

Shadow Removal from a Single Image

Li Xu Feihu Qi Renjie Jiang

Department of Computer Science and Engineering,
Shanghai JiaoTong University, P.R. China
E-mail {nathan.xu, fhqi, blizard1982}@sjtu.edu.cn

Abstract

Shadow detection and removal in real scene images is always a challenging but yet intriguing problem. In contrast with the rapidly expanding and continuous interests on this area, it is always hard to provide a robust system to eliminate shadows in static images. This paper aimed to give a comprehensive method to remove both vague and hard shadows from a single image. First, classification is applied to the derivatives of the input image to separate the vague shadows. Then, color invariant is exploited to distinguish the hard shadow edges from the material edges. Next, we derive the illumination image via solving the standard Poisson equation. Finally, we got the shadow-free reflectance image. Experimental results showed that our method can robustly remove both vague and hard shadows appearing in the real scene images.

1. Introduction

Shadows in images have long been disruptive to computer vision algorithms. They appear as surface features, when in fact they are caused by the interaction between light and objects. This may lead to problems in scene understanding, object segmentation, tracking and recognition. Because of the undesirable effects of shadows on image analysis, much attention was paid to the area of shadow removal over the past decades and covered many specific applications such as traffic surveillance, face recognition, image segmentation and so on [1, 2, 3]. In spite of these extensive studies, however, it is difficult to develop a fully automatic, general-purpose shadow removal system in that shadows and shadings vary from scene to scene and different kinds of shadows may have different physical properties.

A shadow occurs when an object partially or totally occludes direct light from a source of illumination. Shadows can be divided into two sets: self shadows and cast shadows. A self shadow occurs in the portion

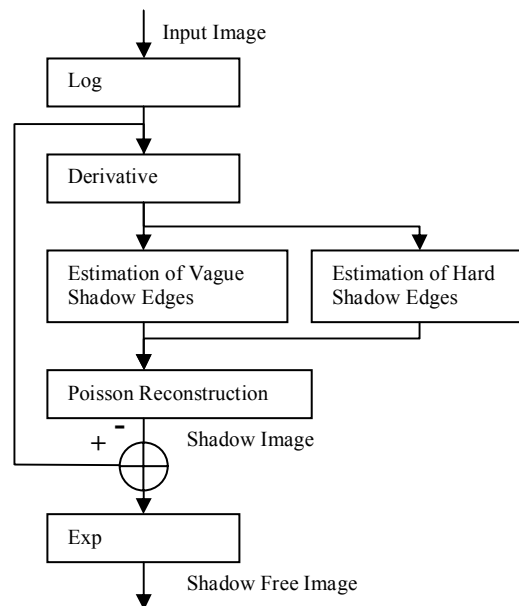


Figure 1. Block diagram for shadow removal system.

of an object whereas a cast shadow is the dark area projected by the object. Cast shadows can be further classified into umbra and penumbra region, which is a result of multi-lighting. One crucial difference between these shadows is their contrast to the background. Usually, self shadows are vague shadows which gradually change intensity and have no clear boundaries. Cast shadows are, on the other hand, hard shadows with sharp shadow boundaries.

In this paper, we proposed a method to remove both vague and hard shadows from a single image. Block diagram of our system is shown in Figure 1. First, we converse the input image to the logarithmic domain referring to the sensitivity of human vision system. Then, we estimate both vague and hard shadow edges on the gradient map of the input. Vague shadow edge estimation is based on the assumption that vague shadow varies slowly across an image whereas changes in materials are rapid. Hard shadow edge detection is conducted based on the color invariant properties.

After both shadow edges are obtained, we combined them and reconstruct the shadow image via solving a Poisson equation. At last, a shadow free reflectance image is derived by extracting the shadow image from the input image. Final shadow and shadow-free images are obtained by exponentiating the logarithmic results. Experimental results showed that our shadow removal system is robust for real scene images consisting both vague and hard shadows of any type.

The rest of the paper is organized as follows: Section 2 reviews related work on this topic. Section 3 introduces the shadow removal system. Section 4 gives experimental results and discussion. Section 5 concludes the paper.

2. Related Work

E.H. Land and his colleagues proposed a model, namely *Retinex*, to separate illumination from reflectance in a given image [4, 5]. Since the illumination images represent the distribution of the lighting and shading, separating it from reflectance is useful to cancel out the effect of shadows. Land's methods could only remove the vague shadows in image since they were based on the following assumptions:

- a. Mondrian world model.
- b. Spatially smooth illumination.
- c. Lambertian surfaces.

Lambertian surfaces assumption is forcing that there is no specular reflectance in image. The Mondrian world and smooth illumination assumptions assume that there is a generally clear signal at each of the boundaries between objects whereas there are no sharp boundaries between shadows and background. Similar work on vague shadow removal can be also found in [6, 7] where authors directly utilize a low-pass filter to separate the illumination images.

Recently, Y. Weiss [8] proposed a system to remove shadows from image sequences and Matsushita et al. [9] extended his work. Their methods also based on a decomposition of images into reflectance and illumination. They used an ML estimation to derive the time-invariant reflectance image from the sequences. Since they make a more generic assumption instead of the Mondrian world and smooth illumination, their method could remove cast shadows as well as vague shadows. However, the system has limitations for that it could only remove moving shadows in image sequences.

Tappen et al. [10] exploit Weiss's reintegrate method to derive the shadow free image from a single input. They trained the classifier to identify shadow edges and propagate the local evidence. The rational

behind the trained classifier is just the spatially smooth illumination assumption and hence the method could only remove vague shadows (shadings on the object).

More recently, Finlayson et al. made a great step in removing cast shadows from single images [11]. They derive scene texture edges from a light-invariant image and by subtracting those edges from the raw input; they obtain a shadow free color image. However, Finlayson's method could only remove hard shadows from scenes lit by the Planckian light.

3. Shadow Removal

Generally, this work is also based on decomposing input images into reflectance image R and the shadow image S (also named illumination image). The raw input image I is the pixel-wise multiplication of the reflectance and shadow, i.e. $I = R \cdot S$. A first step in our work is to converse the raw input into the logarithmic domain: $i = \log I$, $s = \log S$, $r = \log R$, and thereby $i = s + r$. Then, we estimate the shadow image s via combining vague shadow estimation and hard shadow estimation. Before going into the detailed algorithms, we review the assumptions: in our algorithm, we still take the prior that the surfaces are Lambertian and scenes are Mondrian. But we eliminated the assumption of the smooth illumination since we deal with vague and hard shadow independently and each of them has different physical properties.

3.1. Vague shadow estimation

Vague shadow refers to the shadows or shadings in images which have no clear boundaries and gradually changed intensity. In real scenes, self shadows and shadings usually have smooth changes and can be classified as vague shadow. However, even in the cast shadow regions, we can separate the gradually changed components. For instance, the penumbra regions of the cast shadows are more likely to be vague shadow components.

Unlike the Land's work on illumination removal [4], our approach works in the gradient space rather than intensity space. It is mainly because that it is easier to combine the vague and hard shadow removal in the gradient domain. Therefore, we first transfer the raw logarithmic image into gradient domain:

$$G_k(x, y) = \|\nabla i_k(x, y)\| \quad (1)$$

where $i_k(x, y)$ represents the logarithmic image, k denotes the k th color channel: $k = 1, 2, 3$ and G_k is the gradient response. Since vague shadows are expected

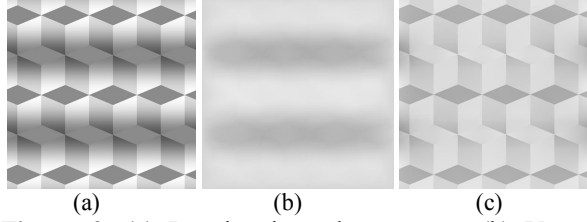


Figure 2. (a) Logvinenko cubes pattern. (b) Vague shadow estimation. (c) Recovered shadow free image.

to be spatially smooth, we clip out the high derivative peaks corresponding to the object boundaries and obtain the vague shadow mask. To determine the clipping threshold, the following criteria is applied:

$$T = \alpha \max_{x,y} \{ \max_k \{ G_k(x,y) \} \} \quad (2)$$

where α is a predetermined value and T represents a threshold. After deriving threshold T , we could do the following clipping operation to obtain a binary vague shadow mask for further process:

$$VS(x,y) = \begin{cases} 1 & \text{if } \max_k(G_k(x,y)) < T \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The $VS(x,y)$ is the binary mask. It could either be applied directly to the derivatives to remove the vague shadows in image or be combined with the hard shadow mask to perform the full shadow removal task. Figure 2 is an example of only applying vague shadow removal to an image: (a) is the Logvinenko cubes pattern from [5]: the image is covered with vertical sinusoidal illumination; (b) is the reconstructed vague shadow image deriving from applying the shadow mask to the derivatives (Details of reconstructing image from derivatives would be discussed in Section 3.3); (c) is the recovered shadow free image using (b). Note that the shadings on the image are totally gone whereas the object details are preserved in the shadow free image.

3.2. Hard shadow estimation

The vague shadow removal process described in the former subsection could be used to eliminate shadows with diffuse boundaries. The rational behind it is the assumption that the scene objects correspond to the sharp details in the image whereas the vague shadow image is expected to be spatially smooth. However, in many outdoor scenes, this may in fact not be the case at all. Cast shadows in the sunlight always have clear boundaries and similar response to the reflectance change in the edge image. To distinguish between edges due to shadows and those due to material changes, a priori information about the scenes and shadows is needed.

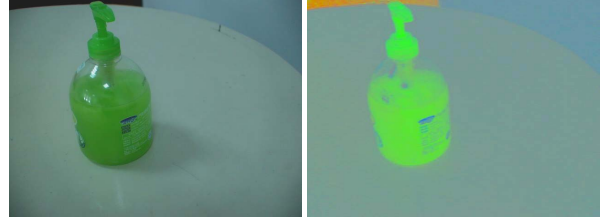


Figure 3. (a) Original image with cast shadows. (b) Color invariant image of (a).

Two kinds of models are exploited in our case to capture the color invariant of the inputs based on the observation that shadows in scenes mainly change the intensity of the surfaces and seldom change the chromaticity.

The first model is similar to the normalized rgb triangle but only take the l_2 -norm as the normalized factor:

$$r' = \frac{r}{\sqrt{r^2 + g^2 + b^2}} \quad (4)$$

$$g' = \frac{g}{\sqrt{r^2 + g^2 + b^2}} \quad (5)$$

$$b' = \frac{b}{\sqrt{r^2 + g^2 + b^2}} \quad (6)$$

where r, g, b is the three color channels of the input image, r', b', g' is the three-band color invariant image which do not contain the shadows. One example of the color invariant map is shown in Figure 3: (a) is the original image with cast shadow on a planer surface. (b) is the result image applying equation (3-5) to the original input. Note that shadows are not spotted in the color invariant map.

The second model is based on the method proposed by G.D. Finlayson et al. [11]. They observed that same colors in an image lie on lines with same slope in the log-chromaticity space in spite of the lighting condition. By projecting the colors in the 2d log-chromaticity space onto the direction orthogonal to the lighting direction, we could get a 1d invariant image. Finlayson's method could be rewritten as follows:

$$inv = \cos(\theta) \cdot \ln\left(\frac{r}{g}\right) + \sin(\theta) \cdot \ln\left(\frac{b}{g}\right) \quad (7)$$

where θ is the projecting direction, inv is the 1d shadow free invariant image.

In our system, we add an exponent operator to the invariant image to get more "pleasing-looking" images. The crucial parameter of this model is the projecting direction θ , Finlayson et al. use a set of pictures taken at same scene under different illumination to perform the calibration. They also point out that different cameras would have different projection directions.



Figure 4. (a) Original scene image with cast shadows from internet. (b) 1-d invariant image of (a).

However, to get such an image set for calibration is not an easy job, especially the images are obtained from the internet. Our experiments showed that although the projecting direction varies from cameras, it only varies a little. Therefore, we could use an approximate direction to derive an approximate invariant image. Figure 4 shows the result of Finlayson's method with an exponent operation. Figure 4 (a) is a scene image with shadows obtained from the internet; (b) is the 1-d shadow free invariant image of (a).

Both of the two models we used here are designed to capture the invariant characteristics of the scene using the color information. They can compensate each other in some cases owing to their different realization of the invariance.

Using the invariant image, hard shadow edge mask could be estimated by comparing the raw input and the invariant image. First, a Gaussian smooth filter is applied to both original image and invariant images. This has the effect to depress the high frequency textures in image. For color images (the smoothed input and the smoothed color invariant image), a transfer to gray scale is conducted by taking the maximum value of three color bands, which is just the definition of Value in the HSV color model. Then, edge detection is performed to both of the smoothed images:

$$E_{ori} = \|\text{edge}(I_{ori})\|, E_{inv(i)} = \|\text{edge}(I_{inv(i)})\| \quad (8)$$

where i is the index of the invariant images, here $i = 1, 2$. The hard shadow edge mask could be derived by pick out the strong edges which appear in the original images but not in the invariant images:

$$HS(x, y) = \begin{cases} 1 & E_{ori}(x, y) > t1 \& \\ & \min_i (E_{inv(i)}(x, y) < t2) \\ 0 & otherwise \end{cases} \quad (9)$$

where $t1, t2$ is the manually set thresholds and $HS(x, y)$ is the estimated hard shadow edge mask. In equation (9), the parameter $t1$ forces the selected shadow edges to be strong edges corresponding to the hard shadows in images and parameter $t2$ picks out edges only belong to shadows. Unlike Finlayson's method, our

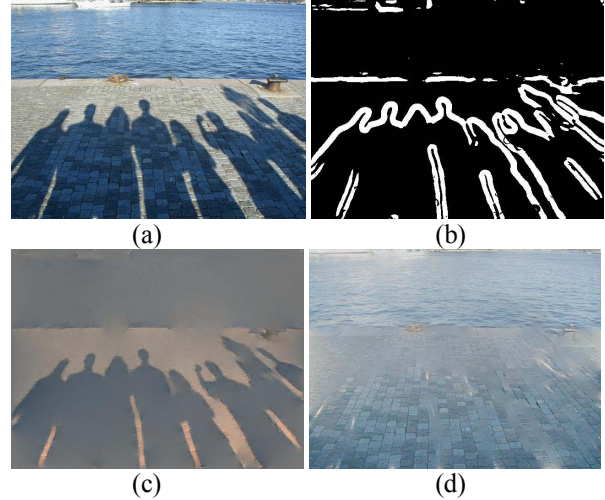


Figure 5. (a) Input image from internet. (b) Hard shadow mask. (c) Reintegrated shadow mask. (d) Recovered shadow free image.

system need no further dilation for shadow edge image since our edge detector is applied to the smoothed image, and hence the shadow edges is complete enough. Moreover, the morphological dilation operator would introduce extra material edges into shadow mask. Figure 5 depicts the results of hard shadow edge estimation: the hard shadow edge mask of (a) is shown in (b) and (c) is the reintegrated shadow image from combining the vague and hard shadow mask. (d) is our recovered shadow-free result.

3.3. Reconstruction

Image reconstruction from gradient information is an approximate inverse problem. Weiss [8] proposed a pseudo inverse function to recover the original image. In our case, we adopt the Poisson equation solution [6] method to reintegrate the shadow images from the derivatives.

We begin by combining the two kinds of shadow edge masks:

$$mask = VS \mid HS \quad (10)$$

Then, apply mask to the gradient field:

$$\nabla i' = mask \cdot \nabla i = (mask \cdot G_x, mask \cdot G_y) \quad (11)$$

We denote the clipped derivatives as G_x' and G_y' and calculated the following scalar:

$$\text{div}G = \frac{\partial G_x'}{\partial x} + \frac{\partial G_y'}{\partial y} \quad (12)$$

Finally, the shadow image could be reconstructed by solving the following well-known Poisson equation:

$$\nabla^2 s = \text{div}G \quad (13)$$

where s is the estimated shadow image. Since the process is done in logarithmic domain, we should add exponent operation to the image:

$$S(x, y) = \exp(s(x, y) - \max_{x,y}(s(x, y))) \quad (14)$$

Extracting shadows from original image:

$$r(x, y) = i(x, y) - s(x, y) \quad (15)$$

$$R(x, y) = \exp(r(x, y) - \max_{x,y}(r(x, y))) \quad (16)$$

where S is the reconstructed shadow image and R is the reconstructed shadow free image. Figure 5 (c) is corresponding to the S and (d) is corresponding to R in equation (14)(16), respectively.

4. Discussion and Results

As deriving a shadow-free image from image sequences has made a great progress, recovering it from a single image remains a difficult problem. It is mainly because the motion information which could help to distinguish between shadows and reflectance becomes useless in a still image. In our method, we begin at estimating shadow edge masks for both vague and hard shadows. In current implementation, the parameter α for vague shadow clipping operation is set to be 0.06. The projecting direction θ for invariant image estimation is chosen to be 43.58° based on a series tests on an arbitrary test set. We chose values of $t1=0.4$ and $t2=0.1$ after the edge magnitudes have been scaled to a range $[0, 1]$ based on the Finlayson's work.

Our system is robust for shadow removal under different environments. Figure 6 gives more examples. The first column is the input; both outdoor and indoor environments are included. Note that all of them come from different cameras. The second column is the reconstructed shadow images; textures of the scene are not included in the shadow images. The last column is the recovered shadow free images; most of vague and hard shadows are removed in these results.

Currently, we developed a demonstration on a 1.8GHz PC using Matlab 6.5. The size of the test data we used in our experiments is 640×480 . For all the test images, the average running time is 5.62 seconds. As a comparison, the standard implementation of McCann *Retinex* [5] with four iterations takes more than 9 seconds for each image on our PCs.

Experimental results showed that our method is effective in removing shadows in images. However, there are some artifacts in the image boundaries since we simply use the Neumann boundary condition for Poisson equation, i.e. we set the derivative in the direction normal to the boundary to zero. A better alternative is to pad the original image with the mirrored image.

Future work includes the further improvement of estimating shadow edges and reconstructing image from gradients. A more challenging task for us is to remove hard shadows from a gray scale image.

5. Conclusion

In this paper, we have described a system for removing shadows from a single scene image, which is an intriguing but challenging problem. Unlike previous attempts on this area, we do not make much assumption on shadows; instead, we proposed a combined method to remove both vague and hard shadows in image. Working in the gradient domain, our system focuses on providing robust shadow edge estimation. The proposed method is remarkably robust and does not require any information of the shadow geometry. Though our current implementation of recovering images would introduce halos into images and the areas near the shadow edge would have artifacts, our shadow removal system still works well for both indoor and outdoor environments and can effectively remove vague and hard, cast and self shadows from a single scene image.

Acknowledgements

The work presented in this paper was partially sponsored by OMRON *Sensing & Control Technology Laboratory*.

References

- [1] J.M. Wang, Y.C. Chung, C.L. Chang, S.W. Chen, "Shadow Detection and Removal for Traffic Images", *Proc. IEEE International Conference on Networking, Sensing and Control*, vol. 1, pp. 649 – 654, 2004
- [2] T. Chen, W. Yin, X.S. Zhou, D. Comaniciu, and T.S. Huang, "Illumination Normalization for Face Recognition and Uneven Background Correction Using Total Variation Based Image Models", *Proc. CVPR*, vol. 2, pp. 532-539, 2005
- [3] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction", *IEEE Trans. Pattern Anal. Machine Intell.* 19(7):721–732, 1997.
- [4] E.H. Land, "The Retinex Theory of Color Vision", *Scientific American*, vol. 237, pp. 108-128, 1977.
- [5] J. McCann, "Lessons Learned from Mondrians Applied to Real Images and Color Gamuts", *Proc. IS&T/SID Seventh Color Imaging Conference*, pp. 1-8, 1999.
- [6] B.K.P. Horn, "Determining Lightness from an Image", *Computer Graphics and Image Processing*, vol. 3, pp. 277~299, 1974.

[7] D.J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and Performance of the Center Surround Retinex", *IEEE Trans. on Image Proc.*, vol. 6, pp.451-462, 1997
 [8] Y. Weiss, "Deriving intrinsic images from image sequences", *Proc. Int. Conf. Computer Vision*, 2001.
 [9] Y. Matsushita, K. Nishino, K. Ikeuchi, M. Sakauchi, "Illumination Normalization with Time-Dependent Intrinsic Images for Video Surveillance", *IEEE Trans. Pattern Anal. Machine Intell.* 26(10):1336-1347, 2004

[10] M. Tappen, W.T. Freeman, and E.H. Adelson, "Recovering intrinsic images from a single image", *IEEE Trans. Pattern Anal. Machine Intell.* vol. 27, no. 9, pp. 1459-1472, 2005
 [11] G.D. Finlayson, S.D. Hordley C. Lu and M.S. Drew, "On the Removal of Shadows from Images", *IEEE Trans. Pattern Anal. Machine Intell.*, Jan, 2006
 [12] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge University Press. 1992.

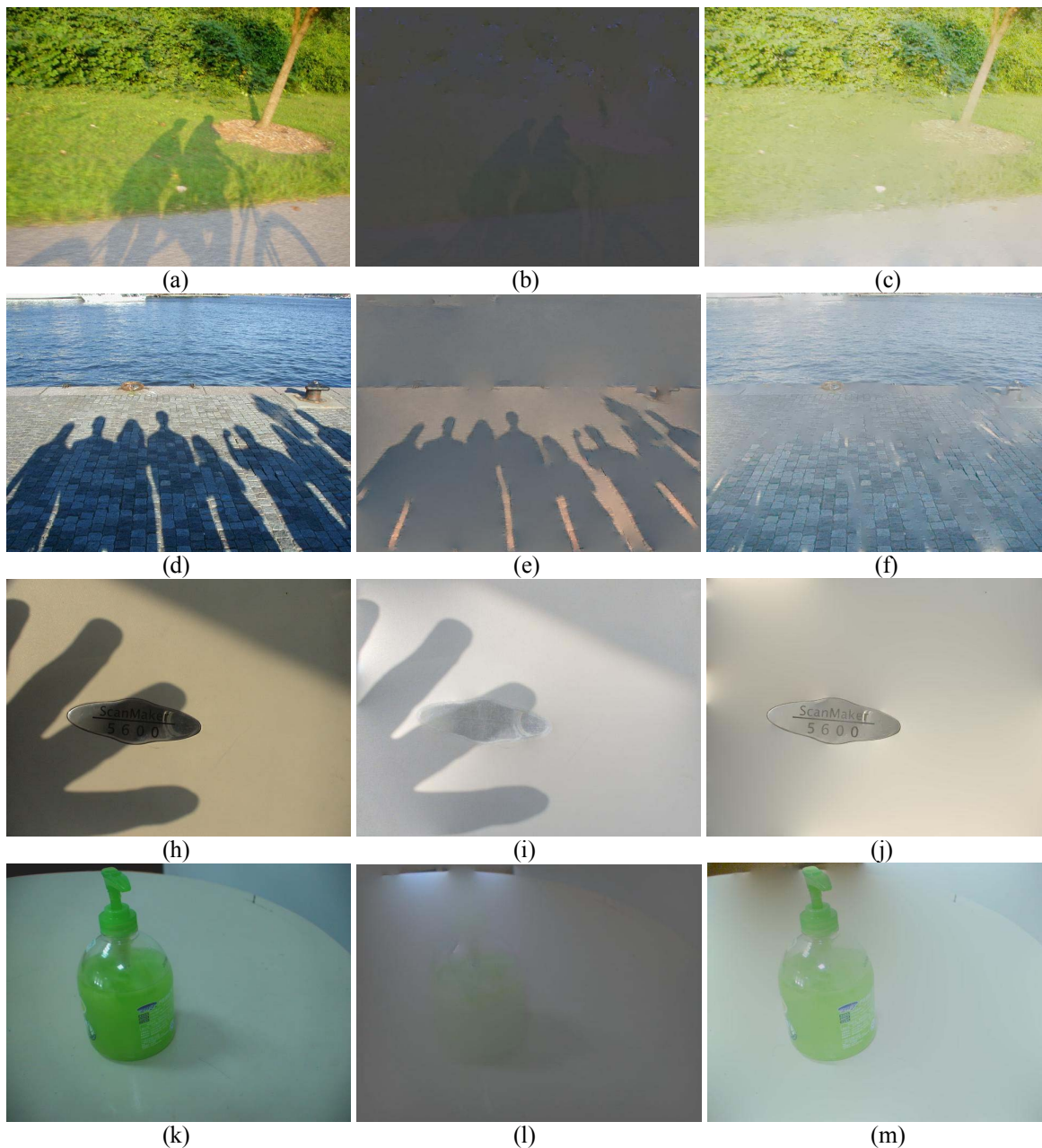


Figure 6. More examples for shadow removal: (a)(d)(h)(k) Original images with shadows. (b)(e)(i)(l) The Reconstructed shadow images based on our method. (c)(f)(j)(m) The recovered shadow free images.