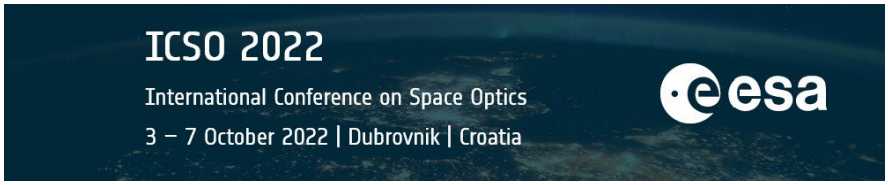
International Conference on Space Optics—ICSO 2022

Dubrovnik, Croatia

3–7 October 2022

Edited by Kyriaki Minoglou, Nikos Karafolas, and Bruno Cugny,



Machine Learning-Based Adaptive Optics for Free-Space Optical Communication: a Training Data Generation Study

International Conference on Space Optics — ICSO 2022, edited by Kyriaki Minoglou, Nikos Karafolas, Bruno Cugny, Proc. of SPIE Vol. 12777, 1277755 · © 2023 ESA and CNES · 0277-786X · doi: 10.1117/12.2690988

Machine Learning-Based Adaptive Optics for Free-Space Optical Communication: a Training Data Generation Study

Vinícius F. Nery[a] and Shinichi Nakasuka[a]

[a]Department of Aeronautics and Astronautics, The University of Tokyo, Tokyo, Japan

ABSTRACT

Recently, free-space optical (FSO) communication systems have gained attention as they pose as a possible solution to cope with the ever-increasing requirements for data transmission. FSO systems have a number of attractive attributes, such as high bandwidth, license-free spectrum, long operational range, security, spatial re-usability, and immunity to electromagnetic interference. On the other hand, optical links suffer from strict pointing and tracking requirements due to narrow beam divergence and channel losses that greatly hinder their reliability. Particularly in the near-Earth satellite optical communications scenario, the atmosphere represents a major challenge to overcome for the establishment of successful and enduring satellite optical links. In that sense, adaptive optics (AO) solutions have been considered for turbulence compensation but given communication requirements solutions are either expensive or sub-optimal. This paper proposes a machine learning-based AO system general framework for atmospheric turbulence compensation. The theoretical foundation for training data generation based on Monte-Carlo phase screens, Zernike polynomials, and optical propagation is presented and validated through simulations. Furthermore, evaluation criteria for the neural network and overall system performance are also discussed.

Keywords: Free-space Optics, Adaptive Optics, Machine Learning, Training Data, Zernike polynomials, Phase Screen Method

1. INTRODUCTION

In recent years, the volume of data being transmitted to and from satellites has increased drastically. Not only mobile Internet (5G) and Internet-of-Things (IoT) applications, but also data-intensive instruments such as multi-band radiometers or hyper-spectral imagers, are increasing greatly the requirements on bandwidth and capacity for satellite system communications. This rapid growth contributed to congestion of the radio-frequency (RF) spectrum and laser communication is a contender to overcome this bottleneck, providing higher throughput communications while reducing the necessary volume, weight and power requirements on the satellite platforms.

Laser communication, or free space optical communication (FSO) utilizes an optical carrier (visible, infrared (IR)) to transfer information from one point to another. For that reason, there are several advantages over RF communication systems such as high bandwidth, license-free band use, long operational range, security, spatial re-usability and immunity to electromagnetic interference.[1] However, the main disadvantage of this type of system is the requirement of tight pointing, acquisition and tracking (PAT) due to narrow beam divergence. Also, due to the wavelength of the optical carrier it is subjected to atmospheric absorption and scintillation (turbulence), which greatly degrade the performance of the system.[2]

The major difficulties imposed by the atmosphere on the performance of FSO communication systems can be roughly summarized in two: cloud blockage and atmospheric turbulence. The former can be mitigated using site diversity where the most appropriate ground station in terms of weather conditions and line of sight (LOS) is selected for beam reception. The latter, however, is inherent to any optical beam propagation in the atmosphere, introducing distortions to its wavefront which in cases of strong or even medium levels can lead to severe drops in performance, coherence loss, and even link failure. To mitigate received beam wavefront distortion, adaptive

Further author information: (Send correspondence to V.F.N)
V.F.N.: E-mail: viniciusfnery@space.t.u-tokyo.ac.jp
S.N.: E-mail: nakasuka@space.t.u-tokyo.ac.jp

optics (AO) is the common approach. In simple terms, AO consists in measuring the incoming wavefront, reconstructing it, and using it to control a deformable mirror (DM) to compensate for distortions. However, the requirements for an optical communication adaptive optics system are more strict than astronomical imaging ones, which can result in excessively expensive or sub-optimal systems.

With the rapid development of artificial intelligence in recent years, an adaptive optics system based on machine learning for FSO communications can be seen as a possible solution for cost and complexity reduction while maintaining efficiency which is desirable for future optical ground stations. Even though there are a few studies at this point on the topic,[3,4] there still is a need to clarify the requirements of such system for the specific scenario of optical communications and the metrics to evaluate such system. In the present paper, we provide an initial model concept and theoretical foundation for a machine learning-based adaptive optics system. Furthermore, since training data is of fundamental importance for the success or failure of neural networks, a detailed discussion on data generation for wavefront detection by using Monte Carlo phase screens and Zernike polynomials is conducted.

2. ATMOSPHERIC TURBULENCE MODEL

Optical waves propagating through the atmosphere are subjected to distortion due to their small wavelengths. Temperature variations along the atmosphere lead to changes in atmospheric density and, hence, index of refraction. This results in significant inhomogeneities in its index profile and as a consequence in a particularly difficult channel to model. Random changes in wind velocity, for example, are a natural result of such temperature gradients, which we perceive as turbulent motion. In the particular case of optical communication, the received beams suffer from intensity and phase fluctuations, i. e., the scintillation effect.

As an attempt to correctly portrait the effects of the atmosphere, numerous models have been proposed in the literature.[5] As the complexity of the atmosphere is beyond the capabilities of deterministic prediction or numerical analysis, statistical analysis is the source for atmospheric models. Among the several dozen of profile models developed from experimental measurements, the Hufnagel-Valley boundary (HVB) model was the one adopted in this work. The HVB model is given by:

$$C_n^2(h) = 5.94 \times 10^{-53} h^{10} e^{-h/1000} \left(\frac{v}{27}\right)^2 + 2.7 \times 10^{-16} e^{-h/1500} + A e^{-h/100}, \tag{1}$$

where h is the height in meters, A is the nominal value of C_n^2 at ground and v is the RMS wind speed. The value of v is determined from

$$v = \left[\frac{1}{15 \times 10^3} \int_{5\times10^3}^{20\times10^3} V^2(h) dh\right]^{1/2}. \tag{2}$$

Here, $V(h)$ is described by the Bufton wind model as

$$V(h) = \omega_g h + v_g + 30 e^{-\frac{h-9400}{4800}}, \tag{3}$$

where v_g is the ground wind speed and ω_g is the slew rate associated with satellite motion relative to an observer on the ground.[6] Since it also encompasses the effects of satellite motion, this model seems to fit well the LEO optical communication scenario.

The described index of refraction parameter C_n^2 is a measure of the local turbulence strength. However, there are other quantities that might be more useful and with more intuitive meanings. This is illustrated by the atmospheric coherence diameter, or Fried parameter, r_0, which considers the optical effect of C_n^2 over a specified propagation length and is defined for a plane-wave source as

$$r_0 = \left[0.423 k^2 \sec(\beta) \int_0^L C_n^2(h) dh\right]^{-3/5}. \tag{4}$$

Here, L is the path length, k is the spatial wave number, and β is the zenith angle. From equation (4) it is noticeable that while C_n^2 is a function of the propagating distance h, the coherence diameter r_0 can express with a single number the optical turbulence effect over the whole propagation path.

For wavefront compensation phase information of the incoming wave is of great importance. Thus, to study the effects of turbulence to the wavefront, the phase power spectral density (PSD) described by the modified Von Kárman model was adopted. The phase PSD is given by

$$\Phi_{\phi}^{mvK}(\kappa) = 0.49 r_0^{-5/3} \frac{\exp\left(-\frac{\kappa^2}{\kappa_m^2}\right)}{(\kappa^2 + \kappa_0^2)^{11/6}}, \quad \text{for} \quad 0 < \kappa < \infty, \tag{5}$$

where $\kappa_m = 5.92/l_0$ and $\kappa_0 = 2\pi/L_0$. The choice for this model relies on its simplicity and inclusion of inner (l_0) and outer (L_0) scale effects.[7]

3. ADAPTIVE OPTICS & MACHINE LEARNING

3.1 Adaptive Optics

To mitigate the effect of atmospheric turbulence, adaptive optics (AO) has been fairly studied. In simple terms this technique consists in a closed loop control system where, based on the distortion of the wavefront detected by sensors, control inputs are generated so actuators can introduce the inverse effect to the incoming beam. A major benefit of AO technology is that it allows coupling of a received laser beam (which has traversed through the atmosphere) into a single-mode fiber.[5]

Conventional methods consist in the use of wavefront sensors (WFS) to measure and analyze the distorted wavefront. Then computers are used to process the information so appropriate control signals can be provided to a fast-steering mirror (FSM) for tip/tilt correction and to a deformable mirror which can correct the wavefront so it returns to its original flat state.[8] A simplified example of such configuration is provided in Fig. 1.

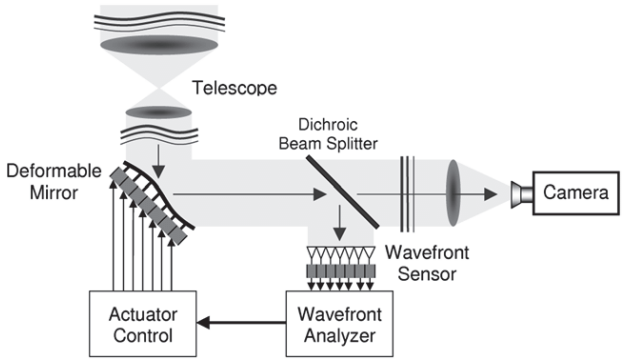


Figure 1. Simple representation of an adaptive optics system.

The main principle behind wavefront correction is known as phase conjugation. This principle consists in multiplying the incoming light field $|E|e^{-j\phi}$ by its complex conjugate $|E|e^{+j\phi}$, thus the name phase conjugation. However, as simple as the idea is, whether it is possible to introduce the correct amount of phase conjugation at the right spot at the right time dynamically is where lies its complexity. Common limitations of AO include availability of a sufficiently strong signal (turbulence and resulting fades are not too strong); WFS, DM, and processor with high enough bandwidth; and adequate spatial correction bandwidth.[9]

3.2 Machine Learning-based AO

In response to the limitations of current AO systems and advancements in AI technology, machine learning (ML) solutions have been proposed in recent literature.[3,4,10] One of the main principles of an AO system is essentially identifying a nonlinear mapping. Given characteristics of the incoming distorted beam at the receiver compute the relationship to the corresponding wavefront aberration and use that information for real-time compensation via available actuators. To better visualize the nature of this relationship, consider a phase (wavefront) distortion

$\Phi(x, y)$ applied to a propagating field $U(x, y, z)$ at the z direction. The corresponding intensity distribution over a distance Δz can be derived from light propagation theory as:

$$I(x, y) = \left| F_t^{-1}\{F_t\{U(x, y, z) \exp[j\Phi(x, y)]\} D(z)\} \right|^2 ,$$
$$D(z) = \exp\left[-i\Delta z \frac{k_x^2 + k_y^2}{2k} \right] , \tag{6}$$

where F_t and F_t^{-1} denotes the Fourier transform and its inverse, and $D(z)$ is the transfer function of Fresnel propagation.[11] It is clear that the relationship in equation (6) is quite complex especially considering that the introduced phase $\Phi(x, y)$ is unknown at first and its value changes dynamically with the atmospheric conditions. In that sense, machine learning solutions have been considered to reshape the wavefront aberration measurement, control and post-processing to either improve performance or reduce complexity of AO systems.[3,4,10] However, ML-based solutions for the optical communication scenario are still limited or very application specific.

To investigate a more general solution applied to optical communication systems a careful study on the training data generation and system framework is of great importance. A schematic of the proposed system can be seen in Fig. 2. The system can be divided in two parts: wavefront detection and wavefront correction. The two parts are represented in figure 2 by the green and yellow colors respectively.

At the detection side a neural network is used to map the distorted input to corresponding aberration. The aberration is given by a mode-limited Zernike polynomial representation and will be used to generate control inputs for the DM. Finally, at the correction side the control-loop is closed and the beam is compensated to minimize an appropriate cost function. Even though a detailed discussion on the neural network architecture is not on the scope of this work, it is impractical and perhaps naive to generate training data without having a network architecture in mind. That said, it is assumed that the neural network is a convolutional neural network (CNN). Given that the focus of this work is to elaborate on the training generation aspect of the system let us focus our discussion from now on to its wavefront detection side.

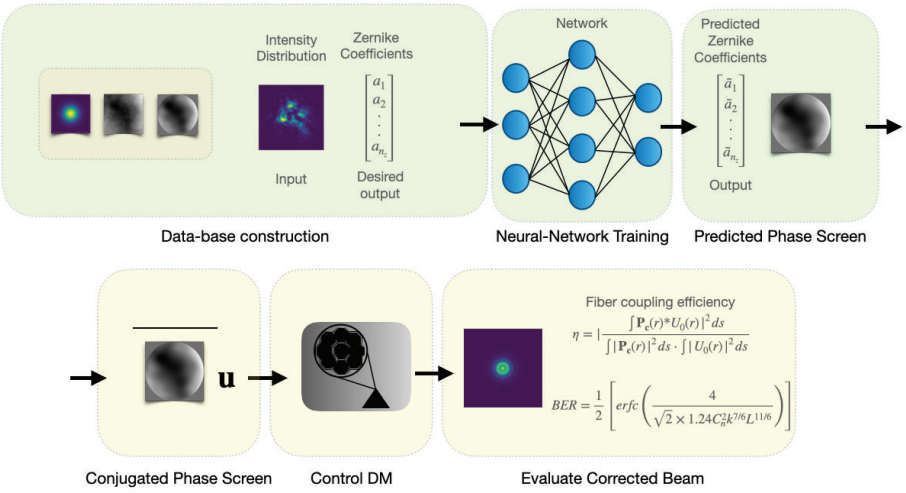


Figure 2. Schematic of the machine learning-based AO system.

It is also important to point out that how to evaluate system performance or neural network performance is not clear from the start and there is not necessarily a correct answer. At Fig. 2 it can be seen at the bottom two possible metrics that can be used to discuss performance. The first one is fiber coupling efficiency, since turbulence causes fluctuations in collected power ($\mathbf{P_c}$) by a fiber optic cable, the coupling efficiency can be expressed as

$$\eta = |\frac{\int \mathbf{P_c}(r)^* U_0(r)|^2 ds}{\int |\mathbf{P_c}(r)|^2 ds \cdot \int |U_0(r)|^2 ds}. \tag{7}$$

On the other hand, since the main goal is improving FSO communication systems, a communication metric such as bit error rate (BER) is also of interest and can be expressed as in[5]

$$BER = \frac{1}{2}\left[erfc\left(\frac{4}{\sqrt{2}\times 1.24 C_n^2 k^{7/6} L^{11/6}}\right)\right]. \tag{8}$$

However, it is important to realize that this expression for the BER is quite general and would need to be reexamined before actual implementation.

4. TRAINING DATA GENERATION

The machine learning algorithm of interest is of the supervised kind. That means training data is of fundamental importance to tune the network's weights and assure accurate predictions. Even though the use of real data from ground stations for training would be ideal to increase system's performance in actual operation scenario, for this study we want to have control of turbulence level that is generated to test the generalization capabilities of the network. For that same reason we consider only one phase-screen at a time, instead of multiple phase screens equally spaced. Therefore, the training data was generated through simulations.

The training data is composed by two sets of data: input data and output data. In the case of our network, as illustrated in figure 2, the input is given by the intensity distributions at the receiver which is represented by an image, while the output is the corresponding phase distortion given in terms of Zernike coefficients. Given a 2D phase map $\phi(x, y)$ and an initial beam, it is possible to generate the input data by equation (6). In the following let us detail how to generate the 2D phase maps which will be referred as phase screens and their representation through Zernike polynomials.

4.1 Gaussian Beam

Optical communication systems use optical waves as carriers, so it is only natural that we consider optical beams as a part of our study for training data generation. Optical wave propagation in the atmosphere can be characterized by solutions of the paraxial Helmholtz equation. The Gaussian beam is a known simple solution and can be described as in Eqs. (9) and (10).

$$E(r,z) = E_0 \frac{W_0}{W(z)} \exp\left[-\frac{r^2}{W(z)^2}\right] \times \exp\left\{-i\left[kz + \frac{kr^2}{2R(z)} + \theta(z)\right]\right\}, \tag{9}$$

where

$$\begin{aligned} W(z) &= W_0\sqrt{1+(\frac{z}{z_0})^2}: && \text{Beam radius,}\\ R(z) &= z[1+(\frac{z}{z_0})^2]: && \text{Radius of curvature,}\\ \theta(z) &= \arctan(\frac{z}{z_0}): && \text{Gouy phase,}\\ z_0 &= \frac{\pi}{\lambda}W_0^2: && \text{Rayleigh range.} \end{aligned} \tag{10}$$

The intensity distribution for a Gaussian beam can be represented as in Fig. 3.

4.2 Phase Screens

Given the stochastic nature of refractive index variations and optical path on the atmospheric channel, turbulence models represent only statistical averages. Thus, to generate individual turbulence realizations we adopt the notion of a thin turbulent layer which carries the statistics of turbulence-induced phase variations. These thin layers are commonly known as phase screens. Assuming that the turbulence-induced phase $\phi(x, y)$ is a Fourier-transformable function, it admits a Fourier-integral representation given by:

$$\Phi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Psi(f_x, f_y)\exp\left[i2\pi(f_x x + f_y y)\right] df_x df_y, \tag{11}$$

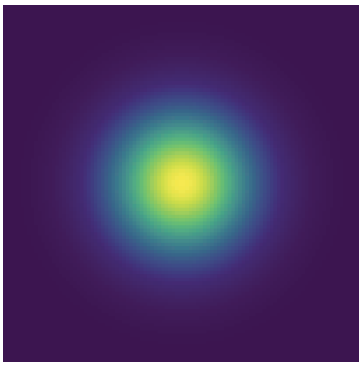


Figure 3. Intensity distribution of a Gaussian beam in 2 dimensions.

where $\Psi(f_x, f_y)$ is the spatial-frequency-domain representation of the phase and $\phi(x, y)$ in cycles/m is a realization of a random process with a power spectral density given by $\Phi_\phi(f)$ defined as in equation (5).[7] However, in terms of actual implementation phase screens must be generated in a finite grid, so a representation in terms of Fourier series is more appropriate and is given by

$$\phi(x, y) = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} c_{n,m} \exp\left[i2\pi(f_{x_n} x, f_{y_m} y)\right]. \tag{12}$$

Here, f_{x_n} and f_{y_m} are the discrete spatial frequencies at x and y directions and $c_{n,m}$ are the Fourier coefficients. The coefficients are determined by filtering a complex Gaussian random number of mean 0 and variance 1 matrix with the phase power spectrum of atmospheric turbulence as represented in equation (13). Thus, the coefficients obey the following relation:

$$\left\langle |c_{n,m}|^2 \right\rangle = \Phi_\phi(f_{x_n}, f_{y_m}) \Delta f_{x_n} \Delta f_{y_m}. \tag{13}$$

This method is known as the power spectrum inverse method.

Finally, Monte Carlo phase screens for different turbulence conditions can be generated by setting a suitable coherence diameter r_0 as in Fig. 4.

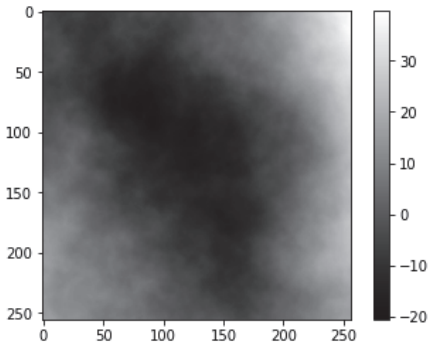


Figure 4. Example of a Monte-Carlo phase screen.

4.3 Zernike Polynomials

There are many ways to represent a wavefront or wavefront aberrations. Given that most of the time we deal with circular apertures, Zernike polynomials are a very convenient representation due to their orthogonality and

completeness over the circle. The definition of Zernike polynomials is given at Eqs. (14) to (16).

$$W(r,\theta) = \sum_{i=1}^{\infty} a_i Z_i(r,\theta), \tag{14}$$

where,

$$Z_i(r,\theta) = \begin{cases} \sqrt{2(n+1)} R_n^m(r)\cos(m\theta), & m \neq 0 \quad i \text{ even} \\ \sqrt{2(n+1)} R_n^m(r)\sin(m\theta), & m \neq 0 \quad i \text{ odd} \\ Z_i = R_n^0(r), & m = 0 \end{cases} \tag{15}$$

and

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s![(n+m)/2-s]![(n-m)/2-s]!} r^{n-2s}. \tag{16}$$

4.4 Mode-limited representation

Even though any wavefront can be written in terms of Zernike polynomials, since there is an infinite number of coefficients that would lead to an infinite number of outputs for the proposed network which is physically impossible. Therefore, instead we consider a mode-limited version. Let n_z be the maximum number of modes (polynomials) considered. A phase screen $\phi(x,y)$ can be represented as:

$$\phi(x,y) = \sum_{i=1}^{\infty} a_i Z_i(r,\theta) = \sum_{i=1}^{n_z} a_i Z_i(r,\theta) + \sum_{i=n_z+1}^{\infty} a_i Z_i(r,\theta). \tag{17}$$

By setting the remaining Zernike terms as the approximation error $\delta\phi$. We can rearrange equation (17) in matricial form as:

$$\mathbf{\Phi} = \mathbf{Z}\mathbf{A} + \delta\mathbf{\Phi}, \tag{18}$$

where $\mathbf{\Phi}$ is a column vector with elements of $\phi(x,y)$ for a grid of N points, $\mathbf{Z}$ is the $N \times n_z$ matrix with elements Z_{ij} and $\mathbf{A}$ is the $n_z \times 1$ mode-limited Zernike coefficients column vector. The coefficients can then be retrieved from the least squares solution of equation (18) as in[7]

$$\mathbf{A} = (\mathbf{Z}^\mathbf{T}\mathbf{Z})^{-1}\mathbf{Z}^\mathbf{T}\mathbf{\Phi}. \tag{19}$$

As it can be seen from Fig. 5, the mode-limited solutions can adequately capture the shape of Monte Carlo phase screens to a certain resolution level with $n_z = 105$ modes. Of course, the number of modes used can be increased but that implies in a larger number of outputs so it is not desirable for this number to be large.

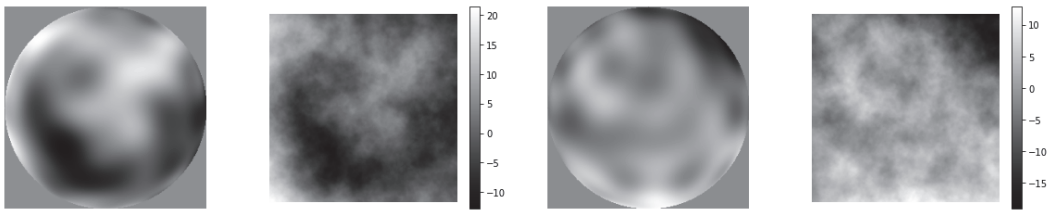


Figure 5. Mode-limited representation of two Monte Carlo phase screens with 105 modes.

4.5 Propagation model

The tools introduced in the previous subsections enable us to correctly generate an atmospheric aberration realization using the phase screen method while retaining its mode-limited Zernike coefficient values. In terms of neural network design, this corresponds to generating the target outputs to be used in the training and validation processes. The inputs of the neural network, on the other hand, are generated by introducing the phase effects of turbulence to a gaussian beam and propagating it through an aperture to an observation plane. The propagation method chosen was the Fresnel diffraction and is given by

$$U(x_0, y_0, z) = \frac{e^{jkz}}{jkz} e^{j\frac{k}{2z}(x_0^2+y_0^2)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x, y, 0) e^{j\frac{k}{2z}(x^2+y^2)} e^{-j2\pi(f_x x + f_y y)} dxdy. \tag{20}$$

Here, $(x, y, 0)$ are the coordinates of a point in the input plane and (x_0, y_0, x) are the coordinates of a point in the observation plane. Furthermore, $f_x = x_0/\lambda z$ and $f_y = y_0/\lambda z$ are the spatial frequencies of the 2-D Fourier transform of $U(x, y, 0)e^{j\frac{k}{2z}(x^2+y^2)}$.[12]

Therefore, given an initial beam described as in equation (9) and a phase screen as described in equation (12) it is possible to obtain a aberrated beam using the propagation model described by equation (20). Moreover, with equation (19) , the coefficients for a Zernike basis are retrieved from the phase screen that generated the aforementioned disturbance. Thus, an input image-desired output pair can be generated to train and validate a neural network.

5. SIMULATION RESULTS & ANALYSIS

In this section, we conduct simulations to validate the theoretical foundation developed thus far for training data generation. The simulation considers a LEO communication scenario. The initial laser beam is generated at the satellite and is propagated through the atmosphere until it is collected by the receiving aperture. At this stage, sensor noise and propagation losses are not considered. The simulation parameters are summarized in table 1.

Table 1. Simulation Parameters.

Symbol	Parameter	Value
d_l	Link distance	500 km
β	Zenith angle	0, 30, 45 deg
v	RMS wind speed	35, 58, 119 m/s
A	Nominal C_n^2	1×10^{-15}, 9×10^{-14}, 1×10^{-13} $\mathrm{m}^{-2/3}$
l_0	Inner scale	0.0001 m
L_0	Outer scale	100 m
w_{ps}	Phase screen width	5.12 m
N	Number of grid points	512
λ	Wavelength	1550×10^{-9} m
I_0	Nominal Tx power	5.0 W
W_0	Beam waist	0.032 m
D	Receiving aperture diameter	0.20 m
d_o	Distance to observation plane	4.5 km

As it can be seen from table 1, different values of β, v, and A were used to compute different atmospheric turbulence conditions. These conditions are incorporated into the coherence diameter r_0 which result in a range of phase screens in the training data generation process.

5.1 Without Atmospheric Disturbance

In the case where there is no disturbance (vacuum propagation) the received beam at the observation plane results in the intensity or diffraction pattern of Fig. 6.

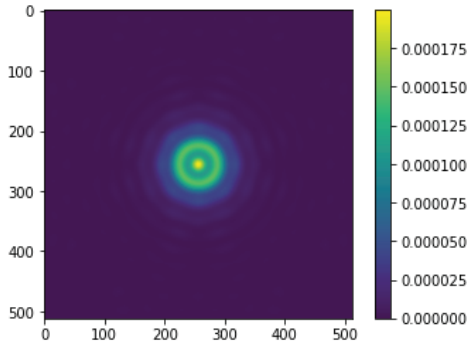


Figure 6. Diffraction pattern of a Gaussian beam at the observation plane after going through an aperture of diameter $D = 20$ cm.

It is noticeable the similarity between the resulting diffraction pattern and the one for plane-wave propagation through a circular aperture. This is a result of the divergence of a Gaussian beam propagating at long distances such as d_l that is seen at the receiver of diameter D as essentially a plane-wave. The radius of the Gaussian beam after propagating $d_l = 500$km at the receiving aperture is $W(d_l) = 7.71$m, in contrast with the aperture diameter $D = 0.20$m.

5.2 With Atmospheric Disturbance

When turbulence is present it is expected that the diffraction pattern at the observation plane should reflect the aberration introduced to the beam's phase. Indeed, Fig. 7 shows the resulting diffraction pattern for different turbulence conditions.

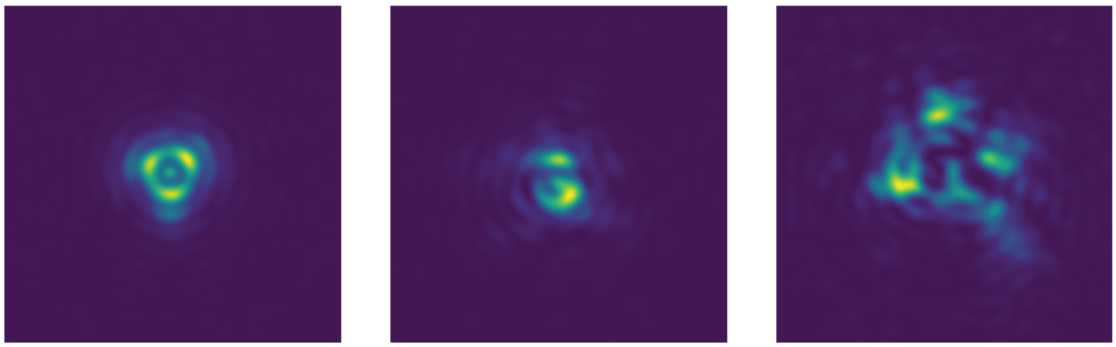


Figure 7. Diffraction pattern for turbulence-aberrated Gaussian beams at the observation plane after going through an aperture of diameter $D = 20$ cm. The left figure corresponds to a weak turbulence scenario, while the one on the middle corresponds to a medium turbulence and the one on the right corresponds to strong turbulence scenarios.

The phase-screen associated to each diffraction pattern and its normalized mode-limited Zernike coefficient values are displayed in Fig. 8. Comparing the diffraction patterns when atmospheric turbulence is considered to the vacuum propagation case the influence of turbulence is clear. Even though the circular shape is somewhat maintained for the weak turbulence scenario, it changes drastically as turbulence intensifies. Furthermore, the

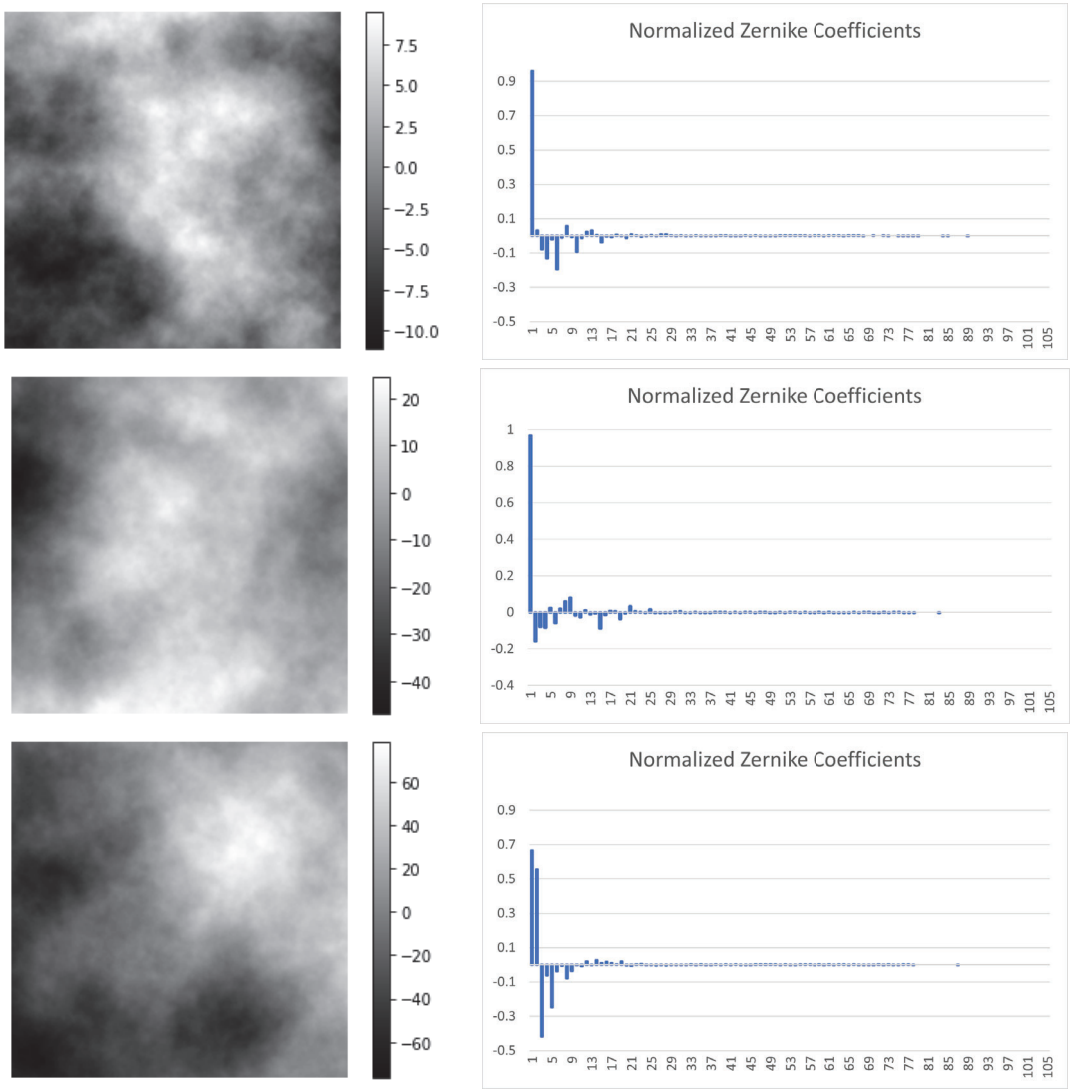


Figure 8. Phase screens associated to each diffraction pattern and their normalized mode-limited Zernike coefficients. At the top is the weak turbulence scenario followed by the medium and strong ones.

intensity becomes spread out so beam power collection would be affected. Observing the diffraction patterns is an efficient way of displaying phase distortion effects since they are the result of interference of different points on the wavefront that travel by paths of different lengths from the telescope's aperture to the observation plane.

By inspecting the Zernike coefficients we can observe that worse turbulence conditions are associated to the number of modes with significant contributions. Even so, higher modes do not seem to contribute much to the distortions introduced by the generated phase screens. This characteristics might be of interest in the training process of a CNN or other neural network models. In particular, considering the choice of the number of modes n_z, it is possible to select the number of modes necessary to achieve a specified percentage of its total root mean square (RMS) value.

6. CONCLUSION & FUTURE WORKS

In this paper, we discussed the concept of a machine learning-based adaptive optics system. To achieve a general system that can be used in a variety of scenarios it is important to care not only for the neural network model itself but how to generate training data for such systems. To that end, this work focused on surveying and collecting the theoretical foundations based on laser propagation through the atmosphere and relevant simulation parameters necessary to construct a database that can be applied to training and validation. Also, comparison between the diffraction patterns generated by unaberrated and aberrated Gaussian beams gave us insights into observing phase distortions effects and their impact in power collection.

It was not in the scope of this work to get into the details of the neural network architecture or even discuss how the generated training data and selected parameters impact its performance. That of course is a natural follow-up from this research. Furthermore, future works also include the integration of the network output and DM control to complete the AO system and optimization to operate at the necessary bandwidth imposed by atmospheric conditions and optical communication requirements.

ACKNOWLEDGMENTS

We would like to express our gratitude to Mr. Takayuki Hosonuma, Mr. Charleston Dale Ambatali and Mr. Ryuichi Hirayama from the Intelligent Space Systems Laboratory (ISSL) in the University of Tokyo for giving us feedback and support over the course of this research. Also, many thanks to prof. Norihide Miyamura from Meisei University for the initial discussions and support to this research topic.

REFERENCES

[1] Khalighi, M. A. and Uysal, M., "Survey on free space optical communication: A communication theory perspective," *IEEE Communications Surveys & Tutorials* **16**, 2231–2258 (2014).

[2] Kaushal, H. and Kaddoum, G., "Optical communication in space: Challenges and mitigation techniques," *IEEE Communications Surveys & Tutorials* **19**, 57–96 (2017).

[3] Wang, K., Tang, M., and et al., J., "Deep learning wavefront sensing and aberration correction in atmospheric turbulence," *PhotoniX* **2:8** (2021).

[4] Liu, J., Wang, P., Zhang, X., He, Y., Zhou, X., Ye, H., Li, Y., Xu, S., Chen, S., and Fan, D., "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," *Opt. Express* **27**, 16671–16688 (2019).

[5] Majumdar, A. K. and Ricklin, J. C., [*Free-Space Laser Communications: Principles and Advances*], Springer, New York, first ed. (2008).

[6] Toyoshima, M., Kuri, T., Werner, K., Toyoda, M., Takenaka, H., Shoji, Y., Takayama, Y., Koyama, Y., Kunimori, H., Jono, T., Yamakawa, S., and Arai, K., "Overview of the laser communication system for the nict optical ground station and laser communication experiments on ground-to-satellite links," *NICT Journal* **59**, 053–075 (2012).

[7] Schmidt, J. D., [*Numerical Simulation of Optical Wave Propagation With Examples in MATLAB*], SPIE Press, Bellingham, Washington (2010).

[8] Hecht, E., [*Optics*], Pearson Education, Harlow, England (2017 (fith edition)).

[9] Willner, A., [*Optical Fiber Telecommunications VII*], Academic Press, Los Angeles, CA (2019).

[10] Guo, Y., Zhong, L., Min, L., Wang, J., Wu, Y., Chen, K., Wei, K., and Rao, C., "Adaptive optics based on machine learning: a review," *Opto-Eletronic Advances* **5**(200082) (2022).

[11] Yang, Z.-Q., hong Yang, L., Gong, L., Wang, L., and Wang, X., "Waveform distortion of gaussian beam in atmospheric turbulence simulated by phase screen method," *Journal of Mathematics* **2022** (2022).

[12] Ersoy, O. K., [*Diffraction, Fourier Optics and Imaging*], John Wiley & Sons, Inc., New Jersey (2007).