Recovery Limits in Pointwise Degradation

Tali Treibitz and Yoav Y. Schechner
Department of Electrical Engineering
Technion - Israel Inst. Technology, Haifa 32000, Israel
ttali@tx.technion.ac.il, yoav@ee.technion.ac.il

Abstract

Pointwise image formation models appear in a variety of computational vision and photography problems. Prior studies aim to recover visibility or reflectance under the effects of specular or indirect reflections, additive scattering, radiance attenuation in haze and flash, etc. This work considers bounds to recovery from pointwise degradation. The analysis uses a physical model for the acquired signal and noise, and also accounts for potential post-acquisition noise filtering. Linear-systems analysis yields an effective cutoff-frequency, which is induced by noise, despite having no optical blur in the imaging model. We apply this analysis to hazy images. 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. The theory considers the pointwise degradation that exists in the scene during acquisition, which fundamentally limits recovery, even if the parameters of an algorithm are perfectly set.

1. Introduction

Many computational vision and photography problems deal with degradation effects that are essentially pointwise. These include specularities over a diffuse reflection [8, 21]; attenuation in flash photography [23]; attenuation and veiling scatter (airlight) in haze [9, 14, 26, 30], dirty windows [10] and other scattering media; semireflections when looking through a window [8], and more. Being pointwise processes, optical blur effects are insignificant in these contexts. An expanding array of methods has been devised to handle such problems. However, what is the recovery *limit*? Are there object features that cannot be effectively recovered, despite the best efforts made by the recovery method? Can we quantitatively assess the recoverability of an object of a certain size and contrast? If some objects are not recovered, is there a point in trying to develop a new method to salvage them, or is their loss fundamental?

This paper approaches these questions. It describes a quantitative bound for recovering objects and features, under pointwise degradation. We seek a bound that depends on basic object and background characteristics, camera parameters, and the physical degradation model. Then, we use this theory to analyze resolution and range limits of dehazing. 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.

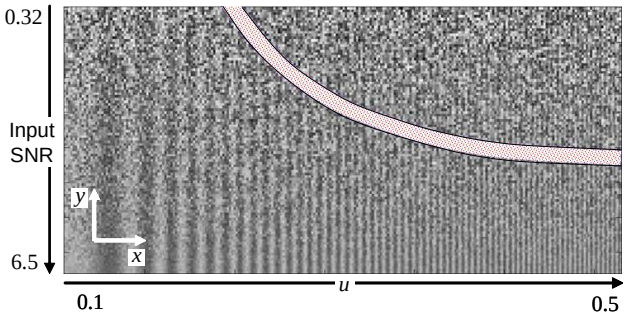

Figure 1. A raw noisy image. The signal is $I = a + s\cos[2\pi u_x x]$. The spatial frequency is $u_x \propto x$. Here a is bias and s is the amplitude. White added noise increases with y. The result is then contrast-stretched. At low frequencies (small x) the pattern is visible even in very low input SNR. Beyond the marked line, on the upper-right corner it becomes nearly impossible to reliably distinguish the signal details under the noise.

Noise limits the recovery. Digital denoising enhances the results [12, 23, 25], but even then, there is a limit. Some works derived recovery-induced amplifications of white noise, concluding that recovery is limited when a signal matches the noise intensity [25, 31]. However, limits based directly on white noise ignore the effective noise suppression possible if the feature of interest is large enough. Consider Fig. 1. Here 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 very low input signal to noise ratio (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 of Fig. 1, it is nearly impossible to reliably distinguish the signal details under the noise. There appears to be a cutoff, around the marked line in Fig. 1, 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 [13], but *noise-limited*.

Noise limits were studied [13] in systems suffering from

optical blur.[1] It attenuates the signal's high frequencies below a threshold set by the noise STD. In pointwise degradations, lack of optical blur implies no such attenuation, hence apparently no effective limits to the resolution. Such analysis cannot account for loss of detail as shown in Fig. 1 and is unsuitable to the degradations dealt with here. In Ref. [32] a cutoff stemmed solely from the typical falloff of signal energy at high frequencies. This cannot explain the effect in Fig. 1, where the input SNR is independent of the frequency u. Refs. [13, 32] do not account for the SNR increase possible by smoothing (implicit by the viewer or explicit by image processing). This leads to underestimation of recoverability. In contrast, this paper derives bounds in problems that involve no optical blur, yet considering the enhancement by potential post-acquisition filtering.

Beyond insights on basic limits, our results can help assess the potential of recovery of typical objects in haze by current cameras. They may also help decide on the required specifications for systems to be proposed.

2. Theoretical Background

2.1. Pointwise Degradation and Range Dependency

Let $l^{\text{object}}(\mathbf{x})$ 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

$$I(\mathbf{x}) = l^{\text{object}}(\mathbf{x})t(\mathbf{x}) + a(\mathbf{x}) + n(\mathbf{x}) \ , \qquad (1)$$

where $t(\mathbf{x})$ and $a(\mathbf{x})$ account for *deterministic* multiplicative and additive effects, respectively. In addition, Eq. (1) includes unbiased uncorrelated *random* noise $n(\mathbf{x})$. Note that a is *non-negative*. We seek recovery of $l^{\text{object}}(\mathbf{x})$. This model fits a wide range of computer vision and computational photography problems:

- In analysis of reflections, l^{object} is the diffuse component and a is the specular component (while $t = 1$) [21].
- A similar distinction exists in the mixture of direct and indirect illumination components [22, 28].
- In semireflections, l^{object} is the scene behind a window of transmittance t, and a is the semireflected layer [8].
- In 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 [10].
- A bright light source near the field of view can contribute an additive component $a(\mathbf{x})$ of flare [15].
- Fixed pattern noise is a deterministic effect [3, 11] of pointwise gain and bias variations, which is modeled by $t(\mathbf{x})$ and $a(\mathbf{x})$ in Eq. (1).

The degradation effects may be distance-dependent. In haze, t is the transmittance of the atmosphere [9, 14, 26, 30]. Its dependency on the object distance $z(\mathbf{x})$ is given by

$$t(\mathbf{x}) = e^{-\beta z(\mathbf{x})} \ . \qquad (2)$$

Here, $\beta \in (0, \infty)$ is the atmospheric attenuation coefficient. The additive component a here is the *airlight* [26], given by

$$a(\mathbf{x}) = a_\infty[1 - t(\mathbf{x})] \ , \qquad (3)$$

where a_∞ is the value of airlight at a non-occluded horizon. Airlight increases with z and dominates the acquired image irradiance[2] at long range (see Fig. 2). There are other distance-dependent pointwise models, including synthetic aperture lighting [17], which may include scatter, and flash photography [1] (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. 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.

The Noise

Photon noise is a fundamental quantum-mechanical effect. It cannot be overcome, regardless of the camera quality. Accounting for this noise component [11, 33], overall the noise variance[3] in the raw image data [29] is

$$\sigma^2 \approx AI(\mathbf{x}) + B, \qquad (4)$$

where $A, B > 0$ and I is the image intensity given in Eq. (1), excluding $n(\mathbf{x})$. The term B encompasses the variance of the signal-independent components of the gray-level noise. As detailed in [11, 27, 29], $B = \rho^2_{\text{read}} + \rho^2_{\text{digit}} + DT$. Here, ρ_{read} is the amplifier readout noise STD, ρ_{digit} is the noise STD of a uniform quantizer, D is the detector dark current and T is the exposure time. Our consequent analysis can be easily applied to other noise models.

Here we understand the second implication of a. Being non-negative, $a(\mathbf{x})$ increases I, sometimes significantly. Thus, it increases the variance of the *random* noise (Eq. 4). This affects all the above mentioned computational photography problems, including imaging in haze.

The noise is spatially white. It 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.

[1] In these contexts, the criteria used are termed minimum-resolvable contrast (MRC) or temperature (MRT) [13].

[2] Contrast is mainly degraded by airlight, rather than blur [31].

[3] The linear relation in Eq. (4) does not hold [18, 19] for cameras having amplifier nonlinearities. However, our fundamental analysis is targeted at recovery that uses high quality cameras. In these cameras, Eq. (4) is typically followed. In low intensities [2], cameras sometime exhibit deviations from Eq. (4). This deviation is negligible in well exposed images. A calibration we have done for a Nikon D100 is consistent with Eq. (4).

2.2. 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

$$m = Mf/(zp) \ . \tag{5}$$

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 $|u_{\rm cutoff}| \leq 0.5$ [1/pixels], then their equivalent effective lower limit size is

$$m \approx 1/(2u_{\rm cutoff}) \tag{6}$$

pixels. Eq. (6) degenerates to $m = 1$ pixel in the upper bound of $u_{\rm cutoff}$. Eq. (6) enables analysis of system resolution in the Fourier domain. According to Eqs. (5,6), once $u_{\rm cutoff}$ is determined, an object at distance z is within the effective resolution of the system if its length is at least

$$M(z) = zp/(2fu_{\rm cutoff}) \ \ [\text{meters}] \ \ . \tag{7}$$

If the CCD resolution is designed to match the lens' optical resolution, and there is no additional blurring effect, the expression in Eq. (7) degenerates to Eq. (5). This yields the geometric bound for a minimum visible object size

$$M_{\rm geometry} = mzp/f \ \ . \tag{8}$$

3. Noise-Limited Resolution

Without noise, even small intensity changes over a background could be stretched to reveal objects and details. Figs. 2a,c demonstrate piecewise contrast-stretch on a synthetic utopian noise-free hazy image. Even in parts that appear blank in Fig. 2a, visibility is retrieved in Fig. 2c. 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. (4) is introduced in Fig. 2b. Now, objects are lost in the parts corresponding to distant regions, despite regional contrast stretch in Fig. 2d. Noise reduction operations affect the signal amplitude. The effect of this operation varies as a function of the signal's spatial frequency.

We note that empirical work had been done about human performance in four specific vision tasks (e.g. object detection and recognition) under noise, culminating in *Johnson tables* and *NIIRS image ratings* [3, 13, 16]. These tasks were associated [13, 16] with the number of resolvable features in an object. This number directly corresponds to the highest spatial frequency in an image of the object. Thus, frequency-domain analysis may partially explain and generalize some aspects of these empirical criteria. More importantly, however, is that our linear-systems analysis is general, and thus applies to computer vision (not human) systems, in contrast to the Johnson tables and NIIRS ratings.

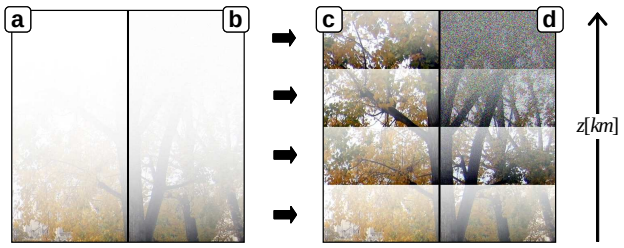

Figure 2. (a) A noise-free hazy image, simulated by linearly changing $z \in [1, 30]$km, $\beta = 0.2\text{km}^{-1}$. Airlight acts as local bias. (b) A slightly noisy version. (c) Regional contrast stretching of (a) reveals the objects and details. (d) Regional contrast stretching of (b) does not recover *small* details at large z, over noise.

3.1. How Much Would Filtering Help?

There are numerous denoising methods. As a basic benchmark, we focus the analysis on linear filtering. 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 Sec. 6.

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

$$s(\mathbf{x}) = S(u) \cos 2\pi ux \ . \tag{9}$$

It is corrupted by white additive noise, whose input STD is σ (Eq. 4). Note that the image in Eq. (9) is composed of a single discrete spatial frequency $\mathbf{u} = (u, 0)$ and its conjugate $(-u, 0)$. The SNR of the raw image is defined as

$$R_{\rm input}(u) = |S(u)|/\sigma \ \ . \tag{10}$$

We study the *signal* at a specific spatial frequency. This indicates the potential for recovering object-features that correspond to a specific size (Eq. 6). 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 post-acquisition processing. In contrast to the signal, the *noise is not treated at different frequencies*. The reason is that in the input, noise is white, irrespective of the feature-size, as illustrated in Fig. 1, and seen in most natural images (as in Fig. 2).

To suppress white noise, consider a flat averaging filter h_W, having a support of $W \times W$ pixels.[4] As we prove in

[4] In principle, better results can be obtained by more sophisticated digital low-pass filters, which may be designed by an array of engineering methods. However, the flat filter we use leads to closed-form expressions, which are useful both for obtaining insights and for a baseline assessment.

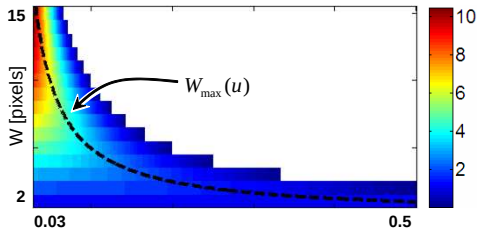

Figure 3. The SNR improvement C, as a function of u and W. The curve of $W_{\max}(u)$ is plotted on top. As u increases, windows are limited to smaller sizes. This limits the ability to suppress noise while maintaining the signal.

the Appendix, applying h_W on $s(\mathbf{x})$ amplifies the SNR by

$$C(u) \equiv \frac{R_{\text{proc}}(u)}{R_{\text{input}}(u)} = \frac{\sin(\pi W u)}{\sin(\pi u)} \quad , \tag{11}$$

where $R_{\text{proc}}(u)$ is the SNR of the processed image. Fig. 3 illustrates $C(u)$ for a range of W. The window size $W_{\max}$ that maximizes the improvement of SNR at u satisfies

$$\frac{\partial C(u)}{\partial W} = 0 \quad \Rightarrow \quad W_{\max} = 1/(2u) \quad . \tag{12}$$

Fig. 3 plots $W_{\max}$. The maximal SNR amplification $C_{\max}(u)$ that can be achieved by spatial averaging is thus

$$C_{\max}(u) \equiv C_{W_{\max}}(u) = 1/\sin(\pi u) \quad . \tag{13}$$

Fig. 4 demonstrates the use of a window of length $W_{\max}$ on the image of Fig. 1. Here, $W_{\max}$ adapts spatially to $u(x)$. Across all frequencies, the signal pattern is better seen in Fig. 4 than in Fig. 1. However, as u increases, the pattern can be observed reliably only at higher values of R_{input} (at smaller y), and is effectively lost at low R_{input}. This is consistent with Eq. (13): a smaller SNR improvement can be achieved at high u, thus requiring a higher input SNR.

3.2. The Cutoff Frequency

3.2.1 Success in a Confidence Interval

Sec. 3.1 showed that post-acquisition filtering may enhance the SNR to R_{proc}. Now, the question is which value of R_{proc} is sufficient? The higher $R_{\text{proc}}(u)$ is, the more *confidence* there is in the recovery of u. The randomness of noise imposes randomness in the success of recovering u in a single incident. Over a large ensemble, the success rate increases with $R_{\text{proc}}(u)$. Following basic probabilistic analysis [6] of confidence intervals, for a minimum success rate ρ, $R_{\text{proc}}(u)$ should be at least $R_{\min}^{\text{output}}$, where

$$R_{\min}^{\text{output}} = \sqrt{2}\,\text{erf}^{-1}(\rho + 1) \tag{14}$$

and erf is the error function [6]. Eq. (14) applies to a Gaussian distribution[5] of the output noise n_{proc}. Hence,

[5]This distribution is reasonable due to averaging by h_W, and also since the number of photons (creating photon noise) per pixel is typically large.

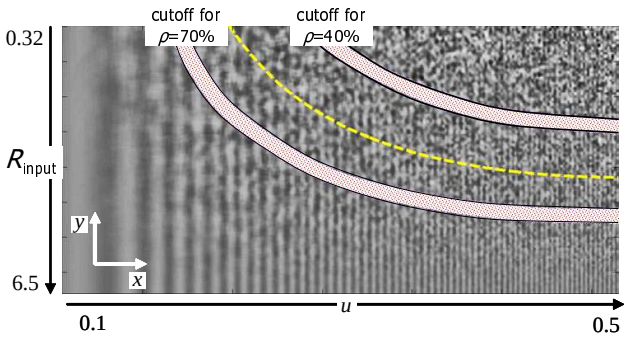

Figure 4. Filtering the image in Fig. 1 with window size $W_{\max}(u)$ improves visibility. Cutoff lines are derived using Eqs. (17,18). Below the line of $\rho = 70\%$, the pattern is clearly seen. Above the upper line, $\rho < 40\%$, and noise dominates. The dashed curve corresponds to the curve plotted in Fig. 1, having $\rho = 50\%$.

$R_{\min}^{\text{output}}$ is the minimal acceptable level of the output SNR, for achieving the desired success rate.

3.2.2 Recovery within a Bound

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 counter-productive, due to the consequent detail loss. In this case, without filtering

$$R_{\text{input}}(u) \geq R_{\min}^{\text{output}} \quad . \tag{15}$$

In moderate SNR, gentle filtering with $W < W_{\max}$ may suffice to bring the output SNR R_{proc} to the acceptable level $R_{\min}^{\text{output}}$. However, at the *limit* of recovery, the signal in u is so low relative to the noise, that $W_{\max}(u)$ should be used. There is no point in trying to use a kernel wider than $W_{\max}$, since it would yield a lower $C(u)$ than a $W_{\max}$-sized kernel (see Eq. 12), while blurring excessively. Hence, to achieve recovery within the required success rate, we hope to have

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

From Eqs. (13,14,16), the recovery rate ρ is achieved if

$$R_{\text{input}} \geq \sin(\pi u) R_{\min}^{\text{output}}. \tag{17}$$

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 $R_{\text{input}} < \sin(\pi u) R_{\min}^{\text{output}}$, the SNR in the original image is too low and cannot be increased to the desired level $R_{\min}^{\text{output}}$, *no matter what* filter size we use. This is a recovery limit, posed by noise.

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

$$u_{\text{cutoff}} = [\arcsin(R_{\text{input}}/R_{\min}^{\text{output}})]/\pi \quad . \tag{18}$$

It is plotted in Fig. 5a. Note in this plot that when Eq. (15) is satisfied, there is no cutoff.

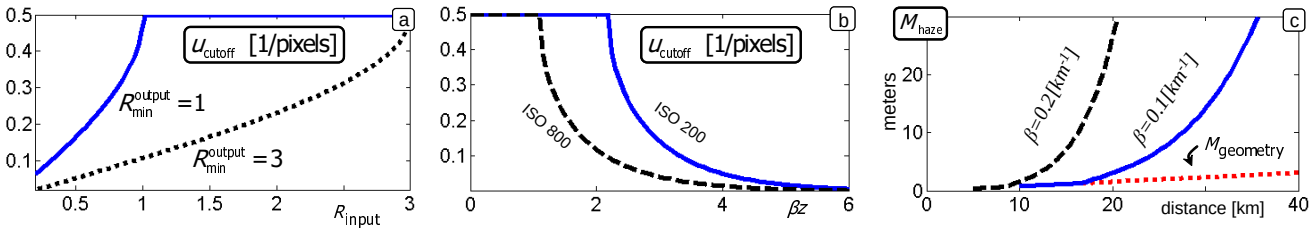

Figure 5. (a) The cutoff frequency as a function of R_{input}, for different values of $R_{\min}^{\text{output}}$. A better input SNR yields a better resolution at the final image. When $R_{\text{input}} < R_{\min}^{\text{output}}$, the image starts to lose reliability. (b,c) Bounds for object and background having $\tilde{E} = 25\%$. (b) As z increases, u_{cutoff} decreases and (c) the minimal discernible object size increases. Plotted for D100 at ISO800 and $\rho \approx 95\%$.

Suppose for a moment that R_{input} is independent of u. The sine function is monotonically increasing. Thus, if $u_{\text{cutoff}} < 0.5$, then considering Eq. (17),

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

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

The cutoff in Eq. (18) is plotted in Fig. 4, for a few values of ρ. Interestingly, the analytical plots correspond to subjective impression about visibility. Below the cutoff line for $\rho \approx 70\%$, the pattern is clearly seen. Above the upper cutoff line, where $\rho \approx 40\%$, noise dominates.

The cutoff frequency sets the *resolution limit* to images with additive white noise, since u_{cutoff} can be used to determine the object size of the least-resolved objects, using Eq. (7). 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.

3.3. Defining the Signal

Typically, there is 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. Due to Eqs. (1,2), 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-vs-background distinction in order to define the *signal* of interest. In Sec. 3.1 and App. A the signal $S(u)$ was the amplitude of a cosine, and random noise was added to that signal. Note that $S(u)$ is *half the difference* between the maximum and minimum values of that cosine. Thus, for consistency, define the signal as

$$S = \tfrac{1}{2}[l^{\text{object}}(u) - l^{\text{back}}(u)]t \quad , \tag{20}$$

in the problem of differentiating an object over a background. What about the image component $a(\mathbf{x})$? Recall from Sec. 2.1, 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 [8, 9, 21, 22, 26, 30]. By itself, this removable non-random component does not decrease the object-vs.-background difference. However (see Sec. 2.1), a increases the photon noise, thus effecting the image SNR.

4. Limitations in Haze

In this section we apply the analysis of Sec. 3 to a specific type of degradation: haze in images.

4.1. SNR in Raw Hazy Images

The noise variance σ^2 is derived in Eq. (4), simply based on the total image intensity I. Using the midway intensity between l^{object} and l^{back}, the mean of the noise variance over the object and its immediate background is

$$\sigma^2(b) = B + A\left\{\tfrac{1}{2}\left(l^{\text{object}} + l^{\text{back}}\right) t(b) + a_\infty[1 - t(b)]\right\}, \tag{21}$$

where $b = \beta z$, based on Eqs. (1-4). In this way, the random noise induced by airlight is incorporated. The signal here is defined in Eq. (20). The SNR R_{input} of a raw hazy image is obtained by using Eq. (20) and σ from Eq. (21) in Eq. (10).

We now introduce some variables that ease the assessment. The saturation graylevel is $V = 2^{\mathtt{B}} - 1$, where $\mathtt{B}$ is the number of bits per pixel. Define the object saliency (without disturbances) as $\tilde{E} = \left|l^{\text{object}} - l^{\text{back}}\right| / V$. 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. (1-4,10) yields the input SNR as a function of b

$$R_{\text{input}}(b) = \frac{e^{-b}\tilde{E}V}{2\sqrt{B + AV\{\tilde{l}e^{-b} + \tilde{a}_\infty[1 - e^{-b}]\}}} \quad . \tag{22}$$

In Eqs. (4,22), the parameters A, B, 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 [27]. In the following, we plot our results for a Nikon D100, whose respective parameters were calibrated in a couple of ISO settings.[6]

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 $R_{\rm input}$ is proportional to it. The value of $\tilde{E}$ is the main input by a user for assessing object visibility in haze. For instance, if a nearby car over a street occupies $\tilde{E} = 20\%$ of the dynamic range, this has prime effect on its distinction in the presence of attenuation (at a distance) and noise. 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 [20]. Thus, as an example, we set $\tilde{E} = 25\%$ in the following plots.

There is a critical optical depth $b_{\rm critical}$, up to which no cutoff frequency exists. This optical depth can be found by solving a quadratic equation based on Eq. (22),

$$R_{\rm input}(b_{\rm critical}) = R_{\rm min}^{\rm output} \quad . \tag{23}$$

At $b < b_{\rm critical}$, Eq. (17) is satisfied $\forall u$. At $b > b_{\rm critical}$, some frequencies start to become unrecoverable.

4.2. 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 $R_{\rm input}(b)$ from Eq. (22) into Eq. (18) yields $u_{\rm cutoff}(\beta z) = \arcsin\left[R_{\rm input}(\beta z)/R_{\rm min}^{\rm output}\right]/\pi$. This cutoff is plotted in Fig. 5b, using $\tilde{E} = 25\%$, calibration data of Nikon D100 at two ISO settings, and $R_{\rm min}^{\rm output} = 2$ in Eq. (22). The value of $u_{\rm cutoff}$ decreases with βz. Moreover, the two ISO settings yield a significant difference in the visible distances.

Using $u_{\rm cutoff}(\beta z)$ in Eq. (7) with $R_{\rm input}(\beta z)$ from Eq. (22) yields the least resolved object length in haze

$$M_{\rm haze}(\beta, z) = \frac{\pi z p}{2f \arcsin\left[R_{\rm input}(\beta z)/R_{\rm min}^{\rm output}\right]}, \tag{24}$$

Eqs. (22,24) depend on the (scene independent) system parameters $\{p, f, A, B, V\}$. They also depend on the scene's z, β and $\tilde{E}$. Fig. 5c plots $M_{\rm haze}$ for different parameter values. Note that $M_{\rm haze} = M_{\rm geometry}$ if $\beta z < b_{\rm critical}$. At larger z or β, there is an abrupt increase of $M_{\rm haze}$.

[6] Here $V = 2^{14} - 1$. For ISO 200, we got $[A, B] = [0.9, 210]$. For ISO 800, we got $[A, B] = [10, 2116]$.

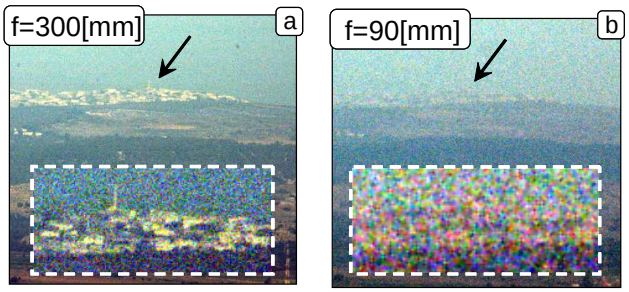

Figure 6. Experiment in haze. The houses in a distant village are clearly seen using $f = 300\text{mm}$. They are not distinguishable with $f = 90\text{mm}$. This apparent resolution cutoff is caused by noise and the lack of pixels needed for effective implicit denoising.

5. Experimental Evidence

Fundamental analysis of bounds is essentially analytical. Nevertheless, it is useful to observe empirical evidence to theoretical conclusions. Fig. 6a shows an image we acquired using a Nikon[7] D100 and $f = 300\text{mm}$, and then contrast-stretched to partially overcome the effect of haze. A regional contrast stretch in a window reveals details of a village at $z = 18\text{km}$, including distinguishable houses ($M \approx 10\text{m}$) and a narrow tower. Thus, these object features *survived any optical blur* that may have been caused by the lens and atmospheric multiple scattering. This is important, since that village was deep within the multiple-scattering range, at $b = 3$. This value was obtained by processing polarimetric images using the method of Ref. [26]. Measurements[8] used in Eq. (24) yield $M_{\rm haze} \approx [3\text{m}\ 4\text{m}\ 8\text{m}]$ corresponding to the RGB channels. This is consistent with the distinct visibility of the individual houses.

Just afterwards, the same camera and zoom lens ($f\# = 8$ in both cases) acquired Fig. 6b. Here, $f = 90\text{mm}$ (image crop was applied). Now, the *houses cannot be distinguished*. Blaming this on atmospheric blur (multiple scattering) is overruled, since these houses were distinguishable in Fig. 6a. Blaming this on pixelization is also overruled, since in these settings (with $f = 90\text{mm}$) $M_{\rm geometry} = 1.6\text{m}$, i.e, much smaller[9] than the buildings' $M \approx 10\text{m}$. We believe the reason for this loss of detail is noise: while the raw σ is the same in both Figs. 6a, 6b (same ISO settings), there are much less pixels in Fig. 6b to implicitly average it out, per object. Indeed, Eq. (24) yields in these settings $M_{\rm haze} \approx [16\text{m}\ 18\text{m}\ 39\text{m}]$ in the RGB channels.

In addition, Fig. 7 shows a scene extracted from the WILD database [20]. On a clear day, visibility exists at

[7] Accounting for 2D subsampling by the CCD's Bayer mosaic, the effective pixel pitch in the respective RGB channels is $p = [2\ \sqrt{2}\ 2] \cdot 7.8\mu\text{m}$.

[8] We measured I directly from the image and then used it to estimate σ using Eq. (4). We measured $l^{\rm back}, l^{\rm object}$ on houses and trees that were located in proximity to the camera at the same time. We set $\rho = 70\%$.

[9] In offline experiments in clear days the houses were distinguishable.

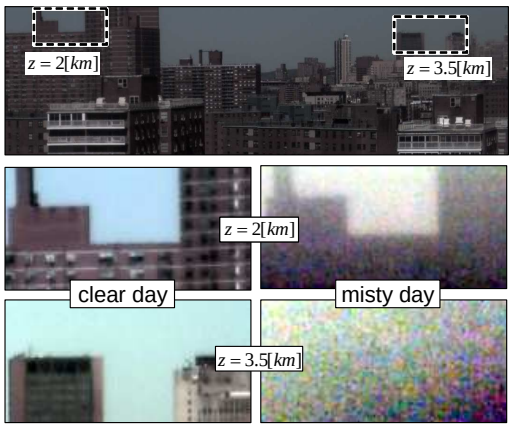

Figure 7. [Top] Clear day scene. [Middle] Small details seen on a clear day at $z \approx 2$km but lost in mist. [Bottom] At $z \approx 3.5$km, visibility in mist quickly worsens: even large buildings are lost.

both $z = 2\text{km}$ and 3.5km for building outlines (large M) and windows (small M). In mist, there is a loss of spatial detail, despite regional contrast stretch: at $z = 2\text{km}$ 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 exacerbated at longer distance: at $z = 3.5\text{km}$ even the building outline is hopelessly obscured by noise.

6. Considering Nonlinear Filtering

This paper deals only with linear filtering. What about nonlinear, anisotropic filtering, e.g., [4, 7]? Also in nonlinear methods, as noise levels rise, details are increasingly lost. Small and low saliency details are lost before the bigger, more salient ones (see [24]). However, the resolution bounds in this case may be somewhat different. How different can they be? Such 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. 1, 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 feature 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. (24). Close to the limit, the signal modulation is very weak (relative to the noise). Then, linear operations may offer a first-order approximation to nonlinear filters that are differentiable. We thus hypothesize that the results we obtained can provide *rules of thumb*, even if more general processing is applied. However, this needs to be verified. Extensive simulations may be required. In addition, Ref. [4] calculates the noise reduction for 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 [12]. The results of our analysis can guide the design of adaptive filtering. Suppose an output $R_{\min}^{\text{output}}$ is desired, given R_{input}. Then a suitable kernel size W can be derived from Eq. (11) for each u, corresponding to the required feature size.

7. Conclusion

This theoretical analysis of bounds dealt with cases where blur in the image formation step can be neglected, and 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 in a desired success rate. We applied this analysis to hazy images. Extension to video is straightforward. There, noise filtering can be done in the spatio-temporal domain, leading to spatio-temporal resolution bounds.

Our framework can benefit other problems of 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.

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, and was conducted in the Ollendorff Minerva Center. Minerva is funded through the BMBF.

A. The Effect of Spatial Averaging

In Sec. 3.1, an image I formed by the model of Eq. (1) is filtered by an averaging filter h_W. The processed image

$$\hat{s} = I * h_W \tag{25}$$

has noise $n_{\text{proc}}(\mathbf{x}) = \frac{1}{W^2}\sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} n(\mathbf{x}_i)$. Here $\Omega(\mathbf{x}) = \{\mathbf{x}_i : \|\mathbf{x}_i - \mathbf{x}\|_\infty < W/2\}$. The noise $n_{\text{proc}}(\mathbf{x})$ is spatially correlated (not white) and its STD is

$$\sigma_{\text{proc}} = \sigma/W \quad . \tag{26}$$

The filter h_W affects each spatial frequency by an amount expressed by Dirichlet's function:

$$H_W(\mathbf{u}) = \text{DTFT}\{h_W(\mathbf{x})\} = \frac{\sin(\pi W u)\sin(\pi W v)}{W^2 \sin(\pi u)\sin(\pi v)}. \tag{27}$$

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

$$H_W(u,0) = \frac{1}{W}\frac{\sin(\pi W u)}{\sin(\pi u)} \quad . \tag{28}$$

Applying h_W on Eq. (9) results in a cosine signal having the same frequency, and whose amplitude is $H_W(u,0)S(u)$. Consider Eqs. (1,10,25,26). Similarly to Eq. (10), the SNR in the processed image $\hat{s}$ is

$$R_{\mathrm{proc}}(u) = \frac{H_W(u,0)|S(u)|}{\sigma/W} \quad . \tag{29}$$

Combining Eqs. (10,28,29) results in Eq. (11).

When $W = 1/u$, the filter h_W is destructive. Then, the frequency u is averaged out, as $R_{\mathrm{proc}} = H_W(\mathbf{u}) = 0$, 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. (29) is $W < 1/u$. In Fig. 3, window lengths outside this domain are in the white region.

References

[1] A. Agrawal, R. Raskar, S. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM TOG*, 24:828–835, 2005.

[2] F. Alter, Y. Matsushita, and X. Tang. An intensity similarity measure in low-light conditions. In *Proc. ECCV*, pages 267–280, 2006.

[3] S. Bobrov and Y. Y. Schechner. Image-based prediction of imaging and vision performance. *J. Opt. Soc. Amer. A*, 24:1920–1929, 2007.

[4] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4:490–530, 2005.

[5] P. Dayan and L. Abbott. *Theoretical Neuroscience*, chapter 4, pages 139–141. MIT press, 2001.

[6] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*, pages 622–624. Wiley, 2nd edition, 2000.

[7] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. IP*, 15:3736–3745, 2006.

[8] H. Farid and E. H. Adelson. Separating reflections from images by use of independent component analysis. *J. Opt. Soc. Amer. A*, 16:2136–2145, 1999.

[9] R. Fattal. Single image dehazing. *ACM TOG*, 27, 2008.

[10] J. Gu, R. Ramamoorthi, P. Belhumeur, and S. Nayar. Dirty glass: Rendering contamination on transparent surfaces. In *Eurographics Symp. on Rendering*, 2007.

[11] G. Healey and R. Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. PAMI*, 16:267–276, 1994.

[12] R. Kaftory, Y. Y. Schechner, and Y. Y. Zeevi. Variational distance-dependent image restoration. In *Proc. IEEE CVPR*, 2007.

[13] N. S. Kopeika. *A System Engineering Approach to Imaging*, chapter 9,19. SPIE press, 1998.

[14] 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 TOG*, 27, 2008.

[15] F. Koreban and Y. Y. Schechner. Geometry by deflaring. In *Proc. IEEE ICCP*, 2009.

[16] J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio. General Image-Quality Equation: GIQE. *App. Opt.*, 36:8322–8328, 1997.

[17] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas. Synthetic aperture confocal imaging. *ACM TOG*, 23:825–834, 2004.

[18] C. Liu, R. Szeliski, S. Kang, C. Zitnick, and W. Freeman. Automatic estimation and removal of noise from a single image. *IEEE PAMI*, 30:299–314, 2008.

[19] Y. Matsushita and S. Lin. Radiometric calibration from noise distributions. In *Proc. IEEE CVPR*, 2007.

[20] S. G. Narasimhan, C. Wang, and S. Nayar. All the images of an outdoor scene. In *Proc. ECCV*, pages 148–162, 2002.

[21] S. Nayar, X. Fang, and T. Boult. Separation of reflection components using color and polarization. *Int. J. Comp. Vision*, 21:163–186, 1997.

[22] 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 TOG*, 25:935–944, 2006.

[23] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama. Digital photography with flash and no-flash image pairs. *ACM TOG*, 23:664–672, 2004.

[24] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Denoising examples. `decsai.ugr.es/~javier/denoise/examples`.

[25] Y. Y. Schechner and Y. Averbuch. Regularized image recovery in scattering media. *IEEE Trans. PAMI*, 29:1655–1660, 2007.

[26] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Polarization-based vision through haze. *App. Opt.*, 42:511–525, 2003.

[27] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. Multiplexing for optimal lighting. *IEEE Trans. PAMI*, 29:1339–1354, 2007.

[28] S. Seitz, Y. Matsushita, and K. Kutulakos. A theory of inverse light transport. In *IEEE ICCV*, pages 1440 – 1447, 2005.

[29] J. Takamatsu, Y. Matsushita, and K. Ikeuchi. Estimating radiometric response functions from image noise variance. In *Proc. ECCV*, pages 623–637, 2008.

[30] R. T. Tan. Visibility in bad weather from a single image. In *Proc. IEEE CVPR*, 2008.

[31] T. Treibitz and Y. Y. Schechner. Active polarization descattering. *to appear, IEEE Trans. PAMI*, 2009.

[32] M. Unser, B. Trus, and A. Steven. A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy*, 23:39–51, 1987.

[33] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM TOG*, 24:756 – 764, 2005.