Simulation of atmospheric turbulence for a qualitative evaluation of image restoration algorithms with motion detection

Claudia S. Huebner*a, Szymon Gladysz
Fraunhofer IOSB, Gutleuthausstrasse 1, 76275 Ettlingen, Germany

ABSTRACT

Remote sensing applications are generally concerned with observing objects over long distances. When imaging over long horizontal paths, image resolution is limited by the atmosphere rather than by the design and quality of the optical system being used. Atmospheric turbulence can cause quite severe image degradation, the foremost effects being blurring and image motion. Recently, interest in image processing solutions has been rising, not least of all because of the comparatively low cost of computational power, and also due to an increasing number of imaging applications that require the correction of extended objects rather than point-like sources only. At present, the majority of these image processing methods aim exclusively at the restoration of static scenes. But there is a growing interest in enhancing turbulence mitigation methods to include moving objects as well. However, an unbiased qualitative evaluation of the respective restoration results proves difficult if little or no additional information on the "true image" is available. Therefore, in this paper synthetic ground truth data containing moving vehicles were generated and a first-order atmospheric propagation simulation was implemented in order to test such algorithms. The simulation employs only one phase screen and assumes isoplanatic conditions (only global image motion) while scintillation effects are ignored.

Keywords: Turbulence simulation, atmospheric turbulence, image restoration, motion detection, image quality

1. INTRODUCTION

1.1 Motivation

Electro-optical systems can be optimized to a great extent with regard to the application they are intended for. However, in remote sensing, image resolution is generally limited by the atmosphere rather than by the design and quality of the optical system being used. When imaging over long horizontal paths, the degree of image degradation is particularly severe since atmospheric turbulence is strongest close to the ground. The foremost image-degradation effects, temporal and spatial blurring, arise from random inhomogeneities in the temperature distribution of the atmosphere, producing small but significant fluctuations in the index of refraction. Light waves propagating through the atmosphere will sustain cumulative phase distortions as they pass through these turbulence-induced fluctuations.

In the past, a number of methods have been suggested for the mitigation of turbulence effects with most related to astronomical applications. But lately, interest in solutions from the field of image processing is rising, not least of all because of the comparatively low cost of computational power, and also due to an increasing number of imaging applications that require the correction of extended objects rather than point-like sources only. At present, the majority of these image processing methods aim exclusively at the restoration of static scenes. But there is a growing interest in enhancing turbulence mitigation methods to include moving objects as well.

However, an unbiased qualitative evaluation of the respective restoration results proves difficult if little or no additional information on the "true image" is available. Therefore, in this paper synthetic ground truth data containing moving vehicles were generated and a first-order atmospheric propagation simulation was implemented in order to test such algorithms. The simulation employs only one phase screen and assumes isoplanatic conditions (only global image motion) while scintillation effects are ignored.

*claudia.huebner@iosb.fraunhofer.de; phone +49 7243 992-252; fax +49 7243 992-299; www.iosb.fraunhofer.de
1.2 Overview

This article is divided into two main parts with obvious emphasis on the first part, outlining the simulation of image data degraded by atmospheric turbulence. Here, the individual steps of the turbulence simulation are described and sample results are given for the cases $C_n^2 = 10^{-14}, 5 \times 10^{-14}, 10^{-13}, 5 \times 10^{-13}, 10^{-12}$; outer scale values are also given (1 m or 10 m). In the second part, a summary is given of the individual components of the turbulence mitigation algorithm to be tested on the simulated data. These include a classic motion compensation and image integration procedure to account for global image motion as well as motion detection and motion estimation methods to additionally consider moving objects by employing local image stacking. To conclude, image quality metrics are discussed in view of the qualitative evaluation of image restoration algorithms with particular regard to deconvolution side-effects.

2. SIMULATION

2.1 "Ground Truth" Data

A 3-D POV-Ray model was created to serve as "ground truth" basis for consequent turbulence simulations. The model consists of a generic military camp in a slightly uneven terrain and contains two moving trucks heading in different directions. From the 3-D model, an RGB image sequence of 2000 frames in BMP format was rendered with a resolution of 1024×768 pixels. The movement of both vehicles lasts over the complete 2000 frames in order to emulate a high speed camera. Their motion trails spanning from the first to the last frame are depicted in Figure 1.

![Fig. 1. Sample frames of the synthetic "ground truth" sequence. (a) First frame with moving vehicles and their tracks marked in green and yellow resp., (b) last frame of sequence with complete motion trails marked by dashed arrows.](image)

This 3-D model was mainly chosen because of its great flexibility with regard to future applications. Sequences with higher resolution can be produced easily. It should be noted, though, that the computation time necessary for the rendering process increases with higher image resolution. It took approximately 2 hrs and 20 min, for instance, to render all 2000 frames at 1024×768 resolution of the sequence used here. Another obvious advantage is that some objects may be removed while others are added. Also, the camera angle can be changed without too much time and effort, or the lighting can be altered.

Another deciding factor was the importance of motion content in our test scene. Specifically, because turbulence mitigation with motion detection was recently the subject of our study [3],[4], it was to test these algorithms that the idea to create a simulation arose in the first place.

2.2 Turbulence Simulation

To test the algorithms a first-order atmospheric propagation simulation was implemented. The simulation employs only one phase screen and assumes isoplanatnic conditions (only global image motion). Scintillation effects are ignored.
Independent phase screens were generated using the classic FFT-based method of McGlamery [1], whereby an array of random numbers is filtered according to the von-Karman spectrum and then inverse Fourier-transformed. This method suffers from inadequate representation of low-spatial-frequency aberrations. Therefore, subharmonics correction is implemented [2] with eight levels of subharmonics. This yielded relatively good accuracy when simulated and theoretical phase structure functions were compared. The outer scale was set either to 1 m or to 10 m. Size of the telescope aperture was set to 10 cm and 256 × 256 arrays were used for phase screens giving pupil sampling of 0.4 mm per pixel. Turbulence strength was chosen between $C_n^2 = 10^{-14}$ and $C_n^2 = 10^{-12}$. Given that in this simulation one monochromatic PSF corresponds to one phase screen, it can be assumed that the images contain only speckles from “frozen” wave fronts and therefore they correspond to very short integration times, of the order of 5 ms.

The phase screens were converted to phasors, multiplied by a circular aperture and embedded in an array of zeros of size 512 × 512 pixels. The field was then Fourier-transformed. Square of the modulus of the result gives the PSF. The (original) pixel scale of the PSFs is 2.44 μrad (Nyquist sampling at the shortest wavelength). Only the central 256 × 256 part of the images is saved.

The simulation is polychromatic: images corresponding to ten wavelengths between 500 and 700nm are generated assuming linear wavefront scaling. The resulting ten PSFs are co-added and the result constitutes the final noise-free PSF. Finally, PSFs were normalized so that their total power was equal to unity.

The assumed parameters of the sensor follow the characteristics of Nikkor ED 800. Since this sensor has pixels of angular size ~ 10.0 μrad, binning of the imaged scene by a factor of 4 was necessary. After the ground truth image was convolved with the PSFs, the resulting image was binned down accordingly.

Figures 2 and 3 show PSF samples created with this method for $C_n^2 = 10^{-14}$, 5·10$^{-14}$, 10$^{-13}$, 5·10$^{-13}$, 10$^{-12}$; outer scale values (1 m or 10 m) are also given in the figures. Outer scale affects mostly image motion, therefore images corresponding to 1 m or 10 m outer scale look similar. The value of $C_n^2$ has the biggest effect on resolution per single frame. Additionally, in order to give an impression of their long term behaviour, 20 PSFs were integrated for each case, respectively, and shown alongside the sample PSFs.

The selected PSFs were cropped to show only the central part as any values outside that area are too small to be discernible in a greyscale image. The scale of the images was not changed, i.e. the resolution in Figures 2 and 3 are both identical.
2.3 Simulation Results

Due to the necessary oversampling of the PSF filter functions, the synthetic images were up-scaled by factor 2 before filtering, using nearest neighbour resizing. The actual filtering was then carried out by means of a convolution with symmetric replication instead of zero padding to account for size differences between images and filter functions. The same PSF was used on all three RGB colour channels since the differences in wavelength between red, blue and green are comparatively small. Afterwards, the filtering result was down-scaled accordingly by factor 4.
Fig. 4. Comparison of (a) "ground truth" and (b)-(f) exemplary simulation results for increasing $C_n^2$, in each case with outer scales 1m and 10m, respectively, as noted.
Some filtering results for all the specified cases, i.e. $C_n^2 = 10^{-14}, \ldots, C_n^2 = 10^{-12}$, outer scales 1m and 10m, respectively, are shown in Figure 4 (spreading over two pages). The results are sorted by increasing turbulence strength, including a section of the unfiltered "ground truth" image for easier comparison. For easier discernability of the differences between the respective simulation results, close-ups were added on the right side. Deterioration is most noticeable in the decreasing resolution of the flag and the tents since here image contrast is highest. The decline between $C_n^2 = 10^{-13}$ and $C_n^2 = 5 \cdot 10^{-13}$ appears to be particularly severe as details like the tent flaps or the flag disappear almost completely.

2.4 Video Simulation

For increased realism with regard to creating a video sequence from the simulations, every 8 frames were integrated into a long exposure in order to imitate a video frame rate. Generally, short-exposure data are preferred as input for image restoration algorithm in order to ensure that the turbulence in each frame is "frozen", i.e. the temporal changes of the temperature variation are so small that the index of refraction can be considered as constant. The number of synthetic frames was purposefully chosen to be large, so that a single (filtered) frame corresponds to a short-exposure image while the average can be considered as long-exposure.

Furthermore, in order to account for electronic or "shot noise", Poisson noise was generated from the images themselves by interpreting each pixel as mean value of a Poisson distribution. In order to also account for additive sensor read-out noise, Gaussian white noise of zero mean and 0.001 variance was added.

This simple noise model is intended to serve only as a first approximation since at this point it is not the aim of the authors to model the noise behaviour of any specific type of detector. Nevertheless, this noise model is likely to be expanded for more accuracy in the future, e.g. to include banding noise and/or fixed pattern noise as are typical for digital sensors. Clearly, this procedure entails additional image degradation, as illustrated in Figure 5, exemplarily on the simulation $C_n^2 = 5 \cdot 10^{-13}$ with outer scale 10m.

![Image](image.png)

Fig. 5. Exemplary results of simulation $C_n^2 = 5 \cdot 10^{-13}$ (outer scale 10m) overlaid with noise. (a) Basic filtering result with Poisson and Gaussian noise; (b) Average of 8 filtered frames with Poisson and Gaussian noise.

3. TURBULENCE COMPENSATION WITH MOTION DETECTION

3.1 Motion Compensated Averaging

The turbulence compensation algorithm to be tested on the simulated data consists of several steps, the first being classic global motion compensation with image integration (MCA – Motion Compensated Averaging) which is in essence the same as normal image integration, the main difference being, that before integrating the next frame of the input sequence it is shifted slightly within a given search space of a number of pixels in every direction such that the input frame best matches a given reference image. Ideally, this reference would be an image unimpaired by atmosphere and optical system. Since normally such an ideal image will not be available, a moving average or temporal median present simple and reasonable substitutes. A detailed description of this part of the algorithm can be found in [5].

Although, in the given case, ground truth data are available by design, the main idea is to test the algorithm's performance and so it will be assumed to be unknown. Given that the simulation assumes isoplanatic conditions, the result can be considered to be a non-distorted image that has been blurred by a Point Spread Function (PSF) of the same size as the pixel motions due to the turbulence.
3.2 Motion Detection

When dealing with atmospherically degraded and noisy data it often is practically unavoidable to apply at least some kind of averaging procedure which has become an integral part of various turbulence compensation schemes for this very reason. Consequentially, occurring motion blur, either real or created by averaging, is amplified proportionately to the number of stacked images. An example for this "ghost-effect" from a sequence degraded by strong atmospheric turbulence ($C_n^2 \sim 10^{-13}$) is shown in Figure 6.

![Image of original frames and ghost-effect in average images]

Fig. 6. Example for ghost-effect caused by motion when averaging. On the left in (a) and (b): original frame and average of 150 frames resp. from before movement accelerates; on the right: frame and average resp. from after.

Motion detection and estimation as detailed in [3] and [4] constitute the next steps in the algorithm. The procedure employs block-matching (BM) aided by image differencing for the segmentation of foreground and background, i.e. moving objects and static scene elements. On the one hand, block-matching enables the detection of moving objects – provided their movements exceed the apparent motion caused by atmospheric turbulence. Objects in the foreground and the background can then be corrected for turbulence separately. On the other hand, BM can be employed to estimate motion vectors between consecutive frames to be used in directed local image stacking (LIS) as illustrated in Figure 7. This has the advantage that object movement is not taken into account when averaging, thus effectively reducing ghost-effects.

![Diagram of local image stacking principle]

Fig. 7. Illustration of local image stacking principle.

3.3 Blind Deconvolution

The final step in the restoration process is a blind deconvolution of the overall result. Essentially, a deconvolution describes the procedure of separating two convolved signals $f$ and $h$. In the spatial domain the blind deconvolution problem takes the general form as illustrated by the equation:

$$g(x, y) = h(x, y) * f(x, y) + n(x, y)$$

where $g$ denotes the blurred image, $h$ the unknown blurring function, generally referred to as Point Spread Function (PSF), $f$ the true image, * the convolution operator and $n$ an equally unknown additive noise component. Unfortunately, the problem is ill-posed: due to the noise the true image can never be recovered even if the exact blurring function is known.
Many attempts at solving this deceptively simple equation for a wide variety of applications can be found throughout literature. An overview of the most popular of these blind image deconvolution algorithms is detailed in [9]. The Lucy-Richardson Deconvolution (LRD) algorithm was developed independently by [10] and [11] and is a nonlinear and basically non-blind method, meaning the PSF, or at least a very good estimate, must be a priori known. It has been derived from Bayesian probability theory where image data are considered to be random quantities that are assumed to have a certain likelihood of being produced from a family of other possible random quantities. The problem regarding the likelihood that the estimated true image, after convolution with the PSF, is in fact identical with the blurred input image, except for noise, is formulated as a so-called likelihood function, which is iteratively maximized. The Iterative Blind Deconvolution (IBD) algorithm, proposed by [12], is mainly a blind version of the LRD algorithm where the PSF needs not to be known, only its support. The IBD uses Expectation Maximization which is an optimization strategy for estimating random quantities that have been corrupted by noise. The likelihood function from the LRD algorithm is maximized iteratively with specified constraints until an estimate for the blurring PSF is retrieved from the data along with the estimate for the true image.

Here, a variation of the IBD-algorithm as proposed in [4] was chosen for the image restoration. Additionally, the Sobel-filtered reference image is used as weighting function in order to amplify the deconvolution effect around edges while reducing ringing effects in homogeneous image regions.

3.4 Image Quality Assessment

As such, there is no simple answer to the question about evaluating image quality. For once, it strongly depends on the intended application. In image or video compression, for instance, it is comparatively easy to decide on an image’s quality. Here, a higher compression rate is equivalent to lower image quality and so-called "full reference" Image Quality Metrics (IQM) can be employed to compare the perfect, i.e. uncompressed reference image to the compressed version.

However, to give an objective qualitative evaluation of the respective restoration results from different turbulence compensation algorithms is much more difficult given that typically little or no additional information on the "ground truth" will be available, meaning it will depend on so-called "no-reference" metrics. Considering that the aim of such image restoration algorithms is to improve the resolution of image data, "good" image quality will consequently refer to high contrast of fine detail. Since the low-pass filter effect of the atmosphere removes the high frequency components, it is only natural to consider metrics that exploit Fourier spectral analysis, image sharpness (intensity squared), grey value gradients (sum or variance of edges) or local intensity variance.

The simulation outlined in Section 2.2 where ground truth data are available by design was originally motivated by the observation made during earlier work [6] that such no-reference IQMs were sensitive to image degrading deconvolution side-effects like ringing, noise amplification, excessive contrast boost, etc., which can affect their performance quite adversely. The existence of ground truth data facilitates matters as the evaluation problem can be reduced to assessing image fidelity, i.e. giving a qualitative measure about the degree of similarity between the restored and the original image. Full-reference metrics can be employed, e.g. simple error summation methods (MSE, RMSE, MAD) although these generally perform unsatisfactorily as discussed in [13], or rather the basic SSIM suggested in [14] which measures structural distortion instead of error energy. It is defined for two images x and y by:

\[
SSIM(x, y) = \frac{2\sigma_{xy} + \sigma_x + \sigma_y}{\sigma_x^2 + \sigma_y^2 + (\sigma_{xy})^2}
\]

while \(\bar{x}\) and \(\bar{y}\) denote the respective mean values, \(\sigma_x\) and \(\sigma_y\) the standard deviations and \(\sigma_{xy}\) the cross-covariance. In the meantime, the SSIM has also been developed [15] for the complex wavelet transform domain (CW-SSIM).

4. CONCLUSION

An unbiased qualitative evaluation of the respective restoration results from different turbulence compensation algorithms proves difficult if little or no additional information on the "ground truth" is available. Therefore, in this paper
synthetic ground truth data containing moving vehicles were generated and a first-order atmospheric propagation simulation was implemented in order to test such turbulence mitigation algorithms by using one of the metrics discussed in the previous section 3.4. The simulation itself has to be considered preliminary as it employs only one phase screen and assumes isoplanatic conditions (i.e. only global image motion) while scintillation effects are ignored.

Consequently, future work will be focused primarily on simulating anisoplanatic conditions but also on generally expanding the simulation by increasing the number of phase screens and diversifying spectral bands (e.g. IR spectrum) as well as taking the difference in wavelengths between the RGB channels of colour images into account.

Since most of the simulation results presented here are in colour, it should be noted that the quality of the results can best be appreciated in the digital version of this paper.

ACKNOWLEDGMENTS

The authors would like to thank their colleague Marcus Hebel for generously providing the 3-D POV-Ray model and helping create the "ground truth" sequence as basis for the simulations presented here.

REFERENCES
