

Article

# Variational Model for Single-Image Reflection Suppression Based on Multiscale Thresholding

Pei-Chiang Shao

Department of Mathematics, Soochow University, Taipei City 111002, Taiwan; shaopi823@gmail.com;  
Tel.: +886-2-2881-9471 (ext. 6709)

**Abstract:** Reflections often cause degradation in image quality for pictures taken through glass medium. Removing the undesired reflections is becoming increasingly important. For human vision, it can produce much more pleasing results for multimedia applications. For machine vision, it can benefit various applications such as image segmentation and classification. Reflection removal is itself a highly illposed inverse problem that is very difficult to solve, especially for a single input image. Existing methods mainly rely on various prior information and assumptions to alleviate the ill-posedness. In this paper, we design a variational model based on multiscale hard thresholding to both effectively and efficiently suppress image reflections. A direct solver using the discrete cosine transform for implementing the proposed variational model is also provided. Both synthetic and real glass images are used in the numerical experiments to compare the performance of the proposed algorithm with other representative algorithms. The experimental results show the superiority of our algorithm over the previous ones.

**Keywords:** multiscale thresholding; reflection removal; reflection suppression; variational model



**Citation:** Shao, P.-C. Variational Model for Single-Image Reflection Suppression Based on Multiscale Thresholding. *Sensors* **2022**, *22*, 2271. <https://doi.org/10.3390/s22062271>

Academic Editors: Shih-Chia Huang, Yan-Tsung Peng, Benjamin C. M. Fung and Cheng Zhang

Received: 20 February 2022

Accepted: 14 March 2022

Published: 15 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

For most multimedia and computer vision applications, the input images are normally assumed to be both clean and clear. However, bad lighting conditions and environments such as photographing behind a window glass or showcase are almost inevitable in our daily life. In these scenarios, the captured images often contain undesired reflections from the objects on the camera side, which severely degrade image quality and affect the performance of multimedia and computer vision applications, such as image segmentation and classification.

Due to the rapid growth of mobile devices, a fast and effective reflection removal algorithm is increasingly desired. Reflection removal is the restoration process that aims essentially at removing or suppressing unpleasing reflections and improving the visibility of the scenes in front of the glass medium. As a preprocessing step, it can benefit various follow-up applications.

Barrow et al. [1] first proposed a linear assumption for an observed degraded image  $Y$  as a superposition of two layers as follows:

$$Y = T + R, \quad (1)$$

where  $T$  and  $R$  are the transmission and reflection layers, respectively. The inverse problem of resolving  $T$  and  $R$  given  $Y$  is highly ill-posed, since there is only one equation for two unknowns per pixel. According to the number of input images  $Y$  used, reflection removal methods can be classified into multiple-image and single-image approaches, where the former makes the problem less ill-posed while the latter is itself highly ill-posed and challenging. Since perfectly separating the transmission and reflection layers is, in general, very difficult even for the case of multiple input images and, in most situations, people tend to care only about the transmission layers, estimating a reasonable  $T$  that might contain

slightly misclassified information from  $R$  can be considered instead. To make the ill-posed problem more tractable for single input image, researchers tried to suppress the reflection in the image (i.e., estimating a reasonable  $T$ ) instead of removing the reflection (i.e., solving an exact separation of  $T$  and  $R$ ) [2,3]. In this paper, we follow the same idea and consider the reflection suppression problem instead.

Multiple-image approaches generally use several degraded images as input, which are captured by a fixed camera. Kong et al. [4] employed a physical formation model and used different angular filters to estimate the optimal reflection layer. Similarly, Schechner et al. [5] focused on various distances using multiple images for reflection removal. Farid and Adelson [6] employed independent component analysis to separate reflections from glossy surfaces or glass using two images with different polarization angles. Agrawal et al. [7] made use of two degraded images captured with and without a flash exposure to identify reflections. Multiple images captured with different camera positions were also considered. Sirinukulwattana et al. [8] exploited varying reflections taken from different viewpoints to smooth reflections and enhance the transmission. Similarly, Li and Brown [9] exploited the subtle changes in the reflection via different viewpoints. Video sequences were also used to separate transmission and reflection layers. Xue et al. [10] exploited motion differences and recovered the transmission and reflection layers. Gai et al. [11] jointly estimated layer motions via a blind separation algorithm. Guo et al. [12] exploited the correlation of the transmission layers across multiple images to separate the two layers. Multiple input images largely reduce the ill-posedness of the reflection removal problem. However, due to fastidious image acquisition requirements, algorithms based on multiple images have limited practical applications.

In most situations, one may not have the chance to capture multiple images of the same scene, which is especially difficult with slightly different camera positions or polarization angles. As a consequence, single-image approaches are becoming practically important. As mentioned above, reflection removal based on a single input image is itself a highly ill-posed problem, and existing methods mainly assume different image priors for the transmission and reflection layers. Levin and Weiss [13] exploited sparse gradient priors with manual labels to distinguish the two layers. Li and Brown [14] assumed that two layers have different smoothness and imposed the smooth and sparse gradient priors over the two layers. Shih et al. [15] exploited the Gaussian mixture model prior for the reflections to solve a deblur optimization problem. Wan et al. [16] exploited the depth of field in a multiscale manner using K-L divergence to identify edges of different layers.

Recently, deep neural networks are also adopted to remove reflection. Fan et al. [17] proposed the first deep neural network to solve the layer separation problem. They exploited edge information to guide the separation process. Wan et al. [18] integrated a gradient inference network with a image inference network as a concurrent network to remove the reflection of a single input image. Yang et al. [19] proposed a cascade deep network to bidirectionally estimate the transmission and reflection layers. More precisely, the network uses for the estimated transmission to estimate reflection and then uses for the estimated reflection to estimate transmission. Zhang et al. [20] proposed a convolutional network with a feature loss, an adversarial loss, and an exclusion loss that enforces the layer separation. The main advantage of deep learning-based approaches for reflection removal is that no handcrafted image priors are necessary. However, a large amount of deep network training time and training datasets is required.

The remainder of this paper is organized as follows. In Section 2, we briefly review two related works for reflection suppression. In Section 3, we present the improved reflection suppression model for a single-input image. The proposed model is based on multiscale hard thresholding, and a direct solver using the discrete cosine transform for implementing the proposed model is also given in Section 3. Numerical experiments are conducted in Section 4 to demonstrate the superior performance of the newly proposed model. Finally, some concluding remarks and discussions are given in Section 5.

## 2. Related Work

As introduced in Section 1, reflection removal is a highly ill-posed inverse problem that is very challenging to solve, especially for the case of a single input image. Arvanitopoulos et al. [2] tried to suppress the reflection instead of removing the reflection. They exploited a Laplacian data fidelity term and a sparse gradient prior, which achieve satisfactory quality for reflection suppression. However, since their model is nonconvex, their algorithms are quite inefficient and require a large number of iterations to converge to a desirable result.

As an improvement of [2], Yang et al. [3] proposed a hard thresholding operation to the gradient of the input image in the Laplacian data fidelity term. Their method also achieves satisfactory quality for reflection suppression, but the algorithm is highly efficient compared with [2]. Since our method is an improvement of [3], we briefly review the works of [2,3] in the following two subsections.

### 2.1. Smooth Regularization on $T$

The algorithm proposed in [2] for reflection suppression mainly relies on the critical assumption that reflection edges (gradients) are smaller in magnitude and less in focus compared to transmission edges. This assumption is reasonable in real-world scenarios since the camera focuses on transmission objects rather than reflection components. Mathematically, this assumption can be expressed as the following formation model [21]:

$$Y = W \circ T + (1 - W) \circ (k_\sigma * R), \quad (2)$$

where  $\circ$  and  $*$  denote the component-wise multiplication and the convolution operation, respectively.  $W$  is the contribution weight of the transmission  $T$  to the observation  $Y$ , and  $k_\sigma$  is a Gaussian blur kernel with standard deviation  $\sigma$  performed on the reflection layer  $R$ . In general,  $W$  is pixel-dependent and depends on the lighting conditions and the distance of the camera to the captured scene. For simplicity, weight  $W$  can be considered as a constant, i.e.,  $W(x) = w, \forall x \in \Omega$ , where  $\Omega$  and  $x = (x_1, x_2)$  denote the image domain and a pixel point, respectively.

Following the successful smoothing model of Xu et al. [22], Arvanitopoulos et al. [2] employed the number of non-zero gradients as the prior information to restrict and regularized output transmission. Their model can eliminate small gradients and simultaneously preserve large image edges. The original minimization problem in [22] is as follows:

$$\min_T \|T - Y\|_2^2 + \lambda C(T), \quad (3)$$

where the following:

$$\|T - Y\|_2^2 := \min_T \int_{\Omega} (T - Y)^2 dx, \quad (4)$$

is the  $L^2$  data fidelity term which keeps the optimal  $T$  close to the observation  $Y$ , and the following:

$$C(T) := \#\{x \in \Omega : |\nabla T(x)| \neq 0\}, \quad (5)$$

is the  $L^0$  regularization term where  $\#$  denotes the cardinality measure applied on a given set (i.e., a measure of the number of elements of the set). The  $L^0$  term regularizes the optimal  $T$  by implicitly performing a hard thresholding of its gradients with threshold  $\lambda$ . Since high frequency details of the image will be eliminated, model (3) is suitable for image smoothing but unsuitable for reflection suppression.

To preserve high frequency details in optimal  $T$ , Arvanitopoulos et al. [2] proposed a Laplacian data fidelity term to replace (4), where the Laplacian of the optimal  $T$  is defined by the following.

$$\mathcal{L}(T) := \text{div}(\nabla T) = \frac{\partial^2 T}{\partial x_1^2} + \frac{\partial^2 T}{\partial x_2^2}. \quad (6)$$

Their data fidelity term based on the Laplacian to enforce the consistency in fine structures is then given as follows:

$$\min_T \|\mathcal{L}(T) - \mathcal{L}(Y)\|_2^2 + \lambda C(T), \quad (7)$$

where regularization term  $C(T)$  is the same as defined in (5). The new model can preserve strong edges and details in transmission  $T$  while eliminating reflection  $R$  simultaneously; thus, it is more suitable for reflection suppression.

## 2.2. Direct Thresholding on $Y$

In [2], the edge information of the optimal  $T$  is preserved by applying the Laplacian differential operator  $\mathcal{L}$ . The reflection component  $R$  is suppressed by applying regularization  $C$  on  $T$ . As the regularization parameter  $\lambda$  increases, more reflections are suppressed in  $T$ . However, as their model defined in (7) is nonconvex, their operator splitting algorithm for minimizing the nonconvex energy functional is highly inefficient and takes a large number of iterations to converge to a desired solution. Therefore, Yang et al. [3] integrated these ideas into a new model formulation, which significantly reduces computation times while retaining satisfactory reflection suppression performances.

Instead of using the hard thresholding operation derived implicitly from the smoothing regularizer  $C$  on  $T$ , Yang et al. [3] adopted the idea from [23,24] to directly integrate the hard thresholding operation for gradients on  $Y$  into the energy functional, which is expressed as follows:

$$\min_T \frac{1}{2} \|\mathcal{L}(T) - \text{div}(\delta_h(\nabla Y))\|_2^2 + \frac{\varepsilon}{2} \|T - Y\|_2^2, \quad (8)$$

where  $\delta_h$  is the hard thresholding operation for gradients and is defined as follows.

$$\delta_h(G(x)) := \begin{cases} G(x), & \text{if } |G(x)| \geq h \\ 0, & \text{if } |G(x)| < h \end{cases}. \quad (9)$$

In the first term of (8), the gradient of the observation  $Y$ , i.e.,  $\nabla Y$ , is thresholded using  $\delta_h$  before taking the divergence operator  $\text{div}$ . Since  $\nabla Y(x)$  with magnitude  $|\nabla Y(x)|$  smaller than  $h$  will be truncated to zero, the model has the ability to suppress reflections in  $T$ . Since the data term also enforces the consistency of  $T$  to  $Y$  via Laplacian operator  $\mathcal{L}$ , fine details of  $Y$  will be preserved simultaneously in  $T$ . Therefore, the  $\delta_h$  in (8) takes over the role of  $C$  in (7) for reflection suppression.

Since the data fidelity term in (8) contains only second-order derivatives of transmission  $T$ , the optimal solution of  $T$  will not be unique, and any optimal transmission  $T$  shifted by a affine function  $A$ , i.e.,  $T + A$ , will also be an optimal transmission. To ensure the uniqueness of the optimal solution, the second term of (8) is incorporated. The model parameter  $\varepsilon$  should be taken as a positive tiny number so as not to cancel out the effect of reflection suppression, which is performed via the first term.

## 3. Proposed Method

In this section, we present an improved version of the model of Yang et al. [3] given in (8) for single-image reflection suppression. The proposed model is based on multiscale hard thresholding, and a direct solver using the discrete cosine transform for implementing the proposed model is also given.

### 3.1. Multiscale Thresholding

The hard thresholding model given in (8) suppresses the reflections by eliminating the gradients of  $Y$  at pixel point  $x$  with magnitude  $|\nabla Y(x)|$  less than the prescribed threshold value  $h$ . For a small  $h$ , the model eliminates only small gradients; thus, only weak reflections will be suppressed while strong reflections remain in the optimal transmission. For a large  $h$ , the model eliminates both small and large gradients; thus, both weak and strong reflections will be suppressed simultaneously. However, when  $h$  is large, many important image edges

in the transmission with their gradient magnitudes smaller than  $h$  will also be suppressed. Therefore, a large  $h$  in (8) usually causes information loss in the optimal transmission. This is the reason why the model should assume that reflection edges are smaller in magnitude and less in focus compared to transmission edges. If this assumption fail to hold, the model will produce either (1) a blurry transmission without any reflections when  $h$  is large or (2) a clear transmission but with many reflections remaining when  $h$  is small. Therefore, a moderate  $h$  is normally picked to trade off between situations (1) and (2) to obtain an acceptable performance.

To suppress both weak and strong reflections while preventing the information loss of important edges in the optimal transmission, we propose a multiscale hard thresholding energy functional to improve the performance of the model of Yang et al. in (8). Beginning with the smallest thresholding scale  $h$ , our model additionally considers another  $N - 1$  countable thresholding scales, i.e.,  $\{2h, 3h, \dots, Nh\}$ . Hence, there are totally  $N$  scales in our model, that is,  $\{h, 2h, 3h, \dots, Nh\}$ , where scale number  $N$  is a prescribed positive integer which can be determined by the user. The model in (8) is first modified as follows.

$$\min_T \frac{1}{2} \sum_{n=1}^N \|\mathcal{L}(T) - \text{div}(\delta_{nh}(\nabla Y))\|_2^2 + \frac{\varepsilon}{2} \|T - Y\|_2^2. \quad (10)$$

When  $N = 1$ , the multiscale energy functional reduces to the energy functional of Yang et al. in (8). When  $N \geq 2$ , the multiscale energy functional has  $N + 1$  terms, and the first  $N$  terms are linearly fused into one with uniform weights  $\frac{1}{2}$ , where each term involves a hard thresholding operation on  $\nabla Y$  with threshold value  $nh$  at  $n$ -th scale. Therefore, the first term in (10) can be viewed as a fusion term to integrate all information of the thresholded Laplacian of  $Y$  at different scales. As  $n$  increases, more and more weak edges are eliminated; thus, only strong edges are left. In other words, the small  $n$  term provides the suboptimal transmission  $T$  with both strong and weak edges while the large  $n$  term provides the suboptimal transmission  $T$  with only strong edges. After adding the second term in (10) to ensure the uniqueness of the optimal solution, the final optimal transmission  $T$  can be viewed as a linear combination of the sub-optimal transmissions of those subfunctionals corresponding to different thresholding scales. The model parameter  $\varepsilon$  is also set to be a positive tiny number just as that provided in (8). In this paper,  $\varepsilon$  will be fixed at  $10^{-6}$ .

To further enhance the preserved edge information, we proposed a pixel-dependent weight function  $\varphi$  to adapt the thresholded gradients pixelwisely for all subfunctionals in (10) to form our final model for reflection suppression:

$$\min_T \frac{1}{2} \sum_{n=1}^N \|\mathcal{L}(T) - \text{div}(\varphi \delta_{nh}(\nabla Y))\|_2^2 + \frac{\varepsilon}{2} \|T - Y\|_2^2, \quad (11)$$

where the adaptive weight function  $\varphi$  is designed to be inversely proportional to  $|\nabla Y|$  as follows.

$$\varphi(x) = \left(1 - \frac{|\nabla Y(x)|}{\max_{x \in \Omega} |\nabla Y(x)|}\right) + \beta, \quad \forall x \in \Omega. \quad (12)$$

The role of shifting  $\beta$  in (12) is to maintain a base level of the adaptive weight function. In this paper,  $\beta$  will be fixed at 1.0.

### 3.2. Solving the Model

Since the energy functional proposed in (11) is convex, a transmission  $T$  is the optimal solution to (11) if and only if it satisfies the Euler-Lagrange equation derived step by step below:

$$\begin{aligned}
& \mathcal{L}^* \left( \sum_{n=1}^N (\mathcal{L}(T) - \text{div}(\varphi \delta_{nh}(\nabla Y))) \right) + \varepsilon(T - Y) = 0 \\
\Rightarrow & \left( \sum_{n=1}^N \mathcal{L}^* \mathcal{L}(T) \right) + \varepsilon T = \mathcal{L}^* \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y \\
\Rightarrow & (N\mathcal{L}^* \mathcal{L} + \varepsilon)T = \mathcal{L}^* \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y \\
\Rightarrow & (N\mathcal{L}\mathcal{L} + \varepsilon)T = \mathcal{L} \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y
\end{aligned} \tag{13}$$

where  $\mathcal{L}^*$  denotes the adjoint operator of  $\mathcal{L}$ . Since  $\mathcal{L}$  is self-adjoint, we have  $\mathcal{L}^* = \mathcal{L}$ . Taking the Fourier cosine transform  $\mathcal{F}_c$  on both sides, we have the following.

$$\begin{aligned}
\Rightarrow & \mathcal{F}_c[(N\mathcal{L}\mathcal{L} + \varepsilon)T] = \mathcal{F}_c[\mathcal{L} \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y] \\
\Rightarrow & (N\mathcal{F}_c(\mathcal{L}) \circ \mathcal{F}_c(\mathcal{L}) + \varepsilon) \circ \mathcal{F}_c(T) = \mathcal{F}_c[\mathcal{L} \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y]
\end{aligned} \tag{14}$$

Denoting  $K = \mathcal{F}_c(\mathcal{L})$  and using component-wise division, we have the following.

$$\mathcal{F}_c(T) = \frac{\mathcal{F}_c[\mathcal{L} \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y]}{NK \circ K + \varepsilon} \tag{15}$$

Taking the inverse Fourier cosine transform  $\mathcal{F}_c^{-1}$ , we arrive at the optimal transmission of (11) as follows:

$$T = \mathcal{F}_c^{-1} \left[ \frac{\mathcal{F}_c[\mathcal{L} \left( \sum_{n=1}^N \text{div}(\varphi \delta_{nh}(\nabla Y)) \right) + \varepsilon Y]}{NK \circ K + \varepsilon} \right], \tag{16}$$

where  $K$  is the Fourier cosine transform of the Laplacian operator, which can be represented, in its finite dimensional matrix form on an  $I \times J$  digital image lattice, as follows:

$$K_{i,j} = 2(\cos(\frac{i\pi}{I}) + \cos(\frac{j\pi}{J}) - 2), \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J-1, \tag{17}$$

and the differential operators  $\mathcal{L}$ ,  $\text{div}$  and  $\nabla$  can be replaced with standard finite difference schemes [25] for digital input images  $[Y_{i,j}]$  for  $0 \leq i \leq I-1$  and  $0 \leq j \leq J-1$ .

The overall algorithm of the proposed multiscale thresholding method for reflection suppression is summarized in (Algorithm 1), and a toy example in Figure 1 shows the effectiveness of the proposed algorithm.

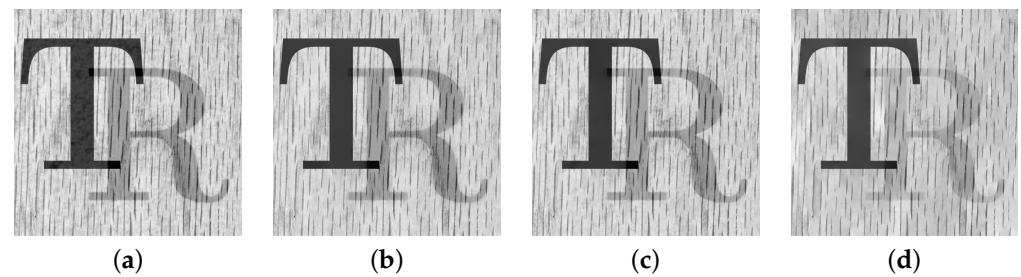
---

**Algorithm 1:** The proposed algorithm for reflection suppression by multiscale thresholding

---

**input** : a degraded rgb color image  $f$   
**set** : parameters  $h, N, \varepsilon = 10^{-6}, \beta = 1.0$   
**for**  $c$  in  $\{r, g, b\}$   
     $Y \leftarrow f^c$   
    compute  $\varphi$  using (12)  
    compute  $T$  using (16) and (17)  
     $g^c \leftarrow T$   
**endfor**  
**output:** the reflection suppression color image  $g$

---



**Figure 1.** Comparison using a toy example. (a) The input test image. (b) The result of algorithm [2]. (c) The result of algorithm [3]. (d) The result of proposed algorithm with  $N = 4$ .

#### 4. Numerical Experiment

In this section, we compare the proposed algorithm with the reflection suppression algorithms of Arvanitopoulos et al. [2] and Yang et al. [3], which are all implemented in MATLAB. Both synthetic and real glass images are tested. All experiments are implemented using MATLAB 2019a executed on a Windows PC with Intel Core i7-4790 processor and 16 GB RAM.

##### 4.1. Synthetic Glass Image

Two pairs of images of size  $512 \times 512$  were blended into one using Equation (2) to synthesize the test images. The constant weight  $W_{i,j} \equiv w$  is set at three different values, which are 0.7, 0.6, and 0.5. The standard deviation  $\sigma = 2$  is used for the Gaussian blur kernel. The model parameters are fixed at  $\lambda = 0.002$  for algorithm [2] and  $h = 0.01$ ,  $\varepsilon = 10^{-6}$ , for both algorithm [3] and the proposed algorithm according to the suggestions in [2,3]. Figure 2a shows two transmission images ‘Lena’ and ‘Clown’ and Figure 2b shows two reflection images ‘Barbara’ and ‘House’. The first test image was synthesized by blending ‘Lena’ with ‘Barbara’ using constant weight  $w$ , and the second test image was synthesized by blending ‘Clown’ with ‘House’ using constant weight  $w$ .



**Figure 2.** Two pairs of images to be synthesized. (a) Two transmission images: ‘Lena’ and ‘Clown’. (b) Two reflection images: ‘Barbara’ and ‘House’. The sizes of these four images are all  $512 \times 512$ .

We begin with  $w = 0.7$ , which corresponds to a relatively weak reflection effect among the three cases. Figure 3 shows the reflection suppression results of all competing algorithms for the two synthetic glass images. It can be found by observing Figure 3d,k that the proposed algorithm with scale number  $N = 1$  simultaneously possesses very significant image enhancement effects than the two competing algorithms in Figure 3b,c,i,j. As the scale number  $N$  increases from 2 to 4 as Figure 3e–g,l–n demonstrates, the proposed

algorithm is able to suppress more and more reflections while preserving the edge structures of the transmission images.

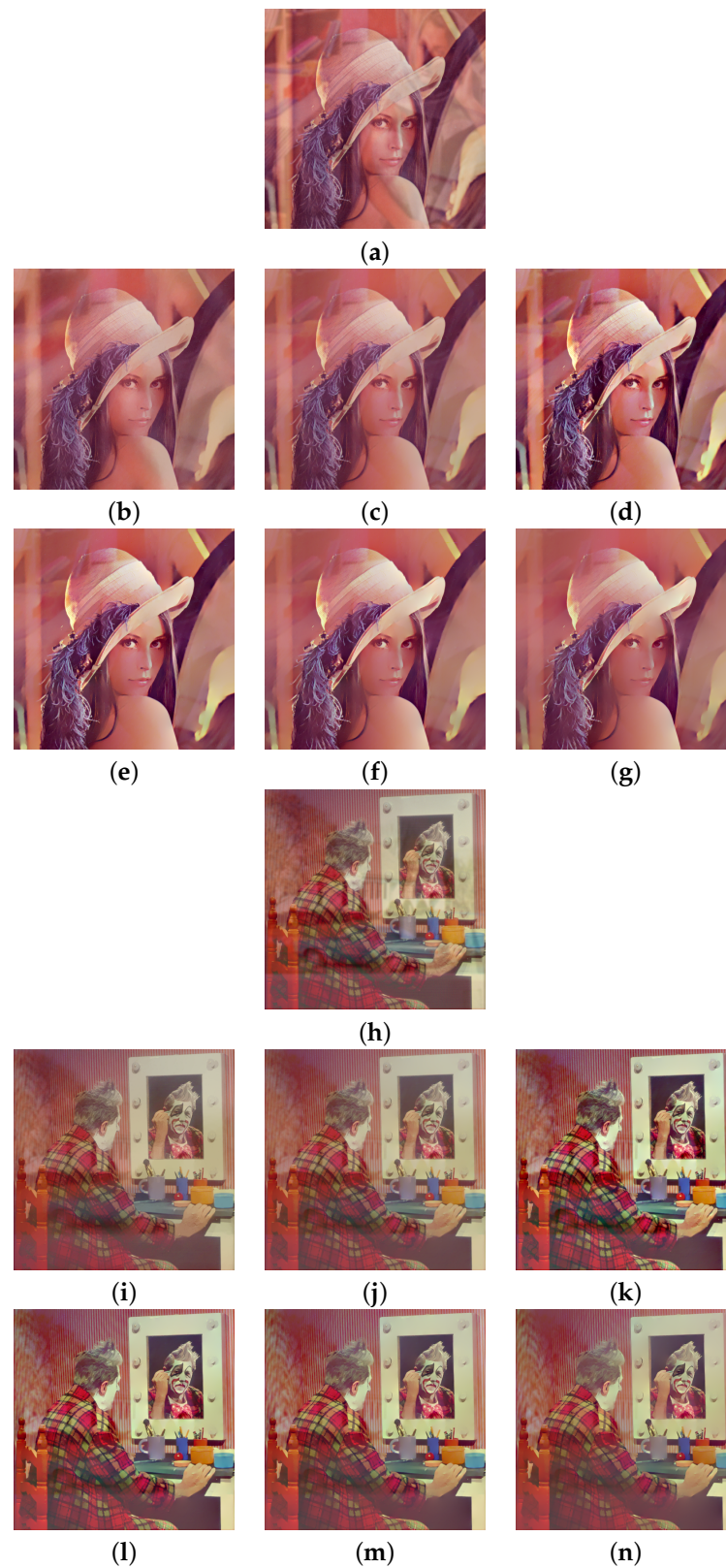
Similar results can be found in Figures 4 and 5, which correspond to the relatively moderate ( $w = 0.6$ ) and strong ( $w = 0.5$ ) reflection effects, respectively. The proposed algorithm with  $N = 1$  in Figures 4d,k and 5d,k still show significant enhancement effects and the reflections are increasingly suppressed as  $N$  increases. To evaluate the competing algorithms quantitatively, we compute the PSNR value for each reflection suppression result using the underlying true transmission image as the reference image. Table 1 shows the PSNR values, and the best value in each row is shown in boldface. It can be seen that the proposed algorithm with  $N = 2$  has the best PSNR values in all cases, and the PSNR values for  $N = 3$  and  $N = 4$  keep very close to that of algorithms [2,3]. This shows that the proposed algorithm is superior to the two representative algorithms. Theoretically, a large  $N$  in (11) can suppress more reflections but obtains lower PSNR values since important edges of the transmission are lost in the same time. On the other hand, a small  $N$  in (11) can get higher PSNR values but suppressing less reflections. Empirical observations from Table 1 suggest that one can set  $N = 2$  for high PSNR values or set  $N = 4$  for suppressing more reflections.  $N = 3$  is another choice that balances these two extreme choices.

We note that algorithms [2,3] produce very similar image quality and PSNR values. However, this is an expected phenomenon. Although the smooth regularization on  $T$  in (7) and the direct thresholding on  $Y$  in (8) have quite different functional forms, their core ideas are rather similar, i.e., “to eliminate small image gradients”. Model (8) directly eliminates the image gradients of  $Y$ , which are less than  $h$  in magnitude. For small  $h$ , less image gradients are eliminated; for large  $h$ , more image gradients are eliminated. On the other hand, model (7) eliminates the image gradients of  $T$  in the order of small magnitude to large magnitude, and the number of total eliminated gradients will be determined by model parameter  $\lambda$ . For small  $\lambda$ , less image gradients are eliminated; for large  $\lambda$ , more image gradients are eliminated. Although tuning parameters  $h$  and  $\lambda$  can produce very similar results and quality, they are still different since the energy functionals, and (7) and (8) are not equivalent. The two main advantages of model (8) compared to model (7) are as follows: (i) tuning the parameter  $h$  is much more intuitive than tuning the parameter  $\lambda$  since  $h$  is directly related to gradient magnitude while  $\lambda$  is not; and (ii) the energy functional (8) is convex in  $T$  while the energy functional (7) is nonconvex in  $T$ , which means that their algorithm will be less efficient.

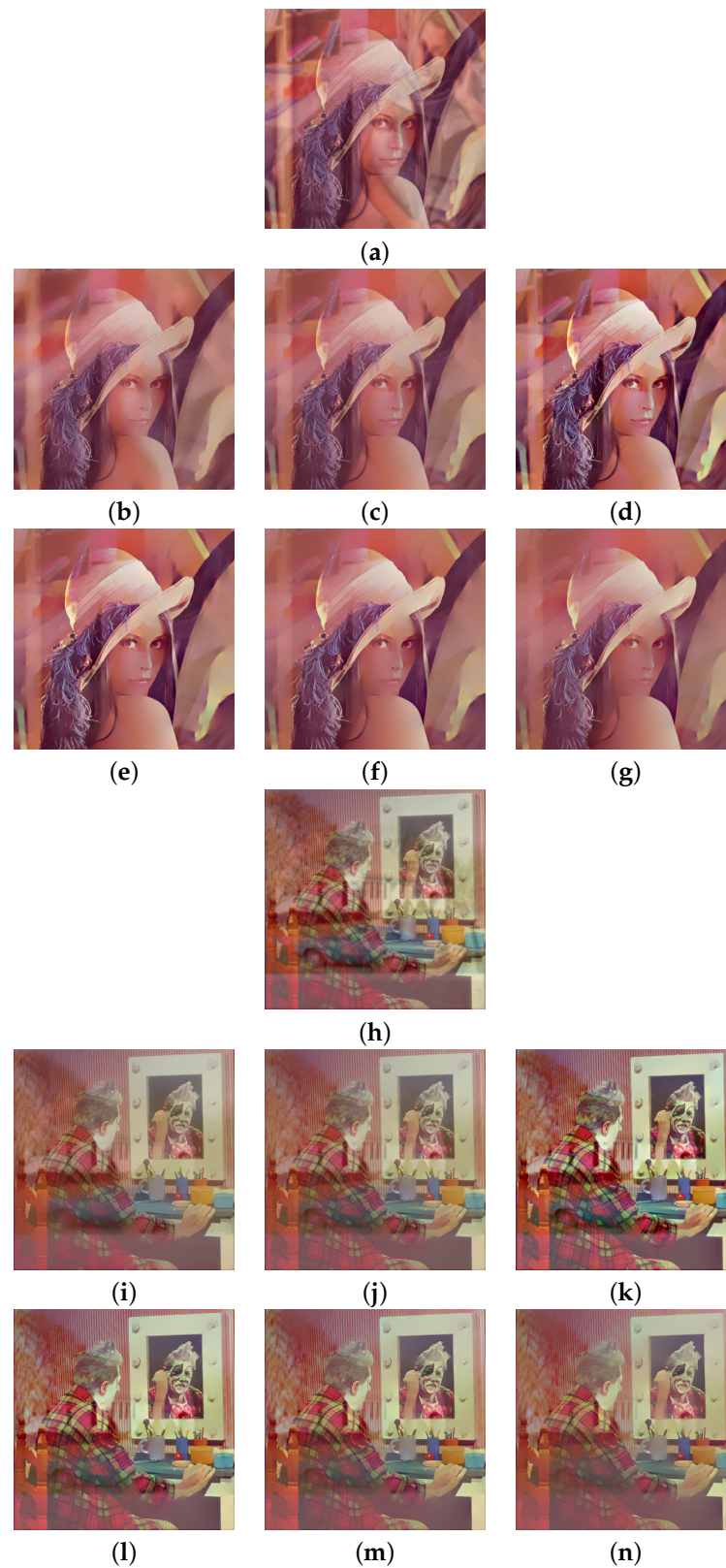
**Table 1.** PSNR values of the reflection suppression results in Figures 3–5. The best value in each row is shown in boldface.

Image	Algorithm [2]	Algorithm [3]	Proposed ( $N = 1$ )	Proposed ( $N = 2$ )	Proposed ( $N = 3$ )	Proposed ( $N = 4$ )
Figure 3a	70.18	70.16	<b>70.73</b>	70.71	69.94	69.14
Figure 3h	66.36	66.37	66.87	<b>67.06</b>	66.82	66.39
Figure 4a	67.96	67.99	68.64	<b>68.65</b>	68.03	67.41
Figure 4h	63.96	63.97	64.31	<b>64.47</b>	64.27	63.96
Figure 5a	66.21	66.22	66.44	<b>66.68</b>	66.28	65.86
Figure 5h	62.07	62.07	62.15	<b>62.35</b>	62.24	62.05

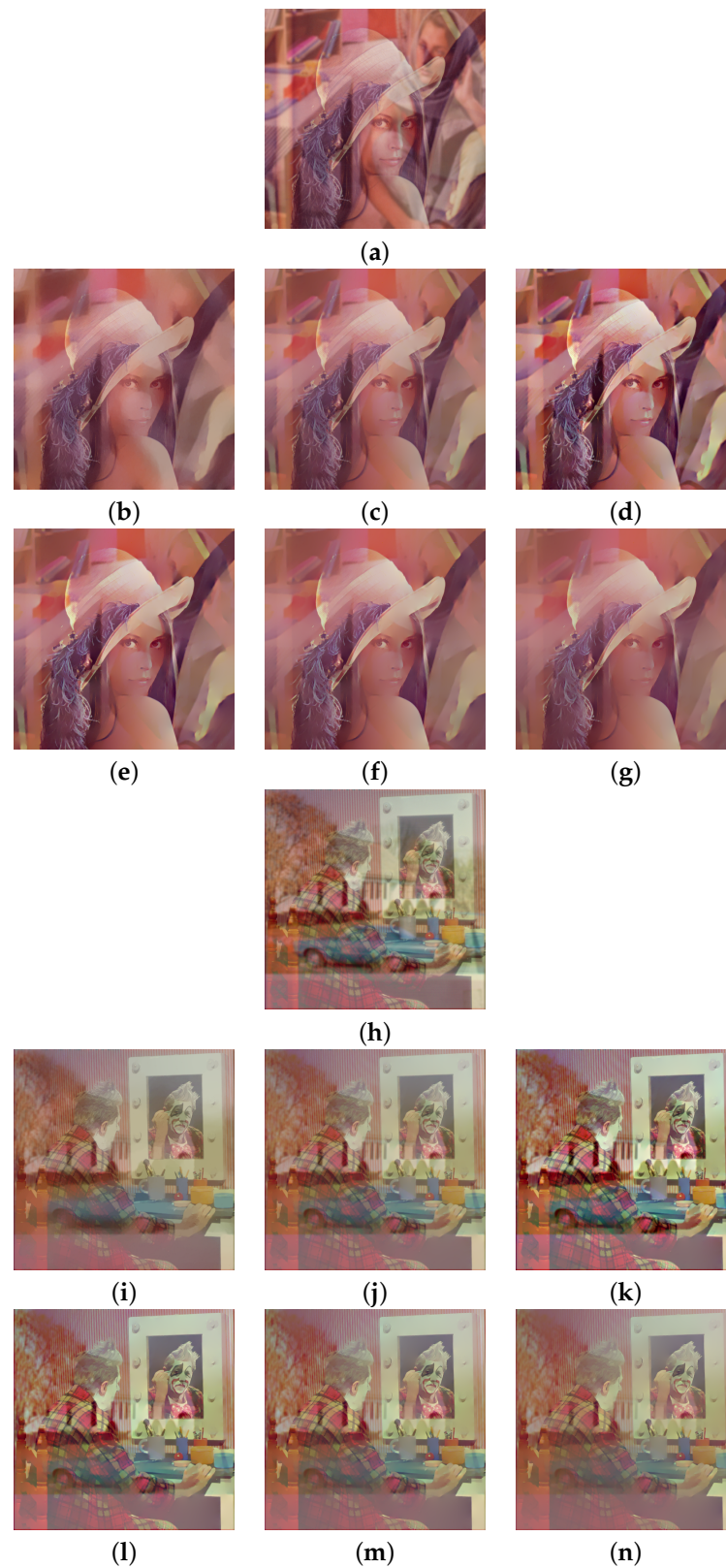




**Figure 3.** Reflection suppression results for synthetic glass images with  $w = 0.7$ . (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.



**Figure 4.** Reflection suppression results for synthetic glass images with  $w = 0.6$ . (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are the results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.



**Figure 5.** Reflection suppression results for synthetic glass images with  $w = 0.5$ . (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.

#### 4.2. Real Glass Image

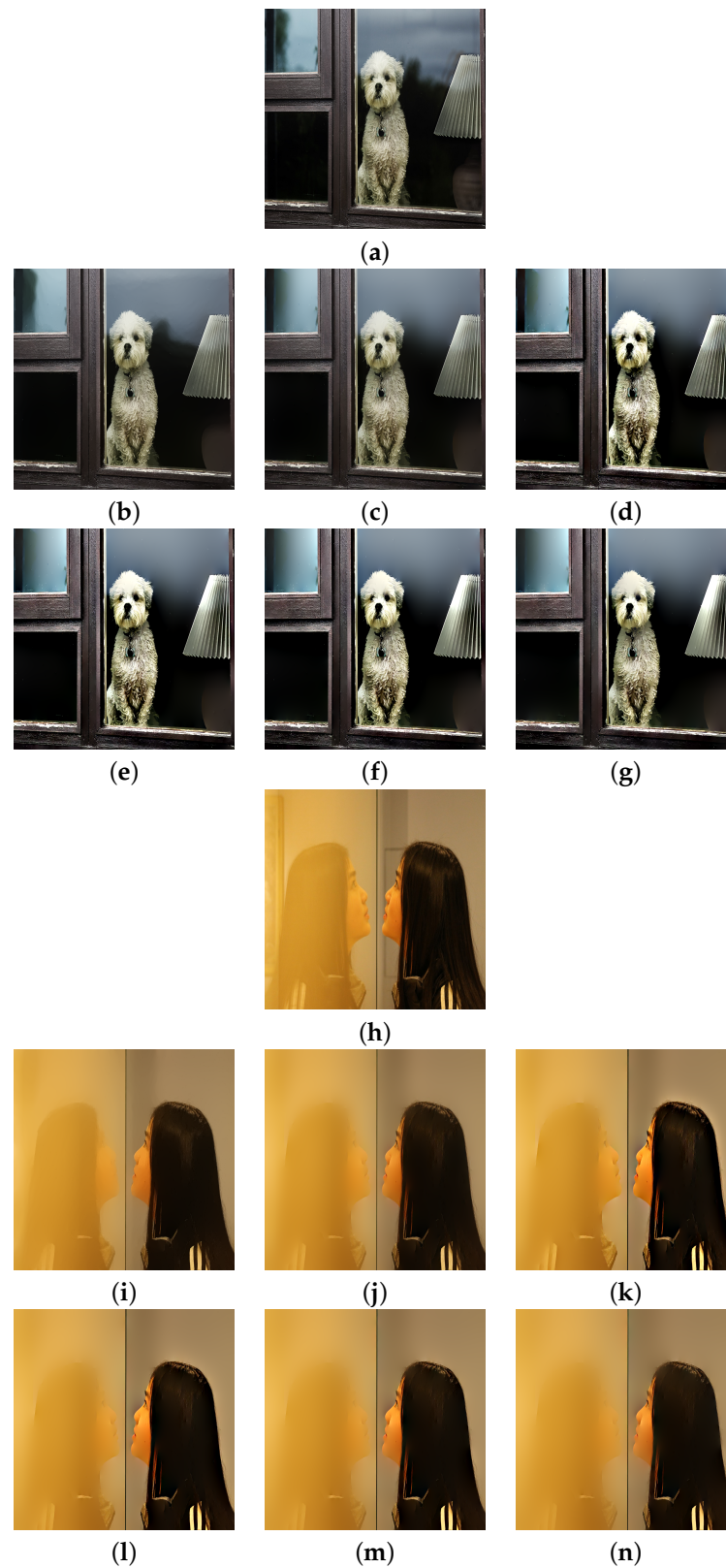
The real glass dataset of Yang et al. [3] is used here. The test images were captured directly using various smartphones with different resolutions. Figures 6–8 show the reflection suppression results of all the competing algorithms for the six real glass input images.

The ‘Dog’ image has some reflections on the window glass. Algorithms [2,3] smoothed out some of them, while the proposed algorithm suppressed all reflections and simultaneously enhanced image details. The ‘Girl’ image has a strong reflection on the left side such that all algorithms can not completely remove it. However, the proposed algorithm with  $N = 4$  suppresses the most reflections among all algorithms. The ‘Tree’ image has an extremely strong reflection that no algorithm can remove it. However, the proposed algorithm still has the best visual performance by enhancing the image contrast. The ‘Office’ image has many small rectangle shape reflections that are hard to suppress, the proposed algorithm with  $N = 4$  still suppresses most of them among all algorithms. The ‘Child’ and ‘Library’ images have many complex reflections and the proposed algorithm with  $N = 4$  still has the best visual performance. Here, we note that the images ‘Office’ and ‘Library’ are relatively challenging among all real examples. The main reason is that the rectangular-shaped window reflections have both bright intensities and sharp edges that violate the critical assumption that reflection gradients are smaller in magnitude and lower in focus compared to transmission gradients. If we want to completely remove those reflections, we need to set large  $h$  and we will lose many transmission edges as a heavy price. Therefore, the reflections of these two images are relatively hard to suppress. We further note that algorithm [3] and the proposed algorithm share the same model parameters  $h$  and  $\varepsilon$  for all input images.

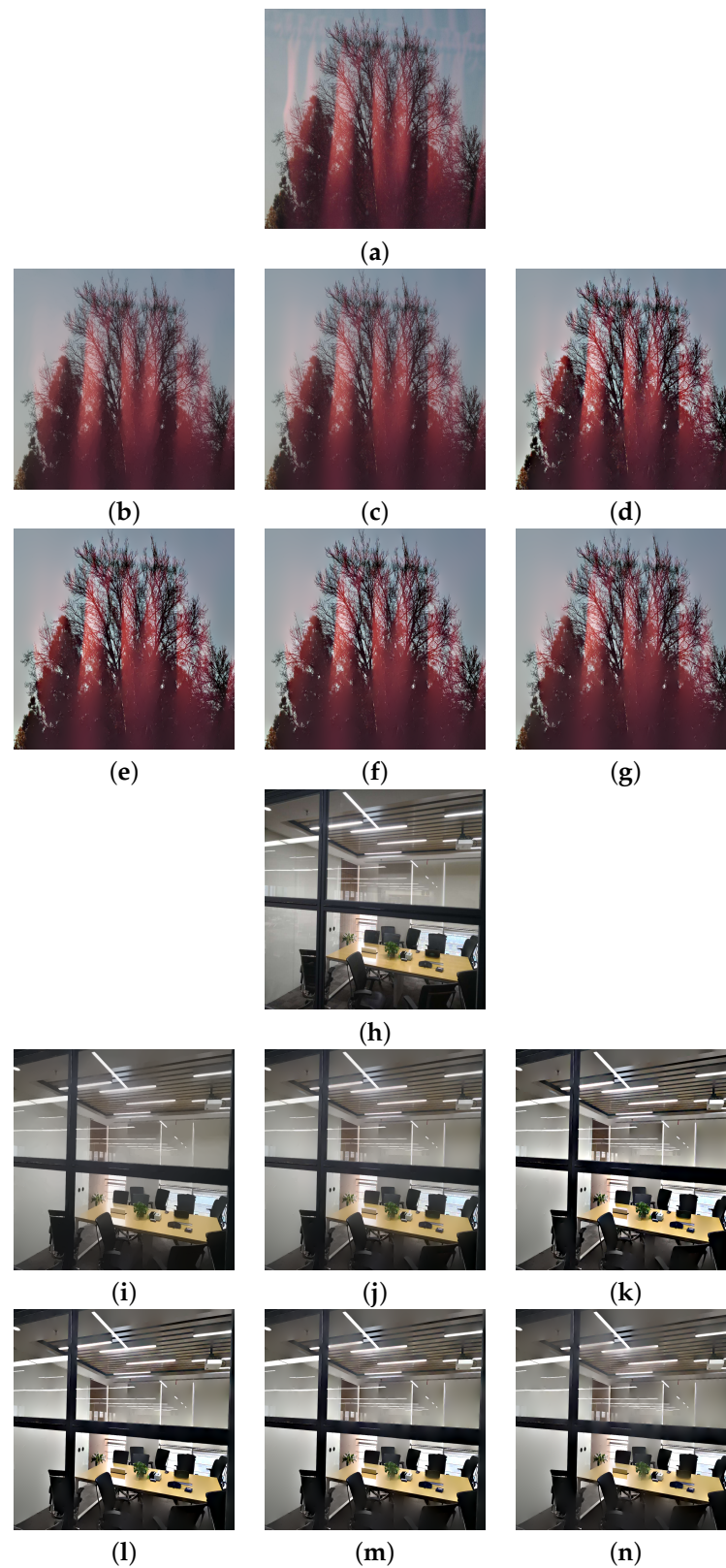
Since we do not have reflection-free solutions for real glass images for computing PSNR values to quantitatively evaluate the algorithms, here, we compare their computation times instead. Table 2 shows the computation time in seconds for all six input images and all competing algorithms. Although the proposed algorithm costs much time than algorithm [3], it is still very efficient compared to algorithm [2].

**Table 2.** Computation time of the reflection suppression results in Figure 6. The best value in each row is shown in boldface.

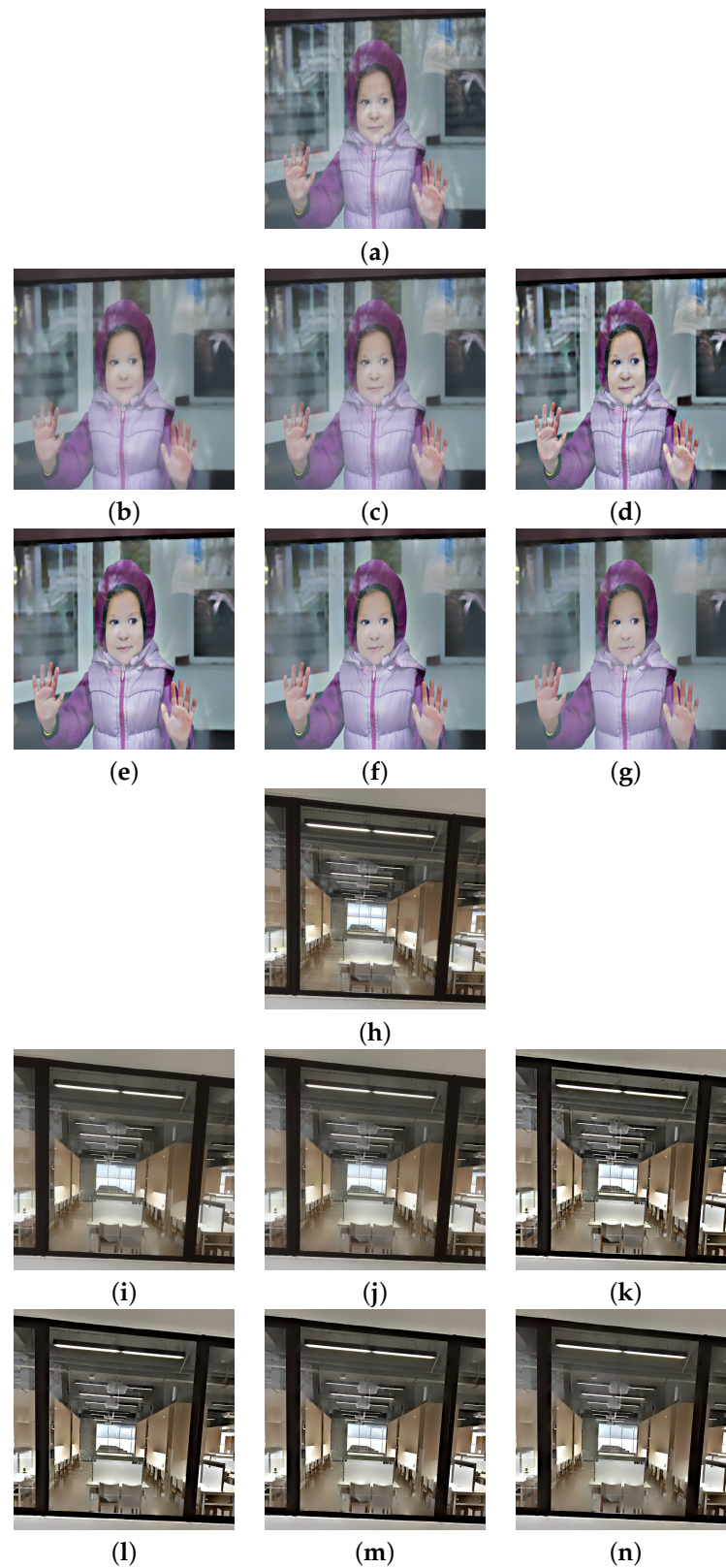
Image	Size	Algorithm [2]	Algorithm [3]	Proposed ( $N = 1$ )	Proposed ( $N = 2$ )	Proposed ( $N = 3$ )	Proposed ( $N = 4$ )
Figure 6a	667 × 1000	445.3	<b>0.682</b>	0.775	0.941	1.062	1.246
Figure 6h	823 × 1200	661.2	<b>0.998</b>	1.145	1.353	1.617	1.785
Figure 7a	880 × 1193	699.8	<b>1.008</b>	1.158	1.402	1.666	1.884
Figure 7h	1080 × 1440	1036.6	<b>1.209</b>	1.388	1.734	2.159	2.472
Figure 8a	1131 × 1698	1377.9	<b>1.888</b>	2.042	2.536	2.987	3.492
Figure 8h	1992 × 2448	3234.2	<b>3.853</b>	4.401	5.476	6.501	7.934



**Figure 6.** Reflection suppression results for real glass images 'Dog' and 'Girl'. (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.



**Figure 7.** Reflection suppression results for real glass images ‘Tree’ and ‘Office’. (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.



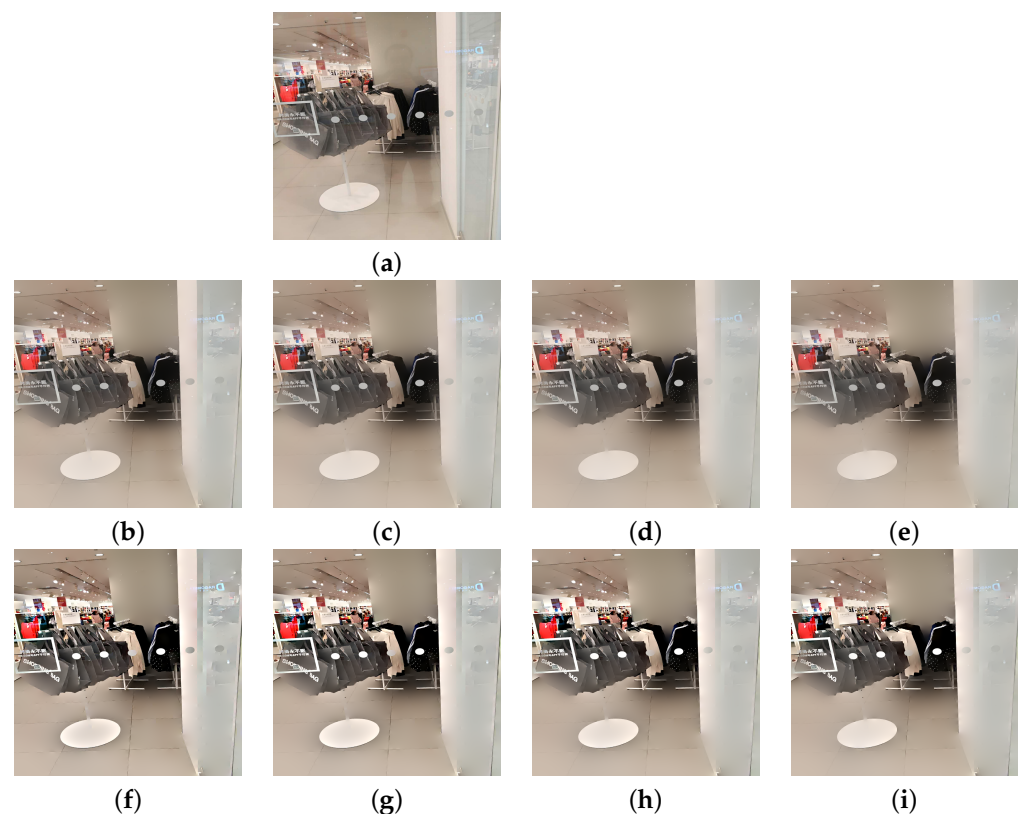
**Figure 8.** Reflection suppression results for real glass images ‘Child’ and ‘Library’. (a,h) Input images. (b,i) The results of algorithm [2]. (c,j) The results of algorithm [3]. (d,k), (e,l), (f,m), and (g,n) are results of the proposed algorithm with  $N = 1, 2, 3, 4$ , respectively.

### 4.3. Ablation Study

Here, we conduct an ablation study for the adaptive weight function  $\varphi$  used in the proposed model and provide an experiment of the proposed algorithm with various values of  $h$  and  $\beta$ .

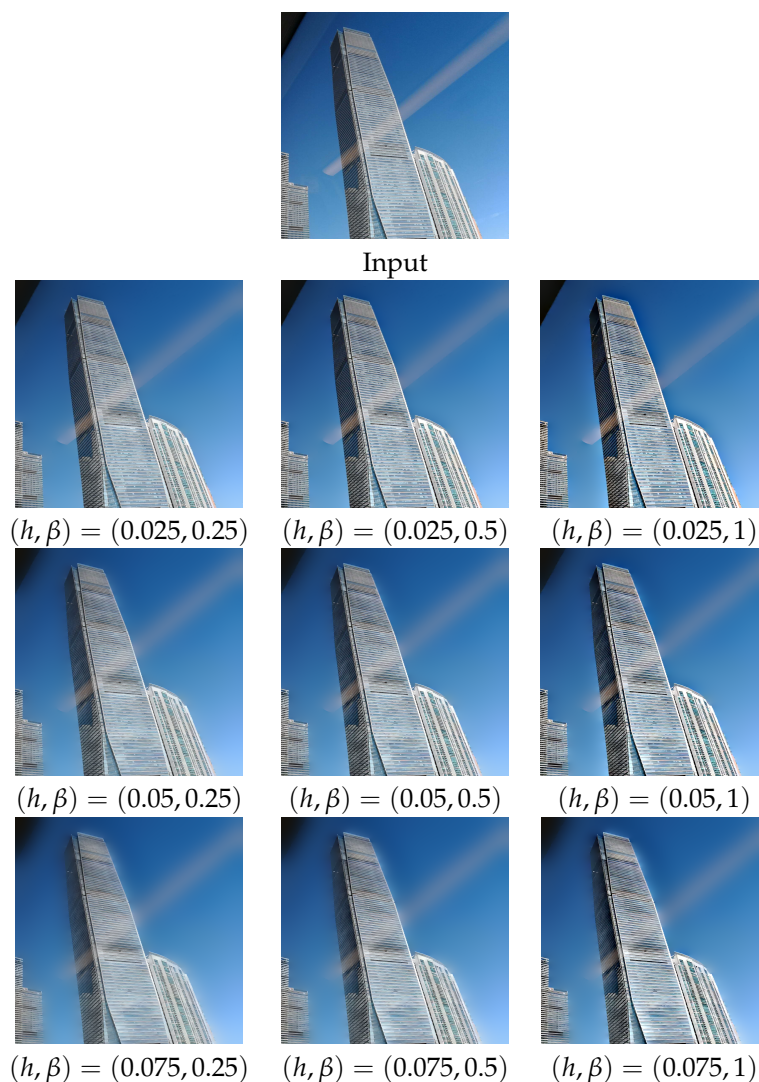
We begin with the ablation study by comparing the reflection suppression results of the proposed algorithm with and without the adaptive weight function  $\varphi$ , i.e., setting  $\varphi$  by (12) and by constant value 1, respectively. Experimental results are provided in Figure 9. From Figure 9b–e, one can observe that, under constant weight  $\varphi = 1$ , more and more reflections are suppressed as  $N$  increases but the transmission edges also become increasingly blurred. Under the adaptive weight  $\varphi$  given in (12), weak transmission edges are enhanced and keep their sharpness as  $N$  increases. The results are shown in Figure 9f–i.

To further elaborate the effects of parameters  $h$  and  $\beta$  to the proposed model, we test the proposed algorithm under  $N = 2$  and  $\varphi$  in (12) with  $h = 0.025, 0.05, 0.075$  and  $\beta = 0.25, 0.5, 1$ . Experimental results are provided in Figure 10. It can be observed that, as the threshold parameter  $h$  increases, the strong reflections are increasingly suppressed but the weak transmission edges on the lower building also become increasingly blurred. However, as enhancement parameter  $\beta$  increases, those weak edges become increasingly enhanced. These two experiments demonstrate the role that adaptive weight function  $\varphi$  played in the proposed variational model.



**Figure 9.** Reflection suppression results for real glass image ‘Mall’. (a) Input image. (b–e) The results of proposed algorithm (under  $\varphi = 1$ ) with  $N = 1, 2, 3, 4$ , respectively. (f–i) The results of algorithm (under  $\varphi$  in (12) and  $\beta = 1$ ) with  $N = 1, 2, 3, 4$ , respectively.





**Figure 10.** Reflection suppression results of the proposed algorithm (under  $N = 2$  and  $\varphi$  in (12)) for real glass image 'Building' with various  $h$  and  $\beta$ .

## 5. Summary and Conclusions

In this paper, we have proposed a both effective and efficient method for suppressing image reflections. Unlike the previous methods that suppressed the reflections only in a single spatial scale, the proposed method can suppress both weak and strong and multiple scale reflections while preserving important edge information via a multiscale hard thresholding framework. A direct solver using the Fourier cosine transform for implementing the proposed method is also given. Experimental results demonstrate that the proposed new algorithm is able to achieve superior performance compared with the previous algorithms. Extending the multiscale idea used in our model for dealing with more challenging situations such as ghosting effect [15,26] should be interesting, and this deserves further studies.

**Funding:** This work was supported by the Ministry of Science and Technology of Taiwan under the grant MOST 110-2115-M-031-002-MY2 (Pei-Chiang Shao).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The algorithm performance is tested on the public datasets, including [https://www.imageprocessingplace.com/downloads\\_V3/root\\_downloads/image\\_databases/standard\\_test\\_images.zip](https://www.imageprocessingplace.com/downloads_V3/root_downloads/image_databases/standard_test_images.zip) (accessed on 1 September 2021), [https://data.mendeley.com/public-files/datasets/3hfzp6vwkm/files/68088adc-c463-417e-bd84-03804d2670ff/file\\_downloaded](https://data.mendeley.com/public-files/datasets/3hfzp6vwkm/files/68088adc-c463-417e-bd84-03804d2670ff/file_downloaded) (accessed on 8 September 2021), and <https://github.com/yyhz76/reflectSuppress/tree/master/figures> (accessed on 15 September 2021).

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Barrow, H.; Tenenbaum, J.; Hanson, A.; Riseman, E. Recovering intrinsic scene characteristics. *Comput. Vis. Syst.* **1978**, *2*, 3–26.
2. Arvanitopoulos, N.; Achant, R.; Susstrunk, S.T. Single image reflection suppression. In Proceedings of the 2017 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4498–4506.
3. Yang, Y.; Ma, W.; Zheng, Y.; Cai, J.F.; Xu, W. Fast single image reflection suppression via convex optimization. In Proceedings of the 2019 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 8141–8149.
4. Kong, N.; Tai, Y.W.; Shin, J.S. A physically-based approach to reflection separation: From physical modeling to constrained optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *36*, 209–221. [[CrossRef](#)] [[PubMed](#)]
5. Schechner, Y.Y.; Kiryati, N.; Shamir, J. Blind recovery of transparent and semireflected scenes. In Proceedings of the 2000 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Hilton Head Island, SC, USA, 13–15 June 2000; pp. 38–43.
6. Farid, H.; Adelson, E.H. Separating reflections and lighting using independent components analysis. In Proceedings of the 1999 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Ft. Collins, CO, USA, 23–25 June 1999; pp. 262–267.
7. Agrawal, A.; Raskar, R.; Nayar, S.K.; Li, Y. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Trans. Graph.* **2005**, *24*, 828–835. [[CrossRef](#)]
8. Sirinukulwattana, T.; Choe, G.; Kweon, I.S. Reflection removal using disparity and gradient-sparsity via smoothing algorithm. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 1940–1944.
9. Li, Y.; Brown, M.S. Exploiting reflection change for automatic reflection removal. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 2432–2439.
10. Xue, T.; Rubinstein, M.; Liu, C.; Freeman, W.T. A computational approach for obstruction-free photography. *ACM Trans. Graph.* **2015**, *34*, 1–11. [[CrossRef](#)]
11. Gai, K.; Shi, Z.; Zhang, C. Blind separation of superimposed moving images using image statistics. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 19–32.
12. Guo, X.; Cao, X.; Ma, Y. Robust separation of reflection from multiple images. In Proceedings of the 2014 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 2195–2202.
13. Levin, A.; Weiss, Y. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 1647–1654. [[CrossRef](#)] [[PubMed](#)]
14. Li, Y.; Brown, M.S. Single image layer separation using relative smoothness. In Proceedings of the 2014 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 2752–2759.
15. Shih, Y.; Krishnan, D.; Durand, F.; Freeman, W.T. Reflection removal using ghosting cues. In Proceedings of the 2015 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3193–3201.
16. Wan, R.; Shi, B.; Hwee, T.A.; Kot, A.C. Depth of field guided reflection removal. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 21–25.
17. Fan, Q.; Yang, J.; Hua, G.; Chen, B.; Wipf, D. A generic deep architecture for single image reflection removal and image smoothing. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 3258–3267.
18. Wan, R.; Shi, B.; Duan, L.-Y.; Tan, A.-H.; Kot, A.C. Crnn: Multi-scale guided concurrent reflection removal network. In Proceedings of the 2018 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 4777–4785.
19. Yang, J.; Gong, D.; Liu, L.; Shi, Q. Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal. In Proceedings of the 2018 European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 654–669.
20. Zhang, X.; Ng, R.; Chen, Q. Single image reflection separation with perceptual losses. In Proceedings of the 2018 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 4786–4794.
21. Schechner, Y.Y.; Kiryati, N.; Basri, R. Separation of transparent layers using focus. *Int. J. Comput. Vis.* **2000**, *39*, 25–39. [[CrossRef](#)]
22. Xu, L.; Lu, C.; Xu, Y.; Jia, J. Image smoothing via L0 gradient minimization. *ACM Trans. Graph.* **2011**, *30*, 1–12.

23. Ma, W.; Morel, J.-M.; Osher, S.; Chien, A. An L1-based variational model for retinex theory and its applications to medical images. In Proceedings of the 2011 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 153–160.
24. Ma, W.; Osher, S. A TV Bregman iterative model of Retinex theory. *Inverse Probl. Imaging* **2012**, *6*, 697–708. [[CrossRef](#)]
25. Hsieh, P.-W.; Shao, P.-C.; Yang, S.-Y. Adaptive variational model for contrast enhancement of low-light images. *SIAM J. Imaging Sci.* **2020**, *13*, 1–28. [[CrossRef](#)]
26. Huang, Y.; Quan, Y.; Xu, Y.; Xu, R.; Ji, H. Removing reflection from a single image with ghosting effect. *IEEE Trans. Comput. Imaging* **2019**, *6*, 34–45. [[CrossRef](#)]