Blind recovery of transparent and semireflected scenes

Yoav Y. Schechner Nahum Kiryati† Joseph Shamir

Department of Electrical Engineering,
Technion - Israel Institute of Technology,
Haifa, Israel 32000
yoavs@tx.technion.ac.il
jsh@ee.technion.ac.il

†Dep. of Electrical Engineering - Systems,
Faculty of Engineering, Tel-Aviv University,
Ramat Aviv, Israel 69978
nk@eng.tau.ac.il

Abstract

We present a method to recover scenes deteriorated by superposition of transparent and semi-reflected contributions, as appear in reflections off windows. Separating the superimposed contributions from the images in which either contribution is in focus is based on mutual blurring and subtraction of the perturbing components. This procedure requires the defocus blur kernels to be known. The use of uncalibrated kernels had previously led to contaminated results. We propose a method for self calibration of the blur kernels from the raw images themselves. The kernels are sought to minimize the mutual information of the recovered layers. This relaxes the need for prior knowledge on the optical transfer function. Experimental results are presented.

1 Introduction

This work treats the common situation in which several (typically two) linearly superimposed contributions exist in a scene. For example [10], looking through a window, the object behind the window (real object [10, 11]) is disturbed by a semi-reflection of another object (virtual object). This situation was described using the term *transparent layers* [2, 20]. Treatment has been based mainly on motion [2, 5, 7, 11, 18, 19] and stereo [3, 18]. Monocular methods were based on polarization [6, 10, 15, 16, 17] or on focus [14].

The focus cue is based on the limited depth of field of the imaging system, and is useful also for separating and recovering three dimensional volumetric specimens acquired by microscopy (as in [1, 8, 9, 12]) or by tomography, where the superposition of transparent layers is inherent. The focus cue can be utilized in common systems with no need for a polarizer. However, the recovery of the layers, as presented in [14], requires the defocus blur kernels to be known. Using inaccurate kernels may leave each of the recovered layers significantly contaminated by its complementary. Evidence to that are seen in the experiment described in [14], where the kernels were assumed but unknown.

This paper presents a method to separate transparent layers without a-priori knowledge of the blur kernels. It presents a method to self-calibrate the defocus point spread function (PSF) given the raw images of the transparent scene. This relieves the need for knowledge of the imaging system properties and their changes during focusing. The method is based on seeking the minimum of the mutual information between the recovered layers. Experimental separation results demonstrate the success of the method.

2 Recovery given the blur kernels

Consider a two-layered scene. Suppose that either manually or by some automatic procedure (as discussed in Ref. [14]), we acquire two images, such that in each image one of the layers is in focus. Let layer f_1 be superimposed on layer f_2. We consider only the slices g_a and g_b, in which either layer f_1 or layer f_2, respectively, is in focus. The other layer is blurred. Modeling the blur as convolution with blur kernels,

$$g_a = f_1 + f_2 * h_{2a} \qquad g_b = f_2 + f_1 * h_{1b} \ . \tag{1}$$

In the frequency domain Eqs. (1) take the form

$$G_a = F_1 + H_{2a}F_2 \qquad G_b = F_2 + H_{1b}F_1 \ . \tag{2}$$

The inverse filtering solution to the problem is

$$\widehat{F}_1 = B(G_a - G_b H_{2a}) \qquad \widehat{F}_2 = B(G_b - G_a H_{1b}) \tag{3}$$

where $B = (1 - H_{1b}H_{2a})^{-1}$. As $H_{1b}H_{2a} \to 1$, $B \to \infty$ hence the solution is inherently unstable in the low frequencies and the recovery of its DC component is ill posed [14]. For $H_{1b}H_{2a} < 1$, the solution can be approximated as a geometric series:

$$\widehat{F}_{1m} = \widehat{B}_m \left(G_a - G_b H_{2a}\right) \quad \widehat{F}_{2m} = \widehat{B}_m \left(G_b - G_a H_{1b}\right) \tag{4}$$

where $\widehat{B}_m = \sum_{k=1}^{m} \left(H_{1b} H_{2a}\right)^{k-1}$. We define the *basic solution* as the result of using $m = 1$. The basic solution has high-pass characteristics. The recovered layers are those given by the basic solution, filtered by $\widehat{B}_m$. Using larger m's improves the balance between the low frequency components to the high ones.

Consider the simulated scene that consists of the image of Lena, as the close object, seen reflected through a window out of which Mt. Shuksan[1] is seen. While any of the layers is focused, the other is blurred by a Gaussian kernel with standard deviation (STD) of 2.5 pixels. The slices in which each of the layers is focused appear in the top of Fig. 1 (all the images in this work are presented contrast-stretched). The basic solution ($m = 1$), shown at the middle row of Fig. 1, removes the crosstalk between the images, but lacks contrast due to the attenuation of the low frequencies (better contrast can be obtained by using a larger m).

Effect of error in the PSF

The method computes $\widehat{F}_{1m} = \widehat{B}_m[G_a - G_b H_{2a}]$. We normally assume (Eq. (2)) that $G_a = F_1 + H_{2a} F_2$ and $G_b = F_2 + H_{1b} F_1$. If the assumption holds,

$$\widehat{F}_{1m} = F_1 (1 - H_{1b} H_{2a})\, \widehat{B}_m \ . \tag{5}$$

Note that, regardless of the precise form of the PSFs, had the imaging PSFs and the PSFs used in the recovery been equal, the reconstruction would have converged to F_1 as $m \to \infty$ when $|H_{1b}|, |H_{2a}| < 1$. In practice, the imaging PSFs are slightly different, i.e., $G_a = F_1 + \widetilde{H}_{2a} F_2$ and $G_b = F_2 + \widetilde{H}_{1b} F_1$ where

$$\widetilde{H}_{1b} = H_{1b} - E_{1b} \ , \quad \widetilde{H}_{2a} = H_{2a} - E_{2a} \ , \tag{6}$$

and E_{1b}, E_{2a} are some functions of the spatial frequency. This difference may be due to inaccurate prior modeling of the imaging PSFs or due to errors in depth estimation. The reconstruction process leads to

$$\begin{aligned} \widetilde{F}_1 &= [F_1 \left(1 - H_{1b} H_{2a}\right) + E_{1b} H_{2a} F_1 - E_{2a} F_2] \widehat{B}_m \\ &= \widehat{F}_{1m} + \widehat{B}_m \left(E_{1b} H_{2a} F_1 - E_{2a} F_2\right) \ . \end{aligned} \tag{7}$$

A similar relation is obtained for the other layer.

An error in the PSF leads to contamination of the recovered layer by its complementary. The larger $\hat{B}_m$ is, the stronger is the amplification of this disturbance. $\hat{B}_m$ monotonically increases with m, within the support of the blur transfer function if $H_{1b} H_{2a} > 0$, as is the case when the recovery PSFs are Gaussians. Note that usually in

[1]Courtesy of Bonnie Lorimer

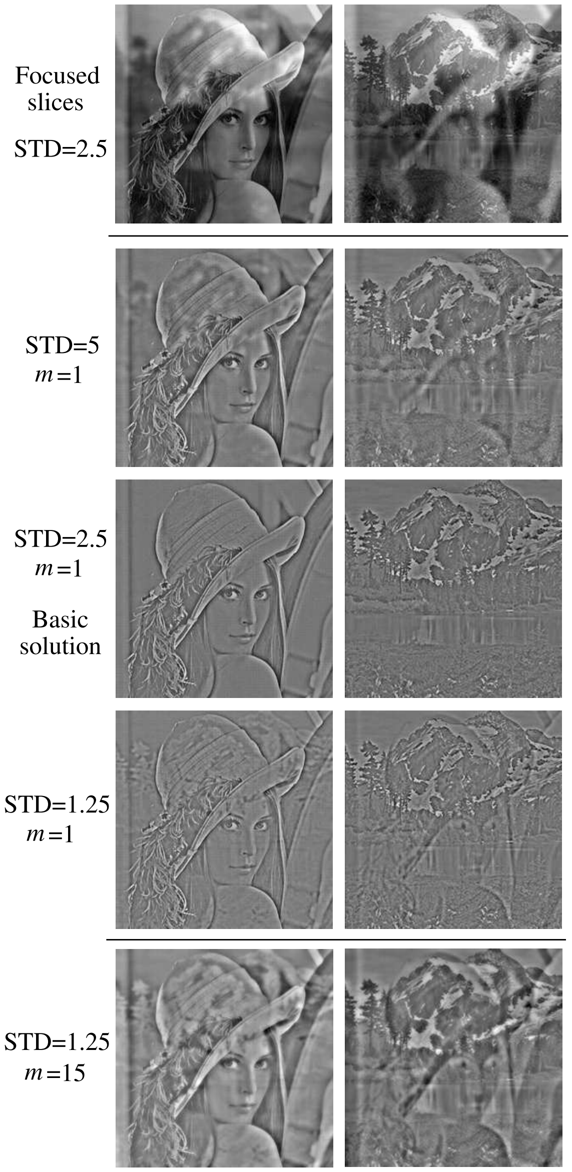

Figure 1. [Top] Simulated focused slices. The defocused layer is blurred by a kernel with STD of 2.5 pixels. [2^{nd} row] Positive traces in the basic solution, if STD=5 is (wrongly) used in the recovery. [3^{rd} row] The basic solution with the correct kernel removes the crosstalk. [4^{th} row] Negative traces if STD=1.25 is (wrongly) used in the recovery. [Bottom] The contamination has increased with m (even though the balance between the low and high frequency components has improved).

the low frequencies (which is the regime of the crosstalk) $H_{1b}, H_{2a} > 0$. Thus, we may expect that the best sense of separation will be achieved using a small m. Actually, the basic solution should provide the least contamination. This is so although the uncontaminated solution obeys $\hat{F} \to F$

as m increases. In other words, decreasing the reconstruction error does not necessarily lead to less crosstalk.

These effects are seen in Fig. 1. When the PSF used in the reconstruction has STD of 5 pixels (instead of the 2.5 pixels used in the image formation), positive traces remain (i.e., brighter areas in one image appear brighter in the other). When the PSF used in the reconstruction has STD of 1.25 pixels, negative traces remain (i.e., brighter areas in one image appear as darker areas in the other). The contamination is slight in the basic solution, but is more noticeable with larger m's, that is, when $\hat{B} \rightarrow B$. So, the separation seems worse, even though each of the images has a better balance (due to the enhancement of the low frequencies).

3 Seeking the blur kernels

The kernels are usually unknown. Even *a-priori* knowledge (if the system is of our design or calibrated) is sometimes inaccurate. We thus wish to achieve self-calibration, i.e., to estimate the kernels out of the images themselves. This will enable blind separation and layers restoration.

To do that, we need a criterion for layer separation. Assume that the statistical dependence of the real and virtual layers is small (even zero). This is reasonable since they usually originate from unrelated scenes. The Kullback-Leibler distance measures how far the images are from statistical independence, indicating their mutual information [4]. Let the probabilities for certain values $\check{f}_1$ and $\check{f}_2$ be $P(\check{f}_1)$ and $P(\check{f}_2)$, respectively. In practice these probabilities are estimated by the histograms of the recovered images. The joint probability is $P(\check{f}_1, \check{f}_2)$. It is estimated by the joint histogram of the images, that is, the relative number of pixels in which $\tilde{f}_1$ has a certain value $\check{f}_1$ and $\tilde{f}_2$ has a certain value $\check{f}_2$ at corresponding pixels. The mutual information is then

$$\mathcal{I}(\tilde{f}_1, \tilde{f}_2) = \sum_{\check{f}_1, \check{f}_2} P(\check{f}_1, \check{f}_2) \log \frac{P(\check{f}_1, \check{f}_2)}{P(\check{f}_1)\ P(\check{f}_2)} \quad . \tag{8}$$

In this approach we assume that if the layers are correctly separated, each of their estimates contains *minimum information* about the other [17]. However, the distance (Eq. 8) depends on the quantization of $\tilde{f}_1$ and $\tilde{f}_2$, and on their dynamic range, which in turn depends on the brightness of the individual layers f_1 and f_2. To decrease the dependence on these parameters, we performed two normalizations. First, each estimated layer was contrast-stretched to a standard dynamic range. Then, $\mathcal{I}$ was normalized by the mean entropy of the estimated layers, when treated as individual images. The self information [4] (entropy) of $\tilde{f}_1$ is $\mathcal{H}(\tilde{f}_1) = -\sum_{\check{f}_1} P(\check{f}_1) \log P(\check{f}_1)$, and the expression for $\tilde{f}_2$ is similar. The measure we used is

$$\mathcal{I}_n(\tilde{f}_1, \tilde{f}_2) = \frac{\mathcal{I}(\tilde{f}_1, \tilde{f}_2)}{[\mathcal{H}(\tilde{f}_1) + \mathcal{H}(\tilde{f}_2)]/2} \quad , \tag{9}$$

indicating the ratio of mutual information to the self information of a layer.

The recovered layers depend on the kernels used. Therefore, the problem of seeking the kernels can be stated as a minimization problem: $[\hat{h}_{1b}, \hat{h}_{2a}] = \arg\ \min_{h_{1b}, h_{2a}} \mathcal{I}_n(\tilde{f}_1, \tilde{f}_2)$. As discussed in Section 2, errors in the kernels lead to crosstalk (contamination) of the estimated layers, which is expected to increase their mutual information.

There are generally many degrees of freedom in the form of the kernels. On the other hand, the kernels are constrained: they are non-negative, they conserve energy etc. To simplify the problem, the kernels can be assumed to be Gaussians. Then, the kernels are parameterized only by their standard deviations. This limitation may lead to a solution that is suboptimal but easier to obtain.

Another possible criterion for separation is decorrelation. It was used for the recovery of semi-reflected layers by independent components analysis in [6], and by polarization analysis in [16]. However, requiring decorrelation between the estimated layers is based on the assumption that the original layers are decorrelated: that assumption is usually only an approximation. Actually, Ref. [17] gives examples for which polarization based separation is better if mutual information is used instead of decorrelation.

To illustrate the use of these criteria, we search for the optimal blur kernels to separate the layers that are superimposed in the images shown in the top row of Fig. 1. Here we simplified the calculations by restricting both kernels to be isotropic Gaussians of the same STD, as these were indeed the kernels used in the synthesis. Hence, the correlation and mutual information are functions of a single variable. As seen in Fig. 2, using the correct kernel (with STD of 2.5 pixels) yields decorrelated basic solutions ($m = 1$), with minimal mutual information ($\mathcal{I}_n$ is plotted). The positive correlation for larger values of assumed STD, and the negative correlation for smaller values, is consistent with the visual appearance of positive and negative traces in Fig. 1. As expected from the theory, in Fig. 1 the crosstalk was stronger with a larger m (e.g., m=15). Indeed, in Fig. 2 the absolute correlation and mutual information are greater for $m = 6$ than for $m = 1$ when the wrong kernel is used.

In a different simulation, the focused slices were synthesized using an exponential imaging kernel rather than a Gaussian, but the STD was still 2.5. The recovery was done with Gaussian kernels. The correlation and mutual information curves (as a function of the assumed STD) were similar to those seen in Fig. 2. The minimal mutual information was however at STD of 2.2 pixels. There was no visible crosstalk in the resulting images.

The blurring along the sensor raster rows may be different than along the columns. This is because beside the optical processes blurring is caused also by interpixel crosstalk

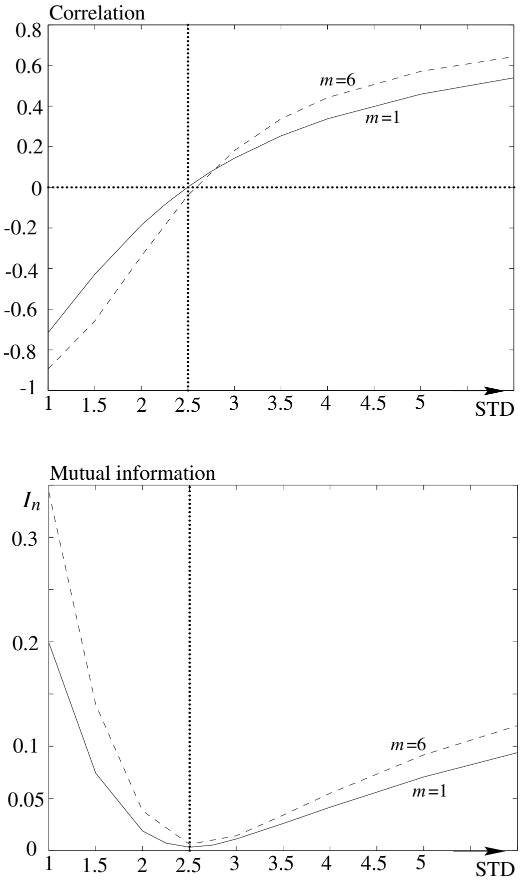

Figure 2. [Solid] At the assumed kernel STD of 2.5 pixels the basic solutions are decorrelated and have minimal mutual information (shown normalized), in agreement with the true STD used in Fig. 1. [Dashed] The absolute correlation and the mutual information are larger for a large value of m.

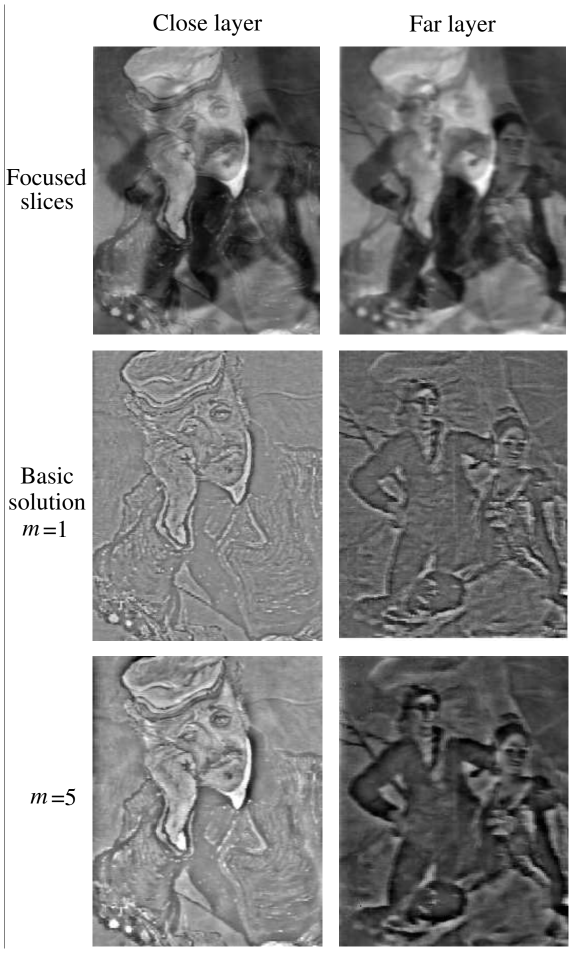

Figure 3. [Top] The slices in which either of the transparent layers is focused. [Middle row] The basic solution ($m = 1$). [Bottom row] Recovery with $m = 5$.

in the sensors, and the raster reading process in the CCD. Moreover, the inter-pixel spacing along the sensor rows is generally different than along the columns, thus even the optical blur may affect them differently. We assigned a different kernel STD to each axis: $\mathrm{STD}^{\mathrm{row}}$ and $\mathrm{STD}^{\mathrm{column}}$. Since two slices are used, there are two kernels, with a total of four parameters. Defining the parameter vector $\mathbf{p} \equiv (\mathrm{STD}^{\mathrm{row}}_{1b}, \mathrm{STD}^{\mathrm{column}}_{1b}, \mathrm{STD}^{\mathrm{row}}_{2a}, \mathrm{STD}^{\mathrm{column}}_{2a})$, the estimated vector $\hat{\mathbf{p}}$ is

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \mathcal{I}_n[\tilde{f}_1(\mathbf{p}), \tilde{f}_2(\mathbf{p})] \quad . \tag{10}$$

There may be numerous parameter combinations that lead to decorrelation in the multi-parameter case, but will not all lead to the minimum mutual information, or to good separation. If $\mathbf{p}$ is N-dimensional, the zero-correlation constraint defines an $N-1$ dimensional hypersurface in the parameter space. It is possible to use this criterion to obtain initial estimates of $\mathbf{p}$, and search for minimal mutual information within a lower dimensional manifold.

Experiment based on two slices

A print of the "Portrait of Doctor Gachet" (by van-Gogh) was positioned closely behind a glass window. The window partly reflected a more distant picture, a part of a print of the "Parasol" (by Goya). The $f\#$ was 5.6. The two focused slices[2] are shown at the top of Fig. 3. The cross correlation between the raw (focused) images is 0.98. The normalized mutual information is $\mathcal{I}_n \approx 0.5$ indicating that significant separation is achieved by the focusing process, but that substantial crosstalk remains.

The optimal parameter vector $\hat{\mathbf{p}}$ in the sense of min-

[2] There was slight magnification with change of focus settings, which was compensated for manually. The magnification is avoided if the imaging system is telecentric [21].

imum mutual information is [1.9, 1.5, 1.5, 1.9] pixels, where STD_{1b} corresponds to the blur of the close layer, and STD_{2a} corresponds to the blur of the far layer. With these parameters, the basic solution ($m = 1$) shown at the middle row of Fig. 3 has $\mathcal{I}_n \approx 0.006$ (two orders of magnitude better than the raw images). Using $m = 5$ yields better balance between the low and high frequency components, but $\mathcal{I}_n$ increased to about 0.02. We believe that this is due to the error in the PSF model, as discussed above.

4 Using a focused slice and a pinhole image

Acquiring one image via a very small aperture ("pinhole camera") somewhat simplifies the algorithms. Since the two images are taken with the same axial positions of the system components, no geometric distortions arise [14]. As shown below, there are also fewer parameters to estimate. The "pinhole" image is described by

$$g_0 = (f_1 + f_2)/a \ , \tag{11}$$

where $1/a$ is the attenuation of the intensity due to the smaller aperture. This image is used in conjunction with one of the focused slices of Eq. (1), for example g_a. The inverse filtering solution is

$$\widehat{F}_1 = S(G_a - aG_0 H_{2a}) \qquad \widehat{F}_2 = S(aG_0 - G_a) \tag{12}$$

where $S = (1 - H_{2a})^{-1}$. As in section 2, S can be approximated by $\widehat{S}_m = \sum_{k=1}^{m} H_{2a}^{k-1}$.

Suppose that, in addition to using H_{2a} in the reconstruction rather than the true imaging transfer function $\widetilde{H}_{2a}$, we inaccurately use the scalar a rather than the true value $\tilde{a}$ used in the imaging process. Let e denote the relative error in this parameter, $e = (a - \tilde{a})/\tilde{a}$. We obtain that

$$\widetilde{F}_1 = \widehat{F}_{1m} - eH_{2a}\widehat{S}_m F_1 - (E_{2a} + eH_{2a})\widehat{S}_m F_2 \ , \tag{13}$$

$$\widetilde{F}_2 = \widehat{F}_{2m} + (E_{2a} + e)\widehat{S}_m F_2 + e\widehat{S}_m F_1 \ , \tag{14}$$

where here $\widehat{F}_{1m}$ and $\widehat{F}_{2m}$ are the results had the imaging defocus kernel been the same as the one used in the reconstruction and had $a = \tilde{a}$. Note the importance of the estimation of a: if $e = 0$ then $\widetilde{F}_2$ (the defocused layer) is recovered uncontaminated by F_1.

Now there is only one filter involved, H_{2a}, since the layer f_1 is focused. There are three parameters to determine: $\mathrm{STD}_{2a}^{\mathrm{row}}$, $\mathrm{STD}_{2a}^{\mathrm{column}}$ and a. The parameter a is easier to obtain as it indicates the ratio of the light energy in the wide-aperture image relative to the pinhole image. Ideally, it is the square of the reciprocal of the ratio of the *f-numbers* of the camera, in the two states. If, however, the optical system is not calibrated, or if there is automatic gain control in the sensor, this ratio is not an adequate estimator of a. a can then be estimated by the ratio of the average values of the images, for example. Such an approximation may serve as a starting point for better estimates.

Figure 4. [Top left] The slice in which the far layer is focused, when viewed with the wide aperture. [Top right] The "pinhole" image. [Bottom] The basic recovery.

Experiment based on a slice and a pinhole image

A print of the "Portrait of Armand Roulin" (by van-Gogh) was positioned closely behind a glass window, which semireflected a more distant picture, a print of a part of the "Miracle of San Antonio" (by Goya). The "pinhole" image was acquired using the state corresponding to the $f\# = 11$ mark on the lens, while the wide aperture image was acquired using the state corresponding to the $f\# = 4$ mark. We have not calibrated the lens, so these marks do not necessarily correspond to the true values. The slice in which the far layer is focused (using the wide aperture) is shown in the top left of Fig. 4. In the "pinhole" image (top right), the presence of the "Portrait" layer is more noticeable, but still it is hardly visible in both of the raw images.

According to the ratio of the $f\#$'s, the wide aperture image should have been brighter than the "pinhole" image by $(11/4)^2 \approx 7.6$. However, the ratio between the mean intensity of the wide aperture image to that of the pinhole image was 4.17. This could be due to poor calibration of the lens by its manufacturer, or because of some automatic gain control in the sensor. We thus added a to the set of parameters to be searched.

We use this case to demonstrate the minimization over a lower dimensional manifold. For each hypothesized pair of blur radii $\mathrm{STD}^{\mathrm{row}}$ and $\mathrm{STD}^{\mathrm{column}}$, the parameter a that led to decorrelation of the basic solution was sought (near the rough estimate based on intensity ratios). Then, the mutual information was calculated over the parameters that cause decorrelation (a manifold of only two parameters). The blur diameters that led to minimal mutual information at $m = 1$ were $\mathrm{STD}^{\mathrm{row}} = \mathrm{STD}^{\mathrm{column}} = 11$ pixels, with the best parameter a being 4.28. The reconstruction results are shown in the bottom of Fig. 4. Their mutual information (normalized) is 0.004. Using a larger m increased the mutual information and the visible crosstalk.

5 Conclusions

This paper presents a method based on focusing to separate transparent layers, as appear in semi-reflected scenes. The focus-based approach is more stable with respect to perturbations and occlusions [13] than methods that rely on stereo or motion. We presented a method for self calibration of the defocus blur kernels given the raw images. It is based on minimizing the mutual information of the recovered layers. Note that since the defocus blur has a similar origin to motion blur [13] and stereo disparity (effectively different mostly in the scale and dimensionality of the kernels), the method described here may possibly be adapted to find the motion PSFs and disparities in transparent scenes.

The PSF was assumed to be shift invariant, meaning that the depth variations within each object are small (however, the semireflecting surface can be inclined at any arbitrary angle). The method can be generalized to more complicated blurring operations by optimizing the mutual information over a higher dimensional parametric space.

In some cases the methods presented here are also applicable to multiplicative layers [19]: If the opacity variations within the close layer are small (a "weak" object), the transparency effect may be approximated as a linear superposition of the layers, as done in microscopy [1, 8, 12]. In microscopy and in tomography, the suggested method for self calibration of the PSF can improve the removal of crosstalk between adjacent slices.

We thank Bonnie Lorimer for the permission to use her photograph of Mt. Shuksan, and Ronen Basri for his useful comments. This research was supported by the Israeli Ministry of Science, by the Ollendorff Fund of the Department of Electrical Engineering, Technion, and by the Tel Aviv University Internal Research Fund.

Yoav Y. Schechner is currently affiliated with the Department of Computer Science, Columbia University, and can be addressed at `yoav@cs.columbia.edu`.

References

[1] D. A. Agard and J. W. Sedat, "Three-dimensional structure of a polytene nucleus", Nature **302**, pp. 676-681 [1983].

[2] J. R. Bergen, P. J. Burt, R. Hingorani and S. Peleg, "A three-frame algorithm for estimating two-component image motion," IEEE Trans. PAMI **14**, pp. 886-895 [1992].

[3] M. Borga and H. Knutsson, "Estimating multiple depths in semi-transparent stereo images," Proc. SCIA, Vol. I, pp. 127-133 [1999].

[4] T. M. Cover and J. A. Thomas, *Elements of information theory*, pp. 12-21 (John Wiley & Sons, New York, 1991).

[5] T. Darrell and E. Simoncelli, " 'Nulling' filters and the separation of transparent motions", Proc. CVPR, pp. 738-739 [1993].

[6] H. Farid and E. H. Adelson, "Separating reflections from images using independent components analysis," JOSA A **16**, pp. 2136-2145 [1999].

[7] M. Irani, B. Rousso and S. Peleg, "Computing occluding and transparent motions," Int. J. Comp. Vis. **12**, pp. 5-16 [1994].

[8] F. Marcias-Garza, A. C. Bovik, K. R. Diller, S. J. Aggarwal and J. K. Aggarwal, "The missing cone problem and low-pass distortion in optical serial sectioning microscopy", Proc. ICASSP, vol-II, pp. 890-893 [1988].

[9] J. G. McNally, C. Preza, J. A. Conchello and L. J. Thomas Jr., "Artifacts in computational optical-sectioning microscopy," JOSA A **11**, pp. 1056-1067 [1994].

[10] N. Ohnishi, K. Kumaki, T. Yamamura and T. Tanaka, "Separating real and virtual objects from their overlapping images," Proc. ECCV vol-II, Lecture notes in CS **1065**, pp. 636-646 [1996].

[11] M. Oren and S. K. Nayar, "A theory of specular surface geometry," Proc. ICCV, pp. 740-747 [1995].

[12] C. Preza, M. I. Miller, L. J. Thomas Jr. and J. G. McNally, "Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections", JOSA A **9**, pp. 219-228 [1992].

[13] Y. Y. Schechner and N. Kiryati, "Depth from defocus vs. Stereo: How different really are they?" Proc. ICPR, pp. 1784-1786 [1998].

[14] Y. Y. Schechner, N. Kiryati and R. Basri, "Separation of transparent layers using focus," Proc. ICCV, pp. 1061-1066 [1998].

[15] Y. Y. Schechner, N. Kiryati and J. Shamir, "Separation of transparent layers by polarization analysis," Proc. SCIA, Vol. I, pp. 235-242 [1999].

[16] Y. Y. Schechner, J. Shamir and N. Kiryati, "Polarization-based decorrelation of transparent layers: The inclination angle of of an invisible surface," Proc. ICCV, Vol. II, pp. 814-819 [1999].

[17] Y. Y. Schechner, J. Shamir and N. Kiryati "Polarization and statistical analysis of scenes containing a semireflector," JOSA A **17**, pp. 276-284 [2000].

[18] M. Shizawa, "On visual ambiguities due to transparency in motion and stereo," Proc. ECCV, pp. 411-419 [1992].

[19] M. Shizawa and K. Mase, "Simultaneous multiple optical flow estimation," Proc. ICPR, pp. 274-278 [1990].

[20] J. Y. A. Wang and E. H. Adelson, "Layered representation from motion analysis," Proc. CVPR, pp. 361-365 [1993].

[21] M. Watanabe and S. K. Nayar, "Telecentric optics for computational vision", Proc. ECCV, vol-II, pp. 439-451 [1996].