Improved Single Image Dehazing using Geometry

Peter Carr Richard Hartley
Australian National University and NICTA
Canberra, Australia

***Abstract*—Images captured in foggy weather conditions exhibit losses in quality which are dependent on distance. If the depth and atmospheric conditions are known, one can enhance the images (to some degree) by compensating for the effects of the fog. Recently, several investigations have presented methods for recovering depth maps using only the information contained in a single foggy image. Each technique estimates the depth of each pixel independently, and assumes neighbouring pixels will have similar depths.**

In this work, we employ the fact that images containing fog are captured from outdoor cameras. As a result, the scene geometry is usually dominated by a ground plane. More importantly, objects which appear towards the top of the image are usually further away. We show how this preference (implemented as a soft constraint) is compatible with the alpha-expansion optimization technique and illustrate how it can be used to improve the robustness of any single image dehazing technique.

***Index Terms*—fog; haze; image; enhancement; optimization; graph cuts; alpha expansion; monotonic;**

I. Introduction

Poor visibility, which arises in foggy weather conditions, hinders the usefulness of outdoor cameras. However, it is possible to compensate for the effects of fog — at least, to some degree — if the optical and depth information about the scene is known [1]. Recent investigations [2], [3], [4] have shown how a reasonably accurate depth map can be estimated automatically from a single foggy input image using a variety of techniques. Although each method uses a different statistical measure to drive the estimation process, they all share a common shortcoming: when the appearance information of a pixel is unreliable, the algorithms are unable to produce a good depth estimate for the corresponding location in the image. As we will explain in Section II, each algorithm assumes neighbouring pixels will have similar depths, and incorporates this regularization into the estimation process.

In foggy conditions, the appearance of an object becomes more similar to that of the fog as the distance between the camera and the object increases. Therefore, the colour dissimilarity between each pixel and the fog gives an indication about its depth. However, it is difficult to estimate the depth of objects which are naturally white or light grey, since their change in appearance within fog is relatively independent of depth. In [2], [3], [4], any errors in the estimated depth maps are not usually visible in the enhanced images, since the solutions are applied to the same data used for the estimation process. However, in other situations, a more accurate depth map may be necessary.

In outdoor surveillance applications, cameras are usually placed high in the air, and the depth of each pixel changes only slightly depending on whether a foreground object is present or not. As a result, one can enhance a series of images using the same depth map, as long as the atmospheric properties remain constant. However, errors in the depth map will become quite obvious, as the results are now being applied to a set of data that is different from that used for the estimation process.

(a) (b) (c) (d) (e)

Fig. 1. A foggy image (a) is enhanced by estimating the optical depth of the scene (b). Here, the probability of a particular depth is based on the contrast of the enhanced image (c). The technique of [2] is unable to handle the appearance of the swan, resulting in: (1) an incorrect depth estimate in (b); and (2) an over-enhanced result in (c). Our geometric prior prefers that the depths of pixels increase as one scans the image from bottom to top. When incorporated into [2], the resulting depth (d) and enhancement (e) of the swan are improved without introducing significant artifacts — i.e., the depth of the tree (which does not adhere to this model) is still mostly correct.

In this work, we focus on the assumption that neighbouring pixels should have similar depths. We show how a stronger prior based on camera geometry can be used to improve the results of any of the single image estimation methods mentioned above (see Figure 1). Our key observation is that weather degradation occurs in outdoor scenes, which means the majority of the images should exhibit the geometry of a camera located above a ground plane. As we will show in Section III, this geometry leads to a simple relationship that objects which appear closer to the top of the image are usually further away. Furthermore, we will show that within the graph-cut based α-expansion energy minimization framework, our trend can be implemented as a preference, and does not always have to hold.

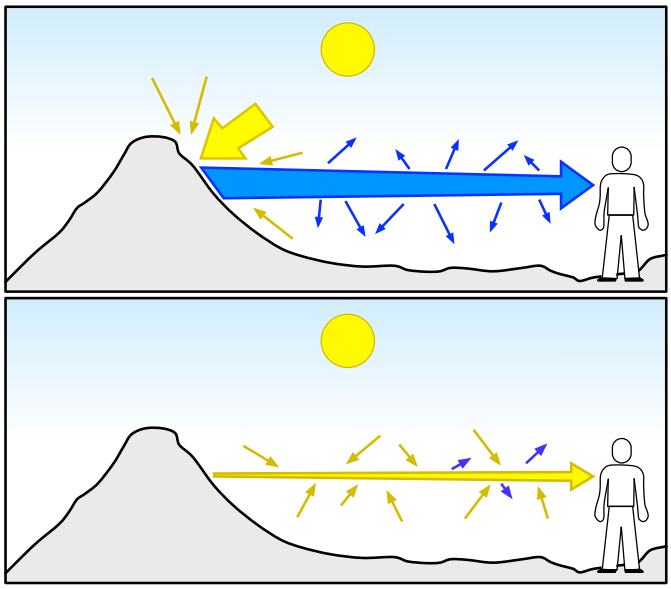

Fig. 2. *The radiance reaching the observer is the sum of the transmitted (top) and airlight (bottom) components [5]. As photons travel through the medium, there is a chance of being scattered, so the radiance of light decreases with distance. This leads to attenuation of the transmitted component. The airlight component arises from photons which were not originally incident on the observer being scatter towards the observer. As with the transmitted case, some of these photons may re-scatter off the path towards the observer.*

II. Background

The presence of aerosols in the lower atmosphere means light may scatter and/or be absorbed while travelling through the medium [1]. This can happen anywhere along the path, and leads to a combination of radiances incident towards the camera (see Figure 2). The *transmitted* component is an attenuated version of the signal that would have reached the camera if no aerosols were present. The *airlight* component arises from ambient light which was scattered towards the camera — i.e., it did not reflect off the corresponding scene surface element.

If the atmospheric and lighting conditions are constant, and the fog is suitably dense to assume the ambient illumination is isotropic [1], [5], the observed image $\mathbf{I}$ is a combination of the transmitted $\mathbf{J}$ and airlight $\mathbf{A}$ components, with the magnitudes determined by the optical distance βd_i of each pixel i:

$$\mathbf{I}_i = e^{-\beta d_i}\mathbf{J}_i + (1 - e^{-\beta d_i})\mathbf{A} \tag{1}$$

To recover the scene radiance $\mathbf{J}$ of a given foggy input image $\mathbf{I}$, one needs to estimate the optical depth of each pixel, as well as the appearance of the horizon (which is equivalent to $\mathbf{A}$). The transmission map $\mathbf{t}$ combines the unknown geometric distance d_i and extinction co-efficient β (the net loss from scattering and absorption) into a single variable:

$$t_i = e^{-\beta d_i} \tag{2}$$

When an object has the same appearance as the fog, $\mathbf{J}_i = \mathbf{A}$ and (1) is satisfied for any transmission value; making it is impossible to estimate the depth of that particular pixel. Furthermore, since the observed radiance $\mathbf{I}_i$ is a linear combination of $\mathbf{J}_i$ and $\mathbf{A}$ and the transmission (2) is limited to the range $[0, 1]$, the observed radiance $\mathbf{I}_i$ must be between these two values. Formally, either $\mathbf{J}_i \geq \mathbf{I}_i > \mathbf{A}$ or $\mathbf{J}_i \leq \mathbf{I}_i < \mathbf{A}$.

Finally, the aerosols of fog are sufficiently large to cause scattering which is independent of wavelength. As a result, $\mathbf{A}$ tends to have a whitish-grey colour, and only the transmission of each pixel and not that of each channel of each pixel (for colour images) needs to be estimated — i.e., a one needs to estimate a scalar value t_i and not a vector $\mathbf{t}_i$.

A. Related Work

Previous methods for enhancing weather degraded imagery required additional input [1], [6], [7], or specialized hardware [8]. None of these methods are suitable for wide-spread implementation in outdoor surveillance networks.

Oakley and Bu [9] developed a method for automatically estimating the amount of airlight within an image. They assumed every pixel was at the same depth and analysed the mean and variance of blocks of pixels. A global cost function was formulated by considering the normalized standard deviation of the entire image. Although the experimental results illustrate that the method works quite well, the assumption that all pixels are at the same depth is certainly not true in the majority of images.

Tan [2] developed a system for estimating depth from a single weather degraded input image. Motivated by the fact that contrast is reduced in a foggy image, Tan divided the image $\mathbf{I}$ into a series of small patches and postulated that the corresponding patch in $\mathbf{J}$ should have a higher contrast (where contrast was quantified as the sum of local image gradients). He employed a Markov Random Field to incorporate the prior that neighbouring pixels should have similar transmission values t_i. The method tends to produce over enhanced images in practice.

Fattal [3] assumed every patch has uniform reflectance, and that the appearance of the pixels within the patch can be expressed in terms of shading and transmission. He considered the shading and transmission signals to be unrelated and used independent component analysis to estimate the appearance of each patch. The method works quite well for haze, but has difficulty with scenes involving fog, as the magnitude of the surface reflectance is much smaller than that of the airlight when the fog is suitably thick.

Recently, He *et al.* [4] employed a model which assumed every local patch in the enhanced image should have at least one colour component near zero. In other words, the work assumed most scenes are made up of either dark or colourful objects. The transmission t_i of each patch was estimated as the minimum colour component within that patch. Instead of using an MRF, the work employed a soft matting algorithm to ensure that neighbouring pixels had similar transmission values.

He *et al.*'s strategy assigns either the minimum or maximum transmission to each pixel. Rearranging (1) and (2), we find

the following expression for the transmission of each pixel:

$$t_i = \frac{\mathbf{A} - \mathbf{I}_i}{\mathbf{A} - \mathbf{J}_i} \tag{3}$$

For a given foggy input image $\mathbf{I}$, both $\mathbf{A}$ and $\mathbf{I}_i$ are fixed. Assuming $\mathbf{J}_i \leq \mathbf{I}_i < \mathbf{A}$, the transmission t_i of a pixel will increase as its estimated scene radiance $\mathbf{J}_i$ increases. The lowest possible transmission occurs when $\mathbf{J}_i = \mathbf{0}$, which corresponds to a black (or a highly saturated colour in a multichannel image when analysing the "dark channel") object a certain distance away — the observed intensity of the pixel is only due to the airlight component. In the rare case when $\mathbf{J}_i \geq \mathbf{I}_i > \mathbf{A}$, the transmission of a pixel will decrease as the scene radiance $\mathbf{J}_i$ increases. In this situation, the minimum possible transmission occurs when $\mathbf{J}_i = \mathbf{1}$, which corresponds to a white object. The object appears darker because of the attenuation effects. As $\mathbf{J}_i$ tends towards $\mathbf{I}_i$, the transmission of the pixel reaches its maximum value of 1.

B. Regularization

The single image dehazing methods cited above use a variety of regularization and optimization methods. Here, we limit our investigation to the graph-cut based α-expansion method employed by [2], as it is able to handle the statistical models of [3], [4] as well and has a good track record with vision-specific energy functions [10].

In this approach, each element t_i of the transmission map is associated with a label x_i, where the set of labels $\mathcal{L} = \{0, 1, 2, \ldots, \ell\}$ represents the transmission values $\{0, \frac{1}{\ell}, \frac{2}{\ell}, \ldots, 1\}$. The most probable labelling $\mathbf{x}^\star$ minimizes the associated energy function [11]:

$$E(\mathbf{x}) = \sum_{i \in \mathcal{P}} E_i(x_i) + \sum_{(i,j) \in \mathcal{N}} E_{ij}(x_i, x_j) \tag{4}$$

Here, $\mathcal{P}$ is the set of pixels in the unknown transmission map $\mathbf{t}$ and $\mathcal{N}$ is the set of pairs of pixels defined over the standard four-connect neighbourhood. The unary functions $E_i(x_i)$ represent the probability of pixel i having transmission t_i associated with label x_i, and can be any one (or a combination of) the methods of [2], [3], [4]. The smoothness term $E_{ij}(x_i, x_j)$ encodes the probability that neighbouring pixels should have similar depths. For simplicity, we consider the linear cost function, which is solvable by α-expansion:

$$E_{ij}(x_i, x_j) = \lambda \left| x_i - x_j \right| \tag{5}$$

Details of the data cost function $E_i(x_i)$ and the value of λ will be discussed in Section V.

III. Outdoor Geometry

In outdoor surveillance, cameras are typically placed high in the air and tilted towards the ground (see Figure 3). The depth of any scene point (such as Q, R, S or T in Figure 3) is the distance between it and the camera's centre of projection $\mathbf{C}$. The depth can also be expressed in the components of distance along the ground and height above the ground. If the camera calibration parameters are known, only one of these two measurements needs to be specified. As long as the scene does not contain any cave-like surfaces, such as the space underneath a bridge, the distance along the ground to the visible scene point is a monotonically increasing function of image plane height (from the bottom of the image to the top).

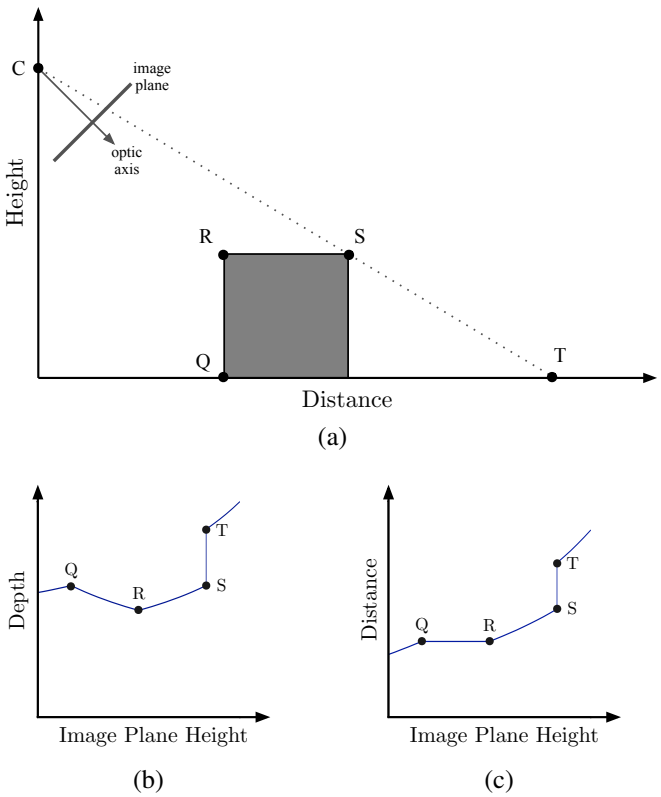

Fig. 3. *The depth of any scene point — i.e., the distance from the centre of projection $\mathbf{C}$ — can be split into distance and height components. If the scene does not contain any cave-like surfaces, the distance of scene points will increase monotonically from the bottom of the image to the top.*

If we consider two pixels i and j, such that j is directly above i, our geometry implies that $d_j \geq d_i$. As a result, the transmission t_j of pixel j, determined by (2), must be less than or equal to the transmission t_i of pixel i (assuming that β is constant over the image). To encourage the pattern $x_j \leq x_i$, we can assign a cost $\tau > 0$ to any pair of labels which violates this trend. The condition can be strictly enforced by setting $\tau = \infty$. The smoothness function (5) now becomes:

$$E_{ij}(x_i, x_j) = \begin{cases} \tau & \text{if } x_i < x_j, \\ \lambda \left| x_i - x_j \right| & \text{otherwise.} \end{cases} \tag{6}$$

A. Related Work

Liu *et al.* [12] examined the problem of estimating coarse 3D scene structure using ordering constraints. The set of labels was limited to $\mathcal{L} = \{$'left', 'right', 'top', 'bottom', 'centre'$\}$, as the underlying scene model assumed five orthogonal planes. The pairwise energy term $E_{ij}(x_i, x_j)$ was an asymmetric Potts model. Infinite penalty terms prevented nonsensical orderings (such as a pixel labelled 'right' being to the left of

a pixel labelled 'left'). The authors were unable to solve the problem using α-expansion, and instead employed alternating horizontal and vertical moves which considered three labels simultaneously.

Ramalingam *et al.* [13] also examined the problem of estimating coarse 3D scene geometry. They, however, used a smaller set of labels $\mathcal{L} = \{$'ground','scene','sky'$\}$. Unlike Liu *et al.*, Ramalingam *et al.* employed soft constraints, since it is possible (although generally unlikely) to have 'sky' below a pixel labelled 'scene'. Their solution used machine learning techniques to establish the likelihood of each triplet of labels. The resulting cost function was a third-order multilabel problem, which was transformed and truncated into a submodular second-order pseudo boolean function and solved using binary graph cuts [10].

IV. Solvability

Each cycle of α-expansion reduces to a binary labelling problem: should the pixel i keep its existing label x_i, or take the label α. The cut of the binary graph determines which vertices should switch to the label α to produce a new configuration $\mathbf{x}'$ having a lower energy. The submodularity condition [14] now involves three labels: the expansion label α, as well as the labels x_i and x_j [15]:

$$E_{ij}(\alpha,\alpha) + E_{ij}(x_i,x_j) \leq E_{i,j}(x_i,\alpha) + E_{i,j}(\alpha,x_j) \quad (7)$$
$$\text{for all } i,j \in \mathcal{P} \text{ and } x_i, x_j, \alpha \in \mathcal{L}$$

By inspection, one can see that the linear distance model (5) adheres to necessary condition for α-expansion. However, the fact that our more complex geometry model (6) also satisfies the above condition is not so straightforward.

We begin by considering the case when $\tau = \infty$. The two variable projection of $E(\mathbf{x})$ will be submodular in every iteration of α-expansion if the current move $\mathbf{x}$ does not have infinite energy [16] — i.e., it does not violate the constraint. However, hard constraints are not useful in practice. In fact, we have already identified that our monotonic trend is violated by several common outdoor scene structures, such as bridges. Furthermore, the likelihood of a violation occurring increases as the camera gets lower to the ground, or if one applies the restriction to depth and not distance (which is necessary for the case of uncalibrated cameras).

Many models for vision applications are determined by the difference between a pair of labels, and not necessarily the labels themselves — i.e., $E_{ij}(x_i,x_j) = f_{ij}(x_i - x_j)$. In this situation, the requirement for α-expansion (7) becomes:

$$f_{ij}(0) + f_{ij}(x_i - x_j) \leq f_{ij}(x_i - \alpha) + f_{ij}(\alpha - x_j) \quad (8)$$

We will now show how a pair of increasing functions which obey the triangle inequality will adhere to the above requirement.

Theorem 1 *Let $f^+(x)$ and $f^-(x)$ be two functions defined for values $x \geq 0$, satisfying the following conditions:*

1) **Monotonicity.** *If $y \geq x \geq 0$, then $f^{\pm}(y) \geq f^{\pm}(x)$.*

2) **Triangle inequality.** *If $q \leq \alpha \leq p$, then*

$$f^{\pm}(0) + f^{\pm}(p-q) \leq f^{\pm}(p-\alpha) + f^{\pm}(\alpha - q).$$

3) **Compatibility.** $f^+(0) = f^-(0)$.

Here $f^{\pm}(x)$ represents either $f^+(x)$ or $f^-(x)$. An energy function:

$$E_{ij}(p,q) = \begin{cases} f^+(p-q) & \text{if } p-q \geq 0 \\ f^-(q-p) & \text{if } p-q < 0, \end{cases} \quad (9)$$

satisfies (7)*, and is solvable α-expansion.*

Proof: We consider the six possible orderings of p, q and α and indicate which of $f^+(x)$ or $f^-(x)$ is to be applied to each term in (8), based on the positivity (or negativity) of the function argument (9).

	0	$p-q$		$p-\alpha$	$\alpha-q$
$q \leq \alpha \leq p$	f^+	f^+	$\leq$	f^+	f^+
$p \leq \alpha \leq q$	f^-	f^-	$\leq$	f^-	f^-
$p \leq q \leq \alpha$	f^+	f^-	$\leq$	f^-	f^+
$q \leq p \leq \alpha$	f^-	f^+	$\leq$	f^-	f^+
$\alpha \leq p \leq q$	f^+	f^-	$\leq$	f^+	f^-
$\alpha \leq q \leq p$	f^-	f^+	$\leq$	f^+	f^-

In the second column (the function applied to argument 0), we may, and do, choose $f^+(x)$ or $f^-(x)$ according to our needs, since $f^+(0) = f^-(0)$ by definition. For the first two lines of the table, the inequality holds because both $f^+(x)$ or $f^-(x)$ adhere to the triangle inequality. For the remaining lines, the inequality holds because both $f^+(x)$ and $f^-(x)$ are monotonic functions. ■

Within the context of Theorem 1, the smoothness function which encourages pixels to get further away (6) as a function of image plane height can be expressed as:

$$f^-(x_i - x_j) = \begin{cases} 0 & \text{if } x_i - x_j = 0, \\ \tau & \text{otherwise.} \end{cases}$$
$$f^+(x_i - x_j) = \lambda\,|x_i - x_j|$$

Both of these functions satisfy the conditions of Theorem 1, and hence our model, which identifies that neighbouring pixels should take similar values and that the value should not normally increase (although this is permissible) when scanning the transmission map from bottom to top, is solvable by α-expansion.

V. Experiments

The enhancement equation (1) is expressed in terms of radiance incident on the camera, so one must first gamma correct the input foggy image $\mathbf{I}$. Since the horizon radiance is assumed to be constant, we estimate $\mathbf{A}$ before attempting to determine the transmission map $\mathbf{t}$.

A. Airlight Estimate

Like [2], we assume that a portion of the image contains pixels which are infinitely far away — i.e., the horizon. If the horizon is not visible, this assumption may still be valid if the fog is suitably thick (as distance here refers to optic distance and not geometric distance). However, our assumption that the camera is positioned right side up, above the ground, and oriented downwards, means that if a region of the image contains sky, it will most likely occur towards the top of the image. Therefore, we only consider the pixels in the upper fifth of the image when estimating $\mathbf{A}$.

If the input is a colour image, we first calculate the corresponding intensity image. The top 1% brightest pixels are identified within the upper fifth of the intensity image. The horizon radiance $\mathbf{A}$ is calculated as the average value of the pixels in the corresponding foggy input image $\mathbf{I}$.

B. Contrast Data Function

Following [2] we quantify contrast using the sum of image gradients within a patch of the enhanced image $\mathbf{J}(x_i)$ for the hypothesized transmission $t_i(x_i)$. The summation is normalized to the size of the patch (5×5 in our experiments):

$$E_i(x_i) = \frac{1}{25} \sum_{i-\mathbf{2}}^{i+\mathbf{2}} \|\nabla \mathbf{J}_i(x_i)\| \tag{10}$$

C. Dark Channel Data Function

Similar to [4] we compute $\hat{x}_i$ as the minimum value of each pixel RGB triplet in the gamma corrected input image $\mathbf{I}$. However, we use a window size of 5×5 and not 15×15. The smoothing process is controlled by the change in appearance of the enhanced pixel, relative to its hypothesized appearance from the dark channel assumption:

$$E_i(x_i) = \|\mathbf{J}_i(\hat{x}_i) - \mathbf{J}_i(x_i)\|^2 \tag{11}$$

D. Reliability

If the "dark channel" of a pixel is higher than 0.9, it has an appearance which is very similar to the fog. In these situations, the values of $E_i(x_i)$ are scaled by a factor of $0.25\times$ for both the contrast and dark channel models, as the measurements are unreliable in both cases.

E. Smoothness Function

Ideally, the two parameters λ and τ should be tuned for each specific application. However, from our experiments we have found that a value of $\lambda = 0.005$ for 32 labels produced adequate results in all trials. The value of τ was either 100λ or 200λ depending on whether the camera was deemed to be close to ground level or suitably high in the air. In future work, we will investigate methods for estimating both of these parameters automatically.

If two neighbouring pixels in the input foggy image $\mathbf{I}$ are quite similar in appearance (defined as a difference of less than 10 intensity units in each channel), the probability of them having the same depth is quite high. Therefore, when this occurs, we increase the cost of the labelling by $10\times$.

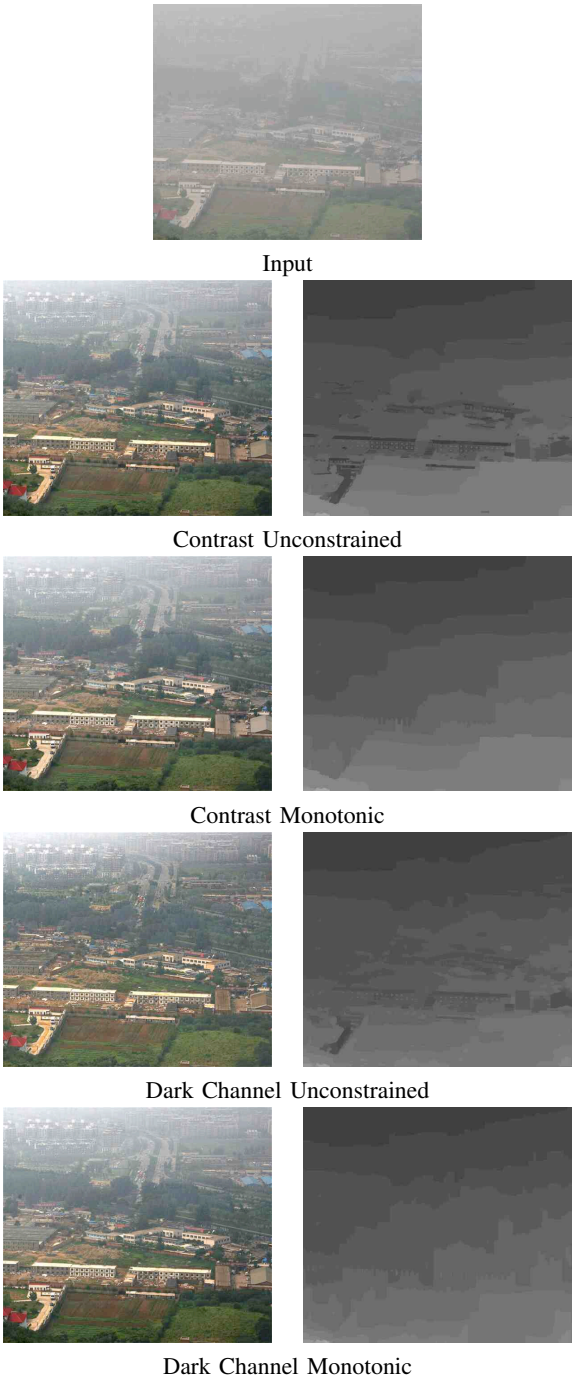

Fig. 4. *A foggy input image is enhanced using the contrast and dark channel data models. Both unconstrained depth maps (which assume that neighbouring pixels should have similar values) contain a significant number of errors. The resulting artifacts in the enhanced images are more severe with the contrast method. The monotonic preference greatly improves the results of the contrast method; less so for the dark channel approach. Since the camera is high above the ground, $\tau = 200\lambda$ was used.*

This formulation minimizes the halo artifacts which arise at significant depth discontinuities in the contrast method [2].

Finally, the assumed ground plane geometry implies depths should change in the vertical direction, but not significantly horizontally. Therefore, we increase the cost by $2\times$ for any label changes in the horizontal direction.

Input

Unconstrained

Monotonic

Fig. 5. *The foggy input image (top) is enhanced using the contrast data measure. In the unconstrained depth map (middle row), the swans are assigned an incorrect depth in an effort to maximize the contrast in their relatively textureless appearance. The monotonic preference (bottom row) is able to correct this mistake. Moreover, no significant errors have been introduced into the image either — i.e., the tree branches at the top of the image are assigned depths which violate the monotonic trend.*

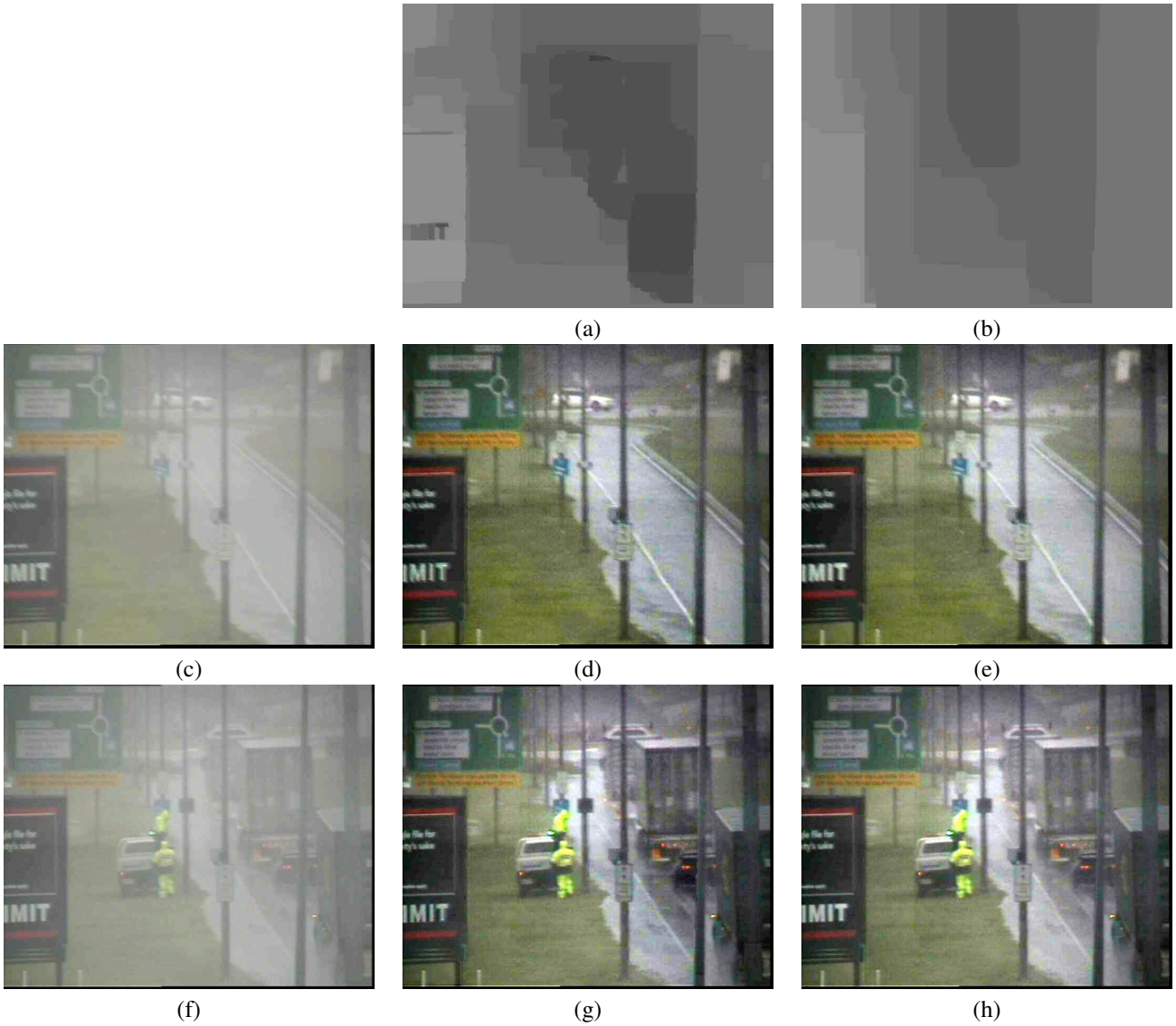

Fig. 6. *Video frames (c) and (f) are enhanced using depth maps (a) and (b) which were estimated without and with monotonicity respectively. Both depth maps were estimated using the contrast data model applied to video frame (c), as the dark channel prior has difficulty with roads. The unconstrained depth map (a) has outliers (due to the appearance of foreground objects and an inherent difficulty with the textureless road), but these do not induce significant artifacts in the enhancement (d). Since the artifacts are not significant, the improvement in the enhancement (e) from using a monotonic depth map (b) is not substantial. However, when the depth map is applied to a video frame (f) captured later, the outliers in (a) over-enhanced the middle vehicle beyond the dynamic range of the image (g). This does not happen in the monotonic correction (h). Although these artifacts may be subtle in a still image, they are quite apparent in a video.*

F. Video

The transportation industry would benefit from automatic fog enhancement technology. For this application, the primary interest is enhancing video — not single images. Fortunately, the image processing routines needed for the enhancement operation (1) are straightforward, and are readily available in most hardware accelerated graphics libraries. As a result, given a depth map, our implementation is able to enhance full-resolution video in real-time using a graphics processing unit.

Since the depth estimation problem is quite involved, our approach applies a single depth map to a series of video frames. During this time, a new estimate of depth is conducted in the background. If the camera is positioned high in the air, the difference in depth between a foreground object and the background behind it is usually small. Therefore, one can typically estimate the depth of a background image and apply this to any video frame without creating significant errors in the enhanced image.

Typically, errors in a depth map do not induce significant artifacts into the enhanced image (see Figures 4 and 5), since the solution is applied to the data used during estimation. Here, the depth map is estimated using one video frame, but applied to another. As a result, the estimation errors caused by various foreground appearances become quite apparent in the enhanced images (see Figure 6).

Fig. 7. *The foggy input image (top) is enhanced using the contrast data measure. The estimated depth of the parked car is wrong (since the appearance of the car is too similar to that of the fog to provide reliable information). The monotonic preference (bottom row) improves the enhancement.*

VI. Conclusions

Previous work on fog enhancement has focused on statistical models for estimating the depth of each pixel. Here, we have explored how the *a priori* camera geometry can be exploited to improve the results of any statistical estimation technique. The expected monotonic trend can be implemented as a soft constraint within an energy minimization framework, which leads to a preference (and not an absolute requirement) that pixels should get further away as one scans the image from bottom to top. Moreover, this preference is fully compatible with the α-expansion algorithm. Finally, our geometric model is not limited just to fog enhancement. One could easily incorporate into other depth estimation techniques, such as stereo disparity estimation.

Acknowledgements

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

[1] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *IJCV*, vol. 48, no. 3, pp. 233–254, 2002.

[2] R. T. Tan, "Visibility in bad weather from a single image," in *CVPR*, 2008.

[3] R. Fattal, "Single image dehazing," in *SIGGRAPH*, 2008.

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

[5] U.S. EPA, *Air Quality Criteria for Pariculate Matter*. U.S. Environmental Protection Agency, 1996, vol. 2.

[6] S. G. Narasimhan and S. K. Nayar, "Interactive (de)weathering of an image using physical models," in *IEEE Workshop on Color and Photometric Methods in Computer Vision*, 2003.

[7] 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 *SIGGRAPH Asia*, 2008.

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

[9] J. P. Oakley and H. Bu, "Correction of simple contrast loss in color images," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 511–522, 2007.

[10] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *PAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.

[11] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, 2001.

[12] X. Liu, O. Veksler, and J. Samarabandu, "Graph cut with ordering constraints on labels and its applications," in *CVPR*, 2008.

[13] S. Ramalingam, P. Kohli, K. Alahari, and P. H. S. Torr, "Exact inference in multi-label CRFs with higher order cliquess," in *CVPR*, 2008.

[14] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *PAMI*, vol. 26, no. 2, pp. 147–159, 2004.

[15] P. Kohli, M. P. Kumar, and P. H. S. Torr, "$\mathcal{P}^3$ & beyond: Solving energies with higher order cliques," in *CVPR*, 2007.

[16] C. Rother, S. Kumar, V. Kolmogorov, and A. Blake, "Digital tapestry," in *CVPR*, 2005.