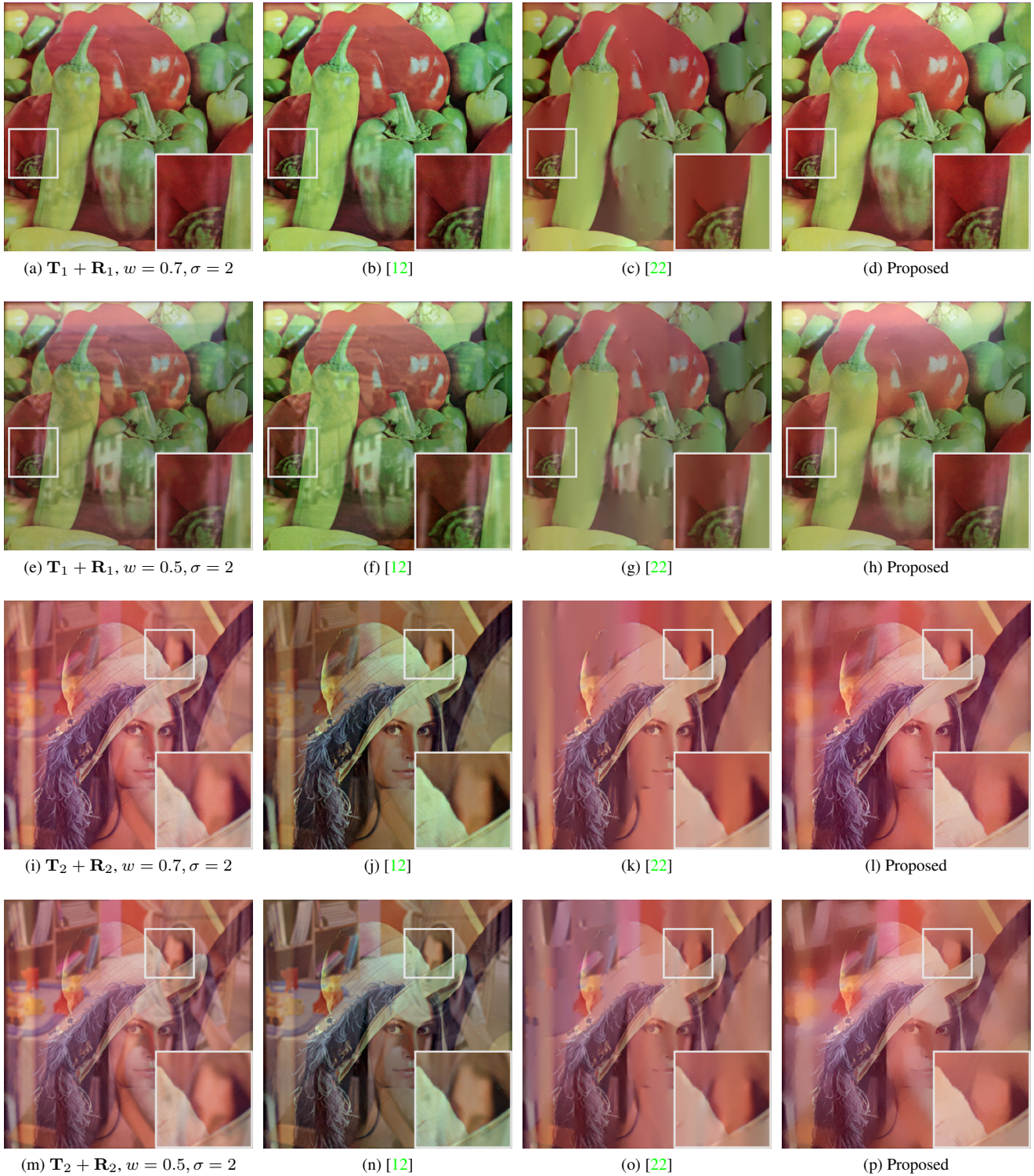(m) $\mathbf{T}_2 + \mathbf{R}_2, w = 0.5, \sigma = 2$ (n) [12]   (o) [22]   (p) Proposed

Figure 5. Comparison of reflection removal methods on synthetic images. We use two blending weights $w = \{0.5, 0.7\}$. Compared to Li and Brown [12] in the second column and Wan et al. [22] in the third column, our method (fourth column) gives superior color reproduction and reflection suppression results. Best viewed on screen.
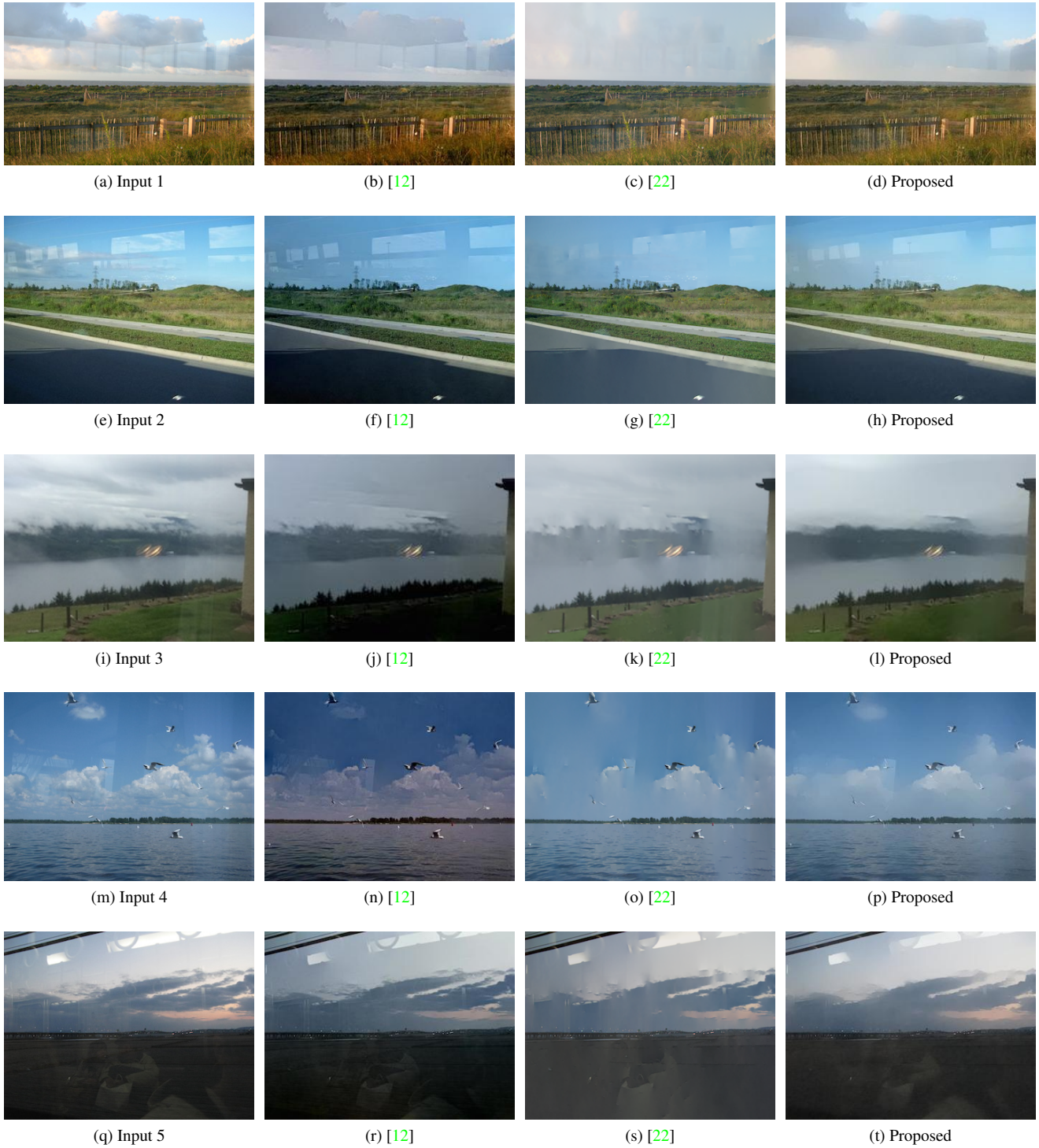
Figure 6. Comparison of reflection removal methods on real-world images taken from the Internet. As in the case of artificial images (Fig. 5), the method of Li and Brown [12] in the second column suffers from poor color reproduction, while the method of Wan et al. [22] in the third column over-smooths the image. Our method in fourth column gives superior color reproduction and reflection suppression results compared to both the state-of-the-art techniques. Best viewed on screen.

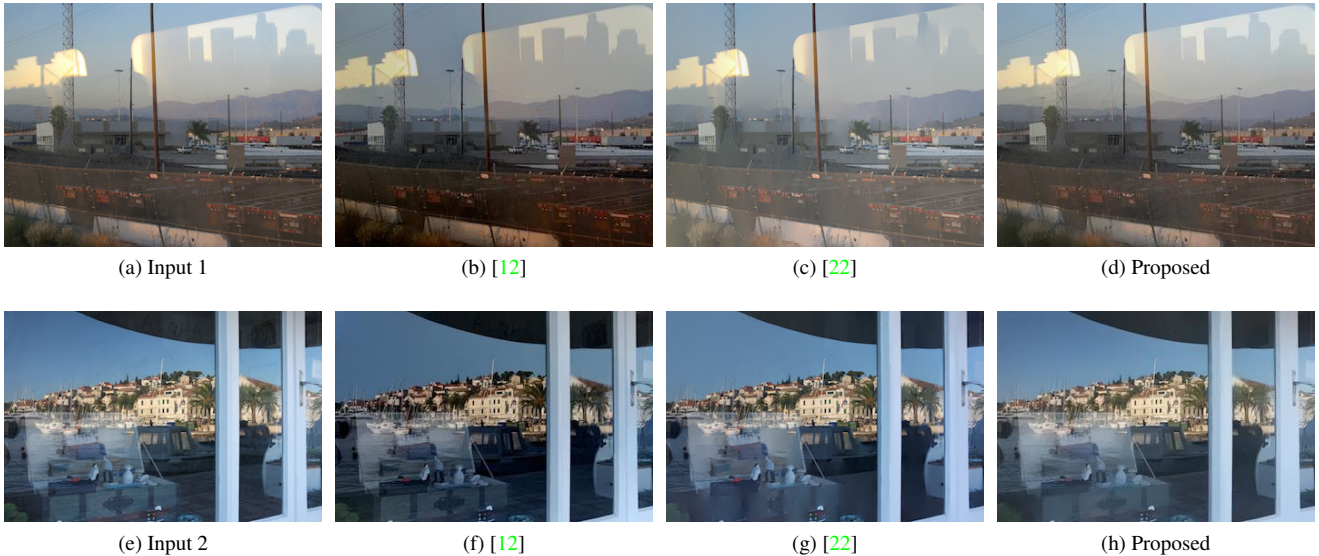| (a) Input 1 | (b) [12] | (c) [22] | (d) Proposed |

| (e) Input 2 | (f) [12] | (g) [22] | (h) Proposed |

Figure 7. Failure cases in reflection removal. In this figure, we show the cases where our technique fails to suppress reflections satisfactorily. We pose our results next to the results from the state-of-the-art [12, 22] to show that they are equally unsuccessful in removing reflections. In these images, the reflection components are too strong to be distinguished from the main transmission layer. Best viewed on screen.

	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Fig. 5a	20.15	0.89	18.33	0.86	**20.37**	**0.93**
Fig. 5e	15.61	0.77	15.63	0.72	**15.72**	**0.81**
Fig. 5i	14.78	0.62	20.17	0.93	**22.35**	**0.96**
Fig. 5m	14.15	0.55	17.76	0.85	**18.09**	**0.89**
	Li and Brown [12]		Wan et al. [22]		Proposed	

Table 1. PSNR and SSIM values for the synthetic experiments. In all cases our algorithm performs better than the state-of-the-art by a significant margin.

Fig. 5). In this case, the reflection component is as strong as the transmission and therefore is hard to suppress or remove. This is to be expected, because all methods rely on the assumption that reflections are weaker and smoother than the transmission signal. However, in real-world scenarios, these assumptions may not hold.

4.2. Real-World Images

We provide results on real-world reflection images downloaded from the Internet. For these images, the comparison can only be visual, since we do not have any ground truth. The results from the methods considered are shown in Fig. 6. Similar to the synthetic experiments, our algorithm is better able to suppress reflections and it preserves colors in the transmission layer without producing additional artifacts. The method of Li and Brown [12] produces dark images, often without maintaing color fidelity. The method of Wan et al. [22] introduces artifacts and over-smooths parts

of the transmission layer. In the cases where the reflections are strong and sharp, we observe that none of the tested algorithms can suppress or remove them, because in such cases the assumptions reflection removal algorithms make are not valid anymore. In Fig. 7 we show two examples of strong reflections, where none of the algorithms succeeds. Note however, that our results are still the closest to the original images and do not contain any additional artifacts.

5. Conclusion

Reflection removal from a single image is a highly ill-posed problem. In order to take into account visual continuity of reflection structures and simultaneously retain important details in the image, we formulate our optimization problem using a Laplacian data fidelity term and an l_0 prior term. We test our approach with experiments on artificial and real-world images and compare our results against the state-of-the-art. Our approach performs better at suppressing reflections than previous single-image reflection removal algorithms.

There is a trade-off between suppressing reflection artifacts and simultaneously retaining high-frequency details. While we clearly outperform the state-of-the-art, we observe that suppressing reflections from a single image remains a hard problem. There is still room to develop a general solution that works for a wide range of images. In particular, it may be interesting to direct future research on visual perception based reflection suppression.

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, 2005. 2

[2] H. G. Barrow and J. M. Tenenbaum. *Recovering Intrinsic Scene Characteristics from Images*. Academic Press, 1978. 1

[3] W. D. Ellis. *A Source Book of Gestalt Psychology*. New York: Harcourt, Brace & World, 1938. 3

[4] H. Farid and E. H. Adelson. Separating reflections and lighting using independent components analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, page 267 Vol. 1, 1999. 2

[5] K. Gai, Z. Shi, and C. Zhang. Blind Separation of Superimposed Moving Images Using Image Statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(1):19–32, 2012. 1, 2

[6] X. Guo, X. Cao, and Y. Ma. Robust Separation of Reflection from Multiple Images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2195–2202, 2014. 1, 2

[7] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980, 2014. 4, 5

[8] N. Kong, Y. W. Tai, and J. S. Shin. A Physically-Based Approach to Reflection Separation: From Physical Modeling to Constrained Optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(2):209–221, 2014. 2, 3

[9] A. Levin and Y. Weiss. User Assisted Separation of Reflections from a Single Image Using a Sparsity Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1647–1654, 2007. 2, 4

[10] A. Levin, A. Zomet, and Y. Weiss. Separating reflections from a single image using local features. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–306–I–313 Vol.1, 2004. 2

[11] Y. Li and M. S. Brown. Exploiting Reflection Change for Automatic Reflection Removal. In *IEEE International Conference on Computer Vision*, pages 2432–2439, 2013. 1, 2

[12] Y. Li and M. S. Brown. Single Image Layer Separation Using Relative Smoothness. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2752–2759, 2014. 2, 4, 5, 6, 7, 8

[13] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear Total Variation Based Noise Removal Algorithms. *Physica D*, 60(1-4):259–268, 1992. 3, 4

[14] B. Sarel and M. Irani. Separating Transparent Layers through Layer Information Exchange. In *European Conference on Computer Vision*, pages 328–341, 2004. 1, 2

[15] Y. Y. Schechner, N. Kiryati, and R. Basri. Separation of Transparent Layers using Focus. *International Journal of Computer Vision*, 39(1):25–39, 2000. 2

[16] Y. Y. Schechner, J. Shamir, and N. Kiryati. Polarization and statistical analysis of scenes containing a semireflector. *Journal of the Optical Society of America A*, 17(2):276–284, 2000. 2

[17] Y. Shih, D. Krishnan, F. Durand, and W. T. Freeman. Reflection Removal using Ghosting Cues. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3193–3201, 2015. 2, 4

[18] S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski. Image-based Rendering for Scenes with Reflections. *ACM Transactions on Graphics*, 31(4):100:1–100:10, 2012. 1, 2

[19] T. Sirinukulwattana, G. Choe, and I. S. Kweon. Reflection removal using disparity and gradient-sparsity via smoothing algorithm. In *IEEE International Conference on Image Processing*, pages 1940–1944, 2015. 2

[20] C. Sun, S. Liu, T. Yang, B. Zeng, Z. Wang, and G. Liu. Automatic Reflection Removal Using Gradient Intensity and Motion Cues. In *ACM Multimedia*, MM '16, pages 466–470, 2016. 1, 2

[21] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 246–253 vol.1, 2000. 1, 2

[22] R. Wan, B. Shi, T. A. Hwee, and A. C. Kot. Depth of field guided reflection removal. In *IEEE International Conference on Image Processing*, pages 21–25, 2016. 2, 4, 5, 6, 7, 8

[23] L. Xu, C. Lu, Y. Xu, and J. Jia. Image Smoothing via L0 Gradient Minimization. *ACM Transactions on Graphics*, 30(6):174:1–174:12, 2011. 3, 4, 5

[24] T. Xue, M. Rubinstein, C. Liu, and W. T. Freeman. A Computational Approach for Obstruction-free Photography. *ACM Transactions on Graphics*, 34(4):79:1–79:11, 2015. 1, 2