

Investigating Haze-relevant Features in A Learning Framework for Image Dehazing

Ketan Tang *

Hong Kong Univ. Science and Technology
tkt@connect.ust.hk

Jianchao Yang

Adobe Research
jiayang@adobe.com

Jue Wang

Adobe Research
juewang@adobe.com

Abstract

*Haze is one of the major factors that degrade outdoor images. Removing haze from a single image is known to be severely ill-posed, and assumptions made in previous methods do not hold in many situations. In this paper, we systematically investigate different haze-relevant features in a learning framework to identify the best feature combination for image dehazing. We show that the dark-channel feature is the most informative one for this task, which confirms the observation of He *et al.* [8] from a learning perspective, while other haze-relevant features also contribute significantly in a complementary way. We also find that surprisingly, the synthetic hazy image patches we use for feature investigation serve well as training data for real-world images, which allows us to train specific models for specific applications. Experiment results demonstrate that the proposed algorithm outperforms state-of-the-art methods on both synthetic and real-world datasets.*

1. Introduction

Haze is an atmospheric phenomenon where fog, dust, smoke and other particles obscure the clarity of the scene. Outdoor images are often contaminated by haze, even on a sunny day. Haze removal, or dehazing, is desired in both consumer photography and computer vision applications, thus has been extensively studied.

Early approaches often require certain types of additional information to be available. For instance, Tan *et al.* [17] assume the scene depth is given. Kopf *et al.* [10] use existing 3D geographic models of the scene for dehazing. Polarized filters are used in [15, 16] to capture multiple images of the same scene, and then different degrees of polarization (DOP) of images are used for haze removal. Nayar and Narasimhan [14, 13] also capture multiple images of the same scene and use the differences of images for estimating the haze properties. While these methods can enhance

the visibility of hazy images, they cannot be applied in applications where additional information or multiple images are not available. Therefore, single image haze removal has been a hot spot of research given its wider application range [5, 18, 8, 12, 19, 2, 7, 6].

Different image priors have been explored for single image dehazing in previous methods. Tan *et al.* [18] directly maximize the contrast of the dehazed image since haze-free images have higher contrast than hazy ones. The results of this method often present severe color distortion as the method is not physically valid. Fattal [5] assumes the transmission and surface shading are locally uncorrelated and estimate the haze by independent component analysis. It is physically valid but the assumption is too strong for a variety of images, thus it tends to under-estimate the haze thickness in practice. He *et al.* [8] propose a dark channel prior for natural outdoor images, which asserts that the local minimum of dark channel (minimum of R, G, B channels) of a haze-free image is close to zero. Dark channel prior is physically sound and generates good results, and [8] has been the state-of-the-art algorithm for general image dehazing. Most recent algorithms improve [8] only in some particular aspects. For example, Tarel and Hautiere [19] replace matting with “median of median along lines” filter for efficiency, Gibson *et al.* [7] with standard median filter, Yu *et al.* [21] with joint bilateral filter; Tarel *et al.* [20] adds planar constraint for road images; Caraffa *et al.* [4] adds planar constraint and noise model.

Despite the remarkable progress on single image haze removal, results generated by state-of-the-art methods are often not satisfactory, as the different priors they use fail on different real-world images. For example, the most successful dark channel prior will break when the input image contains certain structures like a white wall. As we will show later, methods based on the dark channel prior tend to over estimate the thickness of haze, causing result images to be too dark and prone to color distortions, especially in sky regions.

In this paper, we systematically investigated a variety of haze-relevant features in a regression framework based on

*The work was done when Ketan was an intern at Adobe.

Random Forest [3]. In order to learn the regression model, we need to collect haze-free outdoor images with accurate scene depth maps as ground truth, which, unfortunately, are hard to obtain. Instead, we synthesize hazy patches from clean image patches randomly sampled from high quality, haze-free images (not necessarily outdoor images). Surprisingly, such synthetic data turns out to be effective in training the dehazing model. This offers our model the flexibility to learn adaptive models for specific situations, such as heavy haze cases, light haze cases, landscape images, scenery images, and so on.

Based on the regression model and the training strategy, we develop a more robust dehazing algorithm that outperforms previous methods. The learning framework automatically discovers high order relationships between different features for more accurate haze estimation. From the output of the Random Forest regressor, we confirm that the dark-channel feature is the most informative feature for haze estimation, as discovered by He *et al.* [8]. More importantly, we find that other haze-relevant features also contribute noticeably in a complementary way. Experimental results show that by combining multiple features, our proposed approach achieves higher quality results than state-of-the-art methods on both real-world examples and synthetic examples, qualitatively and quantitatively.

2. Problem formulation

The haze image formation model proposed by Koschmieder [11] has been widely used in previous haze removal work [5, 8, 19, 7]:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})), \quad (1)$$

where \mathbf{I} is the observed hazy image, \mathbf{J} is the real scene to be recovered, t is the medium transmission, \mathbf{A} is the global atmospheric light, and \mathbf{x} denotes pixel coordinates. The transmission $t = e^{-kd(\mathbf{x})}$ describes the portion of the light that is not scattered and reaches the camera, where $d(\mathbf{x})$ is the distance from the scene point to the camera, and k is the scattering coefficient of the atmosphere. The problem of haze removal is to estimate \mathbf{J} , t , and \mathbf{A} from a single input image \mathbf{I} . This is difficult as vectors \mathbf{A} , $\mathbf{I}(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ are coplanar and their ending points are collinear in RGB space geometrically [8].

The real scene \mathbf{J} can be recovered if we know atmospheric light \mathbf{A} and transmission t :

$$\mathbf{J} = \frac{\mathbf{I} - \mathbf{A}}{t} + \mathbf{A}. \quad (2)$$

For the time being, we assume we already have a good estimation of the atmospheric light \mathbf{A} . Examining the physical properties of Eqn. (1), we can derive the following con-

straints for t (\mathbf{I} is normalized between 0 and 1):

$$0 \leq t \leq 1, \quad (3)$$

$$0 \leq J^c \leq 1, \forall c \in \{r, g, b\}, \quad (4)$$

where J^c denotes the color channel c . In Eqn. (3) 0 and 1 are the physical bounds for t , and in Eqn. (4) we constrain J^c in $[0, 1]$ to avoid undershoot and overshoot of result image (which would result in texture losing). Plugging in Eqn. (2), the constraints in Eqn. (3, 4) translate to

$$\max \left\{ 1 - \min_c \frac{I^c}{A^c}, 1 - \min_c \frac{\tilde{I}^c}{\tilde{A}^c} \right\} \leq t \leq 1, \quad (5)$$

where $\tilde{I}^c = 1 - I^c$ is the inverse image of I^c and $\tilde{A}^c = 1 - A^c$. Eqn. (5) shows the large ambiguity in identifying the correct transmission t . Therefore, certain regularization or prior on the transmission t is needed for solving this ill-posed problem. One straightforward regularization could be to require that the recovered image \mathbf{J} should follow the natural image statistics, and estimating transmission t can be translated into an energy minimization problem:

$$\min_t \Phi(\mathbf{J}) \quad (6)$$

$$\text{s.t. } \max \left\{ 1 - \min_c \frac{I^c}{A^c}, 1 - \min_c \frac{\tilde{I}^c}{\tilde{A}^c} \right\} \leq t \leq 1, \quad (7)$$

$$\mathbf{J} = (\mathbf{I} - \mathbf{A})/t + \mathbf{A}, \quad (8)$$

where $\Phi(\cdot)$ is an energy function that penalizes solutions far away from the natural image statistics, *e.g.*, total variation. Depending on the particular form of Φ , the regularization term on t can be fairly complicated. Instead of designing a particular form of Φ , He *et al.* [8] assume that the transmission t is locally constant, and the lower bound $1 - \min_c I^c/A^c$ filtered by a “min” filter can serve as a tight lower bound of t . As a result, this method tends to over-estimate the thickness of the haze in real-world photos. Tarel and Hautiere [19] extend this idea with a “median of median along lines” filter on the lower bound of t , which lacks a principled explanation.

Finding a good form of Φ has long been a challenging problem in the area of image restoration. In this paper, instead of proposing yet another form of Φ , we approach the problem from a different angle. Specifically, we implicitly solve the above optimization problem by learning a direct mapping function from the input image \mathbf{I} to the transmission t based on some image features. To do so, we first investigate several haze-relevant features in Sec. 3 and then discuss how to learn the implicit mapping from these features in Sec. 4.

3. Haze-relevant features

As the first step of our approach, we extract a list of image features that are related to the properties of hazy images,

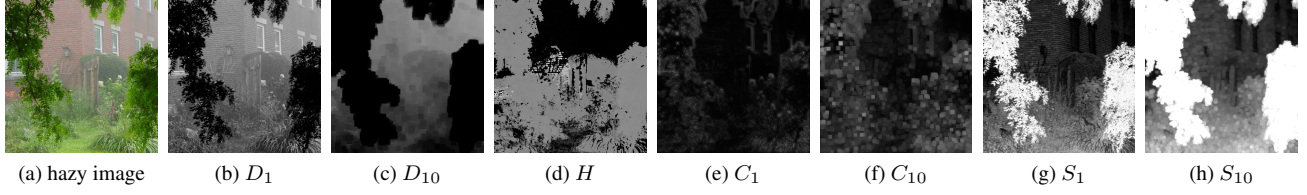


Figure 1: Different feature maps for the input hazy “wall” image. (a) hazy input; (b, c) dark channel feature D_1 and D_{10} ; (d) hue disparity feature; (e, f) local max contrast feature C_1 and C_{10} ; (g, h) local max saturation feature S_1 and S_{10} .

detailed as follows.

3.1. Multi-scale dark channel

The dark channel of an image [8] is defined as the minimum of all pixel colors in a local patch:

$$D_r(\mathbf{x}; \mathbf{I}) = \min_{\mathbf{y} \in \Omega_r(\mathbf{x})} \min_{c \in \{r, g, b\}} I^c(\mathbf{y}) / A^c, \quad (9)$$

where $\Omega_r(\mathbf{x})$ is the patch centered at \mathbf{x} with size $r \times r$. r affects the performance of dark channel feature: while a small r results in a loose bound on t , a large r leads to a too strong local constant assumption on t that will result in over-dehazing. Therefore, we combine $D_r(\mathbf{x}; \mathbf{I})$ with multi-scale r into a multi-scale dark channel feature, $D^k = [D_{r1}, D_{r2}, \dots, D_{rk}]$. In our experiments, we use four scales: $D^4 = [D_{10}, D_7, D_4, D_1]$. Figure 1b and 1c show the dark channel feature D_1 and D_{10} for the “wall” example. As we can see, the dark channel feature has a high correlation to the amount of haze in the image.

3.2. Multi-scale local max contrast

From the image formation model in Eqn. (1), the contrast of the image is reduced by the haze ($t \leq 1$):

$$\sum_{\mathbf{x}} \|\nabla \mathbf{I}(\mathbf{x})\| = t \sum_{\mathbf{x}} \|\nabla \mathbf{J}(\mathbf{x})\| \leq \sum_{\mathbf{x}} \|\nabla \mathbf{J}(\mathbf{x})\|. \quad (10)$$

Based on this observation, Tan’s method [18] enhances the visibility of the image by maximizing the local contrast. In this work, we define the local contrast as the variance of pixel intensities in a local $r \times r$ patch compared to the center pixel. We further compute the local maximum of local contrast values in a $s \times s$ region, as another haze-relevant feature:

$$C_r(\mathbf{x}; \mathbf{I}) = \max_{\mathbf{y} \in \Omega_r(\mathbf{x})} \sqrt{\frac{1}{3|\Omega_s(\mathbf{y})|} \sum_{\mathbf{z} \in \Omega_s(\mathbf{y})} \|\mathbf{I}(\mathbf{z}) - \mathbf{I}(\mathbf{y})\|^2}, \quad (11)$$

where $|\Omega_s(\mathbf{y})|$ is the cardinality of the local neighborhood $\Omega_s(\mathbf{y})$. In our work, s is fixed to be 5 and r is a parameter that will vary. This feature, called multi-scale local max contrast, is formed by stacking C_r with different r . Again, we use four scales in our experiments: $C^4 = [C_{10}, C_7, C_4, C_1]$. Figure 1e and 1f show the contrast feature C_1 and C_{10} for the “wall” image. The correlation between the contrast feature and haze is visually obvious, although is not as strong as the dark channel feature.

3.3. Hue disparity

Hue disparity between the original image and its semi-inverse image has been used to detect haze in [2], where the semi-inverse image is defined as the maximum of the original image and its inverse:

$$I_{si}^c(\mathbf{x}) = \max[I^c(\mathbf{x}), 1 - I^c(\mathbf{x})], c \in \{r, g, b\}. \quad (12)$$

The hue disparity feature is defined as the hue disparity between the original image and its semi-inverse image:

$$H(\mathbf{I}) = |I_{si}^h - I^h|, \quad (13)$$

where the superscript “h” denotes the hue channel of the image in Lch color space. $H(\mathbf{I})$ is normalized to $[0, 1]$ by dividing it by the maximum 360. For haze-free pixels, not all three semi-inverse values will flip from the original ones, which causes large hue change between I_{si} and \mathbf{I} . On the other hand, for hazy pixels, all three semi-inverse values will flip, thus there will not be a hue change. Hue disparity map H is usually noisy and needs to be filtered to reduce noise. As the noise in H is mainly impulse, we apply a median filter to H . Figure 1d shows the hue disparity map H , which is correlated with haze in the image.

3.4. Multi-scale local max saturation

Similar to image contrast, image saturation is also reduced by haze. We define local max saturation, the maximum of pixel-wise saturation values in a local $r \times r$ image patch, as another haze-relevant feature:

$$S_r(\mathbf{x}; \mathbf{I}) = \max_{\mathbf{y} \in \Omega_r(\mathbf{x})} \left(1 - \frac{\min_{c \in \{r, g, b\}} I^c(\mathbf{y})}{\max_{c \in \{r, g, b\}} I^c(\mathbf{y})} \right). \quad (14)$$

Similar to the image contrast feature, we use multi-scale local max saturation in our experiments: $S^4 = [S_{10}, S_7, S_4, S_1]$. Figure 1g and 1h show the local max saturation features S_1 and S_{10} for the “wall” image, which have strong correlations with the amount of haze in the image.

3.5. Observations

By visually examining all the feature maps for the “wall” image shown in Figure 1, we can see that the dark channel and the local max saturation features are most related to the

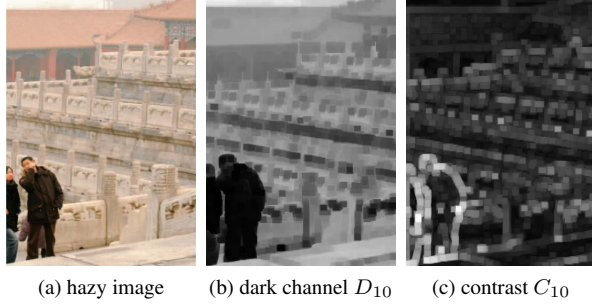


Figure 2: Dark channel prior fails when the scene object has a similar color to the atmospheric light, while the contrast feature can serve as a good complement.

haze thickness. Dark channel is positive correlated with the amount of haze, while color saturation is negative correlated. Hue disparity can serve as a haze detection feature. The local max contrast feature is a little scattered and its correlation with haze is not as strong as other color-based features for this example. However, as we observed through experiments, it is an important complement to other color-based features. In Figure 2, we show the dark channel and contrast feature maps of an example used in [8], where the dark channel prior fails since the color of the marble is similar to the atmospheric light. In this image, the haze thickness of the person and the marble railing right behind him is approximately the same. However, the dark channel feature estimates much more haze for the marble railing than for the person. The contrast feature, instead, can give us better haze estimation in this case.

4. Learning the haze removal model

Based on the above haze-relevant features, we can use Random Forest to learn a regression model for estimating the transmission t for hazy images. Random Forest is chosen for its simplicity and easy analysis of feature importance. Other regression algorithms may also be used. The difficulty for this learning is that it is very hard to get good training data, *i.e.*, outdoor haze-free and hazy image pairs. Another way to prepare the training data, based on the physical property $t(x) = e^{-kd(x)}$, is to synthesize haze-free and hazy image pairs given the accurate outdoor scene depth, which, unfortunately, is also hard to obtain. In this paper, we synthesize our training data without relying on the scene depth as described in the following.

4.1. Training data preparation

Our training data synthesis is based on two assumptions: (1) image content is independent of scene depth or medium transmission, *i.e.*, the same image content can appear at different depth in different images; and (2) depth is locally constant, *i.e.*, image pixels in one small patch tend to have similar depth values. Although we cannot synthe-



Figure 3: Example haze-free training images collected from the Internet.

size the relative depth of all pixels in an image, for a given single image patch, we can assume arbitrary depth based on the above two assumptions. This translates to the following preparation procedure: given a clean patch p_J , the atmospheric light A , and a random transmission $t \in [0, 1]$, we synthesize a hazy patch p_I as $p_I = tp_J + (1 - t)A$. To reduce the uncertainty of variables in learning, we set the atmospheric light $A = [1, 1, 1]$. The clean patches are randomly sampled from haze-free natural images collected from the Internet. Figure 3 shows some example training images. Note that the training images do not have to be landscape or cityscape images, where haze mostly incurs.

Learning the regression model with Random Forest is straightforward: the inputs to Random Forest are haze-relevant features extracted from the synthetic hazy patches and the outputs are their corresponding transmission t 's. However, we know that our synthetic transmission is irrelevant to the local image texture, but our haze-relevant features are closely related to the image content. To break the correlation between our haze-relevant features and the image content, we sort the values within each scale of each feature type before sending them to Random Forest. By doing so, we disturb the image content so the learned models will not be content-specific. Therefore the true relation between features and transmission can be better revealed by the learned random forest model.

4.2. Haze removal with the regression model

Atmospheric light A estimation. The first step of dehazing is estimating the atmospheric light. In [8] the brightest pixel value in the 0.1% pixels with largest dark channel values is taken as A . This method considers only a single pixel, thus it may be affected by noise and result in color distortion. We improve this method by taking the median of all the 0.1% pixels with largest dark channel values. Figure 4 compares the two A estimation methods. We can see that our method does not suffer from color distortion, while previous approaches do.

Testing with Random Forest. Testing with Random Forest is straightforward. First, we apply white balance correction with the estimated A to the input image I . By that the testing patches will have pure white atmosphere light $[1, 1, 1]$, which is consistent with the training procedure. We then partition the image into 5×5 overlapping patches

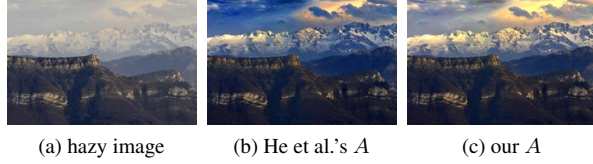


Figure 4: Comparisons of different atmospheric light estimation methods. Note the strong color distortion in the sky region in (b).

and extract haze-relevant features. We aggregate the patch-wise transmission estimations to get a transmission map, on which we further apply guided filter [9] to suppress the blocky artifacts. Finally the real scene can be recovered using Eqn. (2) and white balance restoration.

4.3. Robust postprocessing

With the previously described framework, some recovered image regions might look too dim since the atmospheric light is usually brighter than the scene radiance for some images. For better visual quality, we provide two robust postprocessing options for practical applications.

Adaptive atmospheric light. We notice that for some images the “uniform atmospheric light” assumption does not hold. As a result, the darker regions in these images become too dark after dehazing and image details in them are lost. Based on this observation, we make adaptive small adjustments to the atmospheric light \mathbf{A} according to the input image brightness:

$$\min_{\mathbf{A}} \sum_{\mathbf{x}} \{(\mathbf{Y}_J(\mathbf{x}) - \mathbf{Y}_I(\mathbf{x}))^2 + \lambda(\mathbf{A}(\mathbf{x}) - \mathbf{A}_0)^2\} + \phi(\mathbf{A}), \quad (15)$$

where \mathbf{A}_0 is the initial estimation of the atmospheric light, \mathbf{Y}_J is the illuminance of \mathbf{J} , \mathbf{Y}_I is the illuminance of \mathbf{I} , and $\phi(\cdot)$ is a smoothness regularization. Exactly solving the above optimization is difficult, instead we approximately solve it using a two-step approach: (1) solve \mathbf{A} without the smoothness regularization, which has a closed-form solution; and (2) apply guided filter GF_I to smooth the solution. Let $\beta = 1/t - 1$, we have

$$\mathbf{A}(\mathbf{x}) = \text{GF}_I \left\{ \frac{\beta^2 \mathbf{Y}_I(\mathbf{x}) + \lambda \mathbf{A}_0}{\beta^2 + \lambda} \right\}. \quad (16)$$

Adaptive exposure scaling. Another way to resolve this issue is to adaptively increase the exposure of the image guided by the input:

$$\min_s \sum_{\mathbf{x}} \left\{ \left(1 - s(\mathbf{x}) \frac{\mathbf{Y}_J(\mathbf{x})}{\mathbf{Y}_I(\mathbf{x})}\right)^2 + \lambda(s(\mathbf{x}) - 1)^2 \right\} + \phi(s), \quad (17)$$

where s is the illuminance scaling field. Similarly, we can get a fast approximate solution as

$$s(\mathbf{x}) = \text{GF}_I \left\{ \frac{\mathbf{Y}_J(\mathbf{x}) \mathbf{Y}_I(\mathbf{x}) + \lambda \mathbf{Y}_I^2(\mathbf{x})}{(\mathbf{Y}_J^2(\mathbf{x}) + \lambda \mathbf{Y}_I^2(\mathbf{x}))} \right\}. \quad (18)$$

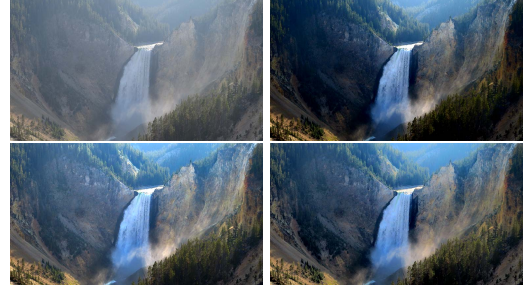


Figure 5: Robust post processing. Top left to bottom right: hazy image, direct dehazing, adaptive atmospheric light, adaptive exposure scaling.

In practice we found that the above two approaches perform similarly. The adaptive atmospheric light works slightly better for weak haze cases, while illuminance scaling works slightly better for heavy haze cases. Figure 5 shows one example where post processing helps with the visual quality of the dehazed results. Adaptive atmospheric light and adaptive exposure scaling performs similarly.

5. Experiment results

We mainly compare our approach with two existing methods: He *et al.* [8], which is considered to be the state-of-the-art method in literature, and Kolor Neutralhazer plugin for Photoshop, which is a well-received commercial software [1]. For [8] guided filter is used to smooth the transmission map. The comparison datasets include synthetic hazy patches, synthetic images from stereos, and real-world images. We evaluate the performance of different algorithms both quantitatively and qualitatively. For Random Forest, we find 200 trees are more than enough for our regression problem. The algorithm is not sensitive to the number of features selected for growing each tree, thus we fix it to be one third of feature dimension. For a fair comparison with the other methods, we do not apply the postprocessing techniques proposed in Sec. 4.3.

5.1. Quantitative results on synthetic patches

We randomly sampled 7000 image patches from our training images, and take 5000 for training and 2000 for testing. For each patch, we uniformly sample 10 random $t \in [0.1, 1]$ to generate 10 hazy patches (we do not use very small t to avoid noise boosting, as in He *et al.* [8]). Therefore, we have in total 50,000 training patches, and 20,000 testing patches.

Effectiveness of regression. Figure 6 plots the predicted transmission vs. the ground truth transmission for our approach and He *et al.*’s method [8] on the testing patches. We can see that in our result the predicted transmission $t_{predict}$ centers around the the 45° line. However, in He *et al.*’s result the predicted transmission $t_{predict}$ is always lower than (sometimes much lower) or equal to the true trans-

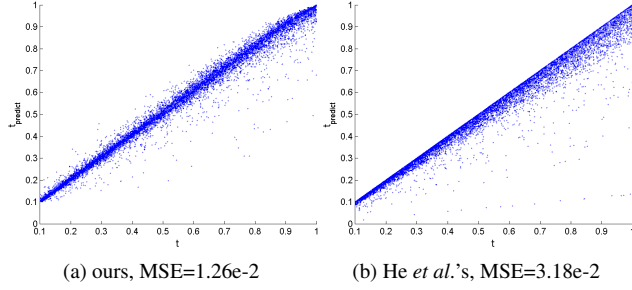


Figure 6: Predicted transmission vs. ground truth for both methods on the test hazy patches.

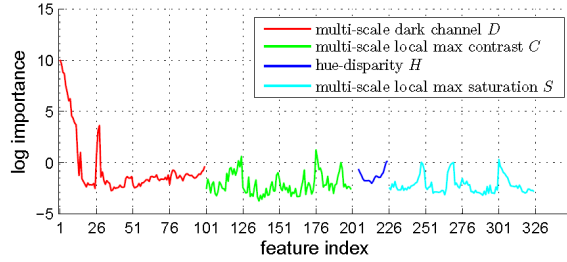


Figure 7: Importance of different features output by the Random Forest regressor.

mission t , which will result in over-dehazing results. As we mentioned before, this is because He *et al.*'s method uses the lower bound of t filtered by a “min” filter. The mean squared error (MSE) of our method is $1.26e-2$, and that of He *et al.*'s method is $3.18e-2$, which is more than twice as large as ours.

Significance of different features. Figure 7 shows the feature “importance” curve output by the learned Random Forest regressor. Note that to show the trend more clearly, we plot the importance score in the log scale. We can see that the multi-scale dark channel features are most important among different features, and the largest scale dark channel features D_{10} are the most important ones, which confirms the dark channel prior. However, our regression model is not relying on a single dark channel value as in [8]; it captures the high-order relationship among the dark channel features in a local patch (5×5), and also the high-order relationship between different kinds of features within the local neighborhood. Therefore, our regression model can do a better transmission estimation. Another observation from the figure is that the importance plot demonstrates some periodic patterns. This is because each kind of features are sorted within its own scales and the regressor tends to rely more on the maximum or minimum statistics of these features.

In Figure 8a, we extensively evaluate the effectiveness of different multi-scale features in terms of test MSE of the Random Forest regressor trained with different feature settings. Symbols D, C, H, S represent the four types of features proposed in Section 3. Superscript 4) represents the

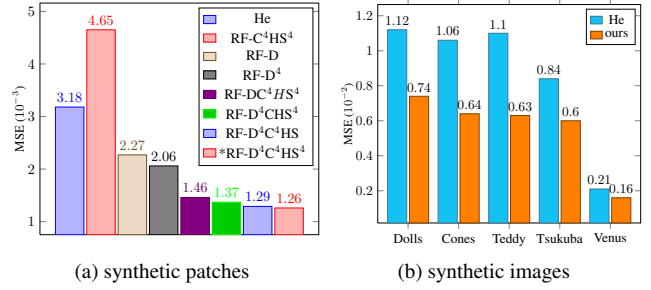


Figure 8: MSE on (a) synthetic test patches and (b) synthetic hazy images from stereo images.

number of scales, while no superscript means a single scale. Training on single scale dark channel features, the model (RF-D) already outperforms He *et al.*'s method. As mentioned earlier, this is because our regressor uses all dark channel features in a local patch rather than only the smallest one as in He *et al.*'s method. Training on features other than the dark channel features, the model (RF- C^4HS^4) is not as well as others. However, by combining single scale dark channel features and all other features, the model (RF- DC^4HS^4) significantly reduces the MSE compared to RF-D or RF- C^4HS^4 , indicating that other features also contribute remarkably in a complementary way to the dark channel features. By adding more scales to each feature type, our model can further reduce the test MSE, and our full model achieves significantly lower MSE compared to He *et al.*'s method.

5.2. Quantitative comparisons on synthetic images

For quantitative evaluation on complete images, we synthesize hazy images from stereo images with known disparity map d . Note that instead of using the physical model (1), we simply define the transmission map as $t = 0.8d$. The reason we can do so is that, the depth of indoor stereo images have very small variation, thus a small scattering coefficient need to be chosen to avoid overly flat haze map. With small scattering coefficient, the exponential function can be well approximated by a linear function. As our method is local, how t is generated does not really affect the conclusion.

The hazy image is then generated as $I = tJ + (1 - t)$, where we assume pure white atmospheric airlight $A = [1, 1, 1]$. Figure 8b shows the MSE of our transmission estimation compared with He *et al.*'s algorithm on five commonly used stereo images. The proposed method outperforms He *et al.*'s approach in all cases.

Figure 9 shows visual comparisons of different methods on two examples: “Dolls” and “Teddy”. He *et al.*'s method tends to overestimate the haze thickness, resulting in darker results and color distortions. Note the color distortion in the yellow rectangle areas in He *et al.*'s results. Kolor has the opposite problem of underestimating the haze thickness,

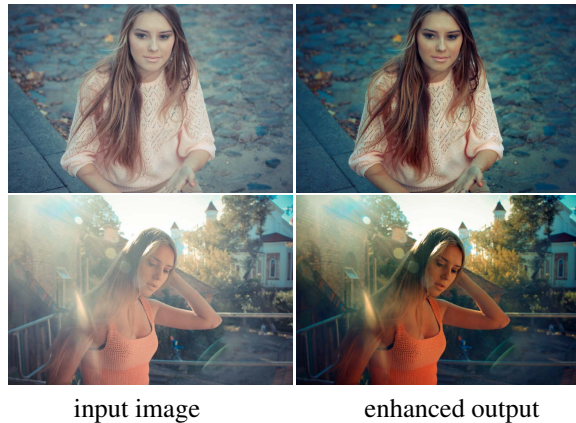


Figure 11: Image enhancement with our light haze model.

and there is still a large amount of haze remaining in its results. In contrast, our results have higher visual quality and less color distortion.

5.3. Visual comparisons on real-world images

Figure 10 shows visual comparisons of different algorithms on a few real-world examples. In these examples, He *et al.*'s algorithm again overestimates the thickness of the haze and produces dark images with noticeable color distortion. Besides over-dehazing, the color distortion of He *et al.*'s results also comes from inaccurate atmospheric light A estimation, which is sensitive to noise. Kolor's method produces results that still have significant remaining haze. Our results recover most scene details and maintain the original colors. More results can be found in the project page¹.

5.4. Image enhancement with light haze model

As dehazing increases image contrast and color saturation, it can be used for enhancing ordinary images. One advantage of our learning-based approach is that we can train specific models for specific applications. For general image enhancement, we train a model for very light haze cases by constraining the training transmission t to be in $[0.5, 1]$. To validate its effectiveness, we collect several images where the main objects are close to the camera, and apply our model for automatic enhancement. Figure 11 shows two such examples, which suggests that the color saturation and contrast have been greatly improved using our algorithm.

6. Conclusion

We have proposed a learning-based approach for robust dehazing. Various haze-relevant multi-scale features are extracted and automatically evaluated. We found dark-channel is the most useful feature, however other features also help significantly for achieving more accurate haze estimation in our learning framework. Experimental results

show that the proposed algorithm performs better than state-of-the-art methods on both synthetic data and real-world images. Our model also has some limitations. First, due to the low signal-to-noise ratio in heavy haze cases, our dehazing model may boost noise, similar to previous works. Second, the model is still based on local cues without knowing the context, and transmission smoothing is needed to suppress blocky artifacts. As future work we want to explore how to use semantic information, e.g. scene structure analysis, and how to combine adaptive denoising for better dehazing.

References

- [1] Kolor neutralhazer plugin for photoshop. <http://www.kolor.com/>. 5
- [2] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert. A fast semi-inverse approach to detect and remove the haze from a single image. *ACCV*, 2011. 1, 3
- [3] L. Breiman. Random forests. *Mach. Learn.*, 45(1), 2001. 2
- [4] L. Caraffa and J. P. Tarel. Markov random field model for single image defogging. In *Intelligent Vehicles Symposium (IV)*, IEEE, 2013. 1
- [5] R. Fattal. Single image dehazing. *ACM SIGGRAPH*, 2008. 1, 2
- [6] K. B. Gibson and T. Q. Nguyen. Fast single image fog removal using the adaptive wiener filter. *ICIP*, 2013. 1
- [7] K. B. Gibson, D. T. Vo, and T. Q. Nguyen. An investigation of dehazing effects on image and video coding. *Image Processing, IEEE Transactions on*, 21(2), 2012. 1, 2
- [8] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. In *CVPR*, 2009. 1, 2, 3, 4, 5, 6
- [9] K. He, J. Sun, and X. Tang. Guided image filtering. *Pattern Analysis and Machine Intelligence, IEEE Trans.*, 2012. 5
- [10] 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. In *ACM SIGGRAPH Asia*, 2008. 1
- [11] H. Koschmieder. Theorie der horizontalen sichtweite. *Beitrage zur Physik der freien Atmosphere*, 1924. 2
- [12] L. Kratz and K. Nishino. Factorizing scene albedo and depth from a single foggy image. In *ICCV*, 2009. 1
- [13] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *Pattern Analysis and Machine Intelligence, IEEE Trans.*, 25(6), 2003. 1
- [14] S. Nayar and S. Narasimhan. Vision in bad weather. In *ICCV*, 1999. 1
- [15] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. In *CVPR*, 2001. 1
- [16] S. Shwartz, E. Namer, and Y. Y. Schechner. Blind haze separation. In *CVPR*, 2006. 1
- [17] K. Tan, K. Tan, and J. P. Oakley. Enhancement of color images in poor visibility conditions. In *ICIP*, 2000. 1
- [18] R. T. Tan. Visibility in bad weather from a single image. In *CVPR*, 2008. 1, 3
- [19] J. P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. In *ICCV*, 2009. 1, 2

¹<http://ihome.ust.hk/~tkt/CVPR2014/cvpr2014.htm>

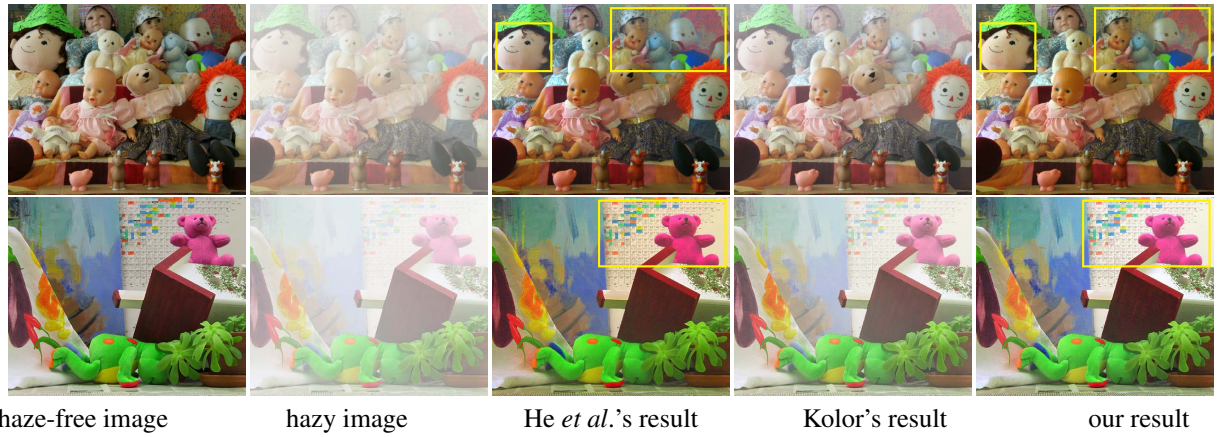


Figure 9: Result on stereo images. Notice the color distortion in the rectangle regions of He *et al.*'s results, and the under-dehazing effect of Kolor's.



Figure 10: Visual comparison of different results on real-world images.

[20] J. P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *Intelligent Transportation Systems Magazine, IEEE*, 4(2), 2012. 1

[21] J. Yu, C. Xiao, and D. Li. Physics-based fast single image

fog removal. In *Signal Processing (ICSP), IEEE 10th International Conference on*, 2010. 1