

Polarization and statistical analysis of scenes containing a semireflector

Yoav Y. Schechner and Joseph Shamir

Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel

Nahum Kiryati

Department of Electrical Engineering-Systems, Faculty of Engineering, Tel-Aviv University, Ramat Aviv 68878, Israel

Received June 3, 1999; revised manuscript received September 22, 1999; accepted September 22, 1999

We present an approach to recover scenes deteriorated by reflections off a semireflecting medium (e.g., a glass window). The method, based on imaging through a polarizer at two or more orientations, separates the reflected and transmitted scenes and determines which is which. We analyze the polarization effects, taking into account internal reflections within the medium. The scene reconstruction requires the estimation of the orientation (inclination and tilt angles) of the transparent (invisible) surface. The inclination angle is estimated by seeking the value that leads to the minimal mutual information of the estimated scenes. The limitations and the consequences of noise and angle error are discussed, including a fundamental ambiguity in the determination of the plane of incidence. Experimental results demonstrate the success of angle estimation and consequent scene separation and labeling. © 2000 Optical Society of America [S0740-3232(00)00802-4] OCIS codes: 100.3190, 260.5430, 150.0150, 100.3020, 150.6910.

1. INTRODUCTION

The situation in which several (typically two) linearly superimposed contributions exist is common in real-world scenes. For example,¹⁻⁴ looking out of a car (or room) window, we see both the outside world (termed the real object⁴⁻⁶) and a semireflection of the objects inside, termed virtual objects (Fig. 1). The separation of the contributions is important for understanding and analysis of the scene (for example, the superposition can certainly confuse autofocusing devices⁴). The mere detection of the phenomenon indicates the presence of a clear, transparent surface in front of the camera at a distance closer than the actual imaged objects.^{4,5}

The term transparent layers has been used to describe situations in which a scene is semireflected from a transparent surface.^{2,7,8} The image is decomposed into depth-ordered layers, each with an associated map describing its intensity (and, if applicable, its depth or motion⁸). We adopt this terminology but stress the fact that we do not deal with imaging through an object with variable opacity. Approaches to reconstructing each of the layers by eliminating the other relied mainly on motion,^{2,7} stereo,⁹ and focus.⁴ An *a priori* assumption of these procedures is that the superimposed layers lie at significantly different optical distances from the camera, and reconstruction of the low-frequency components is ill conditioned⁴ (while the dc component is ill posed). Other fundamental ambiguities in the solutions obtained by motion and stereo were discussed in Refs. 10 and 11.

Suppressing the virtual layer by incorporating a polarizer into the imaging system is a well-known photographic technique.¹ Progress in this direction is marked by advanced polarizing cameras,^{3,12-19} which were applied in particular for removal of specular reflections su-

perimposed on diffuse scattering from opaque surfaces.^{15,18,19} Some previous works attempted to remove the virtual layer by using just the raw output of the polarization analyzer (Fig. 1) in front of the camera.^{3,5} These methods suggested taking several images of the scene at different states of the polarizer and picking one of them as the reconstruction of the real layer. However, polarization filtering eliminates the reflected (virtual) layer only at the Brewster angle^{1,20} (good filtering was demonstrated in Ref. 5 at this angle). Independent-components analysis of polarization-filtered images demonstrated the potential of polarization as an initial step for good separation achieved by signal postprocessing.²¹

A significant disadvantage of the above-indicated works is that they could not determine which of the reconstructed images is of the real object and which is of the virtual one. Obviously, there is an exception when the transparent surface is oriented at the Brewster angle.⁵ Earlier works also did not extract any information about the invisible semireflecting surface itself unless the surface was curved. For this case a method is available⁶ for its geometric analysis and the classification of local features as real or virtual.

This work extends the preliminary analysis that we presented in Refs. 22 and 23. In Ref. 22 we analyzed the physical process and demonstrated an approach for the reconstruction of the contributing layers in case the surface orientation is known. We took into account the implication of internal reflections within double-surfaced reflecting media (e.g., glass windows). The method enables the automatic labeling of the reconstructed layers as virtual or real. Unlike methods that rely on stereo, motion, or defocus, the reconstruction is not ill conditioned at low spatial frequencies and does not require the layers to

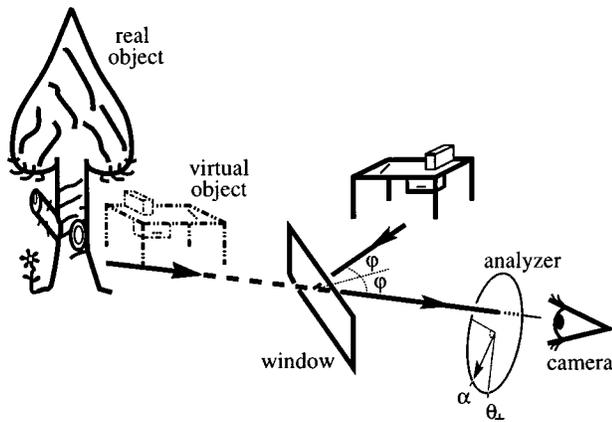


Fig. 1. The image of a real object is partially transmitted through a transparent window inclined at an angle ϕ . The window also creates a virtual image by partially reflecting the image of another object. The combined scene can be viewed through a polarization analyzer (filter) at angle α . The polarization component perpendicular to the plane of incidence is best transmitted for $\alpha = \theta_{\perp}$.

have different depths or motion fields. In Ref. 23 we proposed a method to determine the angle of incidence (AOI) on the transparent surface, which was based on decorrelation of the images. However, in the following we show that this approach may yield erroneous results in case the original images are somewhat correlated, and we propose to use mutual information for this purpose. Furthermore, we analyze the effects of noise and error of the AOI on the reconstruction. The estimation of the azimuthal tilt angle of the plane of incidence (POI) completes the recovery of the orientation of the invisible semireflecting surface in space. We uncover a fundamental ambiguity in the determination of the POI, which is common with transparent scenes, in contrast to reflections off opaque media.

2. IMAGE FORMATION

A. Polarizing Effects

For our analysis we represent the various light intensities by their polarization components, I_{\parallel} (polarization parallel to the POI) and I_{\perp} (polarization perpendicular to the POI). At each interface of the transparent medium, each component has, respectively, reflectivities R_{\parallel} , R_{\perp} and transmissivities T_{\parallel} , T_{\perp} derived from the Fresnel equations.²⁰ However, cases of reflection from double-surfaced media (windows) are by far more common than reflection by a single surface, and internal reflections may be significant. Assuming incoherent light and negligible absorption, for each polarization component the total reflectivity is (Refs. 22 and 23)

$$\tilde{R} = R + T^2 R \sum_{l=0}^{\infty} (R^2)^l = \frac{2}{1+R} R. \quad (1)$$

We assume also that the shift of the image on consecutive reflections is negligible within the significant reflection orders. The transmissivity of each component is given by $\tilde{T} = 1 - \tilde{R}$. When we recall that $R_{\perp} > R_{\parallel}$,¹ it is easy to show also that $\tilde{R}_{\perp} > \tilde{R}_{\parallel}$.

Consider now the influence of the window on the polarization. Since light is generally partially polarized, transmission through the analyzer (Fig. 1) is maximal or minimal for certain analyzing angles α . The intensities at these angles are denoted I_{\max} and I_{\min} , respectively. The degree of polarization of partially linearly polarized light^{16,19} is $(I_{\max} - I_{\min})/(I_{\max} + I_{\min})$. We defined²³ the polarizing effect (PE) of a process as the degree of polarization that it induces on unpolarized incident light. Thus the polarizing effects in a double-surfaced medium are

$$\begin{aligned} \text{PE}_R(\text{window}) &= \frac{\tilde{R}_{\perp} - \tilde{R}_{\parallel}}{\tilde{R}_{\perp} + \tilde{R}_{\parallel}}, \\ \text{PE}_T(\text{window}) &= \frac{\tilde{T}_{\parallel} - \tilde{T}_{\perp}}{\tilde{T}_{\parallel} + \tilde{T}_{\perp}}, \end{aligned} \quad (2)$$

while in a single surface R and T replace \tilde{R} and \tilde{T} . We believe that the polarization effect is a better indicator of the polarization properties of the medium than the ratio of the Fresnel coefficients (used, for example, in Ref. 19), since it is bounded. The PE's of these processes in glass are plotted in Fig. 2. Well-known facts are manifested in this plot. For example, $\text{PE}_R = 1$ at the Brewster angle, at which the parallel component vanishes, and it is zero for AOI $\phi = 0^\circ, 90^\circ$. For most AOI's PE_R is larger than PE_T . Compared with that from a single surface, the PE of reflection from a window is slightly smaller, but the PE of transmission is significantly larger and is almost twice as large at most incidence angles.

Now we consider the relative polarizing effect as the ratio of the PE of transmission and that of reflection. It is easy to show that

$$\frac{\text{PE}_T(\text{window})}{\text{PE}_R(\text{window})} = \frac{\tilde{R}_{\text{av}}}{\tilde{T}_{\text{av}}}, \quad (3)$$

where $\tilde{R}_{\text{av}} = (\tilde{R}_{\perp} + \tilde{R}_{\parallel})/2$ and $\tilde{T}_{\text{av}} = (\tilde{T}_{\perp} + \tilde{T}_{\parallel})/2$ are the reflectivity and the transmissivity of unpolarized light,

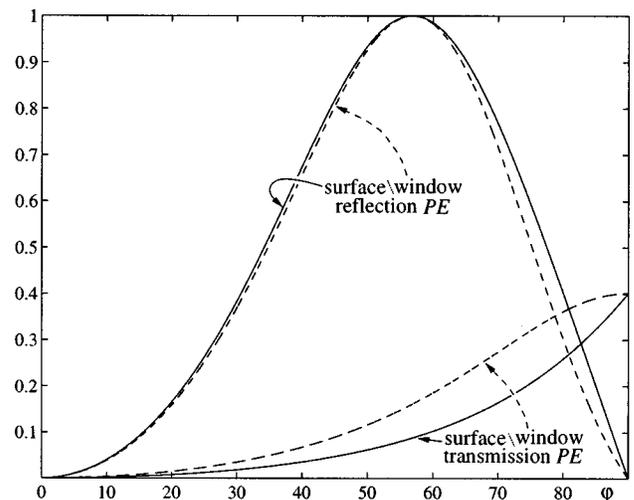


Fig. 2. Polarizing effects (PE's) of reflection and transmission through a single air-glass interface as a function of the angle of incidence (AOI) (solid curves) and PE's of reflection and transmission through a glass window (dashed curves).

respectively. Therefore, neglecting the degree of polarization induced on the transmitted light (PE_T) is equivalent to neglecting the reflection phenomenon (which is invalid in the cases discussed in this work).

B. Imaging

We denote the spatial intensity distribution that is due to the real layer (with no window) by I_T and the spatial intensity distribution that is due to the virtual layer (assuming a perfect mirror replacing the window) by I_R . Let θ_\perp be the orientation of the polarization analyzer for best transmission of the component perpendicular to the POI. Generally, the orientation of the analyzer is α . Assuming initially unpolarized natural light, the observed intensity is

$$f(\alpha) = \left(\frac{f_\perp + f_\parallel}{2} \right) + \left(\frac{f_\perp - f_\parallel}{2} \right) \cos[2(\alpha - \theta_\perp)], \quad (4)$$

where $f_\perp = f(\theta_\perp) = I_R \tilde{R}_\perp / 2 + I_T \tilde{T}_\perp / 2$ and $f_\parallel = f(\theta_\perp + 90^\circ) = I_R \tilde{R}_\parallel / 2 + I_T \tilde{T}_\parallel / 2$.

The intensity is a sinusoidal function of the analyzer's orientation, as with any partially polarized light. However [see Eq. (3)],

$$f_\perp - f_\parallel = PE_R \tilde{R}_{av} I_R - PE_T \tilde{T}_{av} I_T = PE_R \tilde{R}_{av} (I_R - I_T). \quad (5)$$

Thus, if $I_T = I_R$, the amplitude in Eq. (4) is zero, i.e., the light leaving the reflecting medium is unpolarized, even though each of the contributing layers is partially polarized.²⁴ If the real object is brighter than the virtual one ($I_T > I_R$) (e.g., when looking out of the room window during daylight), the intensity $f(\alpha)$ would be minimal at $\alpha = \theta_\perp$. The low PE_T (relative to PE_R) at low and moderate incidence angles is compensated by the high ratio of transmissivity to reflectivity at these angles. Therefore the polarization induced on the transmitted light, rather than on the reflected light, may dominate the determination of the overall polarization.

3. RECONSTRUCTION

A. For a Given Angle of Incidence

Suppose now that the geometry of the setup is known, that is, the plane of incidence (hence θ_\perp) and the angle of incidence φ are known or can be estimated. In this case the estimates of f_\perp and f_\parallel (denoted as \hat{f}_\perp and \hat{f}_\parallel) can easily be obtained. These images are not sensitive to small errors in the estimation of θ_\perp , since $\partial f / \partial \theta_\perp = 0$ at $\alpha = \theta_\perp, \theta_\perp + 90^\circ$. Moreover, in this case \tilde{R}_\perp and \tilde{R}_\parallel are known; thus²³ the estimated intensities can be written as

$$\hat{I}_T(\varphi) = \left[\frac{2\tilde{R}_\perp(\varphi)}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)} \right] \hat{f}_\parallel - \left[\frac{2\tilde{R}_\parallel(\varphi)}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)} \right] \hat{f}_\perp, \quad (6)$$

$$\hat{I}_R(\varphi) = \left[\frac{2 - 2\tilde{R}_\parallel(\varphi)}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)} \right] \hat{f}_\perp - \left[\frac{2 - 2\tilde{R}_\perp(\varphi)}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)} \right] \hat{f}_\parallel. \quad (7)$$

Equations (6) and (7) show that if the incidence is at the Brewster angle (for which $\tilde{R}_\parallel = 0$), \hat{I}_T can be directly as-



Fig. 3. Real layer I_T (top left) and virtual layer I_R (top right), and simulated images at an AOI of 80° (bottom). Although the reflected component in f_\parallel (left) is weak (especially when compared with f_\perp on the right), significant cross talk of the layers exists.

sociated with \hat{f}_\parallel , as demonstrated in Ref. 5. But even at that angle, \hat{I}_R is not proportional to $\hat{f}_\perp - \hat{f}_\parallel$, in contrast to Ref. 5. Nevertheless, operation at the Brewster angle is a rare situation, and one should generally use Eqs. (6) and (7).

Consider the images shown in the top row of Fig. 3 as the true layers in simulating an indoor scene reflected off a window through which an outdoor scene²⁵ is viewed (the results of an experiment with a physical setup are described in Section 4). The bottom row shows the simulated images \hat{f}_\parallel and \hat{f}_\perp (all images are contrast stretched for clarity) for an AOI of 80° . Since this angle is far from the Brewster angle, \hat{f}_\parallel does not suffice to clear the reflected disturbance. However, when the correct AOI was used, the inversion of the image formation process was accurate and correctly separated and labeled the layers, just as shown in the original images in the top row of Fig. 3.

B. Effect of Noise

Suppose that noise n_\perp, n_\parallel is added to the raw images, which thus become $\hat{f}_\perp = f_\perp + n_\perp$ and $\hat{f}_\parallel = f_\parallel + n_\parallel$. The reconstructed images then become

$$\hat{I}_T = I_T + 2 \frac{\tilde{R}_\perp(\varphi)n_\parallel - \tilde{R}_\parallel(\varphi)n_\perp}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)},$$

$$\hat{I}_R = I_R + 2 \frac{\tilde{T}_\perp(\varphi)n_\parallel - \tilde{T}_\parallel(\varphi)n_\perp}{\tilde{T}_\perp(\varphi) - \tilde{T}_\parallel(\varphi)}. \quad (8)$$

If n_\perp and n_\parallel are independent, with zero mean and variance σ^2 , then the variances of the noise in the resulting images are

$$\sigma_T^2 = 2\sigma^2 \left[\frac{1}{\text{PE}_R^2(\varphi)} + 1 \right], \quad \sigma_R^2 = 2\sigma^2 \left[\frac{1}{\text{PE}_T^2(\varphi)} + 1 \right]. \quad (9)$$

The reconstructions become unstable and noise sensitive as the polarizing effects approach zero (Fig. 2), that is, at low or very high incidence angles. Note, however, that the noise amplification at these angles depends on our choice of φ . Therefore, if observation is made at such a problematic AOI, it may be preferable to intentionally insert a more moderate AOI into the reconstruction equations, giving up inversion accuracy for lower noise in the results. At most AOI's the typically low values of PE_T , relative to PE_R , will make σ_R significantly larger than σ_T . We can thus expect the noise in the reconstruction of the virtual layer to be significantly stronger than that in the real layer. At very high AOI, the real layer will be noisier than the virtual one.

The noise can be significantly reduced by common spatial filtering methods (e.g., Wiener filtering). If the noise is white, local blurring can definitely reduce it but will cause reduction of details. Thus the method will separate the low-frequency components of the layers better than the high-frequency ones. This is in contrast to methods that rely on motion, stereo, or focus, for which the high-frequency components are easily resolved while instability occurs at the low frequencies.

C. Effect of Inclination Error

The reflectivities used in Eqs. (6) and (7) depend on the AOI. Suppose that we assume that the AOI is φ , while the true AOI is actually φ_{true} . Then²³

$$\begin{aligned} \hat{I}_T(\varphi) &= (1 - \rho)I_T + \rho I_R, \\ \hat{I}_R(\varphi) &= (1 - \tau)I_R + \tau I_T, \end{aligned} \quad (10)$$

where

$$\rho(\varphi_{\text{true}}, \varphi) = \frac{\tilde{R}_\perp(\varphi)\tilde{R}_\parallel(\varphi_{\text{true}}) - \tilde{R}_\perp(\varphi_{\text{true}})\tilde{R}_\parallel(\varphi)}{\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)}, \quad (11)$$

$$\tau(\varphi_{\text{true}}, \varphi) = \frac{\tilde{T}_\perp(\varphi)\tilde{T}_\parallel(\varphi_{\text{true}}) - \tilde{T}_\perp(\varphi_{\text{true}})\tilde{T}_\parallel(\varphi)}{\tilde{T}_\perp(\varphi) - \tilde{T}_\parallel(\varphi)}. \quad (12)$$

We note that $\rho, \tau = 0$ if and only if $\varphi = \varphi_{\text{true}}$. For a range of assumed values of φ , traces of I_R are left in \hat{I}_T , while for other values of φ the traces are negative.

For a moment suppose that the analyzer is not used during imaging, so the intensity is $I_R\tilde{R}_{\text{av}} + I_T\tilde{T}_{\text{av}}$. If we are interested only in the real layer, we regard $I_T\tilde{T}_{\text{av}}$ as the signal and $I_R\tilde{R}_{\text{av}}$ as the noise. In general, we define the contamination of the real layer as the reciprocal of its signal-to-noise ratio, based on the raw unpolarized image: $\text{contamination}_T = I_R\tilde{R}_{\text{av}}/I_T\tilde{T}_{\text{av}}$. So,

$$\hat{I}_T(\text{without analyzer}) = \tilde{T}_{\text{av}}I_T(1 + \text{contamination}_T). \quad (13)$$

If $I_R\tilde{R}_{\text{av}} = I_T\tilde{T}_{\text{av}}$, the noise is as strong as the signal; thus the latter is 100% contaminated. On the other

hand, if we are interested only in the virtual layer, $I_R\tilde{R}_{\text{av}}$ is the signal and $I_T\tilde{T}_{\text{av}}$ is the noise, so $\text{contamination}_R = I_T\tilde{T}_{\text{av}}/I_R\tilde{R}_{\text{av}}$ and

$$\hat{I}_R(\text{without analyzer}) = \tilde{R}_{\text{av}}I_R(1 + \text{contamination}_R). \quad (14)$$

To get an indication about the effect of an erroneous AOI on the reconstruction results, we rewrite Eqs. (10) as

$$\hat{I}_T(\varphi) = (1 - \rho)I_T(1 + c_T\text{contamination}_T), \quad (15)$$

$$\hat{I}_R(\varphi) = (1 - \tau)I_R(1 + c_R\text{contamination}_R), \quad (16)$$

where we define the contamination coefficients

$$c_T(\varphi) = \frac{\rho(\varphi)}{1 - \rho(\varphi)} \frac{\tilde{T}_{\text{av}}}{\tilde{R}_{\text{av}}}, \quad c_R(\varphi) = \frac{\tau(\varphi)}{1 - \tau(\varphi)} \frac{\tilde{R}_{\text{av}}}{\tilde{T}_{\text{av}}}. \quad (17)$$

Comparing Eq. (13) with Eq. (15), we see that the contamination (relative to the signal) of \hat{I}_T is multiplied by c_T if the reconstruction method is used. The contamination of \hat{I}_R is multiplied by c_R . If there is a 100% contamination in the unpolarized raw image ($I_R\tilde{R}_{\text{av}} = I_T\tilde{T}_{\text{av}}$), Eqs. (17) indicate the relative contamination of each reconstruction.

Suppose now that the error of the AOI is small. To estimate its implication, we differentiate $c_T(\varphi)$ and $c_R(\varphi)$ around $\varphi = \varphi_{\text{true}}$ (where $c_T, c_R = 0$). It can be shown that in this case

$$\frac{dc_T}{d\varphi} = \frac{\tilde{T}_{\text{av}}}{\tilde{R}_{\text{av}}} \frac{1}{\tilde{R}_\perp - \tilde{R}_\parallel} \left(\frac{d\tilde{R}_\perp}{d\varphi} \tilde{R}_\parallel - \frac{d\tilde{R}_\parallel}{d\varphi} \tilde{R}_\perp \right), \quad (18)$$

$$\frac{dc_R}{d\varphi} = \frac{\tilde{R}_{\text{av}}}{\tilde{T}_{\text{av}}} \frac{1}{\tilde{T}_\perp - \tilde{T}_\parallel} \left(\frac{d\tilde{T}_\perp}{d\varphi} \tilde{T}_\parallel - \frac{d\tilde{T}_\parallel}{d\varphi} \tilde{T}_\perp \right). \quad (19)$$

These derivatives, indicating the first-order effect of an error in the AOI, are plotted in Fig. 4. Note that the reconstruction of the virtual layer is quite stable to an error in the AOI. This is so since a small AOI error yields a small c_R , which significantly attenuates the contamination. Thus, if we wish to reconstruct the reflection off a slightly curved window (as a car windshield), we may regard it as having the same AOI (flat along the axial direction) throughout the field of view.

The reconstruction of the real layer is more sensitive to such an error, and the solid curve in Fig. 4 can give us an indication of the performance in such circumstances. For example, consider the case of a 100% contamination in the unpolarized image ($\text{contamination}_T = 1$). If the photograph was taken at $\varphi_{\text{true}} = 30^\circ$ and we erred in our estimation by approximately 3° , we obtain $\approx 20\%$ of contamination ($\approx 7\%$ per degree), which may be noticeable but still is much better than it would be without the reconstruction method. The situation is improved if the reflection is weak. For example, suppose that $\varphi_{\text{true}} = 20^\circ$ and that the reflection is 1/10 the intensity of the transmission ($\text{contamination}_T = 0.1$). We may err by up to approximately 10° (if the first-order approximation is still valid) in the AOI and still get a better image than the unpolarized one.

D. Estimating the Angle of Incidence

Pointwise data analysis provides little information (if any) on the AOI. To estimate it, we need an assumption related to multiple points. Let us assume now 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.²⁶ Let the probabilities for certain values \tilde{I}_T and \tilde{I}_R be $P(\tilde{I}_T)$ and $P(\tilde{I}_R)$, respectively. In practice, these probabilities are estimated by the histograms of the reconstructed images. The joint probability is $P(\tilde{I}_T, \tilde{I}_R)$, which in practice is estimated by the joint histogram of the images, that is, the relative number of pixels in which \tilde{I}_T has a certain value \tilde{I}_T and \tilde{I}_R has a certain value \tilde{I}_R at corresponding pixels. The mutual information is then

$$\mathcal{I}(\hat{I}_T, \hat{I}_R) = \sum_{\tilde{I}_T, \tilde{I}_R} P(\tilde{I}_T, \tilde{I}_R) \log \frac{P(\tilde{I}_T, \tilde{I}_R)}{P(\tilde{I}_T)P(\tilde{I}_R)}. \quad (20)$$

In this approach we assume that if the layers are correctly separated, each of their estimates contains a *minimum* of information about the other. However, the distance [Eq. (20)] depends on the levels that the images \hat{I}_T and \hat{I}_R are quantized to and their dynamic range, which in turn depends on the brightness of the individual layers I_T and I_R . 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-information²⁶ (entropy) of \hat{I}_T is

$$\mathcal{H}(\hat{I}_T) = -\sum_{\tilde{I}_T} P(\tilde{I}_T) \log P(\tilde{I}_T), \quad (21)$$

and similar is that of \hat{I}_R . Thus the measure that we used,

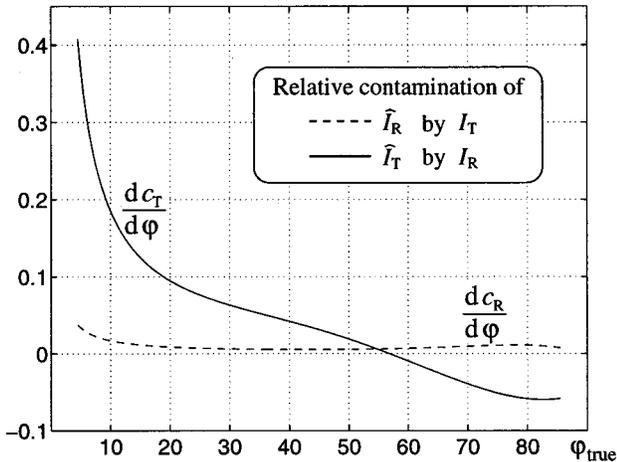


Fig. 4. Relative contamination of each layer, per 1° of error in the AOI, if the reflected contribution is as bright as the transmitted one (after the incidence on a glass window).

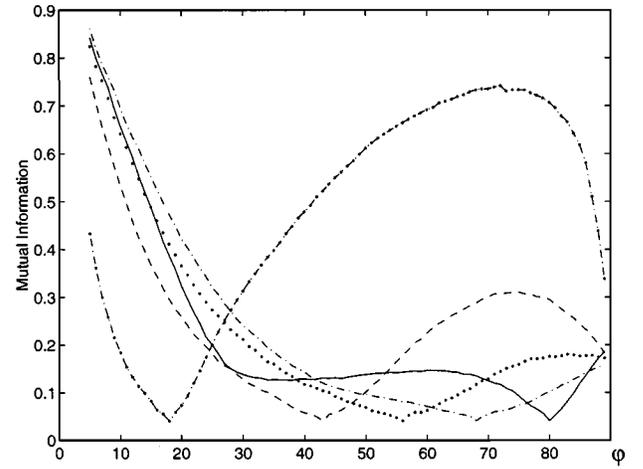


Fig. 5. Mutual information (normalized) of the estimated layers as a function of the assumed AOI φ for $\varphi_{\text{true}} = 18^\circ, 43^\circ, 56^\circ, 68^\circ, 80^\circ$. The minimum mutual information was achieved when $\varphi = \varphi_{\text{true}}$.

$$\mathcal{I}_n(\hat{I}_T, \hat{I}_R) = \frac{\mathcal{I}(\hat{I}_T, \hat{I}_R)}{[\mathcal{H}(\hat{I}_T) + \mathcal{H}(\hat{I}_R)]/2}, \quad (22)$$

indicated the ratio of mutual information to the self-information of a layer. In this approach

$$\hat{\varphi} = \arg \min_{\varphi} \mathcal{I}_n[\hat{I}_T(\varphi), \hat{I}_R(\varphi)]. \quad (23)$$

We tested this approach in several AOI's φ_{true} with the images of Fig. 3. For each test this criterion produced the exact AOI, although the minimum mutual information was not zero. Typical curves of the information measure are plotted in Fig. 5. Note that in the method described above we did not make assumptions about the form of the probability distributions. The only assumption concerns the minimal mutual information, as estimated from the images themselves.

To gain some analytical insight, we make some stronger assumptions. Suppose that the true layers are statistically independent (hence, when correctly separated, $\mathcal{I} = 0$) and are jointly Gaussian. Then linear independence of the reconstructions indicates the states of minimum mutual information. Thus a possible approach is to search for the zero crossing of the cross covariance (or cross correlation) between the estimated images.²³ However, the covariance (for any distribution, not just a Gaussian) is [see Eqs. (10)]

$$\text{Cov}(\hat{I}_T, \hat{I}_R) = \tau(1 - \rho)\text{Var}(I_R) \left[\frac{\text{Var}(I_T)}{\text{Var}(I_R)} - \eta \right], \quad (24)$$

where $\eta(\varphi_{\text{true}}, \varphi) = -[\rho(1 - \tau)]/[\tau(1 - \rho)]$. We note that η monotonically increases with φ and can take any positive value. Thus, besides the desired zero crossing, a zero value of the cross covariance at a wrongly assumed AOI exists. Figure 6 indicates the range of AOI on glass for which $\eta > 0$, leading to the possibility of error. If φ_{true} is smaller than the Brewster angle, the wrong angle of decorrelation will always be at $\varphi > \varphi_{\text{true}}$. This enables the removal of the ambiguity in case *a priori* knowledge indicates that the AOI is smaller than the Brewster

angle. The zero at the wrong φ depends (through η) on the variances of the specific images used. So, if color images are used, each channel will have a different ratio of variances, leading to different wrong crossings. Alternatively, we may analyze distinct parts of the images, so regional estimates of the variance will be different in different patches, leading to different wrong crossings. Another possible way is to estimate the correlation between the images where they are shifted relative to their original position by Δ . Then the wrong zero will appear when $\eta = \text{Cov}[I_T(\mathbf{x}), I_T(\mathbf{x} + \Delta)] / \text{Cov}[I_R(\mathbf{x}), I_R(\mathbf{x} + \Delta)]$. Assuming that this ratio is not constant in Δ , the wrong crossing will be different for different shifts.

Searching for decorrelation may not be reliable enough, since it relies solely on the second moment of the probability distributions (estimated by the finite-size images) and on the assumption that the correlation between the original images is zero. To see a typical case, we consider a simulation based on the images shown in the top row of Fig. 3, with $\varphi_{\text{true}} = 20^\circ$. For each assumed φ , the correlation was calculated, as shown in Fig. 7. The “correct” zero crossing appears at $\varphi = 19^\circ$ (and the second, “wrong” crossing appears at 87°). The estimation of the correlation coefficient based on the layers I_T and I_R shown in the top row of Fig. 3 is ≈ 0.06 , somewhat contradicting the underlying assumption of the decorrelation approach. This deviation leads to the small error in the estimated AOI in that method. Such errors (even of 3°) also occurred in other values of φ_{true} .

In contrast, as plotted in Fig. 5 and in Fig. 7(b), the mutual information criterion yielded precise results. This may also be related to the fact that it exploits the (estimated) probability distributions and does not rely just (or primarily) on their second moment. Therefore mutual information seems to be a more favorable measure.

Let us consider the effect of additive noise expressed in the terms on the right-hand sides of Eqs. (8). If n_\perp and n_\parallel are independent with zero mean and variance σ^2 , then the covariance between the estimated images is the same as Eq. (24) but with the addition of

$$-4\sigma^2 \frac{\tilde{R}_\perp(\varphi)\tilde{T}_\perp(\varphi) + \tilde{R}_\parallel(\varphi)\tilde{T}_\parallel(\varphi)}{[\tilde{R}_\perp(\varphi) - \tilde{R}_\parallel(\varphi)]^2} < 0. \quad (25)$$

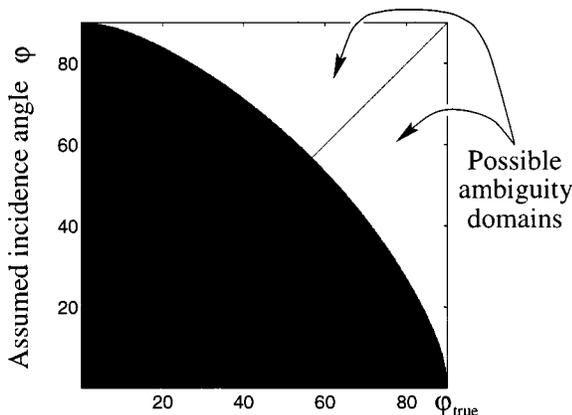


Fig. 6. For each φ_{true} there are domains (white) of assumed angles φ for which a zero of the correlation exists at a wrong angle (besides the correct one).

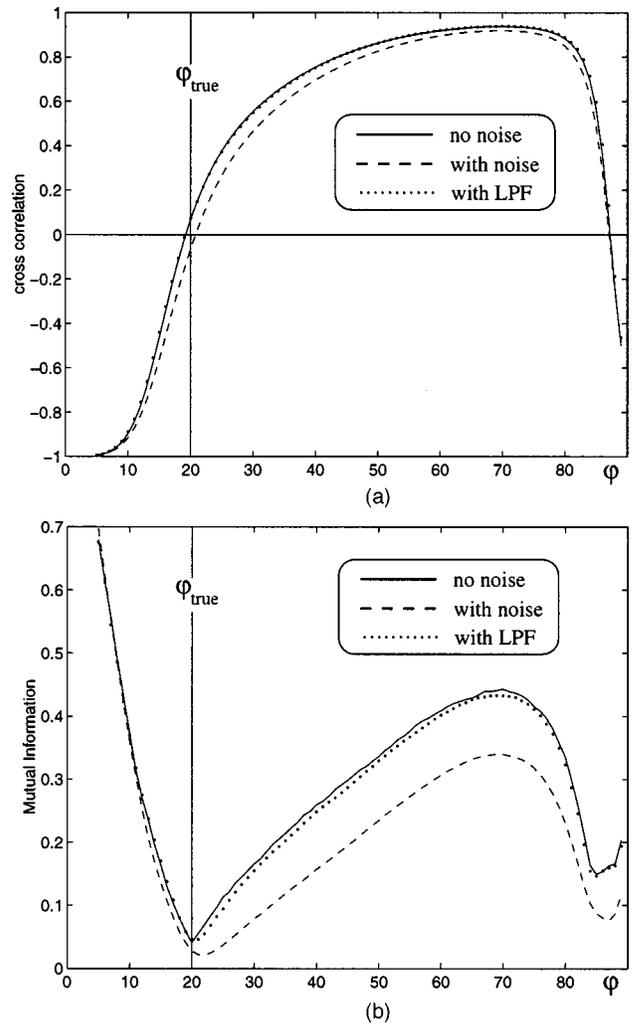


Fig. 7. Decorrelation occurred at $\varphi = 19^\circ$, while the minimum mutual information correctly appeared at 20° . Additive noise increased the estimated angle of the first crossing (as theoretically predicted), while consecutive low-pass filtering (LPF) of the raw images recovered the noiseless results. Similar effects are observed in the minimum mutual information.

Thus noise *decreases* the covariance. We note that the derivative of the covariance with respect to φ is positive at the first crossing and negative at the second. Therefore the noise leads to an *increase* of the first value (angle) of zero crossing and a decrease of the second one. This is demonstrated by the dashed curve in Fig. 7(a), where the noise energy is approximately 2% of the raw images. A way to mitigate this problem is to compensate it by adding the absolute value of relation (25), calculated for each assumed φ , to the measured $\text{Cov}(\hat{I}_T, \hat{I}_R)$ and then look for the zero crossings. This requires an estimate of σ . A more straightforward way is to reduce σ by low-pass filtering the raw images before the estimation of the AOI (and then use this estimated AOI with the raw images). Indeed, in the example plotted in Fig. 7, after we convolved the images with an averaging 5×5 kernel, the original zero crossings were recovered. Note that the angle of minimum mutual information traced a similar response to the noise and to the filtering.

E. Plane-of-Incidence Ambiguity

The orientation of the POI indicates the azimuthal tilt of the transparent (invisible) surface relative to the viewer's coordinate system. This tilt angle is perpendicular to θ_{\perp} . Naturally, θ_{\perp} is important, since once it is determined, \hat{f}_{\perp} and \hat{f}_{\parallel} are unique, enabling the scene recovery. The intensity at each point [Eq. (4)] can be written as

$$f = C + A \cos[2(\alpha - \theta)]. \quad (26)$$

It has 3 degrees of freedom: the dc level C , the amplitude A , and θ . Let $\tan(2\theta) = A_s/A_c$, where we define $A_c \equiv A \cos(2\theta)$ and $A_s \equiv A \sin(2\theta)$. Then $f_i = (1, \cos(2\alpha_i), \sin(2\alpha_i)) \cdot (C, A_c, A_s)$. Thus three measurements of the intensity f_i taken with different general orientations α_i of the analyzer (see Fig. 1)^{12,14,16,19} suffice to determine C , A , and θ .

However, as discussed in Subsection 2.B, if we do not know which object (real or virtual) is brighter at the point measured, there exists a fundamental 90° ambiguity in the determination of the plane of incidence in cases of transparent scenes: It may be associated with either the minimum or the maximum of the analyzer's sinusoidal output. Moreover, if in some parts of the field of view the virtual objects are brighter while in other parts the real objects are, the ambiguity may be unordered, so we cannot associate \hat{f}_{\perp} either with the image of maximal intensity or with the image of minimal intensity. This is in contrast to the underlying assumption in Ref. 5 and the common cases of reflection off opaque objects.^{16,19}

A priori information about the illumination can help eliminate this ambiguity. For example, when fishing, watching the outside dark view from a lit room, or looking into a car from a brightly illuminated environment, we may associate θ_{\perp} with the maximal intensity.

Alternatively, if we have a crude estimation of the orientation of the POI, within a $\pm 45^\circ$ uncertainty interval, the estimation of θ_{\perp} is unique. Suppose that we know that $-45^\circ < \theta_{\perp} \leq 45^\circ$, so we may simply obtain $\theta_{\perp} = 0.5 \tan^{-1}(A_s/A_c)$ and $A = \sqrt{A_c^2 + A_s^2}$. If $A_c < 0$, then $A = -\sqrt{A_c^2 + A_s^2}$. Then

$$\hat{f}_{\perp} = C + A, \quad \hat{f}_{\parallel} = C - A. \quad (27)$$

4. RECONSTRUCTION EXPERIMENT

We imaged a scene, composed of several objects, through an upright glass window. The window semireflected another scene (virtual object). The optical distance between the video camera and both scenes was ≈ 3.5 m. A linear polarizer was rotated in front of the camera between consecutive image acquisitions. For good demonstration quality, five frames were averaged for each analyzer state, but similar results were obtained also for a single frame per state. For the calculation of the polarization components, we used three images that were acquired with the analyzer directed at angles $\alpha = 0^\circ, 45^\circ, 90^\circ$ relative to the assumed vertical direction. Without a polarization analyzer in front of the camera, the image value at each point is $2C$, where C is the dc

component of the sinusoidal function of Eq. (26), calculated at that point. This unpolarized image is shown in the left panel of Fig. 8.

Since we knew that the POI is approximately horizontal, we were able to estimate θ_{\perp} within a $\pm 45^\circ$ uncertainty interval around the vertical direction, and also the amplitude A [Eq. (26)]. Using Eqs. (27), we estimated \hat{f}_{\perp} and \hat{f}_{\parallel} (which are very similar to the raw images taken at $\alpha = 0^\circ, 90^\circ$). As can be seen in the right panel of Fig. 8, the polarizer gives an initial attenuation of the reflected scene, but it remains significant, since the angle of incidence $\varphi_{\text{true}} = 27.5^\circ \pm 3^\circ$ was far from the Brewster angle.

The AOI φ to be estimated was assumed to be between 5° and 85° (close to the singular angles 0° and 90° , the estimation is unstable). For each assumed angle, the cross correlation between the images as well as their mutual information was estimated. The mutual information has a minimum at $\varphi = 25.5^\circ$, while a decorrelation was obtained at $\varphi = 27^\circ$ (Fig. 9). Both these values are within the experimental error of the physical measurement of φ_{true} . A second zero crossing of the correlation coefficient exists at 84° , in agreement with the theory, since for this φ_{true} the threshold for the appearance of this crossing (Fig. 6) is 80° .

Interestingly, the mutual information also has a minimum at 83° . Hence, for this example, the mutual information did not solve the AOI ambiguity problem. This may happen in other cases, as seen in the simulation in Fig. 7. Nevertheless, note that the simulation indicates cases (seen in Fig. 5) where mutual information did solve the ambiguity.

We applied Eqs. (6) and (7) to each point of the images \hat{f}_{\perp} and \hat{f}_{\parallel} by using the estimated AOI $\varphi = 25.5^\circ$. The results are shown in the top row of Fig. 10. The results can be compared with the separate "ground-truth" images shown in the bottom row of this figure. The reconstruction was very successful in most parts of the scenes.

It has been recognized^{16,19} that mechanical rotation of the analyzer leads to small image distortions, causing the creation of false polarization readings at image edges. This technical problem leads to the appearance of false edges in the reconstructions and can be alleviated by using special hardware, namely, liquid-crystal filters that are mechanically stationary.¹⁶ To mitigate this problem, we use an alternative approach based on image processing:²² The raw images are locally aligned to mini-



Fig. 8. Combined scene, seen without an analyzer (left); although the reflected component is smaller in \hat{f}_{\parallel} , the image is still unclear (right).

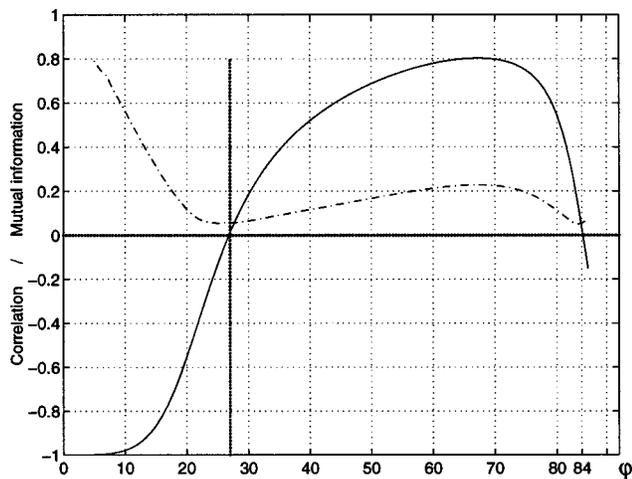


Fig. 9. At the estimated angle 27° , the estimated layers are decorrelated (solid curve); the mutual information of the estimated layers has a local minimum at 25.5° (dotted-dashed curve).

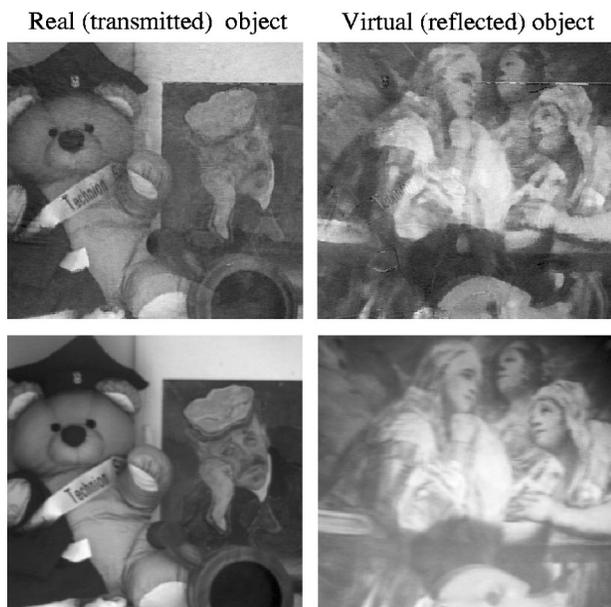


Fig. 10. Reconstructions (top), real object photographed without the interfering glass window (bottom left), and virtual object photographed by removing the objects behind the glass window (bottom right).

mize the local gradients in the resulting reconstructed images (no blurring operator is employed).

5. CONCLUSIONS

Real and virtual objects superimposed by a reflecting surface can be well separated by image processing that follows polarized imaging. This is accomplished by using a proper reflection model (e.g., taking into account the effects of internal reflection within a glass window) and by the inversion of the physical equations of image formation. Using the proper reflectance coefficients allows reconstruction away from the Brewster angle, that is, where the problem cannot be solved by optics alone. Contrary to previous methods of transparent layer separation,

the present procedure enables the unique labeling of the layers as reflected or transmitted. These results can be the basis for useful techniques in professional and amateur still photography, where polarizers are commonly used. The method automatically provides information (the inclination angle) about the transparent (invisible) surface that lies between the camera and the visible objects. This is in addition to the information on the tilt angle of the plane of incidence, for which we indicated a fundamental ambiguity.

We believe that our analysis of the effects of noise indicates the sensitivity of any method that basically relies on polarization for the scene separation, particularly by independent-components analysis of polarization-filtered images. Note that since the method that we presented is more stable at low frequencies, while methods that rely on triangulation for the layer separation are more stable at high frequencies, it may be beneficial to fuse these approaches.

ACKNOWLEDGMENTS

This research was supported by the Eshkol Fellowship of the Israeli Ministry of Science, the Ollendorff Fund, and the Tel-Aviv University Internal Research Fund. We thank Alex Bekker for his help in setting up the experiment and Bonnie Lorimer for the photograph of Mount Shuksan.

Address correspondence to Joseph Shamir at the location on the title page or by e-mail, jsh@ee.technion.ac.il

REFERENCES AND NOTES

1. W. A. Shurcliff and S. S. Ballard, *Polarized Light* (Van Nostrand, Princeton, N.J., 1964).
2. T. Darrell and E. Simoncelli, "Separation of transparent motion into layers using velocity-tuned mechanisms," Media-Lab Tech. Rep. 244 (Massachusetts Institute of Technology, Cambridge, Mass., 1993).
3. H. Fujikake, K. Takizawa, T. Aida, H. Kikuchi, T. Fujii, and M. Kawakita, "Electrically-controllable liquid crystal polarizing filter for eliminating reflected light," *Opt. Rev.* **5**, 93–98 (1998).
4. Y. Y. Schechner, N. Kiryati, and R. Basri, "Separation of transparent layers using focus," in *Proceedings of the International Conference on Computer Vision* (Narosa, Bombay, India, 1998), pp. 1061–1066.
5. N. Ohnishi, K. Kumaki, T. Yamamura, and T. Tanaka, "Separating real and virtual objects from their overlapping images," in *Proceedings of the European Conference on Computer Vision*, Vol. 1065 of Lecture Notes in Computer Science, B. Buxton and R. Cipolla, eds. (Springer, New York, 1996), Vol. II, pp. 636–646.
6. M. Oren and S. K. Nayar, "A theory of specular surface geometry," *Int. J. Comput. Vision* **24**, 105–124 (1997).
7. J. R. Bergen, P. J. Burt, R. Hingorani, and S. Peleg, "A three-frame algorithm for estimating two component image motion," *IEEE Trans. Pattern. Anal. Mach. Intell.* **14**, 886–895 (1990).
8. J. Y. A. Wang and E. H. Adelson, "Representing moving images with layers," *IEEE Trans. Image Process.* **3**, 625–638 (1994).
9. M. Shizawa, "Direct estimation of multiple disparities for transparent multiple surfaces in binocular stereo," in *Proceedings of the International Conference on Computer Vision* (IEEE Computer Society, Los Alamitos, Calif., 1993), pp. 447–454.

10. D. Weinshall, "Perception of multiple transparent planes in stereo vision," *Nature (London)* **341**, 737–739 (1989).
11. M. Shizawa, "On visual ambiguities due to transparency in motion and stereo," in *Proceedings of the European Conference on Computer Vision*, Vol. 588 of Lecture Notes in Computer Science, G. Sandini, ed. (Springer-Verlag, New York, 1992), pp. 411–419.
12. J. E. Solomon, "Polarization imaging," *Appl. Opt.* **20**, 1537–1544 (1981).
13. R. Walraven, "Polarization imagery," *Opt. Eng.* **20**, 14–18 (1981).
14. A. M. Shutov, "Videopolarimeters," *Sov. J. Opt. Technol.* **60**, 295–301 (1993).
15. L. B. Wolff, "Using polarization to separate reflection components," in *Proceedings of the Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society, Los Alamitos, Calif., 1989), pp. 363–369.
16. L. B. Wolff, "Polarization camera for computer vision with a beam splitter," *J. Opt. Soc. Am. A* **11**, 2935–2945 (1994).
17. C. S. L. Chun, D. L. Fleming, W. A. Harvey, and E. J. Torok, "Polarization-sensitive infrared sensor for target discrimination," in *Polarization: Measurement, Analysis and Remote Sensing*, D. H. Goldstein and R. A. Chipman, eds., *Proc. SPIE* **3121**, 55–62 (1997).
18. S. Lin and S. W. Lee, "Detection of specularly using stereo in color and polarization space," *Comput. Vision Image Understand.* **65**, 336–346 (1997).
19. S. K. Nayar, X. S. Fang, and T. Boult, "Separation of reflection components using color and polarization," *Int. J. Comput. Vision* **21**, 163–186 (1997).
20. M. Born and E. Wolf, *Principles of Optics*, 5th ed. (Pergamon, Oxford, UK, 1975), pp. 38–45; J. Shamir, *Optical Systems and Processes* (Society of Photo-Optical Instrumentation Engineers, Bellingham, Wash., 1999), pp. 13–17.
21. H. Farid and E. H. Adelson, "Separating reflections from images using independent components analysis," *J. Opt. Soc. Am. A* **16**, 2136–2145 (1999).
22. Y. Y. Schechner, N. Kiryati, and J. Shamir, "Separation of transparent layers by polarization analysis," in *Scandinavian Conference on Image Analysis*, B. K. Ersboll and P. Johansen, ed. (Pattern Recognition Society of Denmark, Lyngby, Denmark, 1999), Vol. I, pp. 235–242.
23. Y. Y. Schechner, J. Shamir, and N. Kiryati, "Vision through semi-reflecting media: polarization analysis," *Opt. Lett.* **24**, 1088–1090 (1999).
24. This is an example of the principle described in C. F. Bohren, "Maximum degree of polarization of the resultant of two partially polarized incoherent beams," *Appl. Opt.* **26**, 606–607 (1987): The degree of polarization of the superposition of partially polarized incoherent sources is less than (or equal to) that of the beam with the highest degree and may even be zero.
25. Mount Shuksan photograph, courtesy of Bonnie Lorimer. <http://members.aol.com/tetonbon/bonzweb.htm>.
26. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991), pp. 12–21.