Removal of Haze and Noise from a Single Image

Erik Matlin and Peyman Milanfar*

Department of Electrical Engineering
University of California, Santa Cruz

ABSTRACT

Images of outdoor scenes often contain degradation due to haze, resulting in contrast reduction and color fading. For many reasons one may need to remove these effects. Unfortunately, haze removal is a difficult problem due the inherent ambiguity between the haze and the underlying scene. Furthermore, all images contain some noise due to sensor (measurement) error that can be amplified in the haze removal process if ignored.

A number of methods have been proposed for haze removal from images. Existing literature that has also addressed the issue of noise has relied on multiple images either for denoising prior to dehazing[1] or in the dehazing process itself.[2,3] However, multiple images are not always available. Recent single image approaches, one of the most successful being the "dark channel prior",[4] have not yet considered the issue of noise.

Accordingly, in this paper we propose two methods for removing both haze and noise from a single image. The first approach is to denoise the image prior to dehazing. This serial approach essentially treats haze and noise separately, and so a second approach is proposed to simultaneously denoise and dehaze using an iterative, adaptive, non-parametric regression method. Experimental results for both methods are then compared.

Our findings show that when the noise level is precisely known a priori, simply denoising prior to dehazing performs well. When the noise level is not given, latent errors from either "under"-denoising or "over"-denoising can be amplified, and in this situation, the iterative approach can yield superior results.

Keywords: dehazing, denoising, dark channel prior, BM3D, atmospheric light, transmission map, single image

1. INTRODUCTION

Images of outdoor scenes often contain haze, fog, or other types of atmospheric degradation caused by particles in the atmospheric medium absorbing and scattering light as it travels from the source to the observer. While this effect may be desirable in an artistic setting, it is sometimes necessary to undo this degradation. For example, many computer vision algorithms rely on the assumption that the input image is exactly the scene radiance, i.e. there is no disturbance from haze. When this assumption is violated, algorithmic errors can be catastrophic. One could easily see how a car navigation system relying on visual inputs from the scene ahead, which did not take this effect into account could have dangerous consequences.

Among current haze removal research, haze estimation methods can be divided into two broad categories of either relying on additional data or using a prior assumption. Methods that rely on additional information include: taking multiple images of the same scene using different degrees of polarization,[5,6] multiple images taken during different weather conditions,[7] and methods that require user supplied depth information[8] or a 3D model.[9] While these can achieve good results, the extra information required is often not available, and so a more flexible approach is preferable.

Significant progress in single image haze removal has been made in recent years. Tan[10] made the observation that a haze-free image has higher contrast than a hazy image, and was able to obtain good results by maximizing contrast in local regions of the input image. However, the final results obtained by this method are not based on a physical model and are often unnatural looking due to over-saturation. Fattal[11] was able to obtain good results by assuming that transmission and surface shading are locally uncorrelated. With this assumption, he obtains

*This work was supported by AFOSR Grant FA9550-07-1-0365 and NSF Grant CCF-1016018
Further author information:
ematlin@soe.ucsc.edu, milanfar@soe.ucsc.edu

the transmission map through independent component analysis. This is a physically reasonable approach, but this method has trouble with very hazy regions where the different components are difficult to resolve. Lastly, a simple but powerful approach proposed by He et al.[4] uses dark pixels in local windows to obtain a coarse estimate of the transmission map followed by a refinement step using an image matting technique.[12] Their method obtains results on par with or exceeding other state-of-the-art algorithms, and is even successful with very hazy scenes.

Even when the haze content is known, noise can be a major problem when restoring hazy images. The works mentioned above largely avoid the issue, usually by assuming a noise free image or only dehazing images up to a point where noise is negligible. Existing literature that has addressed noise has taken two basic approaches: denoising prior to dehazing, and denoising during dehazing. Recently, Joshi and Cohen[1] took the first approach, addressing noise through image fusion. By taking multiple images of the same hazy scene and using weighted averaging, they obtain a sharp, low noise hazy image that they then dehaze using a variation on the dark channel approach. Schechner and Averbuch[2] use a polarization based method to estimate haze, and address noise by adding a local penalty term proportional to the transmission value as a regularization term in scene radiance recovery. Since the regularization is relatively simple, hazy regions are essentially blurred while non-hazy regions are left sharp. A more sophisticated variant on this theme is proposed by Kaftory and Schechner,[3] which uses a total variation method based on Beltrami flow for regularization. Although effective, the use of complicated PDE methods, requiring minimization over the entire image, is a major drawback. Furthermore, all of the mentioned works rely on multiple images.

This paper addresses the problem of recovering the underlying scene radiance of a **single** noisy, hazy image, with the main contributions as follows. First is an investigation on the effect of noise on estimating haze with the dark channel prior. Next we propose and compare two different methods for scene radiance recovery once a haze estimate has been obtained. The first method is to denoise the image using a state-of-the-art denoising algorithm (BM3D[13]) prior to dehazing, which can be interpreted as a single image adaptation of Joshi and Cohen.[1] The second method is to simultaneously denoise and dehaze using an approach based on iterative non-parametric regression. This can be seen both as an adaptation to a single image as well as a simplification of methods presented in Ref. 2 and Ref. 3.

2. BACKGROUND

There are several methods for estimating the haze contribution in a single image. One of the most successful is the dark channel prior, proposed by He et al.,[4] and is used as the basis for haze estimation in this paper. Here we briefly summarize the approach.

The dark channel prior is derived assuming a noise-free image with the following image formation model:

$$\mathbf{I}(\mathbf{x}) = \mathbf{R}(\mathbf{x})t(\mathbf{x}) + \mathbf{a}_\infty(1 - t(\mathbf{x})) \tag{1}$$

where $\mathbf{x}$ is a pixel location, $\mathbf{I}$ is the observed image, $\mathbf{R}$ is the underlying scene radiance, $\mathbf{a}_\infty$ is the atmospheric light (or airlight), and t is the transmission coefficient. The ultimate goal of haze removal is to find $\mathbf{R}$, which also requires knowledge of $\mathbf{a}_\infty$ and t. From this model, it is apparent that haze removal is an under-constrained problem. In a grayscale image, for each pixel there is only 1 constraint but 3 unknowns; for an RGB color image, there are 3 constraints but 7 unknowns (assuming t is the same for each color channel). Essentially, one must resolve the ambiguous question of whether an object's color is a result of it being far away and mixed with haze, or if the object is close to the observer and simply the measured color.

The key to the dark channel prior is the observation that natural haze-free outdoor images are generally well textured, and contain a variety of colorful objects. As a consequence, most patches will contain one or more pixels with very low intensity in at least one of the color channels. These dark pixels can be attributed to dark objects, shadows, or objects that are primarily a combination of only one or two of the RGB color channels. With this observation in mind, one can construct the so called "dark channel" of an image, which can be expressed mathematically as a minimum value operation in patches around the target pixel:

$$I^{dark}(\mathbf{x}) = \min_{c\in\{r,g,b\}} \left(\min_{\mathbf{y}\in\Omega(\mathbf{x})} (I^c(\mathbf{y})) \right) \tag{2}$$

where $I^{dark}(\mathbf{x})$ represents the "dark channel" of image I at pixel location $\mathbf{x}$, I^c is a color channel of image I, and $\mathbf{y} \in \Omega(\mathbf{x})$ signifies all pixels $\mathbf{y}$ in a local patch around $\mathbf{x}$. If applied to a haze-free image, the above observation yields:

$$I^{dark}(\mathbf{x}) \to 0$$

In contrast, hazy images contain an additive atmospheric light component, yielding:

$$I^{dark}(\mathbf{x}) \to \min_{\mathbf{y}\in\Omega(\mathbf{x})} (\mathbf{a}_\infty(1 - t(\mathbf{x})))$$

Since $t(\mathbf{x})$ is essentially constant in a local window and $\mathbf{a}_\infty$ is usually close to white (meaning it won't significantly affect the dark channel value), the dark channel prior effectively identifies the relative haze content throughout an image. This can aid in estimating the atmospheric light, $\mathbf{a}_\infty$. Following Ref. 4, we estimate this component as the brightest RGB intensities in the hazy image among the pixels corresponding to the top 0.1% brightest dark channel locations, i.e. the regions where haze is most dominant.

Once the atmospheric light is known, in a noise-free image, the transmission map can be estimated as:[4]

$$\hat{t}(\mathbf{x}) = 1 - w \min_{c\in\{r,g,b\}} \left(\min_{\mathbf{y}\in\Omega(\mathbf{x})} \left(\frac{I^c(\mathbf{y})}{a^c_\infty} \right) \right) \tag{3}$$

where the superscript c signifies a specific color channel, i.e. the red, green, or blue color channel of the corresponding parameter. Note that $\mathbf{a}_\infty$ is a vector quantity containing a separate value for each color channel, and so a^c_∞ is a scalar quantity referring to one individual value. Thus, an estimate for the transmission map is obtained by simply subtracting the dark channel of the normalized image from 1. The scaling parameter, w, takes a value from 0 to 1, and corresponds to the amount of haze left in the image[†].

After the initial haze estimate is obtained, a refinement step is required to suppress halo artifacts. He et al.[4] use the Matting Laplacian.[12] Although this is not the quickest solution, it provides visually satisfactory results, and so is the process used in this paper.

Finally, the dehazed image is usually recovered by simple inversion of Eq. (1), solving for $\mathbf{R}$.

3. PROPOSED APPROACH

All images contain some amount of noise due to measurement (sensor) error. Taking this fact into account, noise must be included in the image model for haze formation:

$$\mathbf{Y}(\mathbf{x}) = \mathbf{I}(\mathbf{x}) + \mathbf{n}(x) = \mathbf{R}(\mathbf{x})t(\mathbf{x}) + \mathbf{a}_\infty(1 - t(\mathbf{x})) + \mathbf{n}(x) \tag{4}$$

where the observed image is now $\mathbf{Y}$, and $\mathbf{n}$ is the noise contribution, assumed to be independent and identically distributed (I.I.D.), with zero mean and variance σ^2. Assuming for a moment that the atmospheric light and transmission map are perfectly known, if the scene radiance is recovered by a naive inversion (as done in Ref. 4, the noise contribution is amplified by $1/t(\mathbf{x})$:

$$\begin{aligned}
\hat{\mathbf{R}}(\mathbf{x}) &= \frac{\mathbf{Y}(\mathbf{x}) - \mathbf{a}_\infty(1 - t(\mathbf{x}))}{t(\mathbf{x})} \\
&= \frac{\mathbf{R}(\mathbf{x})t(\mathbf{x}) + \mathbf{a}_\infty(1 - t(\mathbf{x})) + \mathbf{n}(\mathbf{x}) - \mathbf{a}_\infty(1 - t(\mathbf{x}))}{t(\mathbf{x})} \\
&= \mathbf{R}(\mathbf{x}) + \frac{\mathbf{n}(\mathbf{x})}{t(\mathbf{x})}
\end{aligned} \tag{5}$$

Since $t(\mathbf{x})$ is a value between 0 and 1, Eq. (5) implies that except in the special case when haze is absent ($t = 1$), the noise contribution is amplified when recovering the scene radiance. Furthermore, in very hazy regions, where

[†]The scaling parameter is presented here for completeness, in all examples presented here we let $w = 1$.

t is close to 0, the noise contribution can dominate the results. More explicitely, the effective noise variance in the recovered scene radiance, $\hat{\mathbf{R}}$ is proportional to the square of the transmission:

$$\sigma_e^2 \propto \frac{1}{t^2(\mathbf{x})} \tag{6}$$

In addressing this problem for the single image case, the proposed approach includes two major steps: a haze estimation step, and a restoration step. For haze estimation, we build on the dark channel prior. We show that because it is sensitive to noise, denoising is a critical pre-processing step. For restoration we propose and compare two approaches. The first approach is to treat denoising and haze removal as separate processes by denoising the image, prior to dehazing. The second approach is to remove both haze and noise simultaneously via iterative kernel regression.

3.1 Haze Estimation

Eq. (4) reveals a weakness in the dark channel prior. Since the dark channel prior relies on sample minima, it is especially sensitive to outliers. Various approaches can be taken to robustly estimate the dark channel, and by extension the transmission map, considering the presence of noise. A basic method from the field of descriptive statistics is to use quantiles, such as the 10th percentile as an estimate.[14] A more sophisticated approach can be taken by using stochastic approximation[15] to locate local minima, followed by some type of point estimate. However, since other point estimates are needed (such as for the atmospheric light estimate), a still better approach may be to simply denoise the entire image as a pre-processing step. This is the method employed in this paper, with the specific denoising algorithm being the color version of BM3D[13],[16] which uses a block matching and collaborative Wiener filtering scheme for denoising. As this algorithm is currently among the state-of-the-art, results obtained with it should be about as good as one could expect from this haze estimation strategy. From the denoised hazy image, the dark channel, atmospheric light, and transmission map can be estimated. Fig. 1 is an example of the effect of noise on the dark channel of an image. If noise is ignored, the computed dark channel contains significant errors (Fig. 1c). However, if the image is first denoised with BM3D, the resulting dark channel is very similar to that of the noise-free case (Fig. 1d). Fig. 2a shows the dark channel mean squared error (as compared to the dark channel computed on a noise-free image) for the same example image for a range of noise levels.

BM3D requires an estimate of the noise variance as its only tuning parameter. Therefore the question arises of whether it is better to err on the side of over-estimating or under-estimating this parameter. Generally speaking, over-denoising causes an image to become smoother. Since haze is primarily a slowly varying component, one would intuitively conclude that it is better to over-estimate the noise variance. Fig. 2b supports this hypothesis, showing that, for an image with added Gaussian noise ($\sigma_n = 0.1$)[‡], the dark channel estimate is much worse for the "under"-denoised case than for the "over"-denoised case. In the authors' experience, results are similar for the atmospheric light estimate.

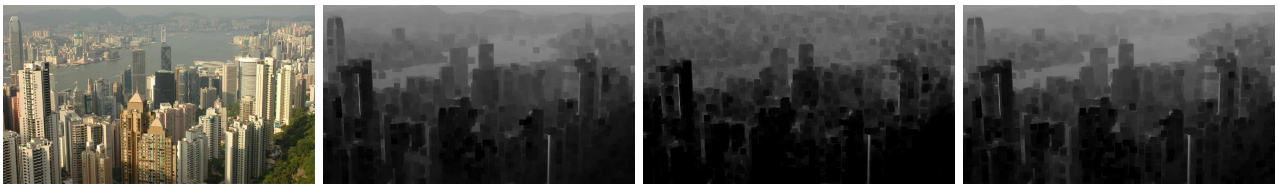
3.2 Restoration

Once the transmission and atmospheric light are estimated, the scene radiance can be recovered from the hazy image. Two possible avenues we explore for this final restoration step are: 1) denoising followed by dehazing, and 2) simultaneous denoising and dehazing via iterative kernel regression. The most significant difference between the two approaches is that the first one treats the haze and noise separately, while the other considers haze and noise together. We expect the first approach to perform well if the denoising step is very good; however, we expect the second approach to be more robust in the case of having to estimate the noise level.

3.2.1 Denoising Followed by Dehazing

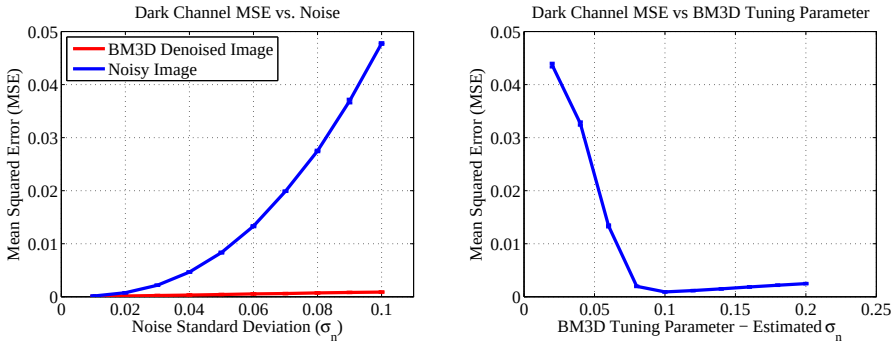
Simply denoising the hazy image prior to performing the dehazing process is a natural approach to handling the problem of noise in scene radiance recovery. It is worth noting that denoising *after* dehazing was considered in Ref. 3 using the Non-Local Means[17] denoising algorithm, with the authors finding that areas of the image

[‡]Note that σ_n is for an image whose intensity values are normalized to 1.



(a) Input Image (b) Dark Channel (from Noise Free) (c) Dark Channel, (from Noisy, $\sigma_n = 0.05$) (d) Dark Channel, (from Denoised, $\sigma_n = 0.05$)

Figure 1: The dark channel computed from the noisy image (c) shows significant differences compared to the dark channel computed from the noise-free image (b). If the noisy image is denoised prior to computing the dark channel, the result is significantly improved (d). Note that σ_n is for an image whose intensity values are normalized to 1.



(a) Dark Channel MSE vs. σ_n (b) Dark Channel MSE vs σ_{BM3D}

Figure 2: (a) The mean squared error of the dark channel computed on a denoised image using BM3D given the exact noise variance is significantly better than if computed on a noisy image for $\sigma_n = 0.01$ to 0.1. (b) The mean squared error of the dark channel is more sensitive to "under"-denoising than to "over"-denoising. In this case, the input image has added Gaussian noise with $\sigma_n = 0.1$. Both results are for the input image shown in Fig. 1a.

containing little haze suffered from over-smoothing. This will be the case with most standard denoising algorithms, since they typically treat the noise level as homogeneous throughout the image, and we know that noise in the recovered scene radiance is inversely proportional to the spatially varying transmission. By contrast, if we denoise the image prior to dehazing, the noise variance is not spatially varying, and so we should expect a standard denoising algorithm to perform properly.

In our restoration scheme, denoising as a pre-processing step is especially convenient considering that it is already necessary for estimating the atmospheric light and transmission map. In the denoising step, we can treat our image model as: $\mathbf{Y} = \mathbf{I} + \mathbf{n}$, with the task being only to estimate $\mathbf{I}$, which encapsulates the hazy image. After the hazy image is denoised, the rest of the dehazing process is exactly the same as in the noise-free case. The complete dehazing procedure is summarized in Algorithm 1.

Algorithm 1 Single Image Haze and Noise Removal via Denoising Followed by Dehazing

1. Estimate $\hat{\mathbf{I}}$ by denoising (e.g. BM3D) input image $\mathbf{Y}$
2. Estimate transmission, t, and atmospheric light, $\mathbf{a}_\infty$ (e.g. dark channel prior) from $\hat{\mathbf{I}}$
3. Dehaze $\hat{\mathbf{I}}$ through inversion of Eq. (1) to obtain the restored image: the scene radiance estimate $\hat{\mathbf{R}}$

3.2.2 Simultaneous Denoising and Dehazing via Iterative Kernel Regression

While the previous approach provides good results when the noise variance is known a priori, we wish to find an alternative approach that is more robust to unknown parameters. For this purpose, we propose a non-parametric kernel regression based method, in which we treat noise and haze together. Before presenting the details of this method, let us first digress for a moment with a brief overview of kernel regression.

Classic parametric regression relies on estimating a set of parameters for a specific signal that a set of data is assumed to represent.[18] A major limitation of this approach is that the estimated signal is limited by the choice of global model. If the underlying signal is sufficiently complicated, natural images being a major example, it may be practically impossible to choose a model that fits all of the data well. Non-parametric methods overcome this limitation by allowing the data to ultimately dictate the structure of the (local) model. In the case of kernel regression, this is known as a *regression function*.[19]

The data model for kernel regression in 2-D is given by:

$$y_i = z(\mathbf{x}_i) + \varepsilon_i, \quad \forall\, \mathbf{x}_i \in \Omega(\mathbf{x}), \quad i = 1, \ldots, N \tag{7}$$

where y_i is a noisy measurement at spatial coordinate $\mathbf{x}_i = [x_{1i}, x_{2i}]^T$, $z(\cdot)$ is the unspecified regression function, ε_i is zero mean I.I.D. measurement noise, and $\Omega(\mathbf{x})$ is a window around the point of interest at coordinate $\mathbf{x}$, containing N samples. In general, the spatial coordinates may be randomly spaced, but for the sake of simplicity we will consider only fixed spacing.

Although the exact form of $z(\cdot)$ may remain unspecified, a local Taylor expansion of this function is given by:

$$\begin{aligned} z(\mathbf{x}_i) &\approx z(\mathbf{x}) + \{\nabla z(\mathbf{x})\}^T (\mathbf{x}_i - \mathbf{x}) + \frac{1}{2}(\mathbf{x}_i - \mathbf{x})^T \{\mathcal{H} z(\mathbf{x})\}^T (\mathbf{x}_i - \mathbf{x}) + \cdots \\ &\approx \beta_0 + \boldsymbol{\beta}_1^T (\mathbf{x}_i - \mathbf{x}) + \boldsymbol{\beta}_2^T \operatorname{vech}\left\{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T\right\} + \cdots \end{aligned} \tag{8}$$

where ∇ is the gradient (2×1) operator, $\mathcal{H}$ is the Hessian (2×2) operator, and $\operatorname{vech}(\cdot)$ is the half-vectorization operator, which lexicographically orders the lower triangular portion of a symmetric matrix in a vector, i.e.:

$$\operatorname{vech}\left(\begin{bmatrix} a & b \\ b & d \end{bmatrix}\right) = \begin{bmatrix} a & b & d \end{bmatrix}^T \tag{9}$$

Therefore for a 2-D image, $\beta_0 = z(\mathbf{x})$, which is the pixel value at the coordinate of interest, and $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of partial derivatives, specifically:

$$\boldsymbol{\beta}_1 = \nabla z(\mathbf{x}) = \left[\frac{\partial z(\mathbf{x})}{\partial x_1}, \quad \frac{\partial z(\mathbf{x})}{\partial x_2}\right]^T \tag{10}$$

$$\boldsymbol{\beta}_2 = \frac{1}{2}\left[\frac{\partial^2 z(\mathbf{x})}{\partial x_1^2}, \quad 2\frac{\partial^2 z(\mathbf{x})}{\partial x_1 \partial x_2}, \quad \frac{\partial^2 z(\mathbf{x})}{\partial x_2^2}\right]^T \tag{11}$$

In contrast to parametric regression, where a global approach is taken, kernel regression involves a local approach, where data are weighted inversely proportional to their distance from the point of interest. Usually this is formulated as a weighted least squares estimation:

$$\min_{\{\boldsymbol{\beta_m}\}_{m=0}^{M}} \sum_{i=1}^{N} \left[y_i - \beta_0 - \boldsymbol{\beta}_1^T (\mathbf{x}_i - \mathbf{x}) - \boldsymbol{\beta}_2^T \operatorname{vech}\left\{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T\right\} - \ldots\right]^2 K_{\mathbf{H}_i}(\mathbf{x}_i - \mathbf{x}) \tag{12}$$

and

$$K_{\mathbf{H}_i}(\mathbf{u}) = \frac{1}{\det(\mathbf{H}_i)} K(\mathbf{H}_i^{-1}\mathbf{u}) \tag{13}$$

where M is the regression order. $K(\cdot)$ is the so-called *kernel function* that penalizes pixels according to their distance from the point of interest, and $\mathbf{H}_i$ is the smoothing matrix (2×2), controlling the strength of this penalty, with the simplest choice being $\mathbf{H}_i = h\mathbf{I}$, where h is known as the *global smoothing parameter* and $\mathbf{I}$

is identity. $K(\cdot)$ is a radially symmetric function with a maximum at zero. The exact choice of function plays only a small role in estimation accuracy[20] compared to the choice of smoothing parameter. So generally a kernel function is chosen given other considerations, such as ease of computation or having convenient properties. A Gaussian function is a typical choice.

There are a variety of methods described in the literature[19–21] for choosing the smoothing parameter. However, since these methods involve dependence on typically unknown parameters, the value is often chosen empirically such that it satisfies some quality measure. Intuitively, in low textured images, a relatively large smoothing parameter is desired, while in highly textured images, a relatively small smoothing parameter is desired.

In classic kernel regression the distance weight is dependent solely on spatial distance. This always results in the estimated pixel being a **linear** combination of the nearby samples. For data such as images, where the underlying structure may be highly complex and non-stationary, this property is a significant drawback. With this fact as motivation, Takeda et al.[22] proposed a locally adaptive regression kernel, where the kernel depends both on the data locations and the data values. More specifically, the smoothing matrix in the regression kernel is redefined as:

$$\mathbf{H}_i = h\mathbf{C}_i^{-1/2} \tag{14}$$

and is called the *steering* matrix, where h is a global smoothing parameter. The matrix, $\mathbf{C}_i$, is estimated from the covariance matrix of the local gradient vectors. A naive estimate is obtained by:

$$\widehat{\mathbf{C}}_i = \mathbf{G}_i^T\mathbf{G}_i = \begin{bmatrix} \sum_i z_{x_1}^2(\mathbf{x}_i) & \sum_i z_{x_1}(\mathbf{x}_i)z_{x_2}(\mathbf{x}_i) \\ \sum_i z_{x_1}(\mathbf{x}_i)z_{x_2}(\mathbf{x}_i) & \sum_i z_{x_2}^2(\mathbf{x}_i) \end{bmatrix} \tag{15}$$

with

$$\mathbf{G}_i = \begin{bmatrix} z_{x_1}(\mathbf{x}_1) & z_{x_2}(\mathbf{x}_1) \\ \vdots & \vdots \\ z_{x_1}(\mathbf{x}_N) & z_{x_2}(\mathbf{x}_N) \end{bmatrix} \tag{16}$$

where $z_{x_1}(\cdot)$ and $z_{x_2}(\cdot)$ are the first derivatives measured along the x_1- and x_2-axes, and N is the number of samples in a local window around the position of interest, $\mathbf{x}_i$.

Finally, the locally adaptive regression kernel is constructed from a Gaussian function and is expressed as the following:

$$K_{\mathbf{H_i}}(\mathbf{x}_i - \mathbf{x}) = \frac{\sqrt{\det(\mathbf{C}_i)}}{2\pi h^2} \exp\left(-\frac{(\mathbf{x}_i - \mathbf{x})^T\mathbf{C}_i(\mathbf{x}_i - \mathbf{x})}{2h^2}\right) \tag{17}$$

With the adaptive kernel regression tools in hand, we apply them to our problem of hazy image restoration. In the above discussion, the underlying signal is corrupted only by noise and we solved for the unknown regression function by applying a local Taylor expansion. However, in the problem of haze removal, the underlying signal is also attenuated according to its transmission value and mixed with the atmospheric light. Furthermore, the transmission map must be estimated from the noisy image. With this in mind, we propose an iterative method based on kernel regression for estimating the scene radiance and transmission map simultaneously. Let us first rewrite Eq. (4) as

$$\mathbf{Y}(\mathbf{x}) = \mathbf{R}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(\mathbf{x}) + \mathbf{n}(\mathbf{x}) \tag{18}$$

where $\mathbf{A}(\mathbf{x}) = \mathbf{a}_\infty(1 - t(\mathbf{x}))$. In contrast to $\mathbf{a}_\infty$, which is the *global* atmospheric light, $\mathbf{A}(\mathbf{x})$ can be considered the *local* atmospheric light, since it describes the atmospheric light contribution on a per pixel basis. Following the non-parametric kernel regression strategy, we can write our estimation problem for each color channel as:

$$\min_{A^c, R^c} \sum_{x_i \in \Omega(\mathbf{x})} \left[Y^c(\mathbf{x}_i) - R^c(\mathbf{x}_i)t(\mathbf{x}_i) - A^c(\mathbf{x}_i)\right]^2 K_{H_i}(\mathbf{x}_i - \mathbf{x}) \tag{19}$$

where $\Omega(\mathbf{x})$ indicates a neighborhood around the coordinate of interest, $\mathbf{x}$, and the superscript, c, indicates a specific color channel of $\mathbf{Y}$, $\mathbf{R}$, or $\mathbf{A}$. K_{H_i} indicates the locally adaptive regression kernel defined above. When considering a color image, the steering matrices are computed on the luminance channel, and are applied to all

color channels. Note also that $t(\mathbf{x}) = 1 - \frac{A^c(\mathbf{x})}{a^c_\infty}$ where again c indicates a specific color channel. We assume that we have an initial estimate of $\mathbf{a}_\infty$.

Since Eq. (19) is a minimization over two unknowns, our strategy is to find the solution iteratively, by decomposing it into two separate minimization problems, and alternating between solving for R^c and solving for A^c. Furthermore, we assume a zeroth order regression model:

$$\min_{R^c} \sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} \left[Y'^c(\mathbf{x}_i) - t(\mathbf{x}_i) R^c(\mathbf{x}) \right]^2 K_{H_i}(\mathbf{x}_i - \mathbf{x}) \tag{20}$$

$$\min_{A^c} \sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} \left[Y''^c(\mathbf{x}_i) - P^c(\mathbf{x}_i) A^c(\mathbf{x}) \right]^2 K_{H_i}(\mathbf{x}_i - \mathbf{x}) \tag{21}$$

where $Y'^c = Y^c - A^c$, $Y''^c = Y^c - R^c$, and $P^c = 1 - \frac{1}{a^c_\infty} R^c$. Eq. (20) and Eq. (21) are simple weighted least-squares problems, and their solutions are:

$$\hat{R}^c(\mathbf{x}) = \frac{\sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} K_{H_i}(\mathbf{x}_i - \mathbf{x}) t(\mathbf{x}_i) Y'^c(\mathbf{x}_i)}{\sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} K_{H_i}(\mathbf{x}_i - \mathbf{x}) t(\mathbf{x}_i)^2} \tag{22}$$

$$\hat{A}^c(\mathbf{x}) = \frac{\sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} K_{H_i}(\mathbf{x}_i - \mathbf{x}) P^c(\mathbf{x}_i) Y''^c(\mathbf{x}_i)}{\sum_{\mathbf{x}_i \in \Omega(\mathbf{x})} K_{H_i}(\mathbf{x}_i - \mathbf{x}) P^c(\mathbf{x}_i)^2} \tag{23}$$

The estimation problem has now been reduced to a set of filtering operations. Note that although the filtering operation itself appears linear, since the steering kernels are computed on the received data, the result is a non-linear filter.

The regression kernel in Eq. (17) is typically used such that the smoothing parameter is kept constant throughout the entire image. This works well when the noise variance is constant; however, in hazy images, the noise variance in the scene radiance changes according to the transmission value. Therefore the smoothing parameter should also be spatially varying in this case. It can be shown[23] that for zeroth order adaptive kernel regression, with the approximation that $\mathbf{C}$ is constant in a local window, the optimal smoothing parameter is:

$$h_{opt} \approx \left(\frac{\sigma^2 (\det(\mathbf{C}))^{5/2}}{2\pi N (\frac{\partial^2 z(\mathbf{x})}{\partial x_1^2} C_{22} - 2 \frac{\partial^2 z(\mathbf{x})}{\partial x_1 \partial x_2} C_{12} + \frac{\partial^2 z(\mathbf{x})}{\partial x_2^2} C_{11})^2} \right)^{1/6} \tag{24}$$

where N is the number of pixels in the local neighborhood, C_{ij} indicates the i,jth element in $\mathbf{C}$ and $z(\mathbf{x})$ is the underlying regression function (i.e. the noise free image).

Note first that Eq. (24) relies on the unknown second derivatives of the noise free image. These can be estimated using higher order kernel regression [§]. Recalling that the effective noise variance, σ^2, is inversely proportional to the square of the transmission, t^2, we also add a $\frac{1}{t^2}$ term. Since the derivative estimates may be prone to errors (as they are estimated from a noisy image), and the noise variance is typically unknown, we add a global smoothing parameter h_{global} to compensate. N is also absorbed into this parameter, since it is a constant. And so we arrive at an expression for a spatially adaptive smoothing parameter:

$$h_{adap}(\mathbf{x}) = h_{global} \left(\frac{(\det(\mathbf{C}))^{5/2}}{t^2(\mathbf{x}) (\frac{\partial^2 z(\mathbf{x})}{\partial x_1^2} C_{22} - 2 \frac{\partial^2 z(\mathbf{x})}{\partial x_1 \partial x_2} C_{12} + \frac{\partial^2 z(\mathbf{x})}{\partial x_2^2} C_{11})^2} \right)^{1/6} \tag{25}$$

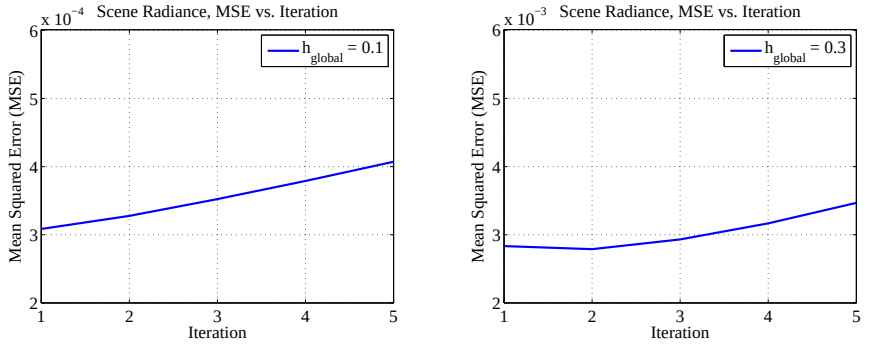
[§]The approach taken here is to estimate the second derivatives using second order classic (non-adaptive) kernel regression with a simple Gaussian kernel on a pilot estimate for the scene radiance.

The complete haze removal procedure is summarized in Algorithm 2. Fig. 3 shows mean squared error results for this procedure carried out on the image shown in Fig. 1a for different levels of added zero-mean Gaussian noise with different global smoothing parameters and number of iterations. If the image is sufficiently denoised in the first iteration, further iterations worsen the result; however, if the image is "under"-denoised, iterations improve the result. Note that a larger global smoothing parameter is required when the noise level is increased.

Algorithm 2 Single Image Haze and Noise Removal via Iterative Kernel Regression

1. Pilot estimate
   - Estimate $\hat{\mathbf{I}}$ by denoising (e.g. BM3D) input image $\mathbf{Y}$
   - Estimate transmission, t, and atmospheric light, $\mathbf{a}_\infty$ (e.g. dark channel prior) from $\hat{\mathbf{I}}$
   - Dehaze $\hat{\mathbf{I}}$ through inversion of Eq. (1) to obtain a pilot estimate of the scene radiance $\hat{\mathbf{R}}_{\mathbf{pilot}}$
2. Estimate 2nd derivatives from $\hat{\mathbf{R}}_{\mathbf{pilot}}$ for adaptive smoothing parameter
3. Iterate between estimate for $\mathbf{R}$ and estimate for $\mathbf{A}$ until the minimum MSE is reached or until some quality measure is maximized. Note that these estimates are computed from the image $\mathbf{Y}$.

   while $Q \leq Q_{max}$ or $MSE \geq MSE_{min}$ **do**
   Estimate $\hat{\mathbf{R}}$ using current $\hat{\mathbf{A}}$ (Eq. (22))
   Estimate $\hat{\mathbf{A}}$ using current $\hat{\mathbf{R}}$ (Eq. (23))
   end while



(a) Scene Radiance MSE vs. Iterations, $\sigma_n = 0.01$

(b) Scene Radiance MSE vs. Iterations, $\sigma_n = 0.05$

Figure 3: The iterative procedure described by Algorithm 2 was performed on the image in Fig. 1a for different levels of added zero-mean Gaussian noise and different global smoothing parameters. If the image is under-smoothed, then iterations can improve the quality of the result. Otherwise, iterations worsen the quality of the result. Note that for a higher noise level, a higher global smoothing parameter is needed.

4. EXPERIMENTAL RESULTS

An example comparing the two approaches with synthetically added noise is shown in Fig. 4. The input image has added zero-mean Gaussian noise with standard deviation, $\sigma_n = 0.01$. Results show that if noise is ignored, the dehazed result suffers from significant noise amplification. By either denoising prior to dehazing, or employing the proposed approach, we are able to significantly improve the quality of the dehazed image. Since the amount of noise is known a priori for this example, we are able to optimize both algorithms such that they minimize the error compared to the noise-free case. Both methods perform well and are visually quite similar.

An example comparing the two approaches on a real noisy, hazy image is shown in Fig. 5. Again, directly dehazing the noisy image results in significant noise amplification in regions with strong haze. Denoising prior to dehazing and the proposed approach are both able to successfully suppress the noise in the dehazed image. However, the proposed approach is able to preserve more detail. Since the noise is not known in this case, we use the quality metric proposed in Ref. 24 to automatically tune our results.

(a) Input (noisy) Image (b) Direct Dehazing (c) Denoise+Dehaze (d) Proposed Method

Figure 4: (a) The input hazy image has added zero-mean Gaussian noise with standard deviation $\sigma_n = 0.01$ for an image whose maximum intensity values are normalized to 1. (b) Noise is considerably amplified if not addressed in the dehazing process. (c) The dehazed result after denoising with BM3D[13, 16] is visually similar in this case to (d) the dehazed result using the proposed iterative kernel regression approach.

(a) Input Image (b) Direct Dehazing (c) Denoise+Dehaze (d) Proposed Method

(e) Direct Dehazing (f) Denoise+Dehaze (g) Proposed Method

Figure 5: Results of directly dehazing a real noisy hazy image (b) show significant amplification of noise in areas of thick haze (e). While denoising with BM3D prior to dehazing suppresses the noise, some detail is lost (c)(f). In this case, the proposed iterative kernel regression method is better able to preserve detail in the dehazed result (d)(g)

5. CONCLUSIONS AND FUTURE WORK

In this paper we have addressed the problem of simultaneously removing haze and noise from a single image, and have made several novel contributions. The first is the adaptation of an existing technique, the dark channel

prior,[4] for estimating haze from a single **clean** hazy image to the case of a single **noisy** hazy image.

We presented two effective methods for final scene radiance recovery. The first method treats haze and noise separately by denoising the image as a pre-processing step to dehazing. The second method considers removing haze and noise simultaneously through an iterative process which alternates between an estimation of the scene radiance and the atmospheric light. We stop these iterations and choose a global smoothing parameter automatically using the quality metric proposed in Ref. 24. In comparing these two methods, we found that when the noise level is known exactly or can be precisely estimated, simply denoising (using a state-of-the-art method) prior to dehazing provides good results. However, the quality of this approach is sensitive to inexact levels of denoising, as would be the case in general practice. The proposed iterative method proved to be more robust, offering visually comparable results to the first method when the noise level is known, and better preserving details when the noise level is estimated.

There are a number of promising directions for future research. One possibility is further study of parameter selection. Another possibility is an extension of the results presented in this paper to different haze estimation methods. Although we have chosen here to employ the dark channel prior, there is no inherent reason why another faster or more accurate method cannot be employed, such as the method proposed in Ref. 25. Finally, extension of haze removal to noisy videos is another topic that deserves attention. He et al. have supplied video results in the supplementary material for Ref. 4, and more recently, Zhang et al.[26] found performance gains by considering the video frames together. Noise, however, was not addressed in their haze removal process, and so is an open problem.

REFERENCES

[1] Joshi, N. and Cohen, M., "Seeing mt. rainier: Lucky imaging for multi-image denoising, sharpening, and haze removal," *IEEE International Conference on Computational Photography (ICCP)* , 1 –8 (March 2010).

[2] Schechner, Y. and Averbuch, Y., "Regularized image recovery in scattering media," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**, 1655–1660 (2007).

[3] Kaftory, R., Schechner, Y., and Zeevi, Y., "Variational distance-dependent image restoration," *CVPR* (2007).

[4] He, K., Sun, J., and Tang, X., "Single image haze removal using dark channel prior," *CVPR* , 1956–1963 (2009).

[5] Schechner, Y. Y., Narasimhan, S. G., and Nayar, S. K., "Instant dehazing of images using polarization," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], **1**, 325332 (2001).

[6] Shwartz, S., Namer, E., and Schechner, Y. Y., "Blind haze separation," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition,*], **1**, 1984–1991 (2006).

[7] Nayar, S. and Narasimhan, S., "Vision in bad weather," in [*Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*], **2**, 820 –827 vol.2 (1999).

[8] Narasimhan, S. G. and Nayar, S. K., "Interactive deweathering of an image using physical models," in [*IEEE IEEE Workshop on Color and Photometric Methods in Computer Vision, In Conjunction with ICCV*], (October 2003).

[9] Kopf, J., Neubert, B., Chen, B., Cohen, M., Cohen-Or, D., Deussen, O., Uyttendaele, M., and Lischinski, D., "Deep photo: Model-based photograph enhancement and viewing," in [*ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2008)*], **27**(5), 116:1–116:10 (2008).

[10] Tan, R., "Visibility in bad weather from a single image," in [*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*], 1 –8 (June 2008).

[11] Fattal, R., "Single image dehazing," *ACM Transactions on Graphics* **27** (August 2008).

[12] Levin, A., Lischinski, D., and Weiss, Y., "A closed-form solution to natural image matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**, 228–242 (2008).

[13] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K., "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Proc.* **16**, 2080 –2095 (2007).

[14] Abebe, A., Daniels, J., McKean, J., and Kapenga, J., [*Statistics and Data Analysis*], Department of Statistics, Wester Michigan University, Michigan (2001).

[15] Kushner, H. and Yin, G., [*Stochastic Approximation and Recursive Algorithms and Applications*], Springer-Verlag, New York (2003).

[16] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K., "Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space," *Proc. IEEE Int. Conf. on Image Process. (ICIP)* (2007).

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

[18] Kay, S. M., [*Fundamentals of Statistical Signal Processing: Estimation Theory*], Prentice-Hall, Inc., Upper Saddle River, NJ, USA (1993).

[19] Wand, M. and M.Jones, [*Kernel Smoothing*], Chapman and Hall (1995).

[20] Hardle., W., [*Applied Nonparametric Regression*], Cambridge Univ. Press, Cambridge, U.K. (1990).

[21] Hardle, W., M.Muller, Sperlich, S., and A.Werwatz, [*Nonparametric and Semiparametric Models*], Springer (2004).

[22] Takeda, H., Farsiu, S., and Milanfar, P., "Kernel regression for image processing and reconstruction," *Image Processing, IEEE Transactions on* **16**, 349 –366 (Feb. 2007).

[23] Matlin, E., *Single Image Haze and Noise Removal*, Master's thesis, Department of Electrical Engineering, University of California, Santa Cruz, June 2011.

[24] Zhu, X. and Milanfar, P., "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *Image Processing, IEEE Transactions on* **19**, 3116 –3132 (Dec. 2010).

[25] Tarel, J.-P. and Hautiere, N., "Fast visibility restoration from a single color or gray level image," in [*Computer Vision, 2009 IEEE 12th International Conference on*], 2201 –2208 (Oct. 2009).

[26] Zhang, J., Li, L., Zhang, Y., Yang, G., Cao, X., and Sun, J., "Video dehazing with spatial and temporal coherence," *The Visual Computer* **27**, 749–757 (2011). 10.1007/s00371-011-0569-8.