

# Resolution loss without imaging blur

Tali Treibitz<sup>1,\*</sup> and Yoav Y. Schechner<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0404, USA

<sup>2</sup>Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel

\*Corresponding author: tali@cs.ucsd.edu

Received February 15, 2012; revised June 5, 2012; accepted June 5, 2012;  
posted June 6, 2012 (Doc. ID 163096); published July 11, 2012

Image recovery under noise is widely studied. However, there is little emphasis on performance as a function of object size. In this work we analyze the probability of recovery as a function of object spatial frequency. The analysis uses a physical model for the acquired signal and noise, and also accounts for potential postacquisition noise filtering. Linear-systems analysis yields an effective *cutoff frequency*, which is induced by noise, despite having no optical blur in the imaging model. This means that a low signal-to-noise ratio (SNR) in images causes resolution loss, similar to image blur. We further consider the effect on SNR of pointwise image formation models, such as added specular or indirect reflections, additive scattering, radiance attenuation in haze, and flash photography. The result is a tool that assesses the ability to recover (within a desirable success rate) an object or feature having a certain size, distance from the camera, and radiance difference from its nearby background, per attenuation coefficient of the medium. The bounds rely on the camera specifications. © 2012 Optical Society of America

OCIS codes: 110.0110, 030.4280.

## 1. INTRODUCTION

Signal-to-noise ratio (SNR) and mean square error are widely used as image quality criteria in image processing and computer vision applications. However, these criteria were shown to be only loosely related to human performance [1]. This motivates research of visibility limits with richer capabilities than SNR. This paper takes a broad look and analyzes the relation between noise and spatial resolution. We study how the ability to distinguish an object from a background depends on spatial frequencies, as well as noise.

Image blur is widely analyzed in terms of the system modulation transfer function (MTF), image defocus, and depth of field. Nevertheless, as we study in this paper, imaging noise by itself also decreases the effective resolution. In Fig. 1c, the noise is substantially stronger than in Figs. 1a and 1b. Thus, the fine details (e.g., small branch) are lost, whereas the coarse details such as the trunk are still visible. The fine details correspond to high spatial frequencies, and the coarse details correspond to lower spatial frequencies. This demonstrates that noise induces an effect resembling a low-pass filter. Digital denoising enhances the results [2,3], but even then, there is a limit, which we seek in this paper. Some works derived recovery-induced amplifications of white noise, concluding that recovery is limited when a signal matches the noise intensity [4]. However, limits based directly on white noise do not account for effective noise suppression possible if the feature of interest is large enough.

In Fig. 2, the input noise standard deviation (STD) is *independent* of the spatial frequency  $u$ : the latter linearly changes with  $x$ , while the former linearly increases with  $y$ . The large features on the left are visible even in a very low input SNR. This may be due to *implicit* smoothing by the viewer's neural system [5], which suppresses noise to reveal the signal. However, on the upper-right corner (Fig. 2), it is very difficult

(if at all possible) to reliably distinguish the signal details under the noise. There appears to be a cutoff, around the marked line, beyond which image signal details are effectively lost. Small features of the signal are visible when the input SNR is high; thus the image is not contrast-limited [6], but *noise-limited*. Moreover, Fig. 3 shows that the energy distribution along the discrete-time frequencies in the original signal is uniform. This rules out the possibility that the visibility differences (between low and high frequencies) are due to the original signal. In fact, the Fourier transform of the continuous signal  $a + s \cos[2\pi u_x x]$  has a constant amplitude [7].

Concepts of visibility under noise were introduced in [8,9]. There, the emphasis is on the minimum resolvable difference between signals. This difference depends on the signal frequency. Previous work has analyzed the effects of noise on detectability for given tasks [10,11], based on defining an ideal observer for each task and conducting psychophysical experiments [12]. Quantitative relations between noise limits and observed spatial frequencies were previously established in systems suffering from *imaging blur* [6,13]. Imaging blur can be caused by the medium (turbulence, scattering), optics (aberrations, diffraction), and detector array (pixel integration, intrapixel electronic crosstalk). In these contexts, the criteria used are termed minimum-resolvable contrast (MRC) or temperature (MRT) [6]. The blur, expressed by the system's MTF, attenuates the signal's high frequencies below a threshold set by the noise STD. In pointwise degradations, lack of imaging (optical) blur implies no such attenuation, hence apparently inducing no effective resolution limits. Such analysis cannot account for the loss of detail shown in Figs. 1 and 2 and is unsuitable to the degradations dealt with here. In [14] a cutoff stemmed solely from the typical falloff of signal energy at high frequencies. This cannot explain the effect in Fig. 2, where the input SNR is independent of the frequency  $u$ . References [15,16] ask a question similar to ours, but under

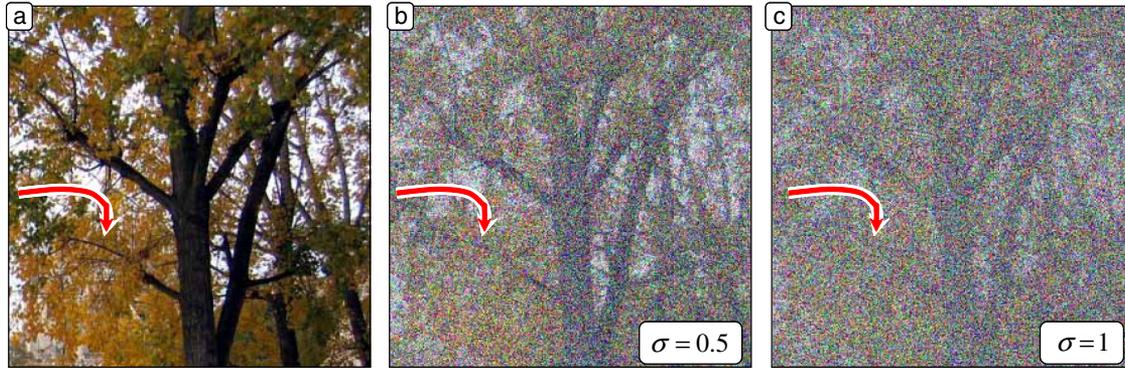


Fig. 1. (Color online) (a) An image of a tree, with a negligible amount of noise. (b), (c) Noise is added to the image with a standard deviation of  $\sigma = 0.5, 1$  (image maximum is 1), respectively. At the higher noise level (c), the fine details (small branch) are lost, whereas the coarse details (trunk) are still visible. The fine details correspond to high spatial frequencies, and the coarse details correspond to lower spatial frequencies. Thus, the noise induces an effect that has resembling consequences to a low-pass filter, even when there is *no blur* in the image formation.

different settings. Reference [15] asks what is the minimum SNR required to discriminate two points separated less than the Rayleigh limit. Reference [16] finds the minimum detectable separation between two point sources at a given SNR, for different blur kernels. However, these studies do not consider absence of optical blur.

In this paper, we establish a theoretical quantitative relation between SNR and detectable spatial frequencies. This means that low SNR reduces the resolution in an image, even with no imaging blur. SNR is often decreased by degradation effects that are essentially pointwise. These include low light conditions, specularities over a diffuse reflection [17,18], semireflections when looking through a window [17], attenuation in flash photography [3], attenuation and veiling scatter (air-light) in haze [19–22], and dirty windows [23]. We derive bounds in problems that involve no optical blur, yet considering enhancement by potential linear postacquisition filtering (implicit by the viewer or explicit by image processing). Then, we use this theory to analyze resolution and range limits of deblurring. Partial results were presented in [24].

## 2. THEORETICAL BACKGROUND

### A. The Noise

Photon noise is a fundamental quantum-mechanical effect. It cannot be overcome, regardless of the camera quality.

Accounting for this noise component [25,26], overall the noise variance in the raw image data [27] is

$$\sigma^2 \approx AI(\mathbf{x}) + B, \quad (1)$$

where  $A, B > 0$  and  $I$  is the image intensity given in Eq. (6), excluding  $n(\mathbf{x})$ . The term  $B$  encompasses the variance of the signal-independent components of the gray-level noise. As detailed in [25,27–30],  $B = \rho_{\text{read}}^2 + \rho_{\text{digit}}^2 + DT$ . Here,  $\rho_{\text{read}}$  is the amplifier readout noise STD,  $\rho_{\text{digit}}$  is the noise STD of a uniform quantizer,  $D$  is the detector dark current, and  $T$  is the exposure time. Equation (1) is consistent with a calibration we have done for a Nikon D100 and a reported calibration of a Point Grey Dragonfly [30].

The linear relation in Eq. (1) does not hold [31,32] for cameras having amplifier nonlinearities. However, our fundamental analysis is targeted at recovery that uses high quality cameras. In these cameras, Eq. (1) is typically followed. In low intensities, cameras sometime exhibit deviations from Eq. (1). This deviation is negligible in well exposed images. In addition, Eq. (1) assumes there is no electronic crosstalk or readout effects between the pixels. Part of our consequent analysis (Sections 3 and 4) is independent of the validity of Eq. (1). At a later stage, we comment on the consequences of other noise models.

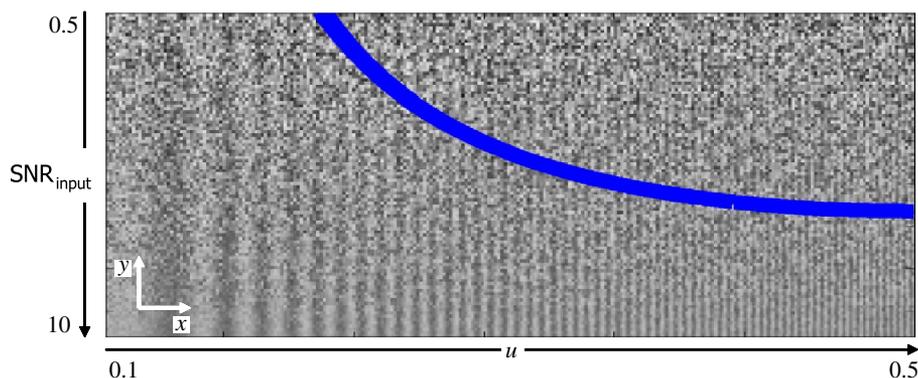


Fig. 2. (Color online) A raw noisy image. The horizontal amplitude change is given by  $a + s \cos[2\pi u_x x]$ . The spatial frequency is  $u_x \propto x$ . Here  $a$  is a bias and  $s$  is the amplitude. White additive noise increases with  $y$ . The result is then contrast stretched. At low frequencies (small  $x$ ) the pattern is visible even in a very low input SNR. Beyond the marked line, on the upper-right corner, it becomes very difficult (if at all) to reliably distinguish the signal details under the noise.

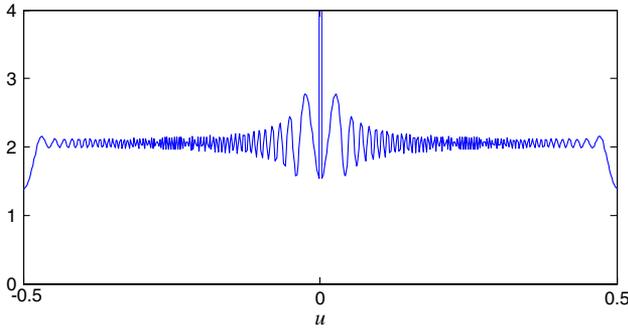


Fig. 3. (Color online) Horizontal discrete-time Fourier transform of  $a + s \cos[2\pi u_c x]$ , which underlies Fig. 2 without noise. Except for the DC component of the image, the energy is rather uniformly distributed across all frequencies. There is no significant or consistent falloff at high frequencies, specifically no  $1/|u|$  falloff. Thus, the visibility loss in Fig. 2 is due to noise, rather than the raw signal.

Light detection noise is spatially uncorrelated (white). Two dimensional detector arrays yield color images via demosaicking, which induces spatial correlation. In this paper, however, we concentrate on monochrome (gray scale) images. White noise may be suppressed by smoothing. Aggressive smoothing suppresses white noise more strongly, but leads to increased blur of objects. This tradeoff of digital blur and output noise yields to a useful conversion, which we exploit: performance limits due to *input noise can be converted to spatial resolution limits*.

### B. System Resolution

Let us observe an object of transversal length  $M$  at distance  $z$ . The camera has focal length  $f$  and pixel pitch  $p$ . Then, the image of the object stretches for  $m$  pixels, where

$$m = \frac{Mf}{zp}. \quad (2)$$

A digital image has a maximum discrete-space frequency of 0.5 [1/pixels]. In the discrete-time Fourier-transform (DTFT) domain, this frequency is reached by a single-pixel object in the image-domain. This is the ultimate resolution of the system (one image pixel, and maximum frequency of 0.5). On the other hand, if image features are effectively limited to a discrete cutoff frequency  $|u_{\text{cutoff}}| \leq 0.5$  [1/pixels], then their equivalent effective lower limit size is

$$m \approx \frac{1}{2u_{\text{cutoff}}} \quad (3)$$

pixels. Equation (3) degenerates to  $m = 1$  pixel in the upper bound of  $u_{\text{cutoff}}$ . Equation (3) enables analysis of system resolution in the Fourier domain. According to Eqs. (2) and (3), once  $u_{\text{cutoff}}$  is determined, an object at distance  $z$  is within the effective resolution of the system if its length is at least

$$M(z) = \frac{zp}{2fu_{\text{cutoff}}}. \quad (4)$$

If the CCD resolution is designed to match the lens' optical resolution, and there is no additional blurring effect, the expression in Eq. (4) degenerates to Eq. (2). This yields the geometric bound for a minimum visible object size

$$M_{\text{geometry}} = mzp/f. \quad (5)$$

### C. Pointwise Degradation and Range Dependency

Let  $I^{\text{object}}(\mathbf{x})$  be the image irradiance of an object acquired at pixel  $\mathbf{x} = (x, y)$  in ideal, undisturbed conditions. The setup, however, may impose pointwise degradation effects. Thus, the measured image is in the form

$$I(\mathbf{x}) = I^{\text{object}}(\mathbf{x})t(\mathbf{x}) + a(\mathbf{x}) + n(\mathbf{x}), \quad (6)$$

where  $t(\mathbf{x})$  and  $a(\mathbf{x})$  account for *deterministic* multiplicative and additive effects, respectively [see Fig. 4]. Signals can be blurred by a medium and the optical system. In this paper, however, we analyze resolution loss caused solely by noise. Thus, we assume a setup with no blur. In addition, Eq. (6) includes unbiased uncorrelated *random* noise  $n(\mathbf{x})$ . Note that  $a$  is *nonnegative*. As such,  $a(\mathbf{x})$  increases  $I$ , sometimes significantly. Thus, it increases the variance of the *random* noise [Eq. (1)]. This affects many imaging problems, as detailed below.

We seek recovery of  $I^{\text{object}}(\mathbf{x})$ . This model fits a wide range of problems:

- In analysis of reflections,  $I^{\text{object}}$  is the diffuse component and  $a$  is the specular component (while  $t = 1$ ) [18].
- A similar distinction exists in the mixture of direct and indirect illumination components [33,34].
- In semireflections,  $I^{\text{object}}$  is the scene behind a window of transmittance  $t$ , and  $a$  is the semireflected layer [17].
- When imaging through a dirty window,  $t$  is the spatially varying transmittance of the dirty window, while  $a$  is the spatially varying scatter by the dirt [23].
  - A bright light source near the field of view can contribute an additive component  $a(\mathbf{x})$  of flare [35].
  - Fixed pattern noise is a deterministic effect [25,36] of pointwise gain and bias variations, which is modeled by  $t(\mathbf{x})$  and  $a(\mathbf{x})$  in Eq. (6).

The degradation effects may be distance-dependent. In haze,  $t$  is the transmittance of the atmosphere [19–22, 37,38]. Its dependency on the object distance  $z(\mathbf{x})$  is given by

$$t(\mathbf{x}) = e^{-\beta z(\mathbf{x})} \quad (7)$$

in a single scattering model. Here,  $\beta \in (0, \infty)$  is the atmospheric attenuation coefficient. The additive component  $a$  here is the *airlight* [21,39], given by

$$a(\mathbf{x}) = a_{\infty}[1 - t(\mathbf{x})], \quad (8)$$

where  $a_{\infty}$  is the value of airlight at a nonoccluded horizon. Airlight increases with  $z$  and dominates the acquired image irradiance at long range (see Fig. 5). Note that contrast is mainly degraded by airlight, rather than blur [4]. There are

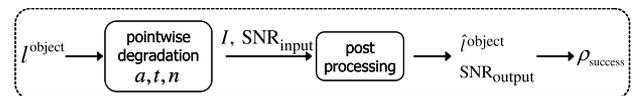


Fig. 4. Image formation and processing flow. The object radiance  $I^{\text{object}}$  is degraded by pointwise effects. The measured image  $I$  is noisy, characterized by  $\text{SNR}_{\text{input}}$ . After postprocessing, the recovery output is characterized by  $\text{SNR}_{\text{input}}$ , leading to a success rate of  $\rho_{\text{success}}$ .

other distance-dependent pointwise models, including synthetic aperture lighting [40], which may include scatter, and flash photography [41] (falloff of object irradiance).

In all the above cases (reflections, flare etc.),  $a(\mathbf{x})$  has two degrading consequences. First, this deterministic component degrades the contrast, and may confuse object appearance (according to [42], human perception is affected by low contrast more than by blur). However, such deterministic disturbances are rather easy to invert by digital subtraction of an estimate of  $a$ , as done by all the above mentioned studies. A second degradation consequence is much more difficult to counter:  $a$  increases the *random* noise, as detailed next.

### 3. NOISE-LIMITED RESOLUTION

Without noise, even small intensity changes over a background can be stretched to reveal objects and details. Figures 5a and 5c demonstrate a piecewise contrast stretch on a synthetic utopian noise-free hazy image. Even in parts that appear blank in Fig. 5a, visibility is retrieved in Fig. 5c. These parts correspond to more distant scene regions, where the accumulated airlight is higher. Yet, actions such as contrast stretching affect both the signal and the noise. Noise following the model in Eq. (1) is introduced in Fig. 5b. Now, objects are lost in the parts corresponding to distant regions, despite regional contrast stretch in Fig. 5d. Noise reduction operations affect the signal amplitude. The effect of this operation varies as a function of the signal's spatial frequency.

Figure 6 shows a scene extracted from the Weather and Illumination Database (WILD) [43]. On a clear day, visibility exists at both  $z = 2$  and 3.5 km for building outlines (large  $M$ ) and windows (small  $M$ ). In mist, there is a loss of spatial detail, despite regional contrast stretch: at  $z = 2$  km, windows (small  $M$ ) are unrecoverable, as contrast stretch simply amplifies the overwhelming noise, yet the building outline (large  $M$ ) is seen well. This detail loss is greatly exacerbated at somewhat longer distance: at  $z = 3.5$  km, even the building outline is hopelessly obscured by noise. Note that the images in the WILD database are color images, demosaicked from a Bayer pattern. To avoid having spatially correlated noise in this example, we used only every second pixel of the green channel.

We note that empirical work had been done about human performance under noise in four specific vision tasks: object detection, orientation, recognition, and identification. This work culminated in the *Johnson tables* and National Image Interpretability Rating Scales (*NIIRS image ratings* [6,36,44]). These tasks were associated [6,44] with the number of resolvable lines in an object. For example, to *detect* a tank

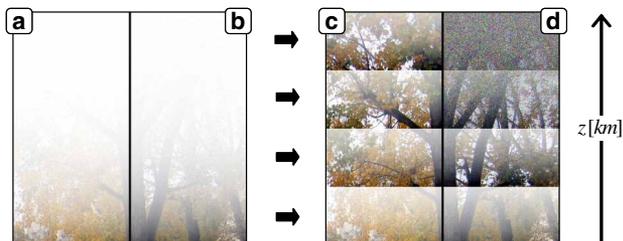


Fig. 5. (Color online) (a) A noise-free hazy image, simulated by  $\beta = 0.2 \text{ km}^{-1}$  and a linearly changing  $z \in [1, 30]$  km. Airlight acts as local bias. (b) A slightly noisy version. (c) Regional contrast stretching of (a) reveals the objects and details. (d) Regional contrast stretching of (b) does not recover *small* details at large  $z$ , over noise.

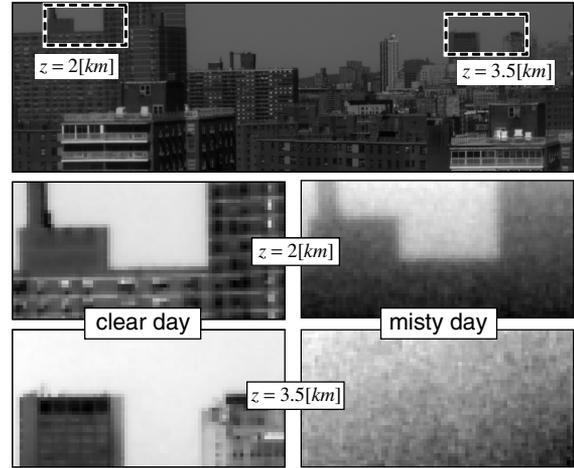


Fig. 6. [Top] Clear day scene. [Middle] Small details seen on a clear day at  $z \approx 2$  km but lost in mist. [Bottom] At  $z \approx 3.5$  km, visibility in mist quickly worsens: even large buildings are lost. Images taken from the WILD database [43]

with 50% success, the tank has to occupy 0.75 line pairs in the field of view [6]. For the tank to be *recognized* as a tank, and not, say, as a truck, its tower has to be visible. The tower is smaller, thus corresponding to higher spatial frequencies than the convex hull of the tank itself. Therefore, the recognition task requires a higher spatial resolution. The number of visible line pairs in an object is related to  $u_{\text{cutoff}}$ . Consider an object of size  $m$  pixels. Then, the number of visible line pairs in this object is

$$\nu = m \cdot u_{\text{cutoff}}. \quad (9)$$

Note that  $\nu = m/2$  in the special case of  $u_{\text{cutoff}} = 0.5$ , i.e., when the resolution is limited only by pixel size (geometry). Following Eq. (9), frequency-domain analysis (which determines cutoff frequencies) may generalize some aspects of empirical criteria such as those in [6,36,44]. More importantly, however, our linear-systems analysis is general, and thus applies beyond human, to computer vision systems, in contrast to the Johnson tables and NIIRS ratings.

#### A. How Much Would Filtering Help?

There are numerous denoising methods. This paper does *not* introduce any new method. As a basic benchmark, we focus the analysis on linear filtering (as in [10]). The main reason is that linear-systems are the basis for frequency-domain analysis, and thus the notion of cutoff frequencies. Moreover, this enables analytic closed-form formulas for bounds, which are intuitive. Extension to nonlinear operations is discussed in Section 6. We start the analysis by considering suppressing white noise by a flat averaging filter. In principle, better results can be potentially obtained by more sophisticated digital low-pass filters, which may be designed by an array of engineering methods. The flat filter we use leads to closed-form expressions, which are useful both for obtaining insights and for a baseline assessment. In Appendix B, we further discuss the effect of a Gaussian filter.

Let  $\mathbf{u} = (u, v)$  be a spatial frequency in units of  $[1/\text{pixels}]$ , where  $u, v \in [-0.5, 0.5]$ . Consider an image signal (similar to [45])

$$s(\mathbf{x}) = \frac{1}{2}S(u) \cos 2\pi u x. \quad (10)$$

Here  $|S(u)|$  is the difference between the maximum and minimum values in this signal.

The image is corrupted by white additive noise, whose input STD is  $\sigma$  [Eq. (1)]. Note that the image in Eq. (10) is composed of a single discrete spatial frequency  $\mathbf{u} = (u, 0)$  and its conjugate  $(-u, 0)$ . Following one of the definitions in [46], the SNR of the raw image (see Fig. 4) is defined as

$$\text{SNR}_{\text{input}}(u) = \frac{|S(u)|}{\sigma}. \quad (11)$$

We study the *signal* at a specific spatial frequency. This indicates the potential for recovering object-features that correspond to a specific size [Eq. (3)]. Eventually, we seek an effective cutoff frequency (resolution) of the overall system, accounting for the pointwise degradations, noise, and the potential smoothing induced by linear postacquisition processing. In contrast to the signal, *noise is not treated at different frequencies*. The reason is that the camera output  $I$  (which is the input for postprocessing) has white noise, irrespective of the feature size, as illustrated in Fig. 2, and seen in most natural images (e.g., Figs. 1 and 5).

To suppress white noise, consider a flat averaging filter  $h_W$ , having a support of  $W \times W$  pixels. As we prove in Appendix A, applying  $h_W$  on  $s(\mathbf{x})$  amplifies the SNR by

$$C(u) \equiv \frac{\text{SNR}_{\text{output}}(u)}{\text{SNR}_{\text{input}}(u)} = \frac{\sin(\pi W u)}{\sin(\pi u)}, \quad (12)$$

where  $\text{SNR}_{\text{output}}(u)$  is the SNR of the processed image (Fig. 4). Figure 7 illustrates  $C(u)$  for a range of  $W$ . We seek the window size  $W_{\text{max}}$  that maximizes the improvement of SNR, similar in spirit to a linear matched filter that was used in [8,9]. The window size  $W_{\text{max}}$  is obtained when  $\sin(\pi W u) = 1$ , i.e., when

$$W_{\text{max}} = \frac{1}{u}(2\kappa + 0.5), \quad \kappa = 0, 1, 2, \dots \quad (13)$$

The maximal SNR amplification  $C_{\text{max}}(u)$  that can be achieved by spatial averaging is then

$$C_{\text{max}}(u) \equiv C_{W_{\text{max}}}(u) = 1/\sin(\pi u), \quad (14)$$

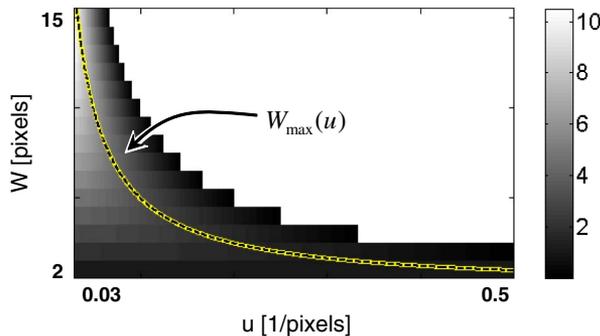


Fig. 7. (Color online) SNR improvement  $C$ , as a function of  $u$  and  $W$ . The curve of  $W_{\text{max}}(u)$  is plotted on top. As  $u$  increases, windows are limited to smaller sizes. This limits the ability to suppress noise while maintaining the signal.

for all  $\kappa$  values that solve Eq. (13). Thus, from now on we set  $W_{\text{max}} \equiv 1/(2u)$  (plotted in Fig. 7). The value of  $C_{\text{max}}(u)$  is plotted in Fig. 8. In the Rose model [47], the SNR is multiplied by the linear dimension of the object to allow for averaging. In small frequencies, this can be considered equivalent to the improvement in Eq. (14). However, for higher frequencies, our model predicts a faster deterioration in SNR.

Figure 9 demonstrates the use of a window of length  $W_{\text{max}}$  on the image of Fig. 2. Here,  $W_{\text{max}}$  adapts spatially to  $u(x)$ . Across all frequencies, the signal pattern is more visible in Fig. 9 than in Fig. 2. Nevertheless, as  $u$  increases, the pattern can be observed reliably only at higher values of  $\text{SNR}_{\text{input}}$  (at smaller  $y$ ), and is effectively lost at low  $\text{SNR}_{\text{input}}$ . This is consistent with Eq. (14): a smaller SNR improvement  $C(u)$  can be achieved at high  $u$ , thus requiring a higher input SNR. How much SNR is actually needed at the output of processing? As discussed in Subsection 4.B, this question is related to the desired *success rate* of detecting a feature. For the moment, let us denote the minimum output SNR that is required as  $\text{SNR}_{\text{output}}^{\text{min}}$ .

## B. The Cutoff Frequency

If the raw image has a high SNR, there is no need to smooth the image: objects at all sizes are reasonably seen despite the noise. Then, smoothing may be counterproductive, due to the consequent detail loss. In this case, without filtering

$$\text{SNR}_{\text{input}}(u) \geq \text{SNR}_{\text{input}}^{\text{min}}. \quad (15)$$

In moderate SNR, gentle filtering with  $W < W_{\text{max}}$  may suffice to bring the output  $\text{SNR}_{\text{output}}$  to the acceptable level  $\text{SNR}_{\text{output}}^{\text{min}}$ . However, at the *limit* of recovery, the signal in  $u$  is so low relative to the noise that  $W_{\text{max}}(u)$  should be used. There is no point in trying to use a kernel wider than  $W_{\text{max}}$ , since it would yield a lower  $C(u)$  than a  $W_{\text{max}}$ -sized kernel [see Eq. (13)], while blurring excessively. Hence, we hope to have

$$C_{\text{max}}(u)\text{SNR}_{\text{input}}(u) \geq \text{SNR}_{\text{output}}^{\text{min}}. \quad (16)$$

From Eqs. (14) and (16),

$$\text{SNR}_{\text{input}} \geq \sin(\pi u)\text{SNR}_{\text{output}}^{\text{min}}. \quad (17)$$

Equation (17) is an *important performance bound*. It dictates a *minimum input SNR*, in order to recover a signal component having spatial frequency  $u$ , at the desired success rate. If  $\text{SNR}_{\text{input}} < \sin(\pi u)\text{SNR}_{\text{output}}^{\text{min}}$ , the SNR in the original image is

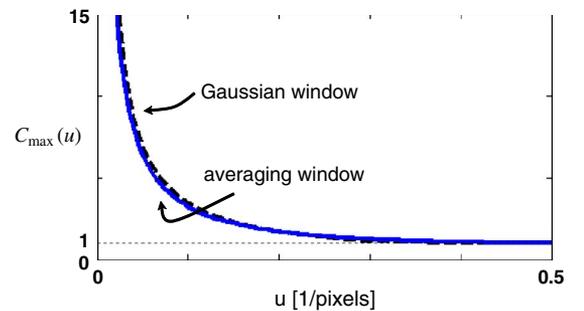


Fig. 8. (Color online) Maximal possible SNR improvement  $C_{\text{max}}$ , as a function of  $u$ . As  $u$  increases,  $C_{\text{max}}$  decreases. The derivation for the Gaussian window appears in Appendix B.

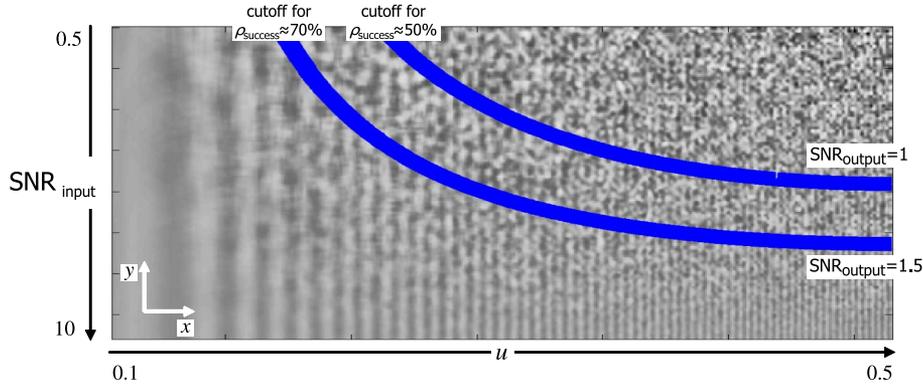


Fig. 9. (Color online) Filtering the image in Fig. 2 with window size  $W_{\max}(u)$  improves visibility. Cutoff lines are derived using Eqs. (17) and (18). Below the line of  $\rho_{\text{success}} = 70\%$ , the pattern is clearly seen. Above the upper line,  $\rho_{\text{success}} < 50\%$ , and noise dominates. The definition of  $\rho_{\text{success}}$  appears in Subsection 4.B.

too low and cannot be increased to the desired level  $\text{SNR}_{\text{output}}^{\min}$ , no matter what filter size we use. This is a recovery limit, posed by noise.

For given values of  $\text{SNR}_{\text{input}}$  and  $\text{SNR}_{\text{output}}^{\min}$ , the lower bound in Eq. (17) is met at a *cutoff* frequency

$$u_{\text{cutoff}} = \frac{1}{\pi} \left[ \arcsin \left( \frac{\text{SNR}_{\text{input}}}{\text{SNR}_{\text{output}}^{\min}} \right) \right]. \quad (18)$$

The cutoff in Eq. (18) is plotted in Fig. 10, for a few values of  $\text{SNR}_{\text{output}}^{\min}$ . Note in this plot that when Eq. (15) is satisfied, there is no cutoff.

Suppose for a moment that  $\text{SNR}_{\text{input}}$  is independent of  $u$ . The sine function is monotonically increasing in the considered interval  $[0, \pi/2]$ . Thus, if  $u_{\text{cutoff}} < 0.5$ , then considering Eq. (17),

$$\forall u > u_{\text{cutoff}}, \quad \text{SNR}_{\text{input}} < \sin(\pi u) \text{SNR}_{\text{output}}^{\min}. \quad (19)$$

Statement (19) is strengthened by the fact that in natural images,  $\text{SNR}_{\text{input}}(u)$  tends to fall off with  $u$  (most of the signal's energy is in low frequencies). Therefore,  $\forall u > u_{\text{cutoff}}$  the minimum desired image quality cannot be achieved.

The cutoff frequency sets the *resolution limit* to images with additive white noise, since  $u_{\text{cutoff}}$  can be used to determine the object size of the least-resolved objects, using Eq. (4). This spatial resolution limit exists, although the degradation model in the raw data  $I(\mathbf{x})$  is pointwise, and no blur affects the raw image formation.

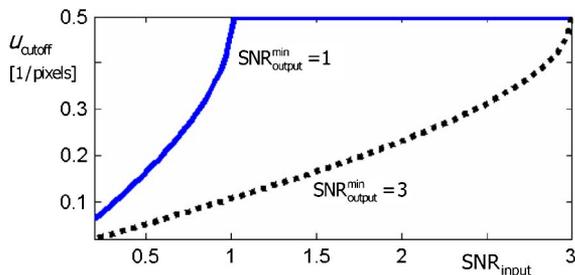


Fig. 10. (Color online) Cutoff frequency as a function of  $\text{SNR}_{\text{input}}$ , for different values of  $\text{SNR}_{\text{output}}^{\min}$ . A better input SNR yields a better output resolution. When  $\text{SNR}_{\text{input}} < \text{SNR}_{\text{output}}^{\min}$ , the image starts to lose reliability.

## 4. RECOVERY WITHIN A BOUND

### A. Defining the Signal

Typically, there is an interest to distinguish objects, e.g., cars, over a nearby background, such as a field, or distinguish finer details, e.g., digits over a license plate. The ability to distinguish an object/detail depends on its spatial size, the radiance difference relative to the background, and the amount of noise. Let the image components  $l^{\text{object}}(u)$  and  $l^{\text{back}}(u)$  correspond to the object and background in ideal, undisturbed conditions. Because of Eqs. (6) and (7), the difference in their image values is  $[l^{\text{object}}(u) - l^{\text{back}}(u)]t$ . These components depend on  $u$ , since an object can be described in different scales: rough, large scale structures correspond to a low  $u$ , while fine-scale details correspond to a high  $u$ .

We use the goal of object-versus-background distinction in order to define the *signal* of interest. In consistency with Subsection 3.A and Appendices A and B,  $S(u)$  is the difference between the two extrema of the signal. Thus, here, define the signal as

$$S = [l^{\text{object}}(u) - l^{\text{back}}(u)]t, \quad (20)$$

in the problem of differentiating an object over a background. What about the image component  $a(\mathbf{x})$ ? Recall from Subsection 2.C that  $a(\mathbf{x})$  is a deterministic component (though it generally varies spatially) and can thus be subtracted from  $I$ , either by contrast stretch or by estimation [17–19,21,22,33]. By itself, this removable nonrandom component does not decrease the object-versus-background difference. However,  $a$  increases the photon noise, thus affecting the image SNR (see Subsection 2.A).

### B. Success in a Confidence Interval

Subsection 3.A showed that postacquisition filtering may enhance the SNR to  $\text{SNR}_{\text{output}}$ . Now, the question is which value of  $\text{SNR}_{\text{output}}$  is sufficient for succeeding in seeing an object feature? Look, for example, at Fig. 11, which shows a clean and a very noisy version of an image of a dark square over a bright background. In the noisy image, the coarse detail (the dark object) is visible. What happens to the fine details? The left edge is still somewhat visible, while the right one is almost completely lost. These different outcomes of noise occur despite the uniformity of the noise variance. This demonstrates that SNR does not give a deterministic answer

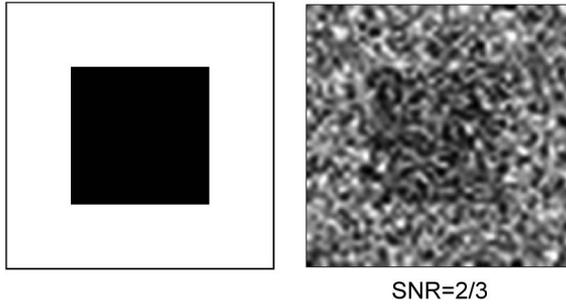


Fig. 11. Randomness of noise imposes randomness in the success of recovering a single object. [Left] A clear square. [Right] Under the same noise level, the bottom and left edges are visible, while the right one is lost.

regarding detection, as the noise is a random process. The higher the  $SNR_{output}(u)$  is, the more *confidence* there is in the recovery of  $u$ . Over a large ensemble, the success rate increases with  $SNR_{output}(u)$ . The randomness of noise imposes randomness in the success of recovering a single object whose characteristic frequency is  $u$ . In this section we discuss this relation.

Let us look at an edge; one pixel is over an object, and the second pixel is of the background (Fig. 12[Left]). However, in reality, both pixels are noisy. Is the edge from the clear image still visible in the noisy image? In the following, assume that the edge is maintained if the intensity difference between the signal and background pixels maintains its sign, regardless of its magnitude (Fig. 12 [Center]). This analysis is targeted at detection bounds at lowest visibility conditions. In this limit, even a difference of one gray level above noise is considered an edge. Cases where the edge reverses its sign are shown in Fig. 12 [Right].

Ignoring the magnitude is reasonable in digital applications, where each intensity difference can be stretched to emphasize the edges. For human vision, however, this assumption may not hold, as some intensity differences are not discerned by the eye. According to Weber’s law [48], the just noticeable difference (JND) is linearly related to the background intensity. At high spatial frequencies, there is a transition from Weber’s law to the De-Vries Rose law, where the JND is proportional to the square root of the background intensity [49]. For analysis of image quality perception, see [48].

The signal  $S$  is defined in Eq. (20). After filtering, the signal becomes  $S_{output} = S * h_W$ . Denote the noise values in the object and background pixels as  $N_{output}^{object}$  and  $N_{output}^{back}$ , respectively. The edge maintains its sign if

$$N_{output}^{back} - N_{output}^{object} < S_{output}. \tag{21}$$

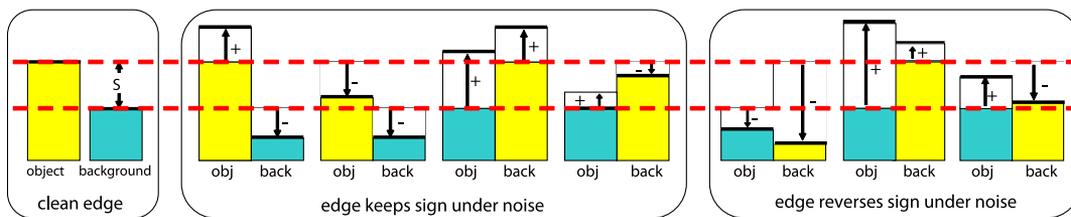


Fig. 12. (Color online) An edge under noise. [Left] Original. [Center] Noise is added to both pixels; however, the edge still keeps its sign. [Right] The edge reverses its sign under noise.

The signal difference is related to  $SNR_{output}^{min}$  similarly to Eq. (11)

$$S_{output} = SNR_{output} \sigma_{output}. \tag{22}$$

where  $\sigma_{output}$  is the noise STD after filtering. Define  $\rho_{maintain}$  as the probability that Eq. (21) holds. From Eqs. (21) and (22),

$$\rho_{sign}^{keep} = \mathcal{P}(N_{output}^{back} - N_{output}^{object} < SNR_{output} \sigma_{output}). \tag{23}$$

Equation (23) is the cumulative distribution function (CDF) of the random variable  $(N_{output}^{back} - N_{output}^{object})$  at the value  $SNR_{output} \sigma_{output}$ . What is the distribution of  $(N_{output}^{back} - N_{output}^{object})$ ? The noise is approximated to be normally distributed with zero mean and variance  $2\sigma_{output}^2$ . (The number of Poisson-distributed photons [creating photon noise] per pixel is typically large. Then, the Poisson distribution is approximated well by a normal distribution.) The CDF  $\eta \sim \mathcal{N}(0, 2\sigma_{output}^2)$  at a value  $\chi$  is

$$\mathcal{P}(\eta < \chi) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{\chi}{2\sigma_{output}} \right) \right]. \tag{24}$$

Combining Eqs. (23) and (24),

$$\rho_{sign}^{keep} = \frac{1}{2} [1 + \operatorname{erf}(SNR_{output}/2)]. \tag{25}$$

The noise can destroy true edges and can also cause the opposite: edges are aliased in locations that had been smooth in the original scene. The probability that an edge in the noisy image is false is

$$\rho_{false} = 1 - \rho_{sign}^{keep}. \tag{26}$$

The total number of edges visible in the noisy image is the number of the true edges that maintained their sign plus the number of the false edges that were created by the noise. Thus, based on Eqs. (25) and (26), the probability that an edge that is visible in the image is a true edge is

$$\rho_{success} = \rho_{sign}^{keep} - \rho_{false} = 2\rho_{sign}^{keep} - 1 = \operatorname{erf}(SNR_{output}/2). \tag{27}$$

Thus, as depicted in Fig. 4, the output SNR determines the success rate in detection. Figure 13 plots Eq. (27).

Equation (27) yields an expression for the minimum required output SNR as a function of the desired success rate

$$SNR_{output}^{min} = 2 \operatorname{erf}^{-1}(\rho_{success}). \tag{28}$$

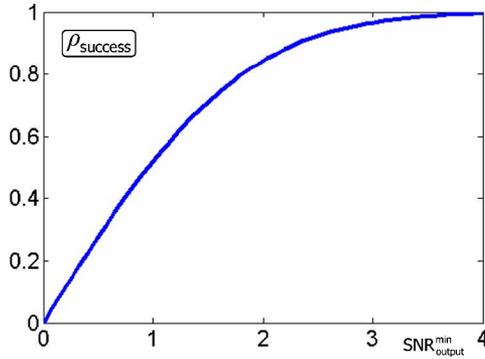


Fig. 13. (Color online) Probability for success as a function of the output SNR (Eq. 27).

Combining Eqs. (18) and (28) yields the cutoff frequency as a function of the desired success rate:

$$u_{\text{cutoff}} = \frac{1}{\pi} \arcsin \left[ \frac{\text{SNR}_{\text{input}}}{2 \text{erf}^{-1}(\rho_{\text{success}})} \right]. \quad (29)$$

The cutoff in Eq. (29) is plotted in Fig. 9, for a few values of  $\rho_{\text{success}}$ . Below the cutoff line for  $\rho_{\text{success}} \approx 70\%$ , the pattern is clearly seen. Above the upper cutoff line, where  $\rho_{\text{success}} \approx 50\%$ , noise dominates.

Equation (29) can be viewed as an equation relating SNR to actual performance under noise:

$$\rho_{\text{success}} = \text{erf} \left[ \frac{\text{SNR}_{\text{input}}}{2 \sin(\pi u)} \right]. \quad (30)$$

Equation (30) differs from other image quality criteria by depending explicitly on the spatial frequency.

### C. Comparison to Image Quality Methods

Recent studies [1,50] have shown that SNR is only loosely related to human performance and perceived image quality, confounded by contrast, brightness, etc. The structural similarity (SSIM) index [1] measures similarity between two images. When a degraded image is compared to a nondegraded version, SSIM can be viewed as a quality measure of the degraded image. The SSIM is composed of three terms comparing luminance, contrast, and structure across small windows in the images. It ranges from 0 to 1, where 1 corresponds to identical images. Figure 14 depicts an experiment comparing Eq. (30) to SSIM. We generated images, each containing a distinct spatial frequency. Then, to each image, white noise was added,

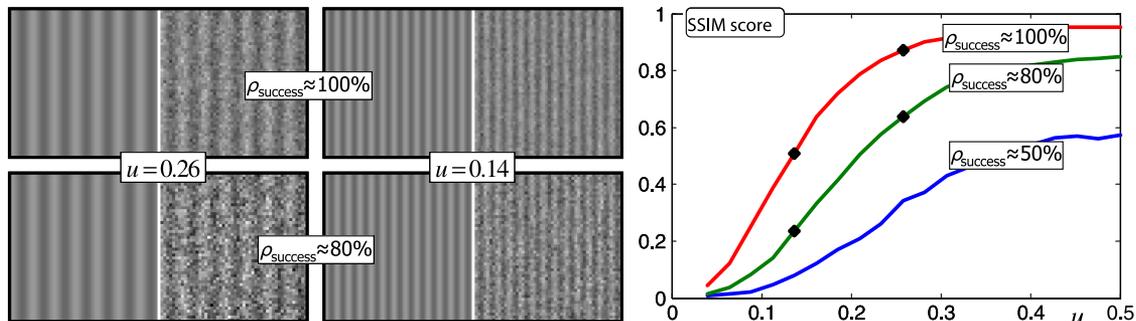


Fig. 14. (Color online) [Left] Examples of four image pairs from our simulation. The noise STD that was added to each frequency  $u$  is calculated to achieve a fixed, constant  $\rho_{\text{success}}$  [Eq. (31)]. Their SSIMs are plotted on the graph [right] using black dots. [Right] SSIM as a function of  $u$  for three values of  $\rho_{\text{success}}$ .

such that  $\rho_{\text{success}}$  is constant across frequencies:

$$\sigma = |S|/[2 \sin(\pi u) \text{erf}^{-1}(\rho_{\text{success}})], \quad (31)$$

i.e., the noise STD  $\sigma$  decreases with  $u$ . We calculated the SSIM score between each clear image and its corresponding noisy version. The SSIM score is monotonic with  $\rho_{\text{success}}$ . However, for constant  $\rho_{\text{success}}$ , SSIM varies with  $u$ , which is monotonic in  $\sigma$ . The SSIM score does not take into account the visibility differences between the different spatial frequencies.

## 5. LIMITATIONS IN HAZE

In this section, we apply the analysis of Section 3 to a specific type of degradation: haze in images. We seek bounds that do not depend on the algorithm, e.g., whether airlight or distances are derived by polarization or an auxiliary map. In other words, optimal operations can achieve a bound, but a suboptimal algorithm or inaccurate model parameters would achieve worse performance.

### A. SNR in Raw Hazy Images

The noise variance  $\sigma^2$  is derived in Eq. (1), simply based on the total image intensity  $I$ . Using the midway intensity between  $l^{\text{object}}$  and  $l^{\text{back}}$ , the mean of the noise variance over the object and its immediate background is

$$\sigma^2(b) = B + A \left\{ \frac{1}{2} (l^{\text{object}} + l^{\text{back}}) t(b) + a_{\infty} [1 - t(b)] \right\}, \quad (32)$$

where  $b = \beta z$ , based on Eqs. (1) and (6)–(8). In this way, the random noise induced by airlight is incorporated. The signal here is defined in Eq. (20). The value of  $\text{SNR}_{\text{input}}$  of a raw hazy image is obtained by using Eq. (20) and  $\sigma$  from Eq. (32) in Eq. (11). For cameras exhibiting nonlinear noise models, the variance in Eq. (32) should be calculated according to the specific model.

We now introduce some variables that ease the assessment. The saturation gray level is

$$V = 2^B - 1, \quad (33)$$

where  $B$  is the number of bits per pixel. Define the object saliency (without disturbances) as

$$\tilde{E} = |l^{\text{object}} - l^{\text{back}}|/V. \quad (34)$$

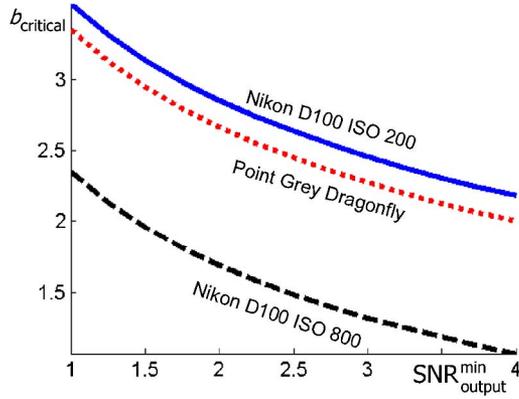


Fig. 15. (Color online) Values of  $b_{\text{critical}}$ , [in Eq. (36)] as a function of  $\text{SNR}_{\text{output}}^{\text{min}}$ , for different cameras. Here,  $\tilde{E} = 25\%$ .

For example,  $\tilde{E} = 1$  for a bright white object (which exploits the full dynamic range) on a black background, or vice versa. Analogously,  $\tilde{l} = (l^{\text{object}} + l^{\text{back}})/(2V)$  is the mid-radiance between the object and background, normalized by the camera's dynamic range. The same normalization is applied to the horizon airlight:  $\tilde{a}_{\infty} = a_{\infty}/V$ . Using these definitions in Eqs. (1) and (6)–(8), Eq. (11) yields the input SNR as a function of  $b$

$$\text{SNR}_{\text{input}}(b) = \frac{e^{-b}\tilde{E}V}{\sqrt{B + AV[\tilde{l}e^{-b} + \tilde{a}_{\infty}(1 - e^{-b})]}}. \quad (35)$$

In Eqs. (1) and (35), the parameters  $A$ ,  $B$ , and  $V$  are scene independent. They depend on the specific camera model and operation mode. They can easily be calibrated or extracted from the camera's specifications [30]. In the following, we calibrated the respective parameters in a couple of ISO settings in a Nikon D100, where  $V = 2^{14} - 1$ . For ISO 200, we got  $[A, B] = [0.9, 210]$ . For ISO 800, we got  $[A, B] = [10, 2116]$ . We also plot results for a Point Grey Dragonfly camera, where  $V = 2^8 - 1$  and  $[A, B] = [0.02, 0.16]$  (data taken from [30]).

The parameters  $\tilde{E}$ ,  $\tilde{l}$ , and  $\tilde{a}_{\infty}$  are scene dependent. For *practical* assessments, assume that in properly exposed images,  $\tilde{l} \approx 0.5$  (nearby objects are at the middle of the camera's dynamic range) and  $\tilde{a}_{\infty} \approx 1$ . In practice, variations around these values of  $\tilde{l}$  and  $\tilde{a}$  are not critical. Nevertheless,  $\tilde{E}$  is important, since  $\text{SNR}_{\text{input}}$  is proportional to it. The value of  $\tilde{E}$  is the main input by a user for assessing object visibility. For instance, if a nearby car over a street occupies  $\tilde{E} = 20\%$  of the dynamic range, this has a prime effect on the car's distinction in the presence of attenuation (at a distance) and noise.

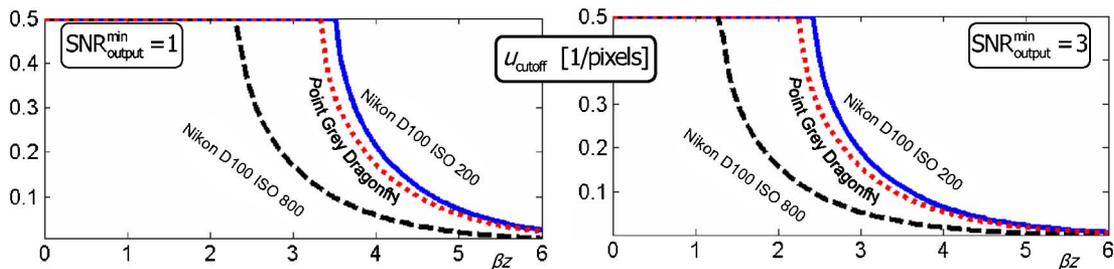


Fig. 16. (Color online) Values of  $u_{\text{cutoff}}(\beta z)$  from Eq. (37) as a function of  $\beta z$ , for different cameras and values of  $\text{SNR}_{\text{output}}^{\text{min}}$ . Here  $\tilde{E} = 25\%$ . When  $\beta z$  reaches the value of  $b_{\text{critical}}$ , the value  $\text{SNR}_{\text{input}}$  drops below the value  $\text{SNR}_{\text{output}}^{\text{min}}$ . Then, there is a frequency cutoff.

We measured typical values for  $\tilde{E}$  in well exposed outdoor images we took. Results ranged between 5–50%. For example, houses in the background of trees had  $\tilde{E} \approx 10\%$ . This is consistent with daylight images in the Columbia WILD database [43]. Thus, as an example, we set  $\tilde{E} = 25\%$  in the following plots.

There is a critical optical depth  $b_{\text{critical}}$ , up to which no cutoff frequency exists. This optical depth can be found by solving a quadratic equation based on Eq. (35),

$$\text{SNR}_{\text{input}}(b_{\text{critical}}) = \text{SNR}_{\text{output}}^{\text{min}}. \quad (36)$$

At  $b < b_{\text{critical}}$ , Eq. (17) is satisfied  $\forall u$ . At  $b > b_{\text{critical}}$ , some frequencies start to become unrecoverable and thus some objects become lost. Figure 15 shows values of  $b_{\text{critical}}$  as a function of  $\text{SNR}_{\text{output}}^{\text{min}}$ , for different cameras.

## B. Resolution Cutoff in Haze

We want to assess the limit that can be achieved, even if denoising by an optimal window size is employed implicitly or explicitly. Plugging  $\text{SNR}_{\text{input}}(b)$  from Eq. (35) into Eq. (29) yields

$$u_{\text{cutoff}}(\beta z) = \frac{1}{\pi} \arcsin \left[ \frac{\text{SNR}_{\text{input}}(\beta z)}{2 \text{erf}^{-1}(\rho_{\text{success}})} \right]. \quad (37)$$

This cutoff is plotted in Fig. 16, using  $\tilde{E} = 25\%$ , calibration data of Point Grey Dragonfly and a Nikon D100 at two ISO settings, in Eq. (35). The value of  $u_{\text{cutoff}}$  decreases with  $\beta z$ . Moreover, the two ISO settings of the same camera yield a significant difference in the visible distances.

Using  $u_{\text{cutoff}}(\beta z)$  from Eq. (37) in Eq. (4) yields the least-resolved object length in haze

$$M_{\text{haze}}(\beta, z) = \frac{\pi z p}{2f \arcsin[\text{SNR}_{\text{input}}(\beta z)/2 \text{erf}^{-1}(\rho_{\text{success}})]}. \quad (38)$$

Equations (35) and (38) depend on the (scene independent) system parameters  $\{p, f, A, B, V\}$ . They also depend on the scene's  $z$ ,  $\beta$ , and  $\tilde{E}$ . Figure 17 plots  $M_{\text{haze}}$  for different parameter values. In both cameras, the ratio  $p/f$  was the same. This corresponds to  $f = 200$  mm,  $p = 7.8 \mu\text{m}$  in the Nikon D100 and  $f = 120$  mm,  $p = 4.7 \mu\text{m}$  in the Dragonfly. The value of  $M_{\text{haze}}$  is plotted for  $\rho_{\text{success}} = 50\%$ . Note that in color cameras that have a single sensor, the effective pixel pitch is larger than the physical one, due to subsampling by the sensor's Bayer mosaic.

Recall the geometric bound for a minimum visible object size  $M_{\text{geometry}}$  [Eq. (5)]. Note that  $M_{\text{haze}} = M_{\text{geometry}}$  if  $\beta z < b_{\text{critical}}$ . As  $\beta z$  increases, the haze increases and thus

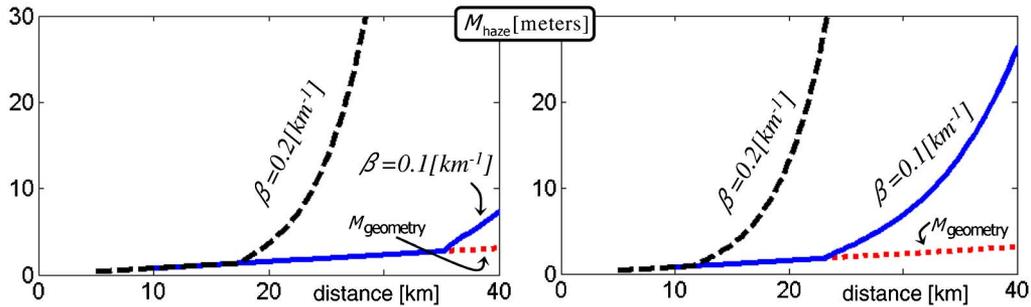


Fig. 17. (Color online) Values of  $M_{\text{haze}}$  from Eq. (38) as a function of  $z$ , for different cameras,  $\text{SNR}_{\text{input}}^{\text{min}} = 1$ ,  $\bar{E} = 25\%$ . [Left] Nikon D100, ISO 200, [Right] Nikon D100, ISO 800. Increasing  $z$  decreases  $\text{SNR}_{\text{input}}$ . When  $\text{SNR}_{\text{input}}$  drops below  $\text{SNR}_{\text{output}}^{\text{min}}$ , there is a frequency cutoff. Then,  $M_{\text{haze}}$  increases above  $M_{\text{geometry}}$ . For example, at ISO 200, when  $\beta = 0.2 \text{ km}^{-1}$ , at  $z = 30 \text{ m}$  the visibility is so bad that only large objects the size of tens of meters can be reliably seen. Changing the ISO setting dramatically changes  $M_{\text{haze}}$ . For example, at  $\beta = 0.2 \text{ km}^{-1}$ ,  $M_{\text{haze}}$  in ISO 800 is five times larger than  $M_{\text{haze}}$  in ISO 200.

the SNR decreases. At  $\beta z = b_{\text{critical}}$ , the SNR decreases below the minimum level and then we start to get a frequency cutoff. Thus, there is an abrupt increase of  $M_{\text{haze}}$ .

## 6. CONSIDERING NONLINEAR FILTERING

This paper deals only with linear filtering. What about nonlinear, anisotropic filtering, e.g., [51–55]? State-of-the-art (e.g., [51–55]) methods are well superior to linear filtering in extracting and preserving image details. Hence, linearity assumptions appear to lead to bounds [Eq. (38)] that are overly pessimistic.

Recently, [56,57] analyzed bounds to image denoising with nonlinear filtering in terms of minimum mean square error. However, also in nonlinear methods, as noise levels rise, details are increasingly lost. Small and low saliency details are lost before the larger, more salient ones (see [58]). Yet, the resolution bounds in this case may be somewhat different. The analysis is beyond the scope of this paper. Still, this section considers some possible aspects.

Frequency-domain analysis (as we have done) assumes system linearity: any image is a superposition of cosines and sines, the eigenfunctions of linear blur operations. In nonlinear filtering, thus, the generality of frequency-domain analysis of bounds would be difficult to apply, if at all. Hence, limits to nonlinear filters should be assessed directly on specific objects, not via frequencies. Moreover, considering Fig. 2, if averaging is performed only in the vertical direction, noise can be substantially reduced without eliminating the dominant horizontal variations and features. This property is exploited by nonlinear anisotropic filters. They are affected by rich regional characteristics, e.g., gradient, curvature, contour length and aspect ratio. Thus, these features and regional parameters should be incorporated into the analysis of bounds of nonlinear filters, in addition to the parameters we used (feature size and saliency).

Despite the complexity and difficulty to assess limits of such filters, the results may not differ greatly from Eq. (38). Close to the visibility limit, the signal modulation is very weak (relative to the noise). Then, linear operations may offer a lower bound to nonlinear filters that are differentiable. Reference [51] calculates the noise reduction of some denoising methods. It may thus serve as a basis for further calculation of resolution limits.

Many nonlinear methods locally adapt smoothing to the scale and contrast of objects [2]. The results of our analysis

can guide the design of adaptive filtering. Suppose a certain  $\text{SNR}_{\text{output}}^{\text{min}}$  is desired, given  $\text{SNR}_{\text{input}}$ . Then, a suitable kernel size  $W$  can be derived from Eq. (12) for each  $u$ , corresponding to the required feature size.

## 7. CONCLUSIONS

We presented a theoretical analysis of resolution bounds, where the degradation is dominantly pointwise. Even then, there can be a cutoff frequency  $u_{\text{cutoff}} < 0.5 [1/\text{pixels}]$ , which bounds the resolution of details that can be recovered at a desired success rate. These resolution bounds should be considered in addition to bounds stemming from blur in the system.

By placing noise and resolution on an equal footing, it becomes possible to analyze tradeoffs that combine blur and noise. For example, in low light conditions, an open aperture improves the SNR but causes blur outside the depth of field. The depth of field can be expanded by closing the aperture, but this decreases the SNR. Following this paper, a reduced SNR can translate to an effective cutoff frequency and thus enable an informed decision about what results in less blur: cutoff because of noise or blur from an open aperture. This analysis requires additional work, to consider also intermediate setups where both blur and cutoff from moderate noise exist. This resembles [59], which analyzes tradeoffs to optimize high dynamic range imaging. Further work is required to generalize this work to color images. Extension to video is straightforward. There, spatio-temporal noise filtering leads to spatio-temporal resolution bounds.

Our framework can benefit other problems of imaging, computer vision, and photography, where pointwise degradations are dominant. There, it may be possible to anticipate the potential recoverability of objects and features in a setting, either for existing recovery methods or for ones to be proposed.

## APPENDIX A: THE EFFECT OF SPATIAL AVERAGING

In Subsection 3.A, an image  $I$  formed by the model of Eq. (6) is filtered by an averaging filter  $h_W$ . The processed image

$$\hat{I} = I * h_W \quad (\text{A1})$$

has noise

$$n_{\text{output}}(\mathbf{x}) = \frac{1}{W^2} \sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} n(\mathbf{x}_i). \quad (\text{A2})$$

Here  $\Omega(\mathbf{x}) = \{\mathbf{x}_i : \|\mathbf{x}_i - \mathbf{x}\|_\infty < W/2\}$ . The noise  $n_{\text{output}}(\mathbf{x})$  is spatially correlated (not white) and its STD is

$$\sigma_{\text{output}} = \sigma/W. \quad (\text{A3})$$

The filter  $h_W$  affects each spatial frequency by an amount expressed by Dirichlet's function (the frequency response is calculated using the DTFT, as the images are discrete signals).

$$H_W(\mathbf{u}) = \text{DTFT}\{h_W(\mathbf{x})\} = \frac{\sin(\pi W u) \sin(\pi W v)}{W^2 \sin(\pi u) \sin(\pi v)}. \quad (\text{A4})$$

Using l'Hospital's rule in the limit  $v \rightarrow 0$  yields

$$H_W(u, 0) = \frac{1}{W} \frac{\sin(\pi W u)}{\sin(\pi u)}. \quad (\text{A5})$$

Applying  $h_W$  on Eq. (10) results in a cosine signal having the same frequency,

$$|S(u)_{\text{output}}| = H_W(u, 0)S(u). \quad (\text{A6})$$

Consider Eqs. (6), (11), (39), and (41). Similarly to Eq. (11), the SNR in the processed image  $\hat{I}$  is

$$\text{SNR}_{\text{output}}(u) = \frac{|S(u)_{\text{output}}|}{\sigma_{\text{output}}} = \frac{H_W(u, 0)|S(u)|}{\sigma/W}. \quad (\text{A7})$$

Combining Eqs. (11), (43), and (45) results in Eq. (12).

When  $W = 1/u$ , the filter  $h_W$  is destructive. Then, the frequency  $u$  is averaged out, as  $\text{SNR}_{\text{output}} = H_W(\mathbf{u}) = 0$ , eliminating by smear details featured by size  $1/u$ . For  $W > 1/u$ , there are negative sidelobes in  $H_W(u, 0)$ , causing reversal of contrast (artifacts). Hence, the domain considered legitimate for Eq. (45) is  $W < 1/u$ . In Fig. 7, window lengths outside this domain are in the blank region.

## APPENDIX B: GAUSSIAN FILTER

Our analysis in Appendix A yields a closed-form analytical expression of the SNR change as a result of using an averaging filter. Yet, a Gaussian filter is often preferred over flat window averaging. The Gaussian filter is defined as

$$h_g[x, y] = \frac{1}{2\pi W_g^2} \exp\left(-\frac{x^2 + y^2}{2W_g^2}\right), \quad (\text{B1})$$

where  $W_g$  is the STD of the Gaussian. The following analysis assumes that the effect of filter truncation is negligible.

How does the Gaussian filter affect the signal? The Fourier transform of a continuous Gaussian is also a Gaussian with an STD that equals  $1/(2\pi W_g)$ . Therefore, the transform of the discrete (sampled) Gaussian is a Gaussian duplicated in the frequency space. Suppose the replicas are mutually well separated (e.g., by three STDs,  $W_g > 0.95$ ). Then in the frequency range  $[-0.5, 0.5]$ ,

$$H_g[u, v] \approx \exp[-2\pi^2(u^2 + v^2)W_g^2]. \quad (\text{B2})$$

Since the raw image noise is uncorrelated, the noise variance in the filtered image is

$$\sigma_{\text{output}}^2 = \sum_{\forall x, \forall y} h_g^2(x, y)\sigma^2. \quad (\text{B3})$$

For high frequencies, we expect using a filter with  $W_g < 0.95$ , where Eq. (47) does not hold. Therefore, we start by a numerical calculation. For a range of values of  $W_g$ , the discrete Fourier transform is calculated. In addition, the noise change ( $\sigma_{\text{output}}/\sigma$ ) is calculated using Eq. (48). Then, the SNR change at a single horizontal frequency  $\mathbf{u} = [u, 0]$  is calculated as in Appendix A

$$C(u, W_g) \equiv \frac{\text{SNR}_{\text{output}}(u)}{\text{SNR}_{\text{input}}(u)}. \quad (\text{B4})$$

For each frequency,

$$C_{\text{max}}^{\text{gaus}}(u) = \max_{W_g} C(u, W_g). \quad (\text{B5})$$

The value of  $C_{\text{max}}(u)$  is plotted in Fig. 8 (dashed). The curve of  $C_{\text{max}}(u)$  corresponding to a Gaussian is very similar to the plot corresponding to a simple averaging window.

The SNR change for the Gaussian filter for  $W_g > 0.95$  can be analytically calculated. Using Parseval's theorem for DTFT,

$$\begin{aligned} (\sigma_{\text{output}}/\sigma)^2 &= \sum_{\forall x, \forall y} h_g^2(x, y) = \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} H_g^2(u, v) du dv \\ &= 1/(2\sqrt{\pi}W_g)^2. \end{aligned} \quad (\text{B6})$$

Combining Eqs. (47) and (51) into Eq. (49),

$$C(u, W_g) = \frac{H_g}{\sigma_{\text{output}}/\sigma} = \exp[-2\pi^2 u^2 W_g^2] (2\sqrt{\pi}W_g). \quad (\text{B7})$$

The STD that yields the maximum possible SNR improvement is

$$\frac{\partial}{\partial W_g} C(u, W_g) = 0 \Rightarrow W_g^{\text{max}} = 1/(2\pi u), \quad (\text{B8})$$

resulting in an SNR improvement of

$$C_{\text{max}}^{\text{gaus}}(u) = \frac{e^{-0.5}}{\sqrt{\pi}u} \approx \frac{1}{\pi u}. \quad (\text{B9})$$

This curve is identical to the numerical calculation in Fig. 8, except for high frequencies. In the low frequencies,  $\sin(\pi u) \approx \pi u$ , and thus  $C_{\text{max}} \approx C_{\text{max}}^{\text{gaus}}$ , as shown in Fig. 8.

## ACKNOWLEDGMENTS

We thank Gal Gur-Arye for her help with the simulations. Yoav Schechner is a Landau Fellow, supported by the Taub Foundation. The work was supported by MAFAT, the Israel Science Foundation (ISF Grant 1031/08) and the Winnipeg Research Fund. It was conducted in the Ollendorff Minerva Center. Minerva is funded through the BMBF. Tali Treibitz is an Awardee of the Weizmann Institute of Science National

Postdoctoral Award Program for Advancing Women in Science and was partially supported by Office of Naval Research Multidisciplinary University Research Initiative Grant #N00014-08-1-0638.

## REFERENCES

- Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
- R. Kaftory, Y. Y. Schechner, and Y. Y. Zeevi, "Variational distance-dependent image restoration," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2007)*, pp. 1–8.
- G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graph.* **23**, 664–672 (2004).
- T. Treibitz and Y. Y. Schechner, "Active polarization descattering," *IEEE Trans. Pattern Anal. Machine Intell.* **31**, 385–399 (2009).
- P. Dayan and L. F. Abbott, *Theoretical Neuroscience* (MIT, 2001), Chap. 4, pp. 139–141.
- N. S. Kopeika, *A System Engineering Approach to Imaging* (SPIE, 1998), Chaps. 9, 10, 19.
- S. W. Smith, *The Scientist & Engineer's Guide to Digital Signal Processing* (California Tech. Publishing, 1997), Chap. 11.
- J. Lloyd, *Thermal Imaging Systems* (Springer, 1975), Chaps. 5, 10.
- O. Schade, Sr., "An evaluation of photographic image quality and resolving power," *J. Soc. Motion Pict. Telev. Eng.* **73**, 81–119 (1964).
- H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," *J. Opt. Soc. Am. A* **7**, 1266–1278 (1990).
- H. Barrett, "NEQ: its progenitors and progeny," *Proc. SPIE* **7263**, 72630F (2009).
- W. Geisler, *Ideal Observer Analysis* (MIT Press, 2003), pp. 825–837.
- I. Cunningham and R. Shaw, "Signal-to-noise optimization of medical imaging systems," *J. Opt. Soc. Am. A* **16**, 621–632 (1999).
- M. Unser, B. L. Trus, and A. C. Steven, "A new resolution criterion based on spectral signal-to-noise ratios," *Ultramicroscopy* **23**, 39–51 (1987).
- M. Shahram and P. Milanfar, "Imaging below the diffraction limit: a statistical analysis," *IEEE Trans. Image Process.* **13**, 677–689 (2004).
- M. Shahram and P. Milanfar, "Statistical and information-theoretic analysis of resolution in imaging," *IEEE Trans. Inf. Theory* **52**, 3411–3437 (2006).
- H. Farid and E. H. Adelson, "Separating reflections from images by use of independent component analysis," *J. Opt. Soc. Am. A* **16**, 2136–2145 (1999).
- S. K. Nayar, X. S. Fang, and T. Boulton, "Separation of reflection components using color and polarization," *Int. J. Comput. Vis.* **21**, 163–186 (1997).
- R. Fattal, "Single image dehazing," *ACM Trans. Graph.* **27**, 72, (2008).
- J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: model-based photograph enhancement and viewing," *ACM Trans. Graph.* **27**, 116 (2008).
- Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," *Appl. Opt.* **42**, 511–525 (2003).
- R. T. Tan, "Visibility in bad weather from a single image," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2008)*, p. 108.
- J. Gu, R. Ramamoorthi, P. Belhumeur, and S. K. Nayar, "Dirty glass: rendering contamination on transparent surfaces," in *Eurographics Symposium on Rendering* (Springer, 2007), p. 159–170.
- T. Treibitz and Y. Y. Schechner, "Recovery limits in pointwise degradation," in *Proceedings of IEEE International Conference on Computational Photography (IEEE, 2009)*, p. 1–8.
- G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Machine Intell.* **16**, 267–276 (1994).
- A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," *ACM Trans. Graph.* **24**, 756–764 (2005).
- J. Takamatsu, Y. Matsushita, and K. Ikeuchi, "Estimating radiometric response functions from image noise variance," in *Proceedings of European Conference on Computer Vision (Springer, 2008)*, pp. 623–637.
- T. J. Fellers and M. W. Davidson, "CCD noise sources and signal-to-noise ratio," *Optical Microscopy Primer (Molecular Expressions™)* (2004).
- S. Inoué and K. R. Spring, *Video Microscopy: The Fundamentals*, 2nd ed (Springer, 1997), Chap. 7, p. 316.
- Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Machine Intell.* **29**, 1339–1354 (2007).
- C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, "Automatic estimation and removal of noise from a single image," *IEEE Trans. Pattern Anal. Machine Intell.* **30**, 299–314 (2008).
- Y. Matsushita and S. Lin, "Radiometric calibration from noise distributions," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2007)*, p. 1–8.
- S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," *ACM Trans. Graph.* **25**, 935–944 (2006).
- S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, "A theory of inverse light transport," in *Proceedings of IEEE International Conference on Computer Vision (IEEE, 2005)*, pp. 1440–1447.
- F. Koreban and Y. Y. Schechner, "Geometry by deflating," in *Proceedings of IEEE International Conference on Computational Photography (IEEE, 2009)*, p. 1–8.
- S. Bobrov and Y. Y. Schechner, "Image-based prediction of imaging and vision performance," *J. Opt. Soc. Am. A* **24**, 1920–1929 (2007).
- R. C. Henry, S. Mahadev, S. Urquijo, and D. Chitwood, "Color perception through atmospheric haze," *J. Opt. Soc. Am. A* **17**, 831–835 (2000).
- K. Tan and J. P. Oakley, "Physics-based approach to color image enhancement in poor visibility conditions," *J. Opt. Soc. Am. A* **18**, 2460–2467 (2001).
- Y. Schechner, D. Diner, and J. Martonchik, "Spaceborne underwater imaging," in *Proceedings of IEEE International Conference on Computational Photography (IEEE, 2011)*, p. 1–8.
- M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic aperture confocal imaging," *ACM Trans. Graph.* **23**, 825–834 (2004).
- A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li, "Removing photography artifacts using gradient projection and flash-exposure sampling," *ACM Trans. Graph.* **24**, 828–835 (2005).
- B. Wells, "MTF provides an image-quality metric," *Laser Focus World* **41** (2005).
- S. G. Narasimhan, C. Wang, and S. K. Nayar, "All the images of an outdoor scene," in *Proceedings of European Conference on Computer Vision (IEEE, 2002)*, pp. 148–162.
- J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, "General image-quality equation: GIQE," *Appl. Opt.* **36**, 8322–8328 (1997).
- P. Roetling, E. Trabka, and R. Kinzly, "Theoretical prediction of image quality," *J. Opt. Soc. Am.* **58**, 342–344 (1968).
- R. D. Fiete and T. A. Tantaló, "Comparison of SNR image quality metrics for remote sensing systems," *Opt. Eng.* **40**, 574–585 (2001).
- A. Burgess, "The rose model, revisited," *J. Opt. Soc. Am. A* **16**, 633–646 (1999).
- B. W. Keelan, *Handbook of Image Quality* (Dekker, 2002), Chaps. 2, 3.
- D. H. Kelly, "Adaptation effects on spatio-temporal sine-wave thresholds," *Vis. Res.* **12**, 89–101 (1972).
- R. Mantiuk, K. Kim, A. Rempel, and W. Heidrich, "Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.* **30**, 40 (2011).
- A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**, 490–530 (2005).

52. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
53. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.* **15**, 3736–3745 (2006).
54. P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.* **18**, 1438–1451 (2009).
55. J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. Image Process.* **12**, 1338–1351 (2003).
56. P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Trans. Image Process.* **19**, 895–911 (2010).
57. A. Levin and B. Nadler, "Natural image denoising: optimality and inherent bounds," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2010), p. 2833–2840.
58. J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Denoising examples," [decsai.ugr.es/~javier/denoise/examples](http://decsai.ugr.es/~javier/denoise/examples).
59. S. Hasinoff, F. Durand, and W. Freeman, "Noise-optimal capture for high dynamic range photography," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2010), p. 553–560.