ESTIMATION OF MULTIPLE ATMOSPHERIC POLLUTANTS THROUGH IMAGE ANALYSIS

Tony Zhang and Robert P. Dick

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, Michigan 48109

ABSTRACT

Multiple atmospheric pollutants, such as $PM_{2.5}$, PM_{10}, and NO_2, degrades air quality in many parts of the world. Fine-grained air pollution data can help combat the problem, but conventional monitoring stations are too expensive to support high spatial resolution; image-based estimates have the potential to improve spatial coverage. We estimate pollutant concentrations from images using the position-and color-dependent properties of scattering and absorption. We are the first to use images to estimate pollutant concentrations in systems with multiple pollutants. We achieve this by considering the differences in scattering and absorption spectra between different pollutants. Our system improves the accuracy of $PM_{2.5}$, PM_{10}, and NO_2 estimation by 22% for single-scene images in Beijing and Shanghai compared to the best existing image-based techniques.

Index Terms— Air Quality, Light Attenuation, Support Vector Regression, Atmospheric Modeling

1. INTRODUCTION

A 2014-2015 study found that only 25 out of 190 cities within China met the National Ambient Air Quality Standards (NAAQS) [1]. The annual average concentration of $PM_{2.5}$ for the 190 cities was 57 $\mu g/m^3$, which is 62.9% above the NAAQS limit [1]. Traditional stationary air quality sensors are sparsely deployed, with only 35 in Beijing [2]. To combat air pollution, it is important to determine how pollutant concentrations vary at high spatial and temporal resolutions. Deploying numerous conventional sensors is unrealistic because they are expensive and require maintenance. Since air quality can be estimated by observing haze effects in images, digital cameras can be used to quantify and analyze pollutants across large areas. Their ubiquity suggests the possibility of improving both spatial and temporal sampling resolutions.

Estimating air quality from images is an active research area. He et al. propose using a dark channel defined as the darkest pixels within the localized patches to estimate airlight and transmission [3]. Liu et al. describe various image features for estimating $PM_{2.5}$ concentration [4]. Liu et al. estimate $PM_{2.5}$ concentration through support vector regression using image features and weather data [5]. Their approach requires manually labeling regions of interest. Li et al. [6] estimate $PM_{2.5}$ concentration using depth and transmission maps, where depth is estimated via deep convolutional neural fields. Liu et al. implemented an image crowdsensing system to obtain $PM_{2.5}$ concentration [7]. The above work neglects absorption and estimates $PM_{2.5}$ concentration only.

This paper presents a method of estimating air quality that considers pollutant spectra, models scattering and absorption in the atmosphere, and considers non-uniform spatial distributions. This is the first system that uses images to estimate the concentrations of multiple pollutants simultaneously, namely $PM_{2.5}$, PM_{10}, and NO_2, and to use images for NO_2 estimation. Our experiments on datasets from Beijing and Shanghai show that considering position-dependent color properties improves accuracy by 22%.

2. VISIBILITY PHYSICS

Previous research on estimating air quality from images uses an atmospheric model to describe an image influenced by haze [8] [9]:

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

$$t(x) = e^{-\beta d(x)}, \tag{2}$$

where x is the location of the pixel, I is the observed image, J is scene radiance (the image without any haze), A is atmospheric light, t is transmission, β is the scattering coefficient, and d is depth. The atmospheric light, commonly known as the airlight, results from scattering and absorption. The color of the original scene shifts towards airlight, the aggregate color of the atmospheric particles and gases, as distance increases.

Eqs. (1) and (2) are based on the following three assumptions: the properties of light attenuation are color-independent, the light attenuation coefficient is influenced by scattering only (absorption is negligible), and the atmosphere is homogeneous. We show that these assumptions do not hold and give a more accurate atmospheric model.

The level of visibility in the atmosphere is highly influenced by pollutants and weather conditions. When light trav-

els through the atmosphere, it encounters particles and gases that affect its path. Light attenuation is caused by both scattering and absorption. Light scattering by particles is the main cause of reduced visibility, but light absorption by both particles and gases accounts for up to 30% of reduced visibility in urban areas [10] [11] [12].

Scattering and absorption are wavelength-dependent. For components of $PM_{2.5}$ and PM_{10} that are smaller than the wavelength of light, relative scattering is inversely proportional to wavelength. In addition, the relative absorptions of $PM_{2.5}$ and PM_{10} are inversely proportional to wavelength [13] [14] [15]. Also, NO_2 absorbs blue light heavily [16].

The atmospheric model in Eq. (1) and (2) demonstrates two mechanisms that influence images: direct attenuation and airlight. The direct attenuation represented by $J(x)t(x)$ causes the intensity of the pixels to decrease in a multiplicative manner and due to scattering and absorption. The $t(x)$ term in direct attenuation models both scattering and absorption since the scene radiance $J(x)$ models neither.

Airlight ($A(1-t(x))$) represents aggregate pollutant color of the atmospheric particles and gases due to scattering and absorption. The effect of the airlight on the light intensities is additive, and the airlight increases as more light get scattered due to particles. Since the airlight implicitly accounts for absorption, $t(x)$ only accounts for scattering.

We extend the atmospheric model from Eqs. (1) and (2) to account for both scattering and absorption, as follows:

$$I_c(x) = J_c(x)t_{1_c}(x) + A_c(1 - t_{2_c}(x)), \tag{3}$$

$$t_{1_c}(x) = e^{-(\beta_{s_c}+\beta_{a_c})d(x)}, \text{ and} \tag{4}$$

$$t_{2_c}(x) = e^{-\beta_{s_c}d(x)}. \tag{5}$$

Eq. (3) incorporates transmission as a function of both scattering and absorption (t_1), and transmission as a function of only scattering (t_2). In Eqs. (4) and (5), β_s is the scattering coefficient and β_a is the absorption coefficient. Additionally, every variable represents color-dependent light attenuation through a subscript c. Wavelength-dependent light attenuation is associated with RGB color channels.

The atmosphere model used in prior work also assumes that the attenuation coefficient (β) is constant for an entire image. In realistic conditions, the density of particles and gases changes as a function of position and altitude, leading to a non-uniform light attenuation coefficient [17] [18]. We explicitly consider this effect.

3. METHODOLOGY

Our technique consists of two main steps: obtaining the transmissivities of scattering and absorption for all three color channels based on Eqs. (3) to (5), and obtaining predicted concentrations for $PM_{2.5}$, PM_{10}, and NO_2 based on the tranmissivities from the prior step.

Algorithm 1: Gradient Descent

Input: $I(x)$, $J(x)$, $d(x)$, A, height, width
Output: β_s, β_a

1 **while** $\mid \beta_a - \beta'_a \mid > \gamma$ *or* $\mid \beta_s - \beta'_s \mid > \sigma$ **do**
2 $\quad \beta'_a = \beta_a$, $\beta'_s = \beta_s$
3 $\quad \hat{I}(x) = J(x)e^{-(\beta_s+\beta_a)d(x)} + A(1 - e^{-\beta_s d(x)})$
4 $\quad \epsilon(x) = (\hat{I}(x) - I(x))/$ (height $\times$ width)
5 $\quad C(x) = \frac{1}{2} \times (\hat{I}(x) - I(x))^2$
6 $\quad \frac{dC(x)}{d\beta_s} = \epsilon(x)d(x)e^{-\beta_s d(x)}(A - J(x)e^{-\beta_a d(x)})$
7 $\quad \frac{dC(x)}{d\beta_a} = -\epsilon(x)d(x)J(x)e^{-(\beta_s+\beta_a)d(x)}$
8 $\quad \beta_s = \beta_s - \alpha \times \sum_x \frac{dC(x)}{d\beta_s}$
9 $\quad \beta_a = \beta_a - \alpha \times \sum_x \frac{dC(x)}{d\beta_a}$
10 **end**

3.1. Obtaining Transmissivities

We used the atmospheric model described in Eq. (3) to (5) to obtain β_{s_c} and β_{a_c} for all colors (c). $I(x)$ is the input image. Using the webcam image dataset and ground truth pollutants, we obtain $J(x)$ by collecting the images with the lowest PM_{10} concentrations and taking the mean of their color intensities so $J(x)$ contains as little air pollution as possible (typically about 5% of the maximum). The depth map is obtained by running a convolutional neural network by Li et al. [19] on $J(x)$. The airlight is estimated using the technique in Berman et al. [20]. Afterward, the only unknown variables in Eq. (3) are β_{s_c} and β_{a_c}.

We use gradient descent to find β_s and β_a by minimizing the cost function $C(x) = \frac{1}{2}(\hat{I}(x) - I(x))^2$, where $\hat{I}(x)$ is the predicted image calculated in Eq. (3) and $I(x)$ is the actual image. The gradient descent algorithm is shown in Algorithm 1. To improve convergence, the algorithm keeps track of the last ten calculated β_{s_c} and β_{a_c} values. If the past beta values are stable, the step size (i.e., learning rate or α) decreases in lines 8 and 9.

In the experimental setup, we split an image into an $n \times n$ grid and obtain β_{s_c} and β_{a_c} for each of the n^2 grid elements using gradient descent. Increasing the number of grid elements to obtain additional light attenuation coefficients (β_{s_c} and β_{a_c}) increases estimation accuracy. This suggests that the per-element transmissions are more accurate than the global aggregate transmission. If that is the case, it is possible to use variance in per-element transmissions to estimate the actual variance in transmissions. Table 1 shows the standard deviation in light attenuation over all grid elements in the image, which increases as n increases.

3.2. Estimation of Pollutant Concentrations

After all coefficients β_{s_c} and β_{a_c} are extracted from each image, we determine their relationships with pollutant concen-

Table 1. Standard Deviation of Light Attenuation with Resolution

	Beijing			Shanghai		
	β_{s_b}	β_{s_g}	β_{s_r}	β_{s_b}	β_{s_g}	β_{s_r}
2×2	0.29	0.37	0.45	0.17	0.16	0.15
4×4	0.34	0.46	0.59	0.24	0.28	0.30
6×6	0.43	0.54	0.64	0.29	0.31	0.33
8×8	0.45	0.58	0.66	0.33	0.35	0.36
10×10	0.49	0.63	0.71	0.34	0.36	0.37

trations using support vector regression. Since $PM_{2.5}$, PM_{10}, and NO_2 have different color-dependent properties for scattering and absorption, it is possible to predict all the pollutants from a single image. Given a dataset $\{(x_1, y_1), ..., (x, y)\}$, we find a regression function $f(x) = w \times \phi(x) + b$ that solves the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}||w||_2^2 + \frac{1}{\lambda}\sum_{n=1}^{m}(\xi_i^+ + \xi_i^-) \\ \text{subject to} \quad & f(x_i) - y_i \leq \epsilon + \xi_i^+, \\ & y_i - f(x_i) \leq \epsilon + \xi_i^-, \text{ and} \\ & \xi_i^+ \geq 0,\ \xi_i^- \geq 0,\ i = 1, \ldots, m. \end{aligned}$$

This formulation uses l_2 regularization in the case of non-linearly separable datasets and outliers, where λ is the regularization parameter. A radial basis kernel function is used.

4. RESULTS AND DISCUSSION

This section describes data collection, experimental evaluation, and findings.

4.1. Data Collection and Experimental Evaluation

The data consist of single-scene images taken in Beijing and Shanghai and their ground truth pollutant concentrations. The images were taken with the same camera location and angle. The Beijing dataset consists of 328 images taken by Yi Zou in 2014 at the Beijing Television Tower [21]. Since the original dataset had varying image sizes, we resize each image in Beijing to 600 by 800. The Shanghai dataset consists of 1,890 images taken from May to December in 2014 at various times and were captured at the Oriental Pearl Tower [22]. We use the $PM_{2.5}$, PM_{10}, and NO_2 data provided by sensor stations within the cities as ground truth [4] [23]. The units for $PM_{2.5}$ and PM_{10} are μg/m^3 and for NO_2 are parts per billion (ppb).

The SVR model uses $6n^2$ features from β values evaluated with two-fold cross validation. The regularization parameter (λ) value for Beijing is $\frac{1}{400}$ and for Shanghai is $\frac{1}{200}$.

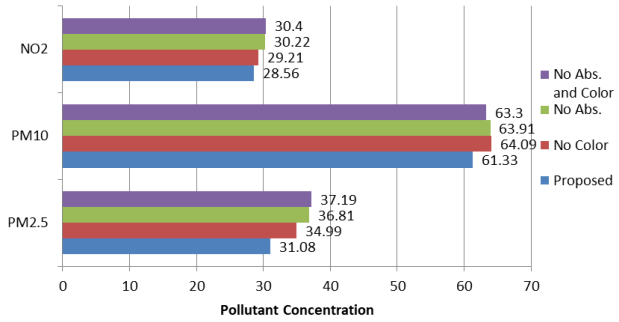

Fig. 1. RMSE for absorption and color (Beijing).[1]

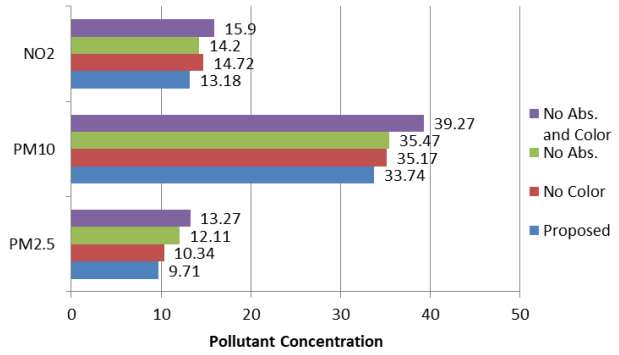

Fig. 2. RMSE for absorption and color (Shanghai).[1]

The two evaluation metrics used are the R^2 coefficient of determination and the root mean squared error (RMSE) between the estimated and ground truth pollutant concentration.

4.2. Effect of Absorption and Color Properties

We evaluate the impact of absorption and color-dependent light attenuation. We use the 6×6 grid size for Beijing and the 8×8 grid size for Shanghai. When absorption is neglected, β_{s_c} is still determined using gradient descent for all colors (c) in Eq. (1) and (2). We also consider neglecting color-dependent properties and find β_s and β_a using gradient descent on a grayscale version of the problem. As shown in Fig. 1 and 2, considering each property generally improves results. Wavelength-dependent scattering and absorption properties can enable analysis of multi-pollutant systems and improve estimation accuracy.

4.3. Effect of Grid Resolution

We evaluate the effect of using multiple grid elements to model light attenuation variation. For the Beijing dataset, the optimal grid size for greatest prediction accuracy of $PM_{2.5}$

[1]Units: $PM_{2.5}$ and PM_{10} are in μg/m^3 and NO_2 is in parts per billion.

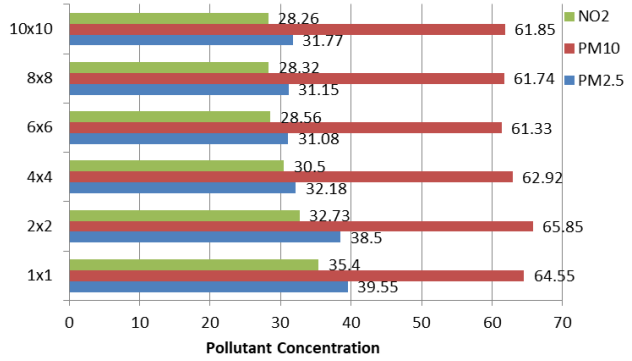

Fig. 3. RMSE for various grid resolutions (Beijing).[1]

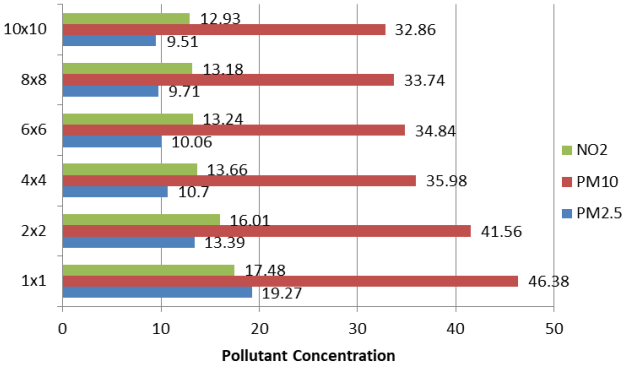

Fig. 4. RMSE for various grid resolutions (Shanghai).[1]

and PM_{10} is 6×6, and 10×10 leads to the best accuracy for NO_2, as shown in Fig. 3. The accuracies for the 8×8 and 10×10 grids are worse than that of 6×6 for $PM_{2.5}$ and PM_{10}. For NO_2, the RMSE keeps decreasing as grid size increases to 10×10, but only slightly from 8×8. For the Shanghai dataset, shown in Fig. 4, the RMSE keeps decreasing as the grid size increases to 10×10. Obtaining β values for an increasing number of grid elements initially rapidly increases accuracy and then levels off. A simple approach to selecting grid resolution would be to use 10×10 for all pollutants, which always enabled accuracy near that of the optimal resolution.

4.4. Discussion

We compare the performance of the proposed approach with the best known existing techniques for estimating air quality via images in Table 2. We use feature selection of coefficients to increase accuracy, eliminating features that increase the root mean square error. We also evaluate our technique with additional weather features, incorporating humidity, temperature, pressure, and wind speed, only available for the Shanghai dataset. For the proposed approach, a grid resolution of 10×10 is used, although it would be possible to increase

Table 2. Comparison of Results with Other Research

Beijing		$PM_{2.5}$ (μg/m^3)	PM_{10} (μg/m^3)	NO_2 (ppb)
Proposed	RMSE	27.53	56.67	24.54
	r^2	0.881	0.525	0.393
Liu et al.	RMSE	38.28	62.51	25.88
	r^2	0.70	0.462	0.328
Li et al.	RMSE	50.67	65.19	29.88
	r^2	0.563	0.349	0.070
Improvement	%	28.08	9.34	5.18
Shanghai		$PM_{2.5}$	PM_{10}	NO_2
Proposed w.o. Weather	RMSE	8.68	28.00	11.67
	r^2	0.917	0.779	0.750
Proposed w. Weather	RMSE	8.32	27.18	11.52
	r^2	0.924	0.794	0.757
Liu et al.	RMSE	13.65	35.46	15.83
	r^2	0.76	0.640	0.548
Li et al.	RMSE	25.66	52.94	19.41
	r^2	0.260	0.208	0.302
Improvement	%	39.05	23.35	27.23

accuracy by tuning grid resolution based on location and pollutant of interest (i.e. these results do not depend on tuning grid resolution). The current approach outperforms Liu et al. [5] and Li et al. [6] by a mean of 22% for all three pollutants. An author of Li et al. [6] indicated that their code is no longer available, so we reimplemented the algorithm described in the paper. A portion of the code from Liu et al. [5] was available, but it was necessary to reimplement other portions.

Various factors influenced prediction accuracy. The distance between the air quality sensors and image sensors is greater than 25 km. All three pollutants may have high spatial and temporal variation so the large distance might introduce error in the ground truth data, i.e., it is possible that the reported error is higher than the actual error.

5. CONCLUSION

We show that using color-dependent features of scattering and absorption enables concurrent estimation of multiple pollutants, namely $PM_{2.5}$, PM_{10}, and NO_2. In addition, we use the position-dependent properties of light attenuation within images to improve prediction accuracy by accounting for non-uniform pollution distribution.

6. ACKNOWLEDGEMENTS

This project was supported, in part, by the National Science Foundation under awards CBET-1240584 and CC-1836230. We would like to acknowledge Qin Lv, Michael Hannigan, and Daven Henze for their advice on this work.

7. REFERENCES

[1] Y. Zhang and F. Cao, "Fine particulate matter (PM2.5) in China at a city level," *Scientific Reports*, vol. 5, October 2015.

[2] F. Huang, X. Li, C. Wang, et al., "PM2.5 spatiotemporal variations and the relationship with meteorological factors during 2013-2014 in Beijing, China," *PloS one*, vol. 10, November 2015.

[3] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, December 2011.

[4] X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, "PM2:5 monitoring using images from smartphones in participatory sensing," in *Proc. IEEE Int. Conf. Comput. Commun. Wkshp. (INFOCOM WKSHPS)*, August 2015, vol. 2015, pp. 630–635.

[5] C. Liu, F. Tsow, Y. Zou, and N. Tao, "Particle pollution estimation based on image analysis," *PloS one*, vol. 11, no. 2, 2016.

[6] Y. Li, J. Huang, and J. Luo, "Using user generated online photos to estimate and monitor air pollution in major cities," in *Proc. Int. Conf. on Internet Multimedia Computing and Service*, New York, NY, USA, 2015, pp. 1–5.

[7] L. Liu, W. Liu, Y. Zheng, et al., "Third-eye: a mobilephone-enabled crowdsensing system for air quality monitoring," in *Proc. Interact. Mob. Wearable Ubiquitous Technol.*, New York, NY, USA, March 2018, vol. 2, pp. 20–26, ACM.

[8] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *Int. J. of Computer Vision*, vol. 48, pp. 233–254, July 2002.

[9] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, February 2000, vol. 1, pp. 598–605.

[10] N. Hyslop, "Impaired visibility: the air pollution people see," *Atmospheric Environment*, vol. 43, pp. 182–195, January 2009.

[11] J. Watson, J. Chow, L.C. Pritchett, L.W. Richards, D.L. Dietrich, et al., "Comparison of three measures of visibility extinction in Denver, Colorado," in *Proceedings of the 82nd Annual Meeting of the Air and Waste Management Association*. Air and Waste Management Association, January 1989.

[12] J Wang, W. Nie, Y. Cheng, Y. Shen, X. Chi, et al., "Light absorption of brown carbon in eastern China based on 3-year multi-wavelength aerosol optical property observations at the SORPES station and an improved absorption angstrom exponent segregation method," *Atmospheric Chemistry and Physics Discussions*, pp. 1–31, January 2018.

[13] A. Hansen, H. Rosen, and T. Novakov, "The aethalometer - an instrument for the real-time measurement of optical absorption by aerosol particles," *Science of The Total Environment*, vol. 36, pp. 191–196, August 1983.

[14] L. Ran, Z. Deng, P.C. Wang, and X. Xia, "Black carbon and wavelength-dependent aerosol absorption in the north china plain based on two-year aethalometer measurements," *Atmospheric Environment*, vol. 142, July 2016.

[15] M Xie, M. Hays, and A. Holder, "Light-absorbing organic carbon from prescribed and laboratory biomass burning and gasoline vehicle emissions," *Scientific Reports*, vol. 7, December 2017.

[16] H Horvath, "Atmospheric light absorption - a review," *Atmos. Environ., Part A*, vol. 27, pp. 293–317, February 1993.

[17] Z. Shi, J. Long, W. Tang, and C. Zhang, "Single image dehazing in inhomogeneous atmosphere," *Optik - International J. for Light and Electron Optics*, vol. 125, pp. 38683875, August 2014.

[18] A. J. Preetham, P. Shirley, and B. Smits, "A practical analytic model for daylight," in *Proc. Conf. on Computer Graphics and Interactive Techniques*, 1999, pp. 91–100.

[19] Z. Li and N. Snavely, "Megadepth: Learning single-view depth prediction from internet photos," in *Computer Vision and Pattern Recognition*, 2018.

[20] D. Berman, T. Treibitz, and S. Avidan, "Air-light estimation using haze-lines," *Proc. Int. Conf. on Computational Photography*, pp. 1–9, 2017.

[21] Y. Zou, "Photos provided by Y. Zou," http://weibo.com/p/1005051000481815/, 2014.

[22] N. Jacobs, N. Roman, and R. Pless, "Consistent temporal variations in many outdoor scenes," March 2007.

[23] C. Song, L. Wu, Y. Xie, J. He, X. Chen, et al., "Air pollution in China: Status and spatiotemporal variations," *Environmental Pollution*, vol. 227, May 2017.