Using Visibility Cameras to Estimate Atmospheric Light Extinction

Nathan Graves and Shawn Newsam
Electrical Engineering & Computer Science
University of California at Merced
ngraves,snewsam@ucmerced.edu

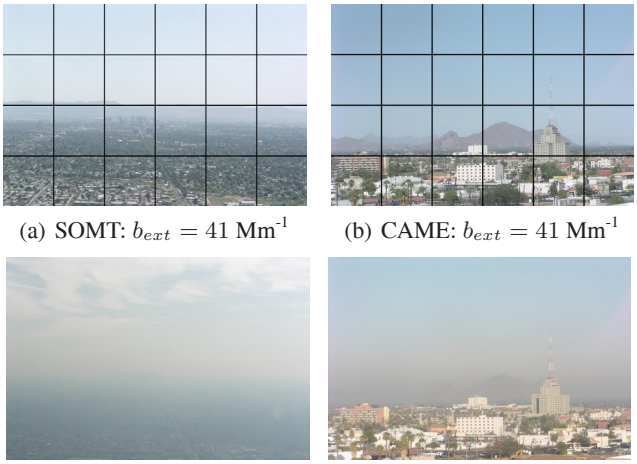
Abstract

We describe methods for estimating the coefficient of atmospheric light extinction using visibility cameras. We use a standard haze image formation model to estimate atmospheric transmission using local contrast features as well as a recently proposed dark channel prior. A log-linear model is then used to relate transmission and extinction. We train and evaluate our model using an extensive set of ground truth images acquired over a year long period from two visibility cameras in the Phoenix, Arizona region. We present informative results which are particularly accurate for a visibility index used in long-term haze studies.

1. Introduction

Quantitative measures of atmospheric visibility are increasingly being used for purposes other than navigation. For example, measures of visibility are being used as indirect estimates of air pollution especially where direct measurements are not available. They are being used to estimate solar irradiance which is important for determining where to situate solar energy farms and for forecasting the near term energy output of existing farms. And, visibility measurements are central to the United States' Environmental Protection Agency's (EPA) goal for improving visual air quality in the Class I Federal areas which include 156 national parks and wilderness areas. In 1977, Congress amended the Clean Air Act with legislation to prevent future and remedy existing impairment of visibility in Class I areas and in 1999, the EPA issued the Regional Haze Rule which mandates that state and federal agencies work together to actually improve the visibility.

Expanding visibility monitoring is key to the EPA's mandates and the agencies charged with monitoring typically use a combination of three techniques. First, they utilize specialized equipment such as transmissometers, which measure light extinction, and nephelometers, which measure light scattering. Second, they use Mie scattering theory to calculate visibility based on measurements of airborne particulates. Finally, and relevant to this work, they deploy networks of visibility cameras. For example, the Interagency Monitoring of Protected Visual Environments (IMPROVE) program has installed and maintains cameras in over two dozen national parks. In addition, regional air quality agencies have deployed visibility camera systems in over 30 cities.



(a) SOMT: $b_{ext} = 41$ Mm^{-1} (b) CAME: $b_{ext} = 41$ Mm^{-1}

(c) SOMT: $b_{ext} = 190$ Mm^{-1} (d) CAME: $b_{ext} = 214$ Mm^{-1}

Figure 1. We investigate methods for estimating the coefficient of extinction b_{ext} using visibility cameras. Shown above are images corresponding to good and poor conditions taken from two such cameras, SOMT and CAME. Ground truth readings from a transmissometer appear in the captions.

This paper focuses on image analysis techniques for deriving quantitative measurements of visibility from such systems. *Visibility cameras are currently used for qualitative purposes only* such as providing visual examples of good and bad days. We feel, however, there is significant opportunity to use these images to derive quantitative measures of visibility perhaps not as accurately as specialized equipment but at much lower cost and possibly even by piggy-backing onto existing web-connected cameras.

We describe methods for estimating the coefficient of light extinction which is a standard measure of atmospheric

978-1-4244-9495-8/10/$26.00 ©2010 IEEE

visibility. We relate this quantity to atmospheric transmission which we estimate based on a standard haze image formation model using measures of local contrast as well as a recently proposed dark channel prior [7]. We derive a log-linear prediction model and perform extensive evaluation using a set of ground truth images acquired over a year long period from two visibility cameras in the Phoenix, Arizona region. Sample images from these two cameras are shown in figure 1. We present informative results which are particularly accurate when mapped to a visibility index used in a multi-year study as part of the EPA Regional Haze Rule.

2. Related Work

There is a sizable body of work on the related problem of improving the fidelity of images taken under hazy or otherwise atmospherically degraded conditions. This includes work by Narasimhan and Nayar on using physics-based models to improve a single image [19, 20] and using multiple images of the same scene but under different conditions [18, 17, 16]; work by Schechner and colleagues on using polarization to improve one or more images [24, 25, 13, 27, 14]; and work by He et al. on using a dark channel prior to dehaze a single image [7]. The objective of this paper, however, is to derive quantitative estimates of atmospheric visibility and so these works are not directly applicable. They can potentially be used to inform the problem as we demonstrate with the dark channel prior.

There is a much smaller body of work on using images to measure atmospheric visibility. Caimi et al. [5] review the theoretical foundations of visibility estimation using image features such as contrast, and describe a Digital Camera Visibility Sensor system, but they do not apply their technique to real data. Kim and Kim [8] investigate the correlation between hue, saturation, and intensity, and visual range in traditional slide photographs. They conclude that atmospheric haze does not significantly affect the hue of the sky but strongly affects the saturation of the sky, but they do not use the image features to estimate visibility. Baumer et al. [3] use an image gradient based approach to estimate visual range using digital cameras but their technique requires the detection of a large number of targets, some only a few pixels in size. This detection step is sensitive to parameter settings and is not robust to camera movement. Also, for ranges over 10 km, they only compare their estimates to human observations which have limited accuracy. Luo et al. [11] use Fourier analysis as well as the image gradient to estimate visibility but they also only compare their estimates to human observations. Raina et al. [22] do compare their estimates to measurements taken using a transmissometer-like device but their approach requires the manual extraction of visual targets. The work by Molenar et al. [12] is closest to the proposed technique in that it is fully automated and the results are compared to transmissometer readings. However, their technique uses a single distant and thus small mountain peak to estimate contrast and thus is very sensitive to camera movement.

In contrast to the works above, our approach is fully automated, does not rely on the detection and segmentation of small targets, is robust to modest camera movement, and performs favorably when compared to ground truth measurements acquired using specialized equipment. We also perform a more thorough investigation into different image features and settings than any of the works above.

3. Background

This section discusses why visibility is reduced by the atmosphere and describes a standard model for the formation of a hazy image that relates atmospheric transmission to the observed image. It then relates transmission to light extinction, the quantity being estimated. Finally it introduces specialized instrumentation for measuring the extinction of light through the atmosphere–transmissometers–and measuring the scattering of light by the atmosphere–nephelometers. These instruments provide the ground truth data for our experiments.

3.1. Why is Visibility Reduced?

Reduced visibility by the intervening atmosphere is mainly due to three first-order processes: 1) light radiating from the scene is absorbed before it reaches an observer; 2) light radiating from the scene is scattered out of the visual pathway of an observer; and 3) ambient light is scattered into the visual pathway of an observer. Absorption and scattering are due to gases and aerosols (particles) suspended in the atmosphere. The combined effect of the absorption and scattering is referred to as the total light extinction. Normally, however, most of the extinction in the atmosphere is due to scattering alone [23] and so in this work we consider the effects of absorption as being negligible.

3.2. Atmospheric Transmission

Atmospheric transmission refers to how well light radiating from a scene is preserved when it reaches an observer. It is a positive scalar quantity ranging from 0 to 1 where larger values indicate improved visibility. Transmission is commonly related to image formation through [28, 6, 15, 17, 7]

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

where x is a two dimensional spatial variable, $I(x)$ is the observed image, $J(x)$ is the scene radiance, A is the ambient (atmospheric) light, and $t(x)$ is the atmospheric transmission. The first term on the right side of eq. 1 is inversely related to the amount of light radiating from the scene that is scattered out of the visual pathway and thus increases with improved transmission. The second term is the amount of

ambient light typically from the sun that is scattered into the visual pathway and thus decreases with improved transmission. In the extremes, the perceived image can either be just the scene radiance or just the scattered ambient light.

3.3. Atmospheric Light Extinction

Atmospheric light extinction is inversely related to transmission through the following exponential equation [26]

$$t(x) = \exp^{-b_{ext}r(x)} \tag{2}$$

where b_{ext} is the extinction coefficient and $r(x)$ is the length of the visual pathway. This assumes the atmosphere is homogeneous along the pathway. We further assume a homogeneous atmosphere throughout a scene. Inverse megameter (Mm^{-1}) is the typical unit of measurement for the extinction coefficient.

3.4. Transmissometer

A transmissometer [1, 10, 4] measures light extinction. It consists of a light source (transmitter) and light detector (receiver), generally separated by a distance of several kilometers, and assesses visibility impairment by measuring the amount of light lost over this known distance. The transmitter emits a uniform light beam of known constant intensity. The receiver separates this light from ambient light, computes the amount of light lost, and reports the extinction coefficient b_{ext}.

3.5. Nephelometer

A nephelometer [1, 21] measures light scattering. It is a compact instrument which measures the amount of light scattered by gases and aerosols in a sampled air volume. It also consists of a transmitter and receiver but configured at an angle so the receiver only receives scattered light. The amount of scattered light is usually integrated over a large range of scattering angles. A nephelometer calculates the scattering coefficient b_{sp} which when added to the absorption coefficient b_{abs} gives the total extinction coefficient $b_{ext} = b_{sp} + b_{abs}$. However, as mentioned above, extinction in the Earth's atmosphere is mostly due to scattering and so we consider b_{ext} and b_{sp} as equivalent.

4. Image Analysis

The goal of this work is to estimate light extinction b_{ext} given an image $I(x)$. We do this by first estimating transmission $t(x)$ from $I(x)$ using eq. 1 and then use eq. 2 to compute b_{ext} (section 5 below on our predictive model discusses how we deal with the unknown value $r(x)$). We investigate two methods for estimating transmission: 1) based on local image contrast as computed in either the spatial or frequency domain; and 2) using a dark channel prior.

4.1. Local Image Contrast

Intuitively, reduced visibility results in an image with less detail especially in the distance. This reduced acuity results from two sources: the objects and their backgrounds become more similar due to increased attenuation and scattering; and the atmosphere acts as a low-pass filter [9], suppressing the higher-frequency image components or details. We use the term local contrast to refer to image acuity and define it as the magnitude of difference in image intensity over a short spatial distance: $C_l(x) = |\nabla_x I(x)|$. The same spatial difference can be computed on the right side of eq. 1 to get

$$\begin{aligned} |\nabla_x I(x)| &= |\nabla_x \left(J(x)t(x) + A(1-t(x))\right)| && (3)\\ &= |\nabla_x J(x)t(x)| && (4)\\ &= t(x)|\nabla_x J(x)| \ . && (5) \end{aligned}$$

Line 4 results from the assumption that the ambient light A is locally constant and line 5 results from the positivity of transmission $t(x)$ and the assumption that it too is locally constant. The quantity $|\nabla_x J(x)|$ is the "true" contrast of the scene when imaged under perfect transmission (section 5 below on our predictive model discusses how we deal with this unknown). This equation shows transmission has the intuitive interpretation as the ratio of the observed contrast to the true contrast.

We now describe two methods for computing local contrast $C_l(x)$.

4.1.1 Contrast in the Spatial Domain

It is natural to consider $|\nabla_x I(x)|$ as the magnitude of the image gradient as computed in the spatial domain. We therefore use Sobel filters to estimate the gradient magnitude at each pixel. To compensate for slight camera movement and other sources of image noise, we compute local contrast in the spatial domain C_{lsd} as the average of the gradient magnitude over an image region Ω:

$$C_{lsd} = \frac{1}{|\Omega|} \sum_{x \in \Omega} |\nabla_x I(x)| \ . \tag{6}$$

Transmission $t(x)$ is assumed constant over this region.

4.1.2 Contrast in the Frequency Domain

A standard way to measure visual acuity is through frequency analysis in the Fourier domain. In particular, the strength or amount of energy in the higher-frequency regions of the Fourier space can be computed by summing the Fourier energy spectral density. Given lower and upper frequencies w_l and w_u, we compute the local contrast in the frequency domain for an image region C_{lfd} as the sum of the square of the magnitude of the two-dimensional

discrete Fourier transform (2D-DFT) $F(u,v)$ in band-pass regions defined by concentric circles centered at the zero-zero or DC frequency:

$$C_{lfd} = \sum_{w_l < \sqrt{u^2+v^2} \leq w_u} |F(u,v)|^2. \quad (7)$$

The cutoff frequencies w_l and w_u can range between 0 and the Nyquist frequency w_{Ny} and determine whether the energy is computed in a low-pass, band-pass, or high-pass region. The DC frequency is never included since it is the average value of an image region and thus not indicative of acuity.

4.2. Dark Channel Prior

We also estimate transmission using a dark channel prior based on the work by He at al. [7] on single image dehazing. The intuition is that one can reasonably expect that any natural image to have a dark region which has very low intensity values in at least one of the color channels when imaged under perfect transmission. Thus, the difference between the observed intensity and the expected low intensity for these image regions–the prior–is indicative of the loss of transmission. He et al. use estimated transmission based on a dark channel prior to perform image correction (dehazing). We use it here to estimate light extinction.

The derivation is as follows [7]. Starting with the haze image formation model, we determine the minimum intensity value in color channel c for an image region Ω:

$$\min_{x\in\Omega}\left(I^c(x)\right) = \min_{x\in\Omega}\left(J^c(x)t(x) + A^c(1-t(x))\right) . \quad (8)$$

Assuming that the transmission and ambient light are constant in the region, this is equivalent to

$$\min_{x\in\Omega}\left(\frac{I^c(x)}{A^c}\right) = t(x)\min_{x\in\Omega}\left(\frac{J^c(x)}{A^c}\right) + (1-t(x)) . \quad (9)$$

Now the minimum is computed with respect to each color channel

$$\min_c\left(\min_{x\in\Omega}\left(\frac{I^c(x)}{A^c}\right)\right) = t(x)\min_c\left(\min_{x\in\Omega}\left(\frac{J^c(x)}{A^c}\right)\right) + (1-t(x)) .$$

Looking more closely at the right hand side of this equation, we realize that based on the dark channel prior, which again assumes there is some region with zero or near-zero haze-free intensity in one of the color channels, that

$$\min_c\left(\min_{x\in\Omega}\left(\frac{J^c(x)}{A^c}\right)\right) = 0 . \quad (10)$$

since A^c is positive. We thus get

$$t(x) = 1 - \min_c\left(\min_{x\in\Omega}\left(\frac{I^c(x)}{A^c}\right)\right) . \quad (11)$$

We estimate the ambient light A^c as the maximum first percentile of pixel intensities in a region just above the horizon. And, in order to be robust to outliers, we compute the minimum ambient-light normalized image intensity for a region $\min_{x\in\Omega}\left(\frac{I^c(x)}{A^c}\right)$ as the minimum first percentile.

5. The Prediction Model

Again, the primary objective is to estimate light extinction b_{ext} given an image $I(x)$. Taking the log of both sides of eq. 2 gives a linear relationship between extinction and transmission

$$\ln t(x) = b_{ext} r(x) . \quad (12)$$

In this case of transmission based on local contrast computed in either the spatial domain C_{lsd} or frequency domain C_{lfd} this becomes

$$\ln C_l = \ln|\nabla_x J(x)| + b_{ext} r(x) \quad (13)$$

where $|\nabla_x J(x)|$ is the true contrast of the scene. Rearranging we get

$$b_{ext} = \frac{\ln C_l}{r(x)} - \frac{\ln|\nabla_x J(x)|}{r(x)} \quad (14)$$

and use linear least squares regression (LLSR) to learn the scaling $\frac{1}{r(x)}$ and offset $-\frac{\ln|\nabla_x J(x)|}{r(x)}$ parameters from a labelled training set.

In the case of transmission based on the dark channel prior, eq. 12 becomes

$$b_{ext} = \frac{\ln t(x)}{r(x)} . \quad (15)$$

However, we found that pure scaling results in poor performance so we include an offset to accommodate for errors in the model–perhaps there is no dark pixel in the image region–and/or errors in the observations such as unreliable estimates of the ambient light. We again use LLSR to learn the scaling parameter and the offset.

6. Dataset

We evaluate our method using an extensive set of images and ground truth extinction readings from the the Arizona Department of Environmental Quality which manages the PhoenixVis.net visibility web cameras website [1]. This website contains live images from six visibility cameras of scenic urban and rural vistas in the Phoenix, Arizona region. Our dataset consists of the following acquired over 2006:

- Digital images of South Mountain (SOMT) captured every 15 minutes.
- Digital images of Camelback Mountain (CAME) captured every 15 minutes.

- The extinction coefficient b_{ext} measured every hour using a transmissometer.
- The scattering coefficient b_{sp} measured every hour using a nephelometer.

The SOMT camera is located on a mountain north of Phoenix and faces south. Figures 1(a) and 1(c) contain examples of good and bad visibility for the SOMT camera. The CAME camera is located on a tall structure in downtown Phoenix and faces north east. Figures 1(b) and 1(d) contain examples of good and bad visibility for the CAME camera. The transmissometer and nephelometer are located in downtown Phoenix and are approximately within the field of view of both cameras.

All images are in the RGB colorspace and have been JPEG compressed at an unknown quality level. The SOMT images measure 1536×1024 pixels. The CAME images are a mix of 1536×1024 and 1230×820 pixels so we transform all CAME images to a common size of 1230×820 using bilinear interpolation. Each image is partitioned using a 6×4 grid and a prediction model is trained and evaluated for each block separately. Figure 1 shows the grid layout for the two scenes.

We only consider images taken at the top of each hour, since this is when the transmissometer and nephelometer readings are made, and during daylight hours, approximately 10 am to 4 pm. This results in a labelled dataset of 8,598 images from the SOMT camera and 7,676 images from the CAME camera.

7. Experiments

We evaluate our method based on how well the learned model is able to predict the (known) extinction coefficient b_{ext} corresponding to an image $I(X)$ using only the image features. We perform five-fold cross-validation to observe how well our method generalizes. The labelled images are randomly partitioned into five equal-sized sets. The model is learned using four of the sets and used to predict the extinction coefficient for the images in the fifth held-out set. We evaluate the accuracy of our model using the coefficient of determination R^2 between the predicted and ground truth values. Let ${b^i_{ext}}'$ and b^i_{ext} be the predicted and true extinction coefficients for image i then

$$R^2 = 1 - \frac{\sum_{i=1}^{n} ({b^i_{ext}}' - b^i_{ext})^2}{\sum_{i=1}^{n} (b^i_{ext} - \overline{b_{ext}})^2} \tag{16}$$

where n is the number of images in the evaluation set and $\overline{b_{ext}}$ is the mean of the true values. R^2 has a maximum value of 1 with higher values indicating a more accurate model.

In order to provide an intuitive feel for the predictions, we also selectively report the mean absolute error (MAE) between the predicted and true values:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |{b^i_{ext}}' - b^i_{ext}| \tag{17}$$

The values of R^2 and MAE reported below are averages over the five training/test splits.

The evaluation is performed on each of the 24 image blocks separately. The visual distance $r(x)$ and transmission $t(x)$ in eqs. 14 and 15 are assumed to be constant over a block. The image region Ω used to compute local contrast in the spatial domain in eq. 6 and the transmission based on the dark channel prior in eq. 11 is taken as an image block. Contrast in the frequency domain is computed by applying the 2D-DFT to an image block.

We perform a series of experiments to: determine which image feature is most effective for predicting the coefficient of extinction; whether the predictions are more correlated with the transmissometer or nephelometer readings; the effect of scene geometry; and the optimal lower and upper cutoff frequencies for the Fourier analysis.

8. Results

The results are summarized in table 1. For each combination of *image feature*–local contrast in the spatial (C_{lsd}) or frequency (C_{lfd}) domain, or dark channel prior; *ground truth readings*–transmissometer or nephelometer; and *scene*–SOMT or CAME–it lists the R^2 and MAE values for the image block that results in the best model as ranked by R^2. The 6×4 image blocks are numbered 1 through 24 in raster-scan order (see figure 1). The values reported for C_{lfd} are the best over a range of lower and upper frequency bounds. We now discuss these results.

Image Features Local contrast consistently outperforms the dark channel prior across scenes and ground truth labelling. Further, contrast computed in the spatial domain using the image gradient generally outperforms contrast computed in the frequency domain using the Fourier energy spectral density. These two contrast features are of course related and will be discussed further in section 8.2 below on frequency bands.

Ground Truth Readings For the SOMT scene, local contrast is a better predictor of the transmissometer than the nephelometer readings while the reverse is true for the dark channel prior. While our model assumes that the effect of absorption is negligible, as is commonly done in atmospheric modelling, the ground truth transmissometer and nephelometer values in our dataset are different indicating there is a non-zero absorption component b_{abs}; i.e., $b_{ext} \neq b_{sp} + b_{abs}$. Table 2 gives the statistics of the ground truth readings. Performing a linear least squares fit between the 8,598 transmissometer and nephelometer readings associated with scene SOMT gives $b_{ext} = 1.18 b_{sp} + 24.6$ with

Table 1. Summary of results for each combination of *image feature*–local contrast in the spatial (C_{lsd}) or frequency (C_{lfd}) domain, or dark channel prior; *ground truth reading*–transmissometer or nephelometer; and *scene*–SOMT or CAME. R^2 and MAE (Mm^{-1}) values are given for the image block that results in the best model as ranked by R^2 (higher is better).

	SOMT						CAME					
	Transmissometer			Nephelometer			Transmissometer			Nephelometer		
	R^2	MAE	block	R^2	MAE	block	R^2	MAE	block	R^2	MAE	block
C_{lsd}	**0.646**	**12.5**	**14**	0.407	10.4	14	0.103	21.6	14	0.199	12.0	14
C_{lfd}	0.549	14.8	21	0.372	10.2	18	0.103	21.5	18	0.169	12.3	18
Dark Channel	0.237	20.3	15	0.309	10.8	16	0.081	21.3	9	0.061	13.0	11

Table 2. The statistics for the ground truth transmissometer and nephelometer readings. All values are in Mm^{-1} except for R^2.

	SOMT (8,598 pts)		CAME (7,676 pts)	
	Trans.	Neph.	Trans.	Neph.
min	9	0	9	0
max	218	153	218	155
mean	57.8	28.2	57.7	28.1
median	50	24	50	24
std dev	31.2	20.1	31.0	19.8
R^2	0.571		0.567	

an R^2 value of 0.571 and an MAE of 15.1 Mm^{-1}. This indicates that there is a nonlinear relation between these two measurements which cannot be accounted for in the scaling and offset parameters of our linear model of b_{ext} and $\ln C_l$ (eq. 14) or $\ln t(x)$ (eq. 15). It also shows interestingly that *once calibrated, the image features can provide a better estimate of light extinction (as measured using the transmissometer) than the nephelometer*.

Returning to table 1, we see that things are reversed for the the CAME scene: local contrast is a better predictor of the nephelometer than the transmissometer readings while the dark channel prior is a better predictor of the transmissometer than the nephelometer readings. We are investigating the reasons for this.

Scenes The image based approach to estimating light extinction is significantly more effective for SOMT than the CAME. This is true for all feature and ground truth combinations. This might be due in part to the different image resolutions particularly for the local contrast approaches (we have not yet done a control experiment in which we analyze lower resolution versions of the SOMT images). It is more likely due to different scene geometry as discussed in the next section.

8.1. Image Regions

Figures 2-5 plot the R^2 values for each of the three image features over all 24 image blocks. Figure 2 contains the results for scene SOMT and the transmissometer readings; figure 3 for scene SOMT and the nephelometer readings; figure 4 for scene CAME and the transmissometer readings; and figure 5 for scene CAME and the nephelometer readings.

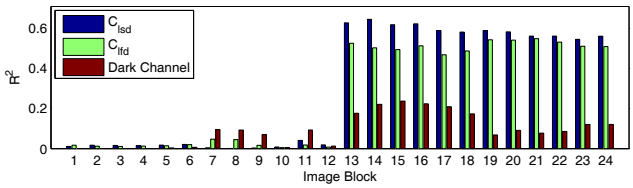


Figure 2. R^2 for different image blocks for SOMT and transmissometer readings.

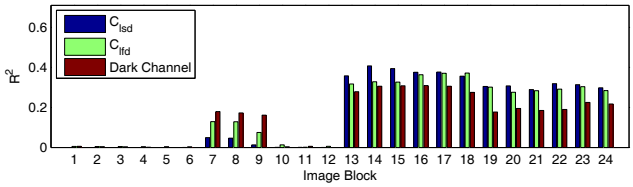


Figure 3. R^2 for different image blocks for SOMT and nephelometer readings.

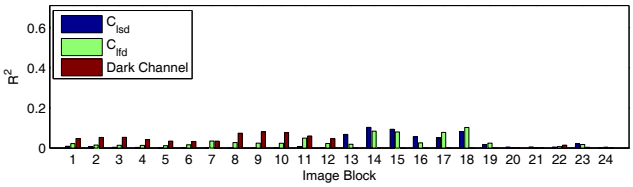


Figure 4. R^2 for different image blocks for CAME and transmissometer readings.

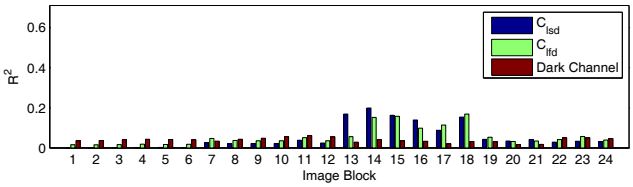
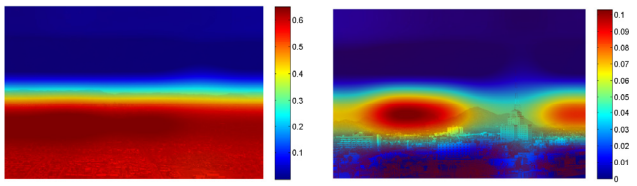


Figure 5. R^2 for different image blocks for CAME and nephelometer readings.

These region level results provide insight into why SOMT is the more effective scene. The R^2 values for all features and ground truth readings are relatively large for all the blocks below or containing the horizon in SOMT. However, the bottom row of blocks (19-24) in CAME which represent the closest parts of the foreground all perform poorly. This foreground region is much closer than any of the SOME regions and thus is too close to the camera



(a) SOMT (b) CAME

Figure 6. Smoothed colormaps of R^2 overlaid on scene images indicating the effect of scene geometry. The distant regions are the most effective in both scenes; the "sweet spot" for CAME, however, is much smaller.

to estimate light extinction. There is simply not enough atmosphere to cause sufficient variation in the image features. This is evident in figures 1(c) and 1(d) in which the foreground regions of the two scenes are affected very differently by a similar increase in light extinction. Further, the lower vantage point of the CAME camera results in a perspective with very little distant scenery in terms of image area. The image features are now extracted from blocks containing sky regions which results in worse performance than SOME.

The effect of different scene geometry is visually depicted in in figures 6(a) and 6(b) using smoothed colormaps of R^2 overlaid on SOMT and CAME images. These results correspond to predicting the transmissometer readings using local contrast in the spatial domain. The distant regions are the most effective in both scenes; the "sweet spot" for CAME, however, is much smaller.

8.2. Frequency Bands

The LLSR fit of local contrast in the frequency domain C_{lfd} is performed for lower w_l and upper w_u cutoff frequencies ranging from 0 to 1 in increments of 0.05 where 1 corresponds to the Nyquist frequency. The R^2 values reported in table 1 and figures 2-5 represent the optimal cutoffs. Figure 7 shows how the optimal frequencies vary by image block for SOMT using the transmissometer readings. In particular, the values for non-sky blocks (7-9 and 13-24) decrease as the scene distance increases. This makes sense because even in relatively good conditions, the atmosphere still acts as a low-pass filter whose attenuation increases with distance and so the higher frequency image signal components for distant scenes do not vary enough to be informative. We also see that w_u never equals the maximum frequency in non-sky regions even when they are close-by. This may be in part due to the low-pass filtering of the atmosphere but is more likely due to JPEG compression which discards the higher frequency signal components.

The two measures of local contrast are of course related since convolution with the Sobel kernels in the spatial domain corresponds to applying a related filter in frequency domain. The fact that contrast in the spatial domain provides the best result overall indicates that the optimal filter for estimating atmospheric light extinction does not have a band-pass frequency response but is more complex. The frequency response of the Sobel kernels might provide a good initial estimate of this filter. This filter would of course need to be tuned to scene distance just like our simple band-pass filters.

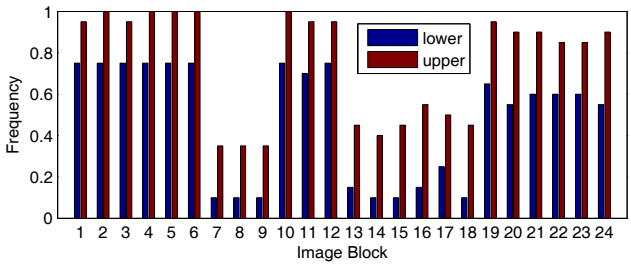


Figure 7. Per block values of the optimal lower w_l and upper w_u frequencies for local contrast in the frequency frequency domain. This is for the SOMT scene and transmissometer readings. A value of 1 corresponds to the Nyquist frequency.

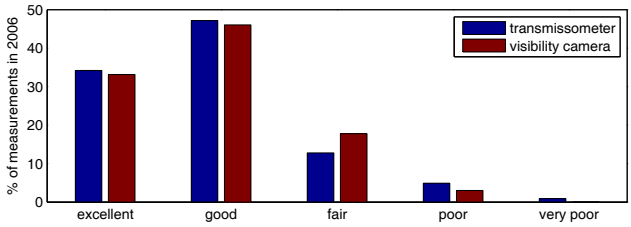


Figure 8. Comparison of visibility index measured using a transmissometer and a visibility camera.

8.3. Visibility Index Based on Deciview Analysis

In 2003, the Arizona Department of Environmental Quality defined a five valued visibility index [2] to track regional visibility conditions over a multi-year period. This index is based on deciviews which are linear with respect to perceived visual changes analogous to how decibels are for sound [21]. Deciview readings DV are derived from transmissometer estimates of light extinction:

$$DV = 10 \ln(b_{ext}/10 \text{ Mm}^{-1}) \; . \tag{18}$$

The visibility index is then determined by binning DV into five ranges corresponding to "excellent", "good", "fair", "poor", and "very poor".

We applied the same deciview conversion and binning to the predictions of our best model, local contrast features in the spatial domain for block 14 of SOMT. Figure 8 compares the predicted index values with those computed using the ground truth transmissometer readings. Note the similarity in the distributions. This fit corresponds to an R^2 value of 0.980 if the values are mapped to 1,...,5.

9. Conclusion

We proposed a method for estimating the coefficient of light extinction using visibility cameras. We used an extensive ground truth dataset to compare a number of model and environmental settings including different image features and the effects of scene geometry. We presented informative results which are particularly accurate when mapped to a visibility index being used in a multi-year study as part of the EPA Regional Haze Rule.

We plan to extend this work by exploring models which do not require ground truth data for calibration. Our long-term goal is to compute visibility index distributions such as in figure 8 using commodity web cameras in a completely automated manner. The results presented in this paper will inform this future work.

Acknowledgements This work was partially funded by the Center for Information Technology Research in the Interest of Society (CITRIS). We would like to thank the Arizona Department of Environmental Quality and Air Resource Specialists, Inc., for providing the images and transmissometer data.

References

[1] Phoenix visibility web camera website managed by the Arizona Department of Environmental Quality (ADEQ). `http://www.phoenixvis.net/`.

[2] Recommendation for a Phoenix Area Visibility Index by the Visibility Index Oversight Committee, March 5, 2003.

[3] D. Baumer, S. Versick, and B. Vogel. Determination of the visibility using a digital panorama camera. *Atmospheric Environment*, 42(11):2593–2602, 2008.

[4] J. Betts. The instrumental assessment of visual range. *Proceedings of the IEEE*, 59(9):1370–1371, September 1971.

[5] F. Caimi, D. Kocak, and J. Justak. Remote visibility measurement technique using object plane data from digital image sensors. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 3288–3291, 2004.

[6] R. Fattal. Single image dehazing. In *ACM SIGGRAPH*, 2008.

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

[8] K. Kim and Y. Kim. Perceived visibility measurement using the HSI color different method. *Journal of the Korean Physical Society*, 46(5):1243–1250, 2005.

[9] V. Krishnakumar and P. Venkatakrishnan. Determination of the atmospheric point spread function by a parameter search. *Astronomy & Astrophysics Supplement Series*, 126:177–181, November 1997.

[10] P. Lee, T. Hoffer, D. Schorran, E. Ellis, and J. Moyer. Laser transmissometer–a description. *Science of The Total Environment*, 23:321–335, 1982.

[11] C.-H. Luo, C.-Y. Wen, C.-S. Yuan, J.-J. Liaw, C.-C. Lo, and S.-H. Chiu. Investigation of urban atmospheric visibility by high-frequency extraction: Model development and field test. *Atmospheric Environment*, 39(14):2545–2552, 2005.

[12] J. V. Molenar, D. S. Cismoski, F. Schreiner, and W. C. Malm. Analysis of digital images from Grand Canyon, Great Smoky Mountains, and Fort Collins, Colorado. In *Regional and Global Perspectives on Haze: Causes, Consequences and Controversies Visibility Specialty Conference*, 2004.

[13] E. Namer and Y. Y. Schechner. Advanced visibility improvement based on polarization filtered images. In *Proc. SPIE 5888: Polarization Science and Remote Sensing II*, pages 36–45, 2005.

[14] E. Namer, S. Shwartz, and Y. Y. Schechner. Skyless polarimetric calibration and visibility enhancement. *Optics Express*, 17(2):472–493, 2009.

[15] S. Narasimhan and S. Nayar. Chromatic framework for vision in bad weather. In *CVPR*, 2000.

[16] S. Narasimhan and S. Nayar. Removing weather effects from monochrome images. In *CVPR*, 2001.

[17] S. Narasimhan and S. Nayar. Vision and the atmosphere. *International Journal on Computer Vision*, 48(3):233–254, July 2002.

[18] S. Narasimhan and S. Nayar. Contrast restoration of weather degraded images. *PAMI*, 25(6):713–724, June 2003.

[19] S. Narasimhan and S. Nayar. Interactive (de)weathering of an image using physical models. In *ICCV Workshop on Color and Photometric Methods in Computer Vision*, 2003.

[20] S. Narasimhan and S. Nayar. Shedding light on the weather. In *CVPR*, 2003.

[21] M. L. Pitchford and W. C. Malm. Development and applications of a standard visual index. *Atmospheric Environment*, 28(5):1049–1054, 1994.

[22] D. S. Raina, N. J. Parks, W.-W. Li, R. W. Gray, and S. L. Dattner. Innovative monitoring of visibility using digital imaging technology in an arid urban environment. In *Regional and Global Perspectives on Haze: Causes, Consequences and Controversies Visibility Specialty Conference*, 2004.

[23] M. G. Ruby and A. P. Waggoner. Intercomparison of integrating nephelometer measurements. *Environmental Science & Technology*, 15(1):109–113, 1981.

[24] Y. Schechner, S. Narasimhan, and S. Nayar. Instant dehazing of images using polarization. In *CVPR*, 2001.

[25] Y. Schechner, S. Narasimhan, and S. Nayar. Polarization-based vision through haze. *Applied Optics, Special issue*, 42(3):511–525, January 2003.

[26] J. Seinfeld and S. Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley, 2006.

[27] S. Shwartz, E. Namer, and Y. Y. Schechner. Blind haze separation. In *CVPR*, 2006.

[28] R. Tan. Visibility in bad weather from a single image. In *CVPR*, 2008.