NB-GTR: Narrow-Band Guided Turbulence Removal

Yifei Xia[1,2]   Chu Zhou[3]   Chengxuan Zhu[3]   Minggui Teng[1,2]   Chao Xu[3]   Boxin Shi[1,2*]

[1] National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

[2] National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

[3] National Key Lab of General AI, School of Intelligence Science and Technology, Peking University

{yfxia, zhou_chu, peterzhu, minggui_teng, shiboxin}@pku.edu.cn, xuchao@cis.pku.edu.cn

Abstract

The removal of atmospheric turbulence is crucial for long-distance imaging. Leveraging the stochastic nature of atmospheric turbulence, numerous algorithms have been developed that employ multi-frame input to mitigate the turbulence. However, when limited to a single frame, existing algorithms face substantial performance drops, particularly in diverse real-world scenes. In this paper, we propose a robust solution to turbulence removal from an RGB image under the guidance of an additional narrow-band image, broadening the applicability of turbulence mitigation techniques in real-world imaging scenarios. Our approach exhibits a substantial suppression in the magnitude of turbulence artifacts by using only a pair of images, thereby enhancing the clarity and fidelity of the captured scene.

1. Introduction

Atmospheric turbulence is a common yet annoying physical phenomenon that degrades the image quality. Recently, learning-based single-image turbulence removal (SITR) methods [5, 25, 26, 30, 39, 40] have been proposed to handle this problem by extracting features from a large amount of training data, offering visually plausible results and high efficiency. However, their performance in handling challenging turbulence scenarios encounters limitations, as accurately distinguishing image content from complex turbulence field using only a single observation remains a significantly challenging task due to the highly ill-posed nature of this problem, which is illustrated in Figure 1.

To alleviate ill-posedness and enhance the overall performance of turbulence removal, some methods propose to restore the turbulence-free image from multiple consecutive degenerate images [3, 10, 18, 22, 23, 27, 42], utilizing the time-varying physical properties of atmospheric turbulence (*e.g.*, the *lucky frame phenomenon* [11], where clearer
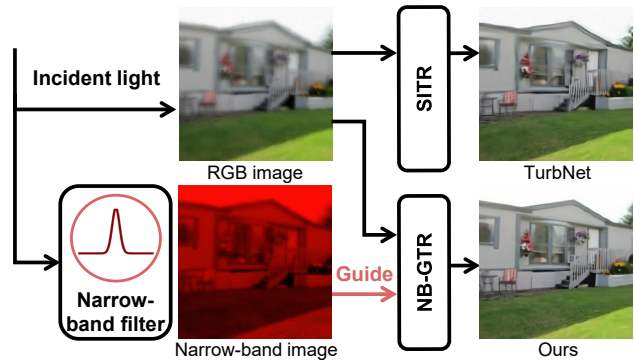

Figure 1. Illustration of the workflow of our narrow-band image guided turbulence removal approach (NB-GTR) and single-image turbulence removal methods (SITR, taking TurbNet [25] as an example). TurbNet [25] processes the degenerated image without any guidance, resulting in inferior performance, while our NB-GTR utilizes an additional captured narrow-band image as a guidance, producing a clearer restoration of the original scene with richer details.

frames occasionally appear in turbulent sequences, and the *zero-mean statistics* [44] of turbulence, indicating a time-dependent zero-mean Gaussian distribution of tilt). With the temporal diversity exploited from multiple consecutive images, these methods usually show better detail recovery ability and higher robustness. However, multiimage approaches require capturing a sequence of frames (usually 10, 12, or 16 [10, 30, 42]) to acquire sufficient information about atmospheric turbulence or to encounter the so-called lucky frames. Moreover, possible motion in practical scenarios can further add to this inconvenience. It is therefore of great interest to explore a trade-off solution that balances the robustness of enhancing results and the convenience of capturing efforts (hopefully using one or two images).

In this work, we propose **NB-GTR**, a **N**arrow-**B**and **G**uided **T**urbulence **R**emoval approach, aiming to remove the turbulence in a single degenerated RGB image under the guidance from an additional narrow-band image cap-

---

*Corresponding author.

tured simultaneously. Specifically, we have observed that narrow-band imaging could remarkably reduce the turbulence in the image plane due to a reduced cumulative effect across the narrower passband of the narrow-band filter, bringing distinctive advantages in facilitating the turbulence removal process. As shown in Figure 1, under the guidance of an additional narrow-band image, our method produces a more authentic result than the state-of-the-art SITR method TurbNet [25]. However, directly feeding the concatenation of the degenerated RGB image and the narrow-band image to existing methods [22, 25, 39] could produce suboptimal results, since the narrow-band image has totally different properties than the RGB one: It essentially loses color information along with high-frequency details and tends to have stronger noise. To address this issue, we cooperate with a two-step fusion strategy and a novel turbulence removal network to better perceive the turbulence field and help turbulence removal. Our NB-GTR not only enjoys the convenience and efficiency of single-image turbulence removal methods but also achieves comparable or even higher performance and robustness than multi-image approaches.

In summary, our contributions to the field of atmospheric turbulence removal are threefold:

1. We demonstrate the substantial effectiveness of narrow-band imaging in reducing turbulence strength, as supported by both theoretical analysis and experimental evidence.

2. We introduce a novel pipeline using a pair of simultaneously shot RGB and narrow-band images. This approach enables us to jointly extract reliable turbulence cues based on their intrinsic relationships, thereby guiding more effective turbulence removal.

3. We present a network specially tailored for the turbulence removal pipeline by the joint optimization of the turbulence extraction and guided turbulence removal module.

Experimental results show that our method achieves state-of-the-art performance in both synthetic and real world scenarios.

2. Related Works

Atmospheric turbulence removal. Conventional turbulence removal methods [1, 2, 12, 23, 29, 45] are based on lucky frame fusion and blind deconvolution, which are time consuming and require a number of input frames. With the rise of deep learning and fast physics-grounded turbulence simulators [4, 24, 32], many learning-based methods have been proposed. Although it is possible to learn the removal of turbulence in a purely end-to-end manner [10, 18], the specific modeling of turbulence is still helpful to an effective turbulence removal. Information about the turbulence field can be modeled by grid deformation [22], diffusion [5] or model uncertainty [39, 40]. With the help of a physics-based turbulence simulator [24], the process of turbulence degradation can be explicitly modeled and removed [17, 25, 42]. Utilizing the unique wavelength dependence of atmospheric turbulence, we present another perspective for effective turbulence modeling and removal.

Wavelength and atmospheric turbulence. Recent works on turbulence simulation and removal [17, 24, 25] mainly adopts the approximation that wavelength dependence of turbulence is visually indistinguishable. However, since the atmosphere acts as a dispersive medium, it has a varying refractive index profile throughout the visual spectrum, as clearly stated by the physics community [6, 14]. Turbulence strength is weaker when observing at a longer wavelength, which is proven by Hardie *et al.* [14] through both theoretical analysis and simulations, opening a door for removing turbulence with its wavelength dependence. Aware of the additional information provided by a set of observations at different wavelengths, some work tackles the problem of facial classification [13], adaptive optics [19, 37], and monochromatic blind deconvolution [16] through atmospheric turbulence with the help of multispectral imaging. Since multispectral imaging is conducted at only a select few wavelengths, these methods are unable to reconstruct a color image, thus leaving a blank in the application of wavelength information for turbulence removal. Our approach addresses this problem by recovering an RGB image under the guidance of narrow-band imaging.

3. Methods

In this section, we discuss our turbulence removal network for RGB and narrow-band images, detailing theoretical motivations and atmospheric turbulence effects on practical imaging in Section 3.1, image fusion strategies in Section 3.2, and the detailed architecture of our network in Section 3.3.

3.1. Background and motivation

Assume an observation of a point source S through the atmosphere, where its angle relative to the optical axis is denoted as $\boldsymbol{\theta} = [\theta_x, \theta_y]^\top$, and its tilt vector is given by $\boldsymbol{\alpha}_\lambda(\boldsymbol{\theta}) = [\alpha_{\lambda,x}(\theta), \alpha_{\lambda,y}(\theta)]^\top$, with the variance of the tilt vector being $T_Z^2(\lambda)$ which can be regarded as a function of wavelength λ. As highlighted by Hardie *et al.* [14], $T_Z^2(\lambda)$ indicates the turbulence-induced warping effect, satisfying $T_Z^2(\lambda) \propto N^2(\lambda)$ at the same temperature and pressure, where $N(\lambda)$ is the atmospheric refractivity given by Edlén's relation, showing that both the magnitude and variance of the atmospheric turbulence-induced tilt tend to decrease as the wavelength increases. When capturing images with a conventional RGB camera, since the color filter array in front of its CMOS sensor usually contains a wide bandpass
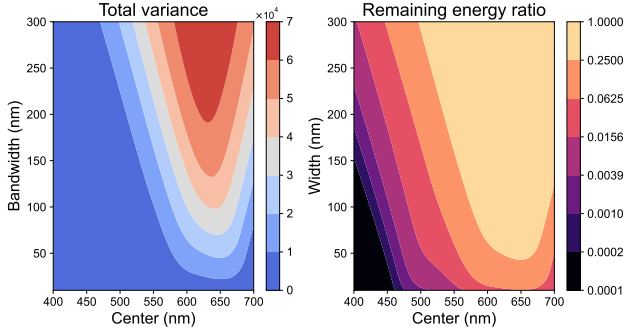
Figure 2. Heatmap of total tilt variance and remaining energy ratio w.r.t. filter bandwidth and central wavelength. To achieve an optimal balance between reducing the observed turbulence intensity and avoiding excessively low residual energy leading to a dim image, the best choice involves selecting a filter with a larger central wavelength and a smaller bandwidth.
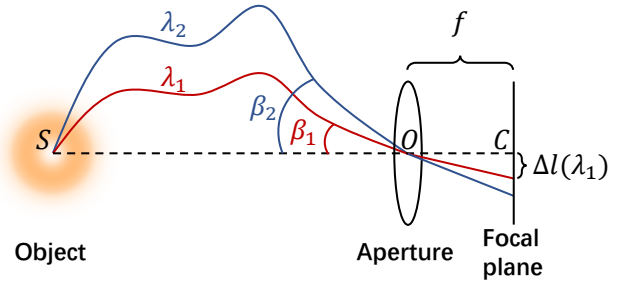

Figure 3. Illustration of the process of determining atmospheric turbulence effects on point source S. Due to atmospheric disturbances, the precise angular offset β_1 of point source S from its original position remains unknown. However, simultaneous observations at two different wavelengths allow the calculation of the difference $\beta_2 - \beta_1$, providing additional information about the atmospheric turbulence at that moment.

(*i.e.*, several hundred nanometers [8]), the total variance of the tilt vector of each band is an accumulation over the entire bandpass, resulting in a significant turbulence strength in the captured RGB image.

Suppose we employ a long-wave narrow-band filter before the camera, where the central wavelength is λ_0 in the red light band, and the bandwidth is $\Delta\lambda$, defining the bandwidth wavelength interval as $\Lambda = (\lambda_0 - \Delta\lambda/2, \lambda_0 + \Delta\lambda/2)$. The response function of this narrow-band filter is given by

$$F(\lambda) = \begin{cases} f(\lambda), & \lambda \in \Lambda \\ 0, & \lambda \notin \Lambda \end{cases}, \tag{1}$$

where $f(\lambda)$ is the response function of the narrow-band filter[1], resulting in a new total variance

$$V(c_i) = \frac{\int_{\lambda \in \Lambda} T_Z^2(\lambda) f(\lambda) c_i(\lambda) \mathrm{d}\lambda}{\int_{\lambda \in \Lambda} f(\lambda) c_i(\lambda) \mathrm{d}\lambda}. \tag{2}$$

Given that the bandwidth $\Delta\lambda$ is typically smaller than $50\,\mathrm{nm}$ and $T_Z^2(\lambda)$ is relatively smaller in the long-wave band, the new total variance V is significantly lower than without the narrow-band filter. Figure 2 shows the relationship between the total variance V and the center wavelength λ_0 and bandwidth $\Delta\lambda$ under a narrow-band filter, indicating minimal total variance under long-wave narrow-band filtering.

3.2. RGB and narrow-band image fusion

Intuition. We have clarified that images captured through a narrow-band filter have relatively lower turbulence statistics. However, the use of a long-wave narrow-band filter results in the loss of more than 90% of the incident light,

---
[1]Specific function of $f(\lambda)$ is detailed in the supplementary materials.

particularly under the filters G and B where almost all energy is blocked. As shown in Figure 1, under the same capture parameters, significantly more noise is present in images captured with narrow-band filters, and the color information is almost lost. Due to the significant modality gap between them, simply combining narrow-band images with RGB images for recovery may mislead the perception of a turbulence field and yield suboptimal results.

To bridge the modality gap, we need an explicit modeling of the turbulence field. Here, we present the *scaling relation*, another important property of a simultaneously captured RGB and narrow-band image pair, from which useful information about the turbulence field can be induced. Before diving into the detailed formulations, we first present an intuitive illustration. The refraction index of the atmosphere is dependent on the wavelength [15], causing the image at different wavelengths of a point source to be separated in the focal plane, as shown in Figure 3. The broadband filter, letting in a wide range of dispersed light, can result in a blurry observation. The narrow-band filter, which focuses at a specific wavelength, tends to provide a clearer image. The shape of the turbulence field in the image pair remains the same, with only a difference proportion to a fixed refractivity ratio as detailed in [14], leading to what we call the scaling relation, which indicates useful information about the turbulence field.

Formulation. For a more formal explanation, consider observing a point source S at wavelength λ_1, its incident light is dispersed a distance $\Delta l(\lambda_1)$ relative to the dispersion free point C. We further denote the turbulence angle as

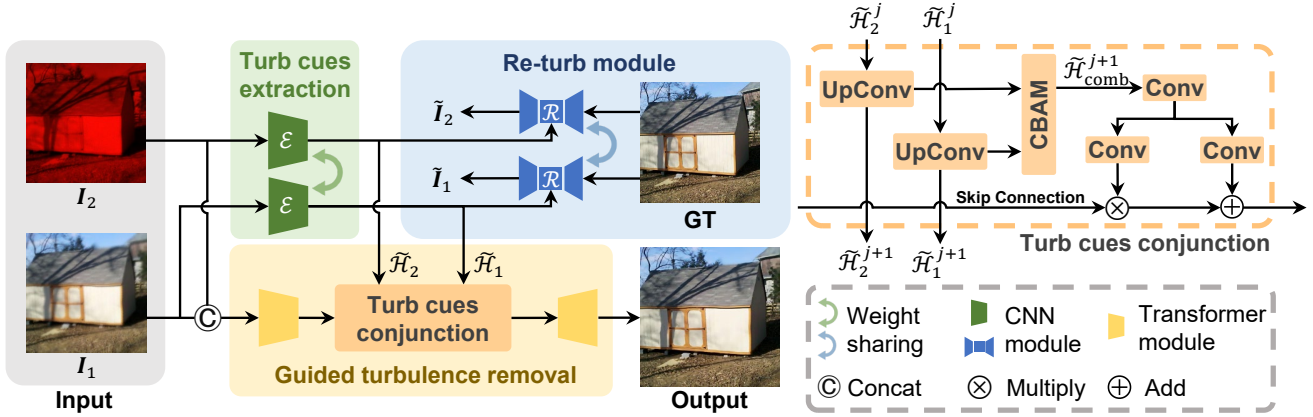$$\beta(\lambda_1) = \frac{\Delta l(\lambda_1)}{f}, \tag{3}$$

Figure 4. Architecture of the NB-GTR network. Turbulence cues are first extracted from the input image pair of an RGB image I_1 and a narrow-band image I_2 by a dual turbulence encoder $\mathcal{E}$ and a re-turb module $\mathcal{R}$ to bridge the modality gap, followed by a conjunction mechanism for enhanced feature interaction and turbulence removal.

where f is the focal length of the imaging system. In the problem of turbulence removal, we are ignorant about where the turbulence free point C is, thus unable to induce $\beta(\lambda_1)$. Luckily, if another observation at wavelength λ_2 is available at the same time, then the turbulence angle can be deduced by subtracting $\beta(\lambda_1)$ from $\beta(\lambda_2)$, following the derivation of Ingensand *et al.* [15]:

$$\beta(\lambda_1) = \nu(\beta(\lambda_2) - \beta(\lambda_1)) = \nu\Delta\beta, \qquad (4)$$

where ν is a wavelength dependent constant.

Although real-world observations are conducted on some bandpass instead of a single wavelength, the direction of turbulence angle would not change, while its amplitude varies. Therefore, Equation (4) holds in real scenarios, with only a different constant ν depending on the unique response function throughout the bandpass. Since the scaling relation holds, given a standard RGB image and its counterpart acquired through a narrow-band filter, it is natural to first induce the turbulence field from the difference between the two images, fully exploiting the temporal synchronicity, and then use the induced turbulence field to assist in the removal of turbulence. The two-stage pipeline, inspired by the scaling relation, translates the differences between RGB and narrow-band images into a detailed turbulence field, enabling effective fusion of these images despite their significant differences.

3.3. Network architecture

In this section, we present a turbulence removal network specifically tailored for a pair of RGB and narrow-band images, as is shown in Figure 4. Using a dual turbulence encoder, the network is able to overcome the modality gap in the image pair, extracting the unique yet interrelated turbulence cues captured simultaneously. The extracted turbu-

lence cues jointly interact with a re-turb and a de-turb module through a specially designed conjunction mechanism. These design enhances turbulence extraction and lead to a more accurate turbulence removal.

Turbulence extraction and re-turb module. Although the scaling relation is conceptually simple, directly applying it in the image space is challenging due to spatially varying turbulence. Therefore, we adopt an adapted light-weight U-Net[31] architecture as a implementation of the turbulence deduction described in Equation (4), which extracts multi-scale spatial features of the turbulence field from the input image pair, formulated as

$$\tilde{\mathcal{H}}_i = \mathcal{E}(I_i), \qquad (5)$$

where $\mathcal{E}$ is the turbulence encoder, $I_i, \tilde{\mathcal{H}}_i$ are the ith turbulent image and its extracted turbulence cues standing for a learnable representation of turbulence field, $i = 1, 2$ corresponds to the RGB and narrow-band image respectively. For a meaningful post-processing of extracted turbulence features, weights of the two encoders of RGB and narrow-band input image are shared. A re-turb module is then introduced to reverse the ground truth image into the turbulent inputs one by one under the assistance of previously extracted turbulence cues respectively, thereby ensuring a robust extraction of turbulence field, which can be implemented by another U-Net with a complete encoder-decoder structure similar to the turbulence encoder to ensure consistency in processing. For parameter efficiency purposes, the parameters are shared in the re-turb module when processing either of the RGB or narrow-band images. The re-turb module integrates turbulence cues isolated earlier through a spatially adaptive modulation [21], thereby informing and guiding the restoration process. This integration plays an important role in ensuring that the turbulence encoder ade-

quately captures the turbulence fields. The re-turbing process can be formulated as

$$\tilde{\boldsymbol{I}}_i = \mathcal{R}(\boldsymbol{J}; \tilde{\boldsymbol{\mathcal{H}}}_i), \tag{6}$$

where $\boldsymbol{J}$ is the turbulence-free reference image, $\mathcal{R}$ is the re-turb module, and $\tilde{\boldsymbol{I}}_i$ represents the re-turbed images.

Guided turbulence removal. By integrating the extracted turbulence cues with the input image pair, a more robust reconstruction result can be achieved, which can be denoted as

$$\tilde{\boldsymbol{J}} = \mathcal{D}(\boldsymbol{I}_1, \boldsymbol{I}_2; \tilde{\boldsymbol{\mathcal{H}}}_1, \tilde{\boldsymbol{\mathcal{H}}}_2), \tag{7}$$

where $\mathcal{D}$ is the de-turb module, $\tilde{\boldsymbol{J}}$ is the reconstructed clean image. Given the complex nature of atmospheric turbulence, the incorporation of a multiscale Transformer module [41] is beneficial to reconstruct a clear image, guided by previously extracted turbulence cues, using the capability of Transformer to handle spatial variations in turbulence. The process integrates the features of the RGB and narrow-band images by embedding them together using overlapping patches. The network employs multi-head self attention blocks (MHSA) at different encoder levels to enhance feature representation while compressing spatial dimensions, a method critical for effective turbulence removal [10]. A key feature of this process is the conjunction of turbulence cues, through which these cues can effectively interact and guide the removal of turbulence. These links merge the cues from the dual turbulence encoder with skip connections at each encoder level, enhancing the feature mix for the up-sampling stage. Here, the network progressively restores the image to its original resolution. Refinement Transformer blocks process the upsampled features, and a final convolutional layer combines wavelength-specific details with spectral data, yielding a turbulence-free image.

Turbulence cues conjunction. At the heart of guided turbulence removal module lies the turbulence cues conjunction mechanism, which provides an effective feature communication between Transformer encoders and turbulence cues. As mentioned in Section 3.2, the characteristics of the turbulence field can be deduced by comparing the differences between two distinct observations. This deduction, although not analytically feasible in the image field, can be effectively performed by employing a Convolutional Block Attention Module (CBAM) mechanism [38] to jointly process the differences between previously extracted turbulence cues. This process can be formulated as follows:

$$\tilde{\boldsymbol{\mathcal{H}}}_{\mathrm{comb}} = \mathcal{F}(\tilde{\boldsymbol{\mathcal{H}}}_1, \tilde{\boldsymbol{\mathcal{H}}}_2), \tag{8}$$

where $\mathcal{F}$ denotes the turbulence deduction in Equation (4), and $\tilde{\boldsymbol{\mathcal{H}}}_{\mathrm{comb}}$ is the combined turbulence cues. This conjunction mechanism ensures that the model can effectively integrate turbulence information from two different types of

image source, offering a more robust and comprehensive representation of turbulence characteristics. Note that the turbulence deduction can be done in multiscale, which is marked by superscripts j in Figure 4, and the turbulences cues are gradually upsampled to match the latent space of skip connections. A convolution layer is then applied for a better feature alignment, followed by a dual convolution layer to refine the merged turbulence features, providing a better feature alignment before integrating them with the skip connections from the de-turb model encoder. The final output of this module is a blend of rescaled encoder features with the latent features extracted by the Transformer encoder, optimizing the input for subsequent processing stages in the network.

In general, the NB-GTR network is designed to fully utilize the spectral characteristics of the dual image capture method, enhancing the process of turbulence removal in both accuracy and efficiency. This is achieved by optimizing both the extraction and removal of turbulence features, allowing the network to effectively identify and use these features for improved turbulence removal.

3.4. Implementation details

Paired turbulence data synthesis. We select 20,000 images for training and 10,000 images for testing from the Places dataset [43]. Each image is first cropped to a size of 512×512 pixels. We use phase-to-space (P2S) transform [24], which is a fast Zernike-based [28] atmospheric turbulence simulator. To generate a pair of atmospheric turbulence images captured by RGB and narrow-band filters at the same time, we need to fix the random field of the generated image and then re-scale it with different Zernike coefficients in P2S following Hardie *et al.* [14]. Specifically, we set the wavelength of the RGB image to $\lambda_{\mathrm{RGB}} = 550$ nm, the central wavelength of the visual spectrum, and the wavelength of the narrow-band filter to $\lambda_{\mathrm{N}} = 680$ nm, at which value a smaller total variance is achieved while ensuring sufficient incoming light according to Figure 2. Scaling coefficients are calculated from the wavelengths chosen according to the experimentally determined relationship between the Zernike coefficients and the wavelengths [35]. The other parameters of the P2S simulator generally follow the selections made in [17][2]. Following the turbulence simulation, we reweight the RGB channels of the narrow-band image by assigning them coefficients of 0.95, 0.04, and 0.01, respectively. This reweighting process is designed to align then the simulation results with the average distribution we determined through actual narrow-band photography. Finally, we introduce noise into the images by first applying Gaussian noise with a standard deviation randomly selected between 0.015 and 0.025 for narrow-band images

---

[2]For more detailed settings of turbulence synthesis parameters, see the supplementary materials.

Input / narrow **Input / RGB** **Ground truth** **Ours** **TurbNet** **N-IDR** **Burstormer** **AT-Net** **FocalNet**
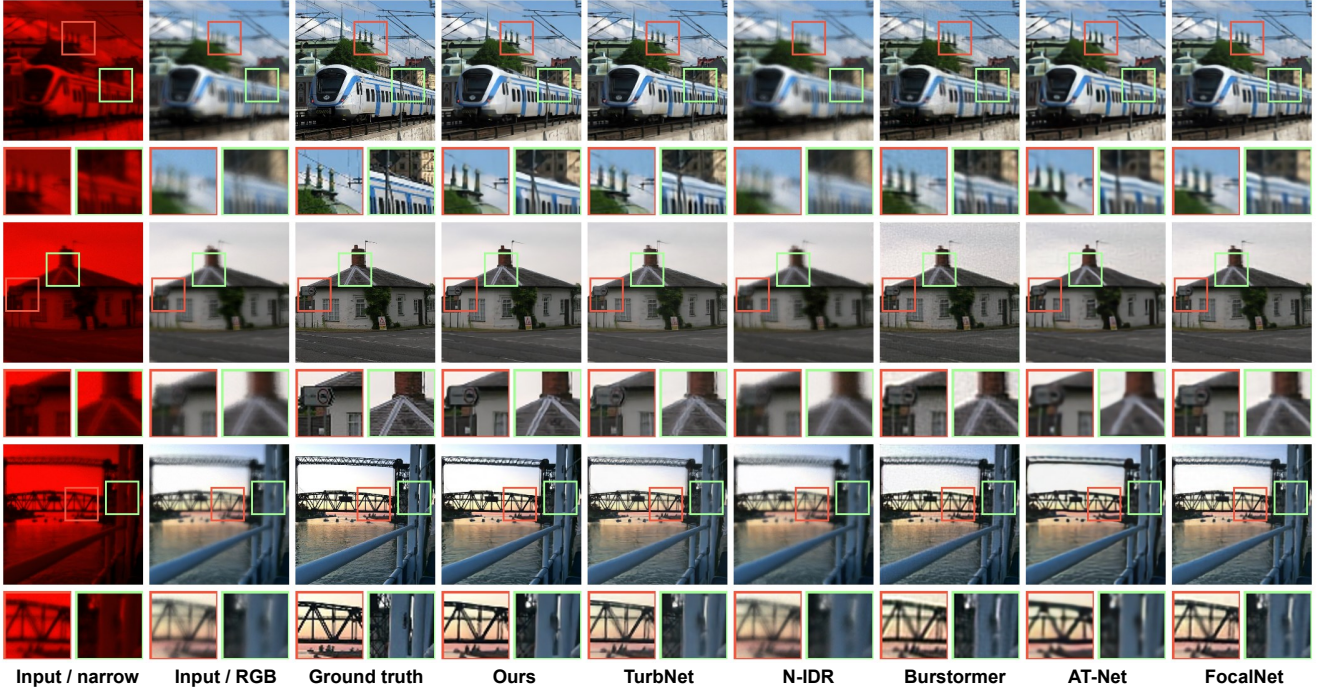
Figure 5. Qualitative comparisons are conducted using synthetic data between our method and three learning-based turbulence removal methods: N-IDR [22], AT-Net [39], and TurbNet [25]. Additionally, two general multi-image restoration methods, Burstormer [9] and FocalNet [7], which share similar problem settings with our method, are included.

and between 0.005 and 0.015 for RGB images to reflect the different noise intensities observed in actual measurements, followed by a similarly scaled quantization noise.

Loss function. The NB-GTR network's loss function comprises consistency and reconstruction losses. This can be represented as

$$L = L_{\text{consist}} + L_{\text{recon}}. \tag{9}$$

Here, L_{consist} refers to the consistency loss which ensures that the encoder $\mathcal{E}$ in Equation (5) is extracting meaningful turbulence cues, and L_{recon} denotes the reconstruction loss which is responsible for generating a clear, turbulence-free image. To be specific, the consistency loss is calculated by comparing the re-turbed results $\tilde{I}_i$ and in Equation (6) with the turbulent input I_i respectively, where subscript i specifies the RGB or narrow-band image, which is formulated as

$$L_{\text{consist}} = \sum_{i=1}^{2} \alpha_i L_1(\tilde{I}_i, I_i), \tag{10}$$

where L_1 stands for the ℓ_1 loss, and α_i is a weight coefficient, which are both set to be 0.4 in our experiment. Besides, the reconstruction loss L_{recon} is defined as the ℓ_1 loss between the output $\tilde{J}$ of de-turb module in Equation (7) and the ground truth J, which can be written as

$$L_{\text{recon}} = L_1(\tilde{J}, J). \tag{11}$$

Training strategy. We implement our method using PyTorch with an AMD EPYC 7532 CPU and 4 NVIDIA 3090 GPUs. The guided turbulence removal module employs MHSA blocks with three attention heads each. The number of blocks in each cascaded level, starting from the top and moving downwards, is arranged as 2, 3, 3, and 4, respectively. All parameters in the network are jointly trained for 300K interations with a batch size of 16. The training is conducted on images of size 224×224, employing horizontal flip and vertical clip for data augmentation. For optimization purposes, the Adam optimizer [20] is employed with hyperparameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. A cosine annealing learning rate scheduler is utilized, decreasing the learning rate from an initial value of 5e^{-5} to 1e^{-6} throughout the training process.

4. Experiments

In this section, we present the performance of our NB-GTR method compared to other methods on synthetic data in Section 4.1 and on real-world data in Section 4.2, respectively, and conduct ablation studies on the modules designed within the NB-GTR network to verify their effectiveness in Section 4.3.

4.1. Evaluation on synthetic data

We compare our NB-GTR with three other turbulence removal methods, namely, N-IDR [22], AT-Net [39] and
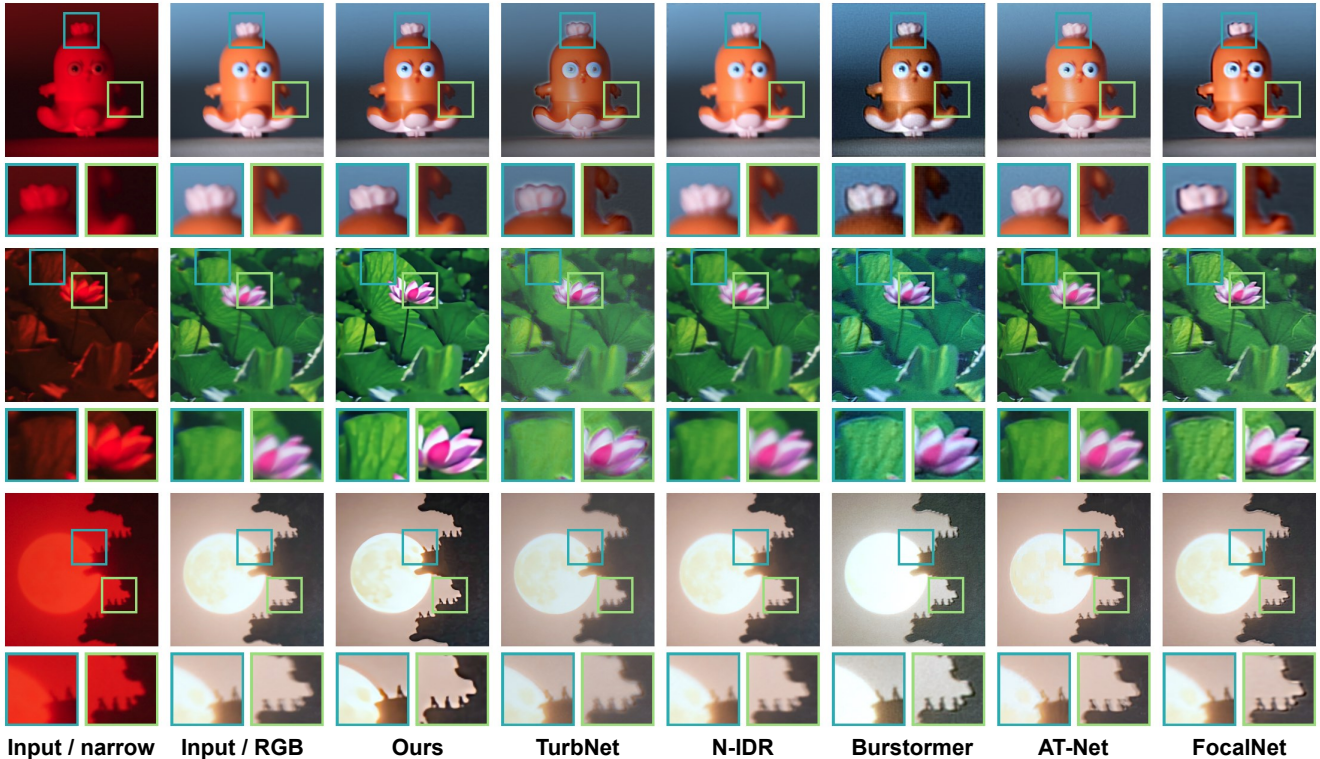
Figure 6. Qualitative comparisons are conducted using real data between our method and three learning-based turbulence removal methods: N-IDR [22], AT-Net [39], and TurbNet [25], along with two general multi-image restoration methods, Burstormer [9] and FocalNet [7]. The initial row is an indoor object, while the next two rows show postcards.

| Input / narrow | Input / RGB | Ours | TurbNet | N-IDR | Burstormer | AT-Net | FocalNet |

TurbNet [25]. In addition, Burstormer [9] and FocalNet [7], which are two general multiimage restoration methods that share problem settings similar to our method, are also compared to ensure a comprehensive evaluation. Our evaluation of synthetic data aims to answer two key questions: *Does the inclusion of an additional narrow-band image significantly aid in turbulence removal from an RGB image?* and *Do the unique designs in NB-GTR effectively utilize the additional narrow-band image for enhanced guidance?*

To address the first question, we conduct two sets of experiments for all comparison methods: one with training solely on RGB images and the other with simultaneous training on both RGB and narrow-band images. Although AT-Net [39] and TurbNet [25] are designed for a single image, slight modifications are made to allow them to accept inputs from two images by merging the images channelwise. Each method is re-trained under the corresponding comparison settings. The experimental results are detailed in Table 1, where the two sets of experiments indicated by *Single* and *Dual* respectively. The results demonstrate the advantage of incorporating an additional narrow-band image with a general performance growth in both PSNR and SSIM [36], two commonly used metrics. For the second question, we observe that NB-GTR consistently outper-

Table 1. Quantitative comparison with three turbulence removal methods N-IDR [22], AT-Net [39], TurbNet [25] and two general image restoration methods Burstormer [9], FocalNet [7] on synthetic data. ↑ indicates the higher, the better throughout this paper. All compared methods are retrained given a single RGB image or a pair of RGB and narrow-band images (denoted as *Single* and *Dual* in the table respectively), with the better result underlined, showing the benefit of an additional narrow-band image.

Methods	PSNR(↑)		SSIM(↑)	
	Single	Dual	Single	Dual
N-IDR [22]	20.77	20.72	0.6265	0.6231
Burstormer [9]	20.10	21.37	0.5672	0.6236
AT-Net [39]	20.75	22.18	0.6008	0.6877
FocalNet [7]	22.44	22.80	0.6998	0.7147
TurbNet [25]	22.41	23.18	0.7167	0.7498
NB-GTR	**25.00**		**0.8199**	

forms the other methods in terms of both PSNR and SSIM metrics when utilizing dual-image input. The quantitative results, as presented in Table 1, clearly demonstrate that NB-GTR not only benefits from the inclusion of the narrow-band image, but also effectively capitalizes on its unique designs to maximize the potential guidance of the additional

narrow-band image.

The visual comparisons presented in Figure 5 highlight the distinct advantages of our NB-GTR method over other approaches. The side-by-side comparison clearly demonstrates that the NB-GTR provides the most accurate reconstruction compared to ground truth. Although methods like FocalNet [7] and Burstormer [9] show proficiency in addressing general image restoration challenges, they fail to effectively handle complex distortions caused by atmospheric turbulence. N-IDR [22], which models turbulence by grid deformation, struggles to make use of the narrowband image modality. AT-Net [39], although it demonstrates reasonable noise reduction capabilities, often fails to fully rectify the warping effects induced by atmospheric turbulence. This reinforces the notion that the unique approach of NB-GTR to leveraging narrow-band image data is particularly effective for turbulence removal in challenging atmospheric conditions.

4.2. Evaluation on real data

Capturing a narrow-band image alongside an RGB image in a single instance can be done using a beam-splitter. Alternatively, this can be approximated with closely aligned parallel cameras. This approach is viable, provided that the separation between the cameras is minor compared to the distance to the observed objects. For our experiments, we employ the latter method for practical convenience, using two Sony α7RIII cameras equipped with Sony FE 70-200 mm f/2.8 GMII lenses. We affixed a narrow-band filter to one lens, characterized by a central wavelength of $\lambda_0 = 680$ nm and a bandwidth of $\Delta\lambda = 40$ nm, which leads to a notable decrease in turbulence intensity using narrow-band imaging[3]. Turbulence data are collected by installing heat stoves in the office, which can generate authentic atmospheric turbulence with controllable strength. It is a technique commonly adopted to evaluate the removal of turbulence in previous papers such as TSR-WGAN [18], TurbNet [25], and the Vorontsov-Carhart method [33]. Following SIR[2] [34], we capture photographs of indoor objects and postcards to ensure a variety of scenes.

As shown in Figure 6, the visual results support our quantitative data, showing a noticeable difference in the restored image quality. The output of NB-GTR notably exhibits less turbulence, rendering the images visually clearer and sharper, especially in areas with severe distortion. This benefit is particularly evident compared to other techniques such as AT-Net [39] and Burstormer [9], which shows less effectiveness in mitigating atmospheric turbulence. Our visual comparative analysis confirms that NB-GTR is practical and reliable for image restoration in challenging atmospheric scenarios, using only a pair of images.

---

[3]A video captured in real conditions is available in the supplementary materials.

Table 2. Quantitative comparison of ablation studies.

Model setting	PSNR($\uparrow$)	SSIM($\uparrow$)
W/o returb	24.89	0.8141
W/o encoder+returb	24.87	0.8142
W/o narrow	24.97	0.8165
W/o conjunction	24.93	0.8148
Complete model	**25.00**	**0.8199**

4.3. Ablation studies

In our ablation studies, detailed in Table 2, we systematically examine the influence of different components in the NB-GTR network on its performance. Initially, we test the model without the re-turb mechanism (W/o returb), observing a minor drop in both PSNR and SSIM compared to the full model. This underlines the importance of the re-turb mechanism in improving image quality, especially in reducing noise and maintaining details. Next, we evaluate the model lacking both the encoder and re-turb (W/o encoder+returb), which shows an even bigger performance drop, indicating that the combined effect of each component is crucial for an optimized result. Furthermore, we evaluate the model without the narrow-band turbulence encoder and its conjugate (W/o narrow), emphasizing the value of narrow-band imaging in managing complex atmospheric distortions. The absence of the conjunction mechanism (W/o conjunction) also results in lower performance. Finally, the complete NB-GTR model, integrating all components, achieves the highest PSNR and SSIM, demonstrating how the individual components work together to enhance the overall efficiency of the design.

5. Conclusion

In our study, we highlight the significant role of narrowband imaging in reducing the strength of atmospheric turbulence, as evidenced through both theoretical analysis and experimental validation. We propose a novel approach, NB-GTR, that utilizes a pair of simultaneously captured RGB and narrow-band images, leveraging their intrinsic relationships to extract reliable turbulence cues for more effective turbulence removal. Our network, designed specifically for narrow-band imaging guided atmospheric turbulence removal, achieves state-of-the-art performance.

Limitations. A limitation of our method is the additional cost incurred by requiring a narrow-band filter for data acquisition. Additionally, the parallel camera capturing approach we employ cannot achieve the simultaneity offered by beam-splitting, yet opting for the latter results in a more intricate capture setup.

References

[1] Nantheera Anantrasirichai, Alin Achim, Nick G Kingsbury, and David R Bull. Atmospheric turbulence mitigation using complex wavelet-based fusion. *IEEE Transactions on Image Processing*, 22(6):2398–2408, 2013. 2

[2] Nantheera Anantrasirichai, Alin Achim, and David Bull. Atmospheric turbulence mitigation for sequences with moving objects using recursive image fusion. In *Proc. of International Conference on Image Processing*, 2018. 2

[3] Wai Ho Chak, Chun Pong Lau, and Lok Ming Lui. Subsampled turbulence removal network. *arXiv preprint arXiv:1807.04418*, 2018. 1

[4] Nicholas Chimitt and Stanley H Chan. Simulating anisoplanatic turbulence by sampling intermodal and spatially correlated zernike coefficients. *Optical Engineering*, 59(8):083101–083101, 2020. 2

[5] Hyungjin Chung, Jeongsol Kim, Sehui Kim, and Jong Chul Ye. Parallel diffusion models of operator and image for blind inverse problems. In *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1, 2

[6] JE Craig and WC Rose. Wavelength effects on images formed through turbulence. In *Proc. of Flow Visualization and Aero-Optics in Simulated Environments*, 1987. 2

[7] Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Focal network for image restoration. In *Proc. of International Conference on Computer Vision*, 2023. 6, 7, 8

[8] Jason Deglint, Farnoud Kazemzadeh, Daniel Cho, David A Clausi, and Alexander Wong. Numerical demultiplexing of color image sensor measurements via non-linear random forest modeling. *Scientific reports*, 6(1):28665, 2016. 3

[9] Akshay Dudhane, Syed Waqas Zamir, Salman Khan, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Burstormer: Burst image restoration and enhancement Transformer. *arXiv preprint arXiv:2304.01194*, 2023. 6, 7, 8

[10] Hamidreza Fazlali, Shahram Shirani, Michael BradforSd, and Thia Kirubarajan. Atmospheric turbulence removal in long-range imaging using a data-driven-based approach. *International Journal of Computer Vision*, 130(4):1031–1049, 2022. 1, 2, 5

[11] David L Fried. Probability of getting a lucky short-exposure image through turbulence. *Journal of the Optical Society of America*, 68(12):1651–1658, 1978. 1

[12] Jérôme Gilles, Tristan Dagobert, and Carlo De Franchis. Atmospheric turbulence restoration by diffeomorphic image registration and blind deconvolution. In *Proc. of International Conference on Advanced Concepts for Intelligent Vision Systems*, 2008. 2

[13] Yilu Guo, Hong Song, Hongbo Liu, He Wei, Ping Yang, Shuyue Zhan, Hangzhou Wang, Hui Huang, Ningfang Liao, Quanquan Mu, et al. Model-based restoration of underwater spectral images captured with narrowband filters. *Optics Express*, 24(12):13101–13120, 2016. 2

[14] Russell C Hardie, Michael A Rucci, Santasri Bose-Pillai, Richard Van Hook, and Barry K Karch. Modeling and simulation of multispectral imaging through anisoplanatic atmospheric optical turbulence. *Optical Engineering*, 61(9):093102–093102, 2022. 2, 3, 5

[15] H Ingensand and B Boeckem. A high-accuracy alignment system based on the dispersion effect. In *Proc. of International Workshop on Accelerator Alignment*, 1997. 3, 4

[16] Harry R Ingleby and Donald R McGaughey. Parallel multiframe blind deconvolution using wavelength diversity. In *Proc. of Image Reconstruction from Incomplete Data III*, 2004. 2

[17] Ajay Jaiswal, Xingguang Zhang, Stanley H Chan, and Zhangyang Wang. Physics-driven turbulence image restoration with stochastic refinement. In *Proc. of International Conference on Computer Vision*, 2023. 2, 5

[18] Darui Jin, Ying Chen, Yi Lu, Junzhang Chen, Peng Wang, Zichao Liu, Sheng Guo, and Xiangzhi Bai. Neutralizing the impact of atmospheric turbulence on complex scene imaging via deep learning. *Nature Machine Intelligence*, 3(10):876–884, 2021. 1, 2, 8

[19] Xizheng Ke and Xiaozhan Chen. Correcting wavefront distortion of dual-wavelength beams due to atmospheric turbulence with a correction coefficient. *Optics and Photonics Journal*, 10(4):64–77, 2020. 2

[20] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6

[21] Dasong Li, Yi Zhang, Ka Chun Cheung, Xiaogang Wang, Hongwei Qin, and Hongsheng Li. Learning degradation representations for image deblurring. In *Proc. of European Conference on Computer Vision*, 2022. 4

[22] Nianyi Li, Simron Thapa, Cameron Whyte, Albert W Reed, Suren Jayasuriya, and Jinwei Ye. Unsupervised non-rigid image distortion removal via grid deformation. In *Proc. of International Conference on Computer Vision*, 2021. 1, 2, 6, 7, 8

[23] Zhiyuan Mao, Nicholas Chimitt, and Stanley H Chan. Image reconstruction of static and dynamic scenes through anisoplanatic turbulence. *IEEE Transactions on Computational Imaging*, 6:1415–1428, 2020. 1, 2

[24] Zhiyuan Mao, Nicholas Chimitt, and Stanley H Chan. Accelerating atmospheric turbulence simulation via learned phase-to-space transform. In *Proc. of International Conference on Computer Vision*, 2021. 2, 5

[25] Zhiyuan Mao, Ajay Jaiswal, Zhangyang Wang, and Stanley H Chan. Single frame atmospheric turbulence mitigation: A benchmark study and a new physics-inspired Transformer model. In *Proc. of European Conference on Computer Vision*, 2022. 1, 2, 6, 7, 8

[26] Kangfu Mei and Vishal M Patel. LTT-GAN: Looking through turbulence by inverting gans. *IEEE Journal of Selected Topics in Signal Processing*, 17(3):587–598, 2023. 1

[27] Nithin Gopalakrishnan Nair and Vishal M Patel. Confidence guided network for atmospheric turbulence mitigation. In *Proc. of International Conference on Image Processing*, pages 1359–1363, 2021. 1

[28] Robert J Noll. Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America*, 66(3):207–211, 1976. 5

[29] Omar Oreifej, Xin Li, and Mubarak Shah. Simultaneous video stabilization and moving object detection in turbu-

lence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):450–462, 2012. 2

[30] Shyam Nandan Rai and CV Jawahar. Removing atmospheric turbulence via deep adversarial learning. *IEEE Transactions on Image Processing*, 31:2633–2646, 2022. 1

[31] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Proc. of Medical Image Computing and Computer-Assisted Intervention*, 2015. 4

[32] Armin Schwartzman, Marina Alterman, Rotem Zamir, and Yoav Y Schechner. Turbulence-induced 2D correlated image distortion. In *Proc. of International Conference on Computational Photography*, 2017. 2

[33] Mikhail A Vorontsov and Gary W Carhart. Anisoplanatic imaging through turbulent media: image recovery by local information fusion from a set of short-exposure images. *Journal of the Optical Society of America*, 18(6):1312–1324, 2001. 8

[34] Renjie Wan, Boxin Shi, Ling-Yu Duan, Ah-Hwee Tan, and Alex C Kot. Benchmarking single-image reflection removal algorithms. In *Proc. of International Conference on Computer Vision*, pages 3922–3930, 2017. 8

[35] Fang Wang, Qiyuan Zhang, Haoyu Wang, Dayong Zhu, Huaikang Zhu, Wenxin Jia, and Sen Han. An improved fitting method for predicting the zernike coefficient–wavelength curves. *IEEE Photonics Journal*, 13(2):1–17, 2021. 5

[36] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 7

[37] Joseph Winocur. Dual-wavelength adaptive optical systems. *Applied Optics*, 22(23):3711–3715, 1983. 2

[38] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proc. of European Conference on Computer Vision*, 2018. 5

[39] Rajeev Yasarla and Vishal M Patel. Learning to restore images degraded by atmospheric turbulence using uncertainty. In *Proc. of International Conference on Image Processing*, 2021. 1, 2, 6, 7, 8

[40] Rajeev Yasarla and Vishal M Patel. CNN-based restoration of a single face image degraded by atmospheric turbulence. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(2):222–233, 2022. 1, 2

[41] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient Transformer for high-resolution image restoration. In *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. 5

[42] Xingguang Zhang, Zhiyuan Mao, Nicholas Chimitt, and Stanley H Chan. Imaging through the atmosphere using turbulence mitigation Transformer. *arXiv preprint arXiv:2207.06465*, 2022. 1, 2

[43] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6):1452–1464, 2017. 5

[44] Xiang Zhu and Peyman Milanfar. Stabilizing and deblurring atmospheric turbulence. In *Proc. of International Conference on Computational Photography*, 2011. 1

[45] Xiang Zhu and Peyman Milanfar. Removing atmospheric turbulence via space-invariant deconvolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):157–170, 2012. 2