Automatic Recovery of the Atmospheric Light in Hazy Images

Matan Sulami   Itamar Glatzer   Raanan Fattal   Mike Werman

Hebrew University of Jerusalem

Abstract

Most image dehazing algorithms require, for their operation, the atmospheric light vector, A, which describes the ambient light in the scene. Existing methods either rely on user input or follow error-prone assumptions such as the gray-world assumption. In this paper we present a new automatic method for recovering the atmospheric light vector in hazy scenes given a single input image. The method first recovers the vector's orientation, $\hat{A} = A/\|A\|$, by exploiting the abundance of small image patches in which the scene transmission and surface albedo are approximately constant. We derive a reduced formation model that describes the distribution of the pixels inside such patches as lines in RGB space and show how these lines are used for robustly extracting $\hat{A}$.

We show that the magnitude of the atmospheric light vector, $\|A\|$, cannot be recovered using patches of constant transmission. We also show that errors in its estimation results in dehazed images that suffer from brightness biases that depend on the transmission level. This dependency implies that the biases are highly-correlated with the scene and are therefore hard to detect via local image analysis. We address this challenging problem by exploiting a global regularity which we observe in hazy images where the intensity level of the brightest pixels is approximately independent of their transmission value. To exploit this property we derive an analytic expression for the dependence that a wrong magnitude introduces and recover $\|A\|$ by minimizing this particular type of dependence.

We validate the assumptions of our method through a number of experiments as well as evaluate the expected accuracy at which our procedure estimates A as function of the transmission in the scene. Results show a more successful recovery of the atmospheric light vector compared to existing procedures.

1. Introduction

Haze and other types of atmospheric particles scatter light and reduce the visibility of the scene in mid- and long-ranged photography. Multiple events of light scatter-

Figure 1. Dehazed images obtained using the true atmospheric light vector A, using a vector with a wrong orientation (yet the correct magnitude $\|A\|$), and using vectors with the correct orientation, $A/\|A\|$, yet magnitudes which are too small and large.

ing in the scene introduce a layer of ambient light, known as airlight [10]. This layer blocks the direct scene transmission and reduces the image contrast. Many image dehazing methods rely on the following image formation model

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

where $I(\mathbf{x})$ is the input hazy image, $J(\mathbf{x})$ is the haze-free scene radiance image and $\mathbf{x} = (x, y)$ denotes pixel coordinates. Assuming that $I(\mathbf{x})$ and $J(\mathbf{x})$ are three color channel RGB images, Eq. (1) applies to each color channel independently. The transmission function, $t(\mathbf{x})$, is assumed by most methods to be the same in each color channel. Thus, it corresponds to a scalar function describing the visibility in the scene with values between zero (no visibility) and one (clear sight). The atmospheric light color vector A is an RGB vector that describes the intensity of the scattered light in the scene at each color channel. The total contri-

bution of the ambient light, the *airlight*, is described by the term $(1 - t(\mathbf{x}))A$ in Eq. (1). Finally, we note that Eq. (1) applied to radiometrically-linear images $I(\mathbf{x})$.

In order to dehaze a given image I by recovering J, according to Eq. (1), both the transmission t and the atmospheric light vector A must be estimated. While significant progress was made in estimating the transmission from a single image [21, 2, 4], the atmospheric light vector is either selected manually or computed using basic procedures that follow the gray-world assumption and take the average value of I or take its brightest pixel as the estimated A. Inaccurate estimation of A undermines the accuracy at which the transmission is estimated as well as the accuracy of the recovered image J. Fig. 1, shows the effects of using inaccurate atmospheric light vector orientation, $\hat{A} = A/\|A\|$, and magnitude $\|A\|$, when recovering J.

In this paper we describe a new method for recovering the atmospheric light vector A automatically from a single input hazy image I. The new method first recovers A's orientation, $\hat{A}$, based on a simplified formation model that describes pixels inside small image patches. This model assumes that both the transmission and the surface albedo functions are constant inside each patch. By finding a small number of patches obeying this model, $\hat{A}$ is recovered by simple geometric operations in the RGB colorspace. In Section 3.1 we derive the localized haze model, derive a procedure for robustly finding image patches that obey it and explain how it is used for recovering $\hat{A}$.

While patches of constant transmission and albedo are useful for recovering $\hat{A}$, we show that the magnitude of the atmospheric light vector, $\|A\|$, cannot be recovered from such patches. A simple analysis shows that the use of wrong magnitude leads to false variations in the brightness of the output dehazed image. These multiplicative biases depend on the transmission level of each pixel. As shown in Fig. 1, the use of a large $\|A\|$ makes the closer regions (of low transmission) overly bright and the farther ones too dark. The reverse effect takes place when using $\|A\|$ which is too small. While these biases are visually disturbing, they are hard to detect by local spatial measurements since the brightness biases are correlated with the transmission and the latter is correlated with the image content; the transmission function $t(\mathbf{x})$ is typically smooth where $J(\mathbf{x})$ is, and it is discontinuous along the edges in $J(\mathbf{x})$.

In order to cope with the challenging task of recovering $\|A\|$, we searched for a global image statistic that shows low variability over ensembles of natural hazy scenes and, at the same time, exhibits sensitivity to the dependency between pixel brightness and transmission. The function describing the brightest pixel as function of transmission level fits this bill and allows estimating $\|A\|$ more accurately than existing procedures, which often neglect this aspect when recovering A. In Section 3.2 we derive an analytic expression for the dependence between the pixel brightness and transmission as function of the error in the estimated $\|A\|$. We then derive a procedure that minimizes this particular type of dependence for recovering $\|A\|$.

We validate the assumptions of our method through a number of experiments as well as obtain an estimate for the expected accuracy of this procedure as function of the lowest transmission in the scene.

Finally, in Section 4 we report the evaluation of our method both on synthetic images, for which the true atmospheric light vector is known, as well as on real images acquired under hazy conditions. The results show a greater accuracy in the recovery of A than existing methods.

2. Previous Work

Here we briefly describe how the atmospheric light vector A is estimated by existing dehazing methods. We start with methods that use additional scene data beyond the hazy input image.

Tan and Oakley [20] assume the scene depth is given at every pixel and recover A according to the image formation model in Eq. (1). Given two images of a scene, each taken under different weather conditions, Narasimhan and Nayar [14] estimate the orientation of the atmospheric light vector, $\hat{A}$, from the difference between the two images. Given two or more images, taken at different polarization angles, Shwartz et al. [19] obtain scene measurements that differ in the amount of haze. Given several user selected pixels or regions that correspond to the same surface color at different depths, A is recovered. In a similar setting of polarized photography, Schechner et al. [18] estimate A from user-provided pixels that most resemble the atmospheric light color, such as the sky.

The following methods estimate A from a single hazy image. Narasimhan and Nayar [15] describe several user-interactive algorithms for deweathering images where the user provides a selection of pixels belonging to the same surface color at different depths or coarse scene depth information. Fattal [2] also describes a user-assisted approach where the user provides several image patches that correspond to different surface albedos. Given these patches, $\hat{A}$ is estimated by intersecting RGB planes that contain the patches' pixels. The success of this approach depends on the ability to find regions in the image where both the surface shading and transmission vary sufficiently.

In the context of fully-automated single-image procedures, the dark-object subtraction method [1] assumes a constant transmission in the scene, i.e., $t(\mathbf{x}) = t$, and extracts the uniform layer of ambient airlight from the darkest object in the scene. Oakly and Bu [16] also assume a constant transmission and extract the layer of haze based on an observation that in natural images the local sample mean of the pixel intensity is proportional to its standard deviation.

Additive shifts from this law provide an estimate for the airlight in the scene. Both approaches are limited to scenes where the haze optical depth is constant in which case dehazing can be performed by only recovering the airlight $(1 - t)A$ and without the need to explicitly recover A.

Tan [21] assumes that the brightest pixels in the image are saturated with haze and uses them to estimate A. In the dark-channel method [4, 24] the airlight A is estimated from the brightest pixels which are also bright in the dark channel. Tarel and Hautière [22] first white-balance the image according to the gray-world assumption and then dehaze the image using a pure-white airlight vector $A = [1, 1, 1]$.

In hazy scenes most of the surfaces are illuminated by the ambient light created from light scattering. Hence, estimating A corresponds to estimating the dominant illuminant in the scene. A large body of work exists on the recovery of the illuminant in non-hazy scenes. We mention here a few examples. Lee [12], as well as Lehmann and Palm [13], use specular highlights to estimate the scene illuminant. The color coordinates of the specular highlight fall on the straight line connecting the illuminant color point and the surface color point. Therefore, in scenes containing two or more surfaces of different color the illuminant chromaticity is recovered by intersecting the lines. Finlayson et al. [3] find a plausible illuminant by checking whether the image pixels could be produced by a set of plausible light sources. This method does not recover the brightness of the illuminant. Finally, Zhou and Kambhamettu [25] use plastic spheres that reflect the illumination in their highlights and find the light sources by stereo vision.

While all these methods achieve remarkable results over clear-sighted scenes, they are not adapted to the setting we are focusing on. Namely, specular highlights are not likely to appear in hazy scenes where the dominant illuminant is an ambient light which is spread across the entire scene. Similarly, in our work we avoid any intervention in the scene and assume a single *given* input image.

3. New Method

Our method recovers the atmospheric light vector A in two stages. At the first step we estimate its orientation, $\hat{A} = A/\|A\|$, from small patches of $I(\mathbf{x})$ that admit a simplified form of Eq. (1). In Section 3.1 we derive this reduced model, explain how patches obeying it are found and how they are used for recovering $\hat{A}$. In the second stage we describe a global image prior which is sensitive to dependencies created from the use of a wrong $\|A\|$ when dehazing an image. In Section 3.2 we describe a simple procedure for recovering $\|A\|$ by minimizing these dependencies.

3.1. Estimating the Atm. Light Orientation $A/\|A\|$

The visibility in light-scattering scenes depends mainly on the distance $d(\mathbf{x})$ between the camera and the surfaces
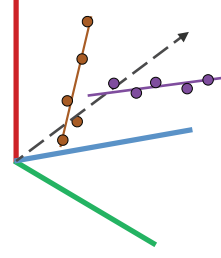

Figure 2. Pixels from two patches are shown in RGB space (brown and purple circles). While these lines do not intersect one another, both intersect the ray passing through the origin in the direction of $\hat{A}$ (dashed line). This geometric constraint provides us a canididate for $\hat{A}$ from every pair of patches found.

as well as on the amount of light scattering in the medium, measured by the scattering coefficient β. Under these considerations, the scene transmission is given by

$$t(\mathbf{x}) = \exp\left(-\int_0^{d(\mathbf{x})} \beta(\mathbf{r}_\mathbf{x}(s)) ds\right), \qquad (2)$$

where $\mathbf{r}_\mathbf{x}(\cdot)$ parameterizes the ray projected to the $\mathbf{x}$ pixel and $\beta(\cdot)$ describes the scattering coefficient in the 3D scene. Intuitively, Eq. (2) accumulates, in a multiplicative manner, the amount of visibility along the viewing ray of each pixel due to light scattering.

Images of natural environments typically capture distinct objects, each with its own limited range of depth values. Indeed, studies of range images support a collage model that describes the world as piecewise smooth regions, see [7]. Furthermore, the density of water droplets, dust and other aerosols that determine β vary smoothly in the atmosphere due to the diffusion processes that govern them [5]. Hence, according to Eq. (2), the transmission $t(\mathbf{x})$ is also expected to be piecewise smooth. Therefore, it is likely that in many small image patches the transmission is approximately constant.

The direct scene transmission $J(\mathbf{x})$ can be factored into the product $J(\mathbf{x}) = l(\mathbf{x})R(\mathbf{x})$ where R is an RGB vector describing the chromaticity of the light reflected from the surface (hence we assume $\|R\| = 1$), and l is a scalar describing the intensity of the reflected light. While this model applies to more general surfaces, in case of diffuse surfaces R corresponds to the surface reflectance coefficients, or albedo, and l to the incident light projected on the surface (up to an arbitrary factor between the two). For simplicity we refer to R as the surface reflectance or albedo and to l as the surface shading or brightness.

Similarly to the scene transmission, statistics of natural images suggest that RGB pixels can be viewed as a composition of distinct objects according to the dead-leaves model [11]. See also [8].

input image   projected lines and $\hat{A}$ lines estimated in different patches
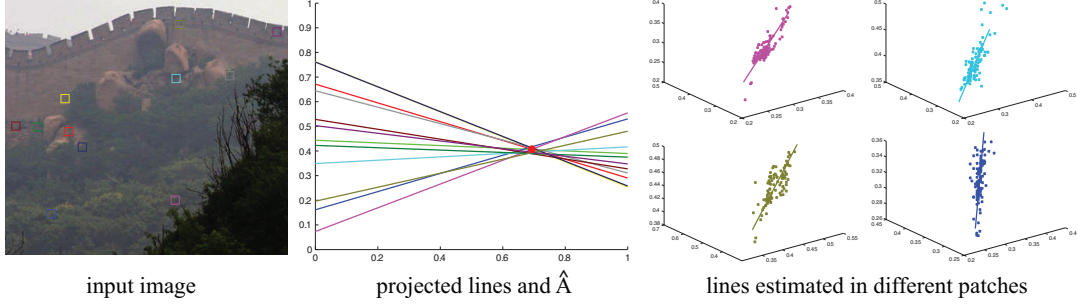
Figure 3. Example patches. Left image shows several patches found to obey our reduced model. The graphs in the middle show the patch lines visualized by their projection onto a plane perpendicular to the estimated $\hat{A}$. The red dot shows the intersection of the plane with $\hat{A}$. At the right we see the pixels and the lines found in four patches.

Being a component of $I(\mathbf{x})$, the surface albedo R inherits this property (arguably to a greater extent as it lacks the shading component). Thus, similarly to the scene transmission, it is likely that the surface albedo is approximately constant in many small patches of the image. This assumption is successfully used in the context of image dehazing in [18, 15] and [2].

Altogether, these observations motivate us to derive the following simplified formation model for small image patches,

$$I(\mathbf{x}) = t_i l'(\mathbf{x}) R_i + (1 - t_i) A = l(\mathbf{x}) R_i + c_i A, \quad (3)$$

where t_i is the transmission at pixels $\mathbf{x} \in \Omega_i$ of the i-th patch Ω_i and R_i is the reflectance vector in the patch. We absorb the patch transmission into the shading function and denote the latter by l.

We explain below how we find small image patches that obey Eq. (3) and proceed by explaining how the atmospheric light's orientation $\hat{A}$ is estimated once such patches are found.

According to Eq. (3) the pixels of such patches lie on a line in RGB space. Unlike the lines recovered in [12] and [13] from specular highlights in non-hazy scenes, the patch lines we recover are not expected to intersect. Nevertheless, the line of each patch intersects the ray that passes through the origin and whose orientation is $\hat{A}$, i.e., $s\hat{A}$, where the intersection occurs at $l = 0$ and $s = c_i \|A\|$. Thus, by extracting the lines in two or more patches and considering the planes containing the lines and the origin, we obtain an estimate for $\hat{A}$ which is given by the planes' intersection. This geometric scenario is illustrated in Fig. 2. Note that this procedure cannot recover the magnitude of the atmospheric light vector, $\|A\|$. This follows from the fact that c_i and A appear in a product form in Eq. (3), allowing them to exchange a non-zero multiplicative factor and its inverse.

While the minimal requirement of this procedure is two lines, we use multiple patches to recover $\hat{A}$ by the following robust procedure. We intersect the planes defined by every pair of patches found and obtain a list of candidate orientation vectors $\hat{A}_j$. We then compute the distances d_{ij} between the i-th patch line and the j-th ray computed as the Euclidean distance between the nearest points on the line and the ray. We then select the ray with the lowest median distance, i.e., $\operatorname{argmin}_j(\operatorname{median}_i\{d_{ij}\})$, as the orientation of the atmospheric light vector that we output.

The line of the i-th patch is defined the line passing through the centroid of $I(\mathbf{x})$ inside $\mathbf{x} \in \Omega_i$ in the direction defined by the largest principal component (PCA) of these pixels. In order to avoid outliers, we perform the PCA twice: we first use all the patch pixels and then recompute it after discarding the 20% pixels in Ω_i which are farthest from the initial line estimated.

The procedure for recovering $\hat{A}$ in [2] requires the user to identify patches in which both the shading $l(\mathbf{x})$ and the transmission $t(\mathbf{x})$ functions vary sufficiently (the albedo is assumed constant). However, the piecewise smoothness of the scene radiance and transmission, discussed earlier, make it hard to find patches in which the pixels form well-defined two-dimensional planes in RGB space. For example, when examining natural outdoor images, between 80% and 97% of the patches show a one-dimensional pixel distribution (defined by $\lambda_2 < \lambda_1/10$ where $\lambda_1 \geq \lambda_2$ are the largest eigenvalues of the patch PCA). In contrast, our approach requires only the shading function to vary and hence uses more patches to estimate $\hat{A}$ more accurately.

Finding Patches that Obey Model Eq. (3). Image patches which are well modeled by Eq. (3) consist of pixels that lie on a line in RGB space, whose orientation is the reflectance vector R. We identify such patches in $I(\mathbf{x})$ by making sure they meet a list of conditions. It is assumed here that $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the three principal eigenvalues computed by the robust PCA we described above.

1. *Positive principal component.* The strongest principal axis should correspond to the line's orientation, i.e., the reflected light vector R. The latter is expected to have

non-negative coordinates. Thus, we discard patches in which the largest principal component has coordinates with different signs.

2. *Single large principal eigenvalue.* In order to achieve a reliable estimation of the patch line orientation we favor patches with significant largest principle eigenvalue λ_1. Hence, we require the patches to have $\lambda_1 > \tau_1$ and explain below how the threshold τ_1 is set.

3. *Rank-one PCA matrix.* Since we expect the pixels to lie on a line, there should be a single significant principal component. Hence, we require $\lambda_1/\lambda_2 > \tau_2$ and explain below how τ_2 is set.

4. *Patch lines must not pass through the origin.* Lines that pass through the origin either intersect $s\mathrm{A}$ or describe patches in which no airlight is present and hence are irrelevant to its estimation. In both cases these lines cannot be used to define planes and are discarded. Specifically, we require the Euclidian distance d between the origin and the closest point on the patch line to obey $d > \tau_3$ and explain below how τ_3 is set.

5. *Patches do not contain an edge.* Eq. (3) assumes the patch is composed of pixels with the same reflectance and transmission values. According to the discussion above, both these functions are expected to be piecewise smooth. Therefore we must avoid considering patches that contain strong discontinuities. We use a Canny edge-detector and skip patches containing edge pixels.

We run this search over the entire image and use small windows of 10-by-10 pixels as candidate patches. We initiate each of the thresholds τ_1, τ_2 and τ_3 such that 50 patches exceed it (by sorting the patches according to each condition independently). We then check how many patches exceed *all* three thresholds and if this does not yield 10 patches we reduce each threshold by 3% in turn until this number of patches are found. In order to avoid patches with very similar lines, once a patch is accepted we exclude from this search all the patches whose lines form less than 15 degrees with the orientation of the accepted line.

Ultimately, the patches that pass all thresholds are used for estimating $\hat{A}$ as explained above. Fig. 3 shows the patches found by this procedure and the estimated orientation of the atmospheric light vector.

3.2. Estimating the Atm. Light Magnitude $\|A\|$

As we saw above, the magnitude of the atmospheric light vector cannot be recovered using the patch model in Eq. (3) since A is defined up to a non-zero multiplicative scalar. This is not the case in the global image formation model,
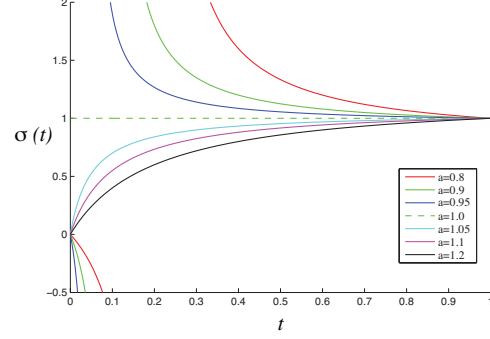

Figure 4. Graphs show the expected bias in pixel brightness, σ, as function of the transmission level, t, when dehazing an image using aA.

Eq. (1), and as shown in Fig. 1, $\|A\|$ is a meaningful quantity with a major impact over the resulting dehazed image. Let us proceed by investigating the effect of multiplying the *correct A* by a factor $a > 0$ in Eq. (1). By keeping $J(\mathbf{x})$ decomposed to the shading-albedo product, we obtain the following equation

$$I(\mathbf{x}) = t_a(\mathbf{x})l_a(\mathbf{x})R_a(\mathbf{x}) + \big(1 - t_a(\mathbf{x})\big)aA, \quad (4)$$

where t_a, l_a and R_a correspond to the transmission, shading and albedo functions obtained for every value of a respectively. Eq. (4) does not have a unique solution, however there is one property that successful image dehazing should achieve and that can guide our analysis. Namely, the airlight component $\big(1-t_a(\mathbf{x})\big)aA$ should match the true airlight $\big(1-t(\mathbf{x})\big)A$, even when $a \neq 1$. For example, this is the solution which will be obtained in [2] at regions of constant albedo and independent variation in $t(\mathbf{x})$ and $l(\mathbf{x})$ for the following reason. In order to fully satisfy Eq. (4) (as most methods do) the recovered R and A must span the correct two-dimensional subspace. Hence, any deviation from the correct airlight, $\big(1-t(\mathbf{x})\big)A$, must be balanced by the coefficients $l(\mathbf{x})$ and $t(\mathbf{x})$ which will become dependent. This solution cannot be reached by [2] which imposes no-correlation between the two variables. Another example is the method in [4] which also recovers the airlight term at regions where the dark-channel prior is met (despite not knowing $\|A\|$).

Assuming this is indeed the case, the use of a wrong magnitude ($a \neq 1$) leads to the following transformed transmission estimate

$$\big(1-t_a(\mathbf{x})\big)aA = \big(1-t(\mathbf{x})\big)A \implies t_a(\mathbf{x}) = \big(t(\mathbf{x})-1\big)/a+1. \quad (5)$$

Equating the airlight terms means that the direct transmission terms must also balance, i.e., $t_a(\mathbf{x})l_a(\mathbf{x})R_a(\mathbf{x}) = t(\mathbf{x})l(\mathbf{x})R(\mathbf{x})$. The latter implies that $R_a(\mathbf{x}) = R(\mathbf{x})$ since these are the only terms defining the linear subspace in
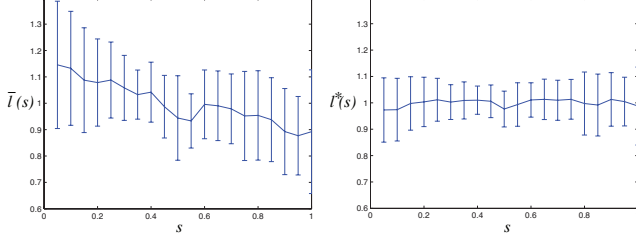
Figure 5. Graphs show the mean and standard-deviation (vertical bars) of $\bar{l}(s)$ (left) and $l^*(s)$ (right) estimated from an ensemble of 32 images.

Figure 6. Graphs show the upper (blue) and lower (red) bounds within which we expect to estimate a assuming variablilty bounds of $\pm 10\%$ in $l^*(s) = k$ for an image containing the lowest transmission level t (horizonal axis of this plot). The expected accuracy is approximately the difference between these graphs. In this case $a = 1$ is the correct solution.

which the balance holds (recall that $\|R\| = \|R_a\| = 1$). As a consequence we get the following transformation in the shading function

$$l_a(\mathbf{x}) = \sigma_a\big(t(\mathbf{x})\big)l(\mathbf{x}), \quad \text{where} \quad \sigma_a(t) = \frac{ta}{t+a-1}. \quad (6)$$

Thus, the use of aA results in the dehazed image $J_a(\mathbf{x}) = l_a(\mathbf{x})R(\mathbf{x}) = \sigma_a(t(\mathbf{x}))l(\mathbf{x})R(\mathbf{x})$ that differs from $J(\mathbf{x}) = l(\mathbf{x})R(\mathbf{x})$ by a scalar factor $\sigma_a(t(\mathbf{x}))$, i.e., in its brightness. Moreover, this bias *depends* on the transmission level $t(\mathbf{x})$ at each pixel. As shown in Fig. 4, an underestimated magnitude, $a < 1$, causes the output image to become brighter at regions of low transmission (yet $t > 1-a$, otherwise $\sigma_a < 0$), and an overestimated magnitude, $a > 1$, makes the low-transmission pixels appear darker. Please note that these expected results are confirmed with the experiment shown in Fig. 1.

These biases in the image brightness are hard to detect locally since they depend on the transmission and the latter correlates with the content of $J(\mathbf{x})$; pixels of distinct objects typically correspond the same surface color and depth range and hence transmission values and, on the other hand, boundaries between objects correspond to edges in J and discontinuities in t due to jumps in depth. In order to cope with this difficulty, we introduce a new regularity that we observe in natural hazy images. Consider the following function that outputs the top 1% brightness value found at each transmission level s

$$l^*(s) = \text{percentile}\{l(\mathbf{x}) \,|\, t(\mathbf{x}) = s\}. \quad (7)$$

In Fig. 5 we plot the average and standard-deviation of $l^*(s)$ at each transmission value. These statistics where gathered from an ensemble of 32 images in which the transmission is known[1]. For reference we also plot the average

---

[1]Due to the lack of such ground-truth data (hazy images I and their A), we conducted this test on natural outdoor scenes in which the disparity $d(\mathbf{x})$ is known (computed from stereo image pairs using [6]). The transmission used for this purpose was computed by $t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}$ with β set such that the most distant object receives $t = 0.01$. We also normalized the overall brightness of each image such that the average of $l^*(s)$ will be the same for all the images. The same was done for $\bar{l}(s)$.
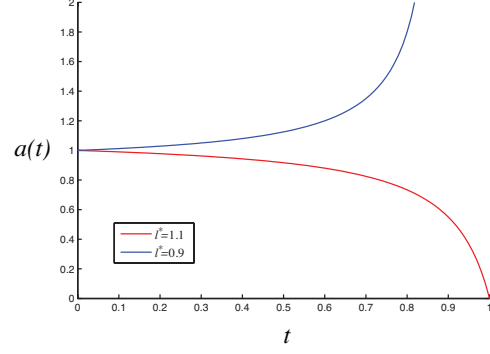
and standard-deviation of the brightness at each transmission value, $\bar{l}(s) = \text{average}\{l(\mathbf{x})|t(\mathbf{x}) = s\}$.

The graphs indicate that $\bar{l}(s)$ exhibits a greater dependence on the transmission level s than $l^*(s)$ and the lower standard-deviation of the latter implies that its invariance is more consistent across the different images. We attribute this experimental result to the presence of large objects in the image which occupy a narrow range of transmission values (or depths) and bias the corresponding transmission bins of $\bar{l}(s)$. The extremal events that determine $l^*(s)$ are equally affected by all the objects in a particular transmission range, regardless of their size.

In practice we compute the percentile statistics from the image which we dehazed with a wrong $\|A\|$, i.e., $J_a(\mathbf{x})$. This provides us with

$$l_a^*(s) = \text{percentile}\{l_a(\mathbf{x}) \,|\, t_a(\mathbf{x}) = s\}, \quad (8)$$

since we do not know the correct pixel transmission level $t(\mathbf{x})$, but recover $t_a(\mathbf{x})$. However, according to Eq. (5) and Eq. (6), l_a^* and l^* are related by

$$l_a^*\big((s-1)/a + 1\big) = \sigma_a(s)l^*(s). \quad (9)$$

While the relations also hold in case of the average statistic $\bar{l}(s)$, given $l^*(s)$'s higher consistency and invariance to s we use the latter for recovering $\sigma_a(s)$ through what we call *max-brightness transmission invariance* image prior that assumes $l^*(s) = k$ for some constant k. Thus, according to Eq. (8), we attribute variations of the form of $\sigma_a(s)$ in the observed l_a^* to errors in the estimated magnitude of the atmospheric light vector.

Thus, we recover a by the following procedure. We dehaze the image $I(\mathbf{x})$ by some dehazing method using the atmospheric light vector $\hat{A}$ computed at the first stage of
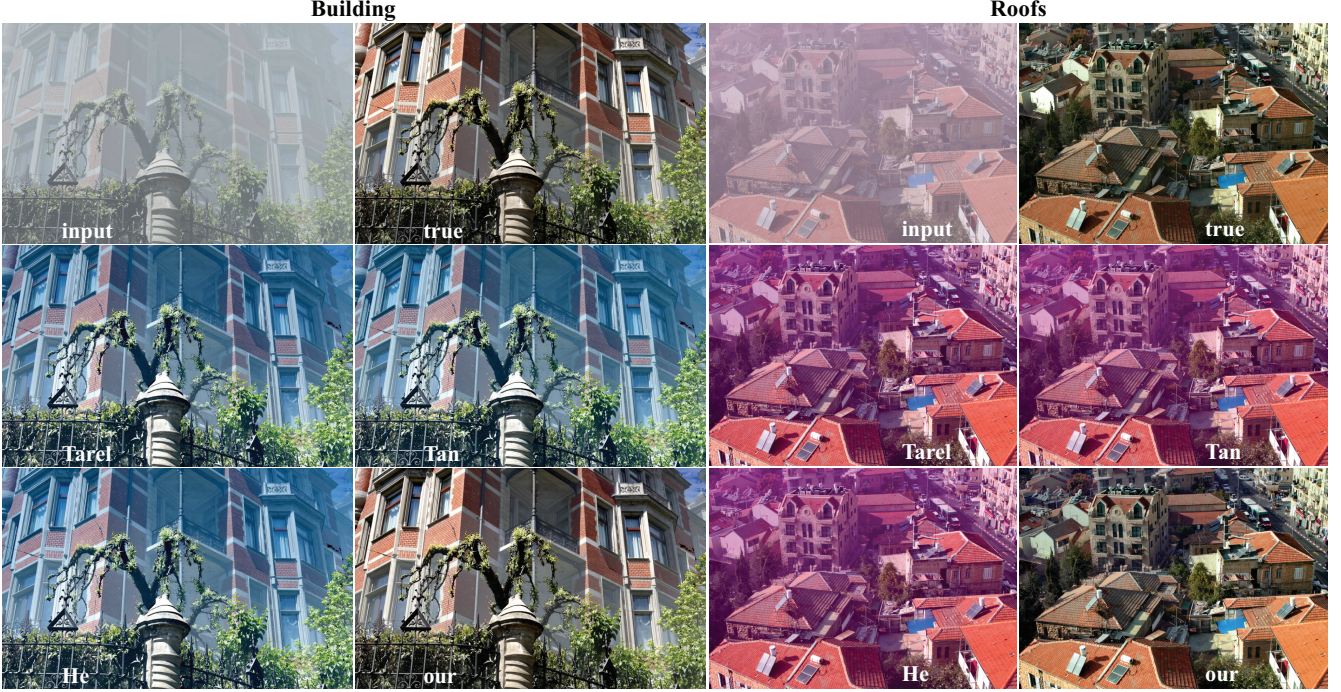
Figure 7. Comparison with known ground-truth A. Images dehazed using the atmospheric light vectors estimated by the methods of He et al. [4] (labeled He), Tan [21] (labeled Tan), Tarel [22] (labeled Tarel), our method, and the true vectors used for synthesizing the input.

our method. Since $\hat{A}$ is a unit vector, $\|\hat{A}\| = 1$, it is likely that $a \neq 1$. We then compute l_a^* from the resulting image and estimate a by $\tilde{a}$ obtained by solving

$$\min_{\tilde{a},k} \sum_s \left(l_a^*\big((s-1)/\tilde{a}+1\big) - \sigma_{\tilde{a}}(s)k \right)^2. \quad (10)$$

Finally, we take $\hat{A}/\tilde{a}$ as our output atmospheric light vector magnitude. In order to avoid the excessively-large (and small) values seen in Fig. 4 for the cases of $a < 1$, we increase the norm of $\hat{A}$, which we use when computing l_a^* in Eq. (10), by steps of 0.2 until no negative values are found (and hence no excessively large values as well). We solve Eq. (10) using Matlab's fminsearch starting from the initial guess $\tilde{a}=1$ and $k=1$.

Accuracy. Despite the relatively low variability seen in Fig. 5, individual images can exhibit a considerable dependence between l and t. In such cases our max-brightness transmission invariance will not be valid and will lead to a wrong estimation of $\|A\|$—one that will lower the dependence of l on t. Thus, for example, if low-transmission (far) regions of the scene should be brighter than the rest of the scene, the use of our A will lower the intensity of these regions. In other words, the error we produce in the estimation of $\|A\|$ tends to compress the intensity levels in the image. As demonstrated in Fig. 1, this dynamic-range compression of the image may be preferable over the opposite scenario. We shall see more such examples in Section 4.

Finally, in order to assess the expected accuracy of the estimated magnitude $\|A\|$, under the assumption of $\pm 10\%$ variability in $l^*(s)$, we compute the maximal and minimal values of $a(t)$ that can be obtained within the observed the envelope of $l^*(s)$ shown in Fig. 5. However, since the effect of a wrong $\|A\|$ leads to larger biases in pixels of low transmission values, as seen in Fig. (4), this analysis greatly depends on the minimal t value found in an image. Therefore, we compute the maximal and minimal $a(t)$ allowed by solving Eq. (6) for $\sigma_a(t) = 0.9$ and $\sigma_a(t) = 1.1$ respectively *for every* value of t. As shown in Fig. 6, the accuracy of a increases as t decreases, and in case of of an image containing pixels with at most $t = 0.5$, we expect an accuracy of about $\pm 10\%$ in $\|A\|$.

4. Results

We conducted two types of experiments in order to evaluate the accuracy of our method: we applied it on images with known ground-truth A and on real-world hazy images which serve as benchmark for dehazing. As our method does not require any user input, we applied the automatic procedures described in [21, 4] and [22] for comparison. As part of this evaluation we used the A estimated by the different method to dehaze images and used the method of [4] for this purpose. Please note however that the goal of this paper is the recovery of A and not the image dehazing (i.e., transmission estimation) hence the results should be judged

	Orientation				Magnitude				l_∞ Endpoint Error			
	He	**Tan**	**Tarel**	**Our**	**He**	**Tan**	**Tarel**	**Our**	**He**	**Tan**	**Tarel**	**Our**
Mansion	6.092	6.419	5.344	**0.148**	0.433	0.458	0.696	**0.196**	0.304	0.325	0.427	**0.131**
	4.059	4.397	2.180	**0.083**	0.079	0.105	0.082	**0.068**	0.127	0.149	0.088	**0.043**
	3.134	2.831	3.145	**0.193**	0.318	0.344	0.592	**0.115**	0.224	0.235	0.380	**0.068**
Church	3.647	6.073	3.490	**0.028**	0.147	0.208	0.292	**0.094**	0.149	0.232	0.220	**0.060**
	5.862	7.402	5.228	**0.142**	0.352	0.476	0.628	**0.111**	0.250	0.357	0.396	**0.077**
	2.051	2.472	1.389	**0.182**	0.204	0.266	0.393	**0.194**	0.144	0.187	0.237	**0.120**
Bikes	1.071	1.212	0.899	**0.102**	**0.113**	0.212	0.513	0.116	0.080	0.133	0.299	**0.072**
	2.177	2.642	2.308	**0.129**	**0.100**	0.202	0.484	0.156	**0.091**	0.153	0.295	0.098
	3.246	4.536	3.642	**0.114**	**0.085**	0.239	0.548	0.181	**0.094**	0.193	0.329	0.117
Dolls	1.295	2.318	1.010	**0.229**	**0.217**	0.287	0.663	0.382	**0.139**	0.196	0.389	0.233
	3.318	3.745	2.561	**0.283**	0.037	0.101	0.127	**0.011**	0.073	0.101	0.116	**0.009**
	3.912	3.289	4.941	**0.729**	0.210	0.276	0.615	**0.191**	0.176	0.197	0.421	**0.124**
Moebius	4.231	4.231	2.221	**0.443**	0.141	0.141	0.390	**0.131**	0.150	0.150	0.236	**0.080**
	3.316	3.316	8.486	**0.288**	**0.031**	0.031	0.230	0.082	0.087	0.087	0.306	**0.057**
	3.139	3.139	3.847	**0.932**	0.071	0.071	0.295	**0.003**	0.100	0.100	0.221	**0.019**
Reindeer	0.536	1.454	0.752	**0.041**	0.079	**0.004**	0.029	0.049	0.057	0.028	**0.027**	0.032
	0.470	2.587	0.689	**0.156**	**0.010**	0.163	0.306	0.103	**0.013**	0.132	0.191	0.066
	0.825	1.023	1.055	**0.045**	0.039	0.089	0.311	**0.000**	0.039	0.072	0.190	**0.001**
Trees	2.858	2.099	3.615	**0.785**	**0.027**	0.055	0.291	0.066	0.071	0.066	0.224	**0.049**
	3.884	2.241	3.475	**0.220**	**0.034**	0.063	0.330	0.043	0.086	0.065	0.232	**0.027**
	4.231	4.644	4.215	**1.060**	0.091	0.080	0.095	**0.051**	0.148	0.148	0.148	**0.043**
Village	3.915	4.629	3.708	**0.142**	**0.228**	0.271	0.440	0.351	**0.183**	0.220	0.286	0.231
	1.747	2.028	1.879	**0.519**	**0.243**	0.291	0.531	0.281	**0.168**	0.201	0.336	0.179
	3.726	**3.538**	4.184	3.549	**0.360**	0.444	0.818	0.399	**0.243**	0.281	0.510	0.258
Roofs	5.092	5.056	4.790	**3.239**	0.151	0.154	0.163	**0.126**	0.173	0.178	0.170	**0.115**
	4.521	4.709	4.377	**0.783**	0.384	0.387	0.582	**0.245**	0.285	0.291	0.390	**0.156**
	4.545	4.535	4.405	**1.129**	**0.458**	0.461	0.677	0.507	**0.314**	0.314	0.437	0.318
Mean	3.218	3.576	3.253	**0.581**	0.172	0.218	0.412	**0.158**	0.147	0.178	0.278	**0.103**
Median	3.318	3.316	3.490	**0.220**	0.141	0.208	0.393	**0.116**	0.144	0.178	0.286	**0.077**

Table 1. Estimation error of $\hat{A}$ (left four columns), the magnitude $\|A\|$ (middle four rows), and the endpoint error of A (eight columns on the right). Every test contains four columns reporting the error produced by the methods of He et al. [4] (labeled He), Tan [21] (labeled Tan), Tarel [22] (labeled Tarel) and our method. In each cell we report the results obtained on the same image using three different A vectors which were sampled randomly. The lowest error obtained in each category is emphasized in boldface.

Figure 8. Comparison over natural images. The images shown were dehazed using the atmospheric light vectors estimated by the methods of He et al. [4] (labeled He), Tan [21] (labeled Tan), Tarel [22] (labeled Tarel), and our method.

also by the authenticity of the colors in the recovered scene and the plausibility of its brightness.

Our Matlab implantation takes around 10 seconds to recover the orientation of the atmospheric light vector from a 640-by-480 pixels image. The magnitude requires another 3 seconds and the time to compute $t(\mathbf{x})$ using the dark-channel method [4]. Recent implementation of this method requires a fraction of a second to process a one-megapixel

image.

Before we go into the details of the evaluation let us note that at the supplemental material, accompanying the paper, we provide additional results and comparisons. The code and all the data used and created in the evaluation will be available online once the paper is published.

Evaluation over images with known A**.** Databases of images $I(\mathbf{x})$ and their ground-truth A are unfortunately unavailable to us and acquiring such data is quite challenging. As an alternative to the use of purely computer-generated low-resolution dataset, used in [23], we simulated hazy images from pairs of hazy-free RGB images, which we consider as $J(\mathbf{x})$, and their corresponding disparity maps $d(\mathbf{x})$. Assuming a constant scattering coefficient β in space, we computed a spatially-varying transmission maps by $t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}$ and set β such that the most distant object in the scene receives $t = 0.1$. We gathered 9 pairs of images and disparity maps from three different sources: the high resolution scene reconstruction from lightfields of [9], the Middlebury database [17], as well as by collecting high-resolution stereo image pairs and recovering the disparities using [6]. Each of these pairs was used to produce three hazy images according to Eq. (3) using three randomly-sampled atmospheric light vectors. The orientation of these vectors is sampled uniformly from the $10°$ cone around $[1, 1, 1]$, and their magnitude is sampled uniformly from the interval $[0.8, 1.8]$.

Table 1 reports the errors produced by the different methods when estimating $\hat{A}$ (in absolute angles), $\|A\|$ (absolute value) and A (endpoint l_∞ norm). The data shows a considerable improvement by our method in estimating $\hat{A}$, where the average error is about fifth of the error produced by the competing method'. The low median error indicates a high success rate. The accuracy improvement in estimating $\|A\|$ is less dramatic where the average error is lower than the competitors' and this improvement is somewhat higher for the median error. In less than 30% of the images the method of He et al. achieves a more accurate estimation of $\|A\|$. Finally, our method achieves the most accurate estimates of A(in l_∞ norm) in two thirds of the images which boils down to a %30 reduction on average next to the method of He et al. and almost %50 reduction in the median error.

Fig. 7 shows the results obtained on two images participating in this test. Wrong estimations of $\|A\|$ are identified by the blue and purple color-shifts. In these examples, it appears that errors in A also lead to lower haze removal, i.e., over estimations of the transmission. Additional examples from this test set are available at the supplemental material.

Evaluation over real-world hazy images. In the second experiment we applied and compared our method on a set of real images acquired under hazy conditions. In this experiment there is no firm ground-truth criterion and the results should be judged by the authenticity of the colors in the scene and the plausibility of its brightness. Fig. 8 shows the resulting dehazed images produced by the different methods on two benchmark images. In case of the Highway Image our method achieves a more accurate $\hat{A}$ which makes the distant building appear more neutral. However, it also appears as if it underestimated $\|A\|$ and therefore this distance regions appears brighter. As we mentioned in Section 3.2, the type of error produced by our in $\|A\|$ tends to compress the dynamic range of the image. For some applications this outcome can be beneficial as it tends to emphasize the details of the scene and increase its visibility. In case of the Aerial Image the green fields appear more uniform in our result, suggesting a more accurate $\|A\|$.

Failure cases. Fig. 9 shows several cases in which our method fails. The Wheat Image contains a single surface albedo and hence our method cannot obtain enough constraints over $\hat{A}$. The New York Image is rich in detail and its resolution does not provide us with a sufficient of number of mono-albedo patches. In case of Village 1 and Roofs 3 our method produces relatively high l_∞ (beyond 0.3, see Table 1). This error, again, leads to a more dynamically-compressed output compared to the ground-truth solution.

5. Conclusions

We presented an automatic method to recover the atmospheric light vector given a single hazy image. The method recovers the vector's orientation based on a simplified model that describes simple geometric configurations of pixels inside small mono-albedo patches. We described a robust procedure for finding patches that obey this model as well as explained how they are used for recovering the atmospheric light's orientation.

We showed how errors in the magnitude of the atmospheric light leads to biases in the pixels' brightness which depend on the transmission level. We explained the difficulties this dependency creates when trying to detect these biases via localized analysis. To overcome this issue, we introduced a global image regularity, the max-brightness transmission invariance, which is sensitive to dependence between the pixel brightness and the transmission. We described a simple procedure for estimating magnitude of the atmospheric light using this new image prior. Our study is accompanied with a statistical analysis of the regularity which estimates its consistency and the expected accuracy at which $\|A\|$ can be estimated given the amount of transmission in the scene.

Limitations. While we presented many evidences for the accuracy improvement our method achieves, there are many cases in which its underlining assumptions do not hold. As we showed in the previous section, the lack of sufficient patches obeying our simplifies model, in Eq. (3), prevents us from obtaining an estimate for $\hat{A}$. Moreover,

Figure 9. Failure cases of our method. Wrong $\hat{A}$ in the Wheat and New York images leads to severe color-shifts. In both the Village and Roofs images, $\|A\|$ is inaccurate and leads to a more dynamically-compressed dehazed image.

in images where $l^*(s)$ varies significantly, the accuracy at which we estimate $\|A\|$ is compromised. We expect that both these scenarios become less likely as the image resolution increases.

References

[1] P. S. Chavez. An improved dark-object subtraction technique for atmonspheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24:450–479, 1988.

[2] R. Fattal. Single image dehazing. In *ACM SIGGRAPH 2008 papers*, SIGGRAPH '08, pages 72:1–72:9, New York, NY, USA, 2008. ACM.

[3] G. D. Finlayson, S. D. Hordley, and I. Tastl. Gamut constrained illuminant estimation. *International Journal of Computer Vision*, 67:93–109, 2006.

[4] K. He, J. Sun, and X. Tang. single image haze removal using dark channel prior. pages 1956–1963, 2009.

[5] W. Hinds. *Aerosol technology: properties, behavior, and measurement of airborne particles*. Wiley-Interscience. Wiley, 1999.

[6] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 35(2):504 – 511, 2013.

[7] J. Huang, A. Lee, and D. Mumford. Statistics of range images. *Proceedings. IEEE Conference on Computer Vision and Pattern Recognition*, pages 324–331, 2000.

[8] J. Huang and D. Mumford. Statistics of natural images and models. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1, pages 637–663, 1999.

[9] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. Gross. Scene reconstruction from high spatio-angular resolution light fields. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH)*, 32(4):73:1–73:12, 2013.

[10] H. Koschmieder. Theorie der horizontalen sichtweite. In *Beitr. zur Phys. d. freien Atm.*, pages 171–181, 1924.

[11] A. B. Lee, D. Mumford, and J. Huang. Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision*, 41:35–59, 2001. 10.1023/A:1011109015675.

[12] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *J. Opt. Soc. Am. A*, 3(10):1694–1699, Oct 1986.

[13] T. M. Lehmann and C. Palm. Color line search for illuminant estimation in real-world scene. *J. Opt. Soc. Am. A*, 18(18):2679?–2691, Oct 2001.

[14] S. Narasimhan and S. Nayar. Chromatic framework for vision in bad weather. In *Proceedings. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 598–605, 2000.

[15] S. G. Narasimhan and S. K. Nayar. Interactive (De)weathering of an Image using Physical Models. In *ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV)*, October 2003.

[16] J. P. Oakley and H. Bu. Correction of simple contrast loss in color images. *IEEE Transactions on Image Processing*, 16(2):511–522, 2007.

[17] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision*, 47(1-3):7–42, Apr. 2002.

[18] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. pages 325–332, 2001.

[19] S. Shwartz, E. Namer, and Y. Y. Schechner. Blind haze separation. CVPR '06, pages 1984–1991, Washington, DC, USA, 2006. IEEE Computer Society.

[20] K. Tan and J. P. Oakley. Enhancement of color images in poor visibility conditions. *Proceedings of International Conference on Image Processing*, 2:788–791, 2000.

[21] R. T. Tan. Visibility in bad weather from a single image. *Proceedings of IEEE CVPR*, 2008.

[22] J.-P. Tarel and N. Hautière.   Fast visibility restoration from a single color or gray level image.   In *Proceedings of IEEE International Conference on Computer Vision (ICCV'09)*, pages 2201–2208, Kyoto, Japan, 2009. http://perso.lcpc.fr/tarel.jean-philippe/publis/iccv09.html.

[23] J.-P. Tarel, N. Hautière, A. Cord, D. Gruyer, and H. Halmaoui.   Improved visibility of road scene images under heterogeneous fog.   In *Proceedings of IEEE Intelligent Vehicle Symposium (IV'2010)*, pages 478–485, San Diego, California, USA, 2010. http://perso.lcpc.fr/tarel.jean-philippe/publis/iv10.html.

[24] B. Xie, F. Guo, and Z. Cai. Improved single image dehazing using dark channel prior and multi-scale retinex. In *Intelligent System Design and Engineering Application (ISDEA), 2010 International Conference on*, volume 1, pages 848–851, oct. 2010.

[25] W. Zhou and C. Kambhamettu.   A unified framework for scene illuminant estimation. *Image Vision Comput.*, 26:415–429, March 2008.