

Computational Correction of Eye Aberrations: A Physical Modeling Approach with Zernike Polynomials and Deep Learning

Kebin Contreras and Jorge Bacca

Universidad Industrial de Santander, Bucaramanga, Colombia

jbacquin@uis.edu.co

Abstract—A computational correction strategy for eye aberrations, specifically targeting astigmatism, is proposed. This method computationally designs a transformed image that allows individuals with astigmatism to perceive the original scene with improved clarity. To achieve this, a convolutional neural network (CNN) is trained to minimize the error between a well-defined reference image and an image generated by convolving the transformed image with a simulated point spread function (PSF) representative of a person with astigmatism, modeled using the second-order Zernike polynomials Z_2^2 and Z_2^{-2} . Additionally, upsampling techniques are applied to simulate zoom effects in the scene. Quantitative results demonstrate that bicubic scaling with a factor of 2 significantly enhances the perceived visual acuity of the person, increasing SSIM from 0.86 to 0.98 and PSNR from 25.8 dB to 37.7 dB under moderate distortion (severity 1.0). The model was trained on images from the KITTI dataset with simulated aberrations and evaluated on the DIV2K dataset, confirming its generalization to out-of-distribution data. These results highlight the effectiveness of integrating physics-based optical modeling with deep learning for ocular aberration correction.

Index Terms—Optical aberrations, astigmatism, Zernike polynomials, computational correction, point spread function, image restoration, physical modeling, computer vision.

I. INTRODUCTION

Visual quality in human vision is essential for everyday tasks, as even minor optical degradations can significantly affect perception and decision-making [1]. Aberrations such as defocus, coma, and astigmatism degrade the quality of visual input, distorting spatial information and reducing the ability to resolve fine details. These imperfections impact the clarity of the image seen by the eye, leading to blurry or distorted vision, which can be particularly detrimental in activities that require high precision, such as reading or driving [2–4].

To quantitatively model visual aberrations in the human eye, Zernike polynomials jointly with Fourier transforms serve as powerful mathematical tools to represent wavefront distortions in the human eye. These polynomials decompose the wavefront of the eye into a series of orthogonal basis functions over a unit disk, providing a precise description of the optical imperfections of the eye [3]. This method plays a crucial role in diagnosing vision impairments and designing corrective treatments, such as eyeglasses or contact lenses, by identifying specific aberrations in the optical system of the eye. In particular, Zernike polynomials are used to model

common eye defects, such as astigmatism, by representing the horizontal and vertical components of the aberration with second-order modes Z_2^1 and Z_2^{-1} represent horizontal and vertical astigmatism, respectively (Figure 1) [3].

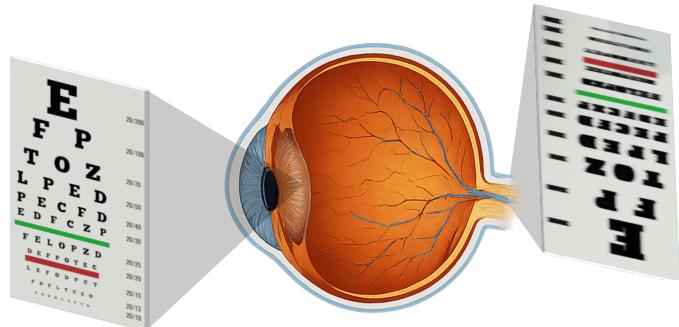


Figure 1: Schematic representation of optical degradation in a visual system caused by astigmatism. Due to the anatomical configuration of the eye, the visual scene is inverted when projected onto the retina.

Historically, the correction of optical aberrations has been achieved through the physical design of corrective optical elements, such as cylindrical lenses or aspheric surfaces tailored to specific requirements [5]. In ophthalmology, the use of these corrective lenses faces several challenges, including physical constraints related to the geometry of the eye, tear film stability, and limitations in customizing lenses for individual patients. These factors can undermine the effectiveness of purely optical corrections, particularly for patients with complex or dynamic forms of astigmatism [5]. As a result, some patients may avoid wearing corrective glasses or opt for invasive treatments such as surgery.

To address these limitations, computational strategies have been proposed that digitally modify the input image. These strategies include the use of optical filters integrated into displays or pre-compensatory transformations that anticipate the distortions introduced by the visual system [6, 7]. While these approaches remain an area of active research, they aim to adapt visual content so that, after passing through an aberrated medium such as the human eye, the perceived image is corrected and enhances visual clarity.

Recent studies have explored the application of deep learning techniques for image restoration in the presence of optical

This work was supported by MINCIENCIAS AMSUD 2023, CTO No. 112721-110-2024 under project "Computational ultrasound imaging: from theory to applications".

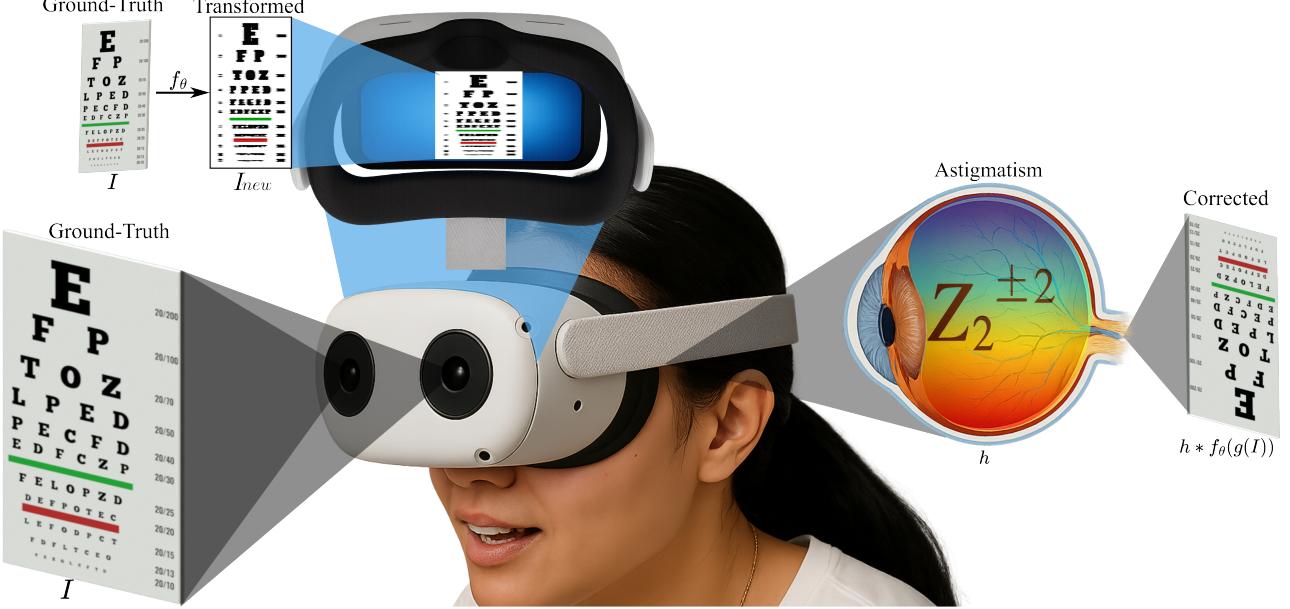


Figure 2: Illustration of a potential implementation of the proposed method for astigmatism correction. A camera captures the scene, and the image is processed by a convolutional neural network trained to generate a transformed version of the image, optimized for sharp viewing by a person with astigmatism.

distortions [8]. However, most existing methods rely solely on data-driven approaches and fail to incorporate physically interpretable models of aberrations. Specifically, there is a notable gap in research regarding astigmatism correction through wavefront representations grounded in optical theory, such as Zernike polynomials. This gap underscores the need for integrating physical modeling with computational learning techniques to develop more targeted and effective solutions for aberration correction [9].

This work proposes a computational strategy for correcting optical astigmatism by generating a transformed image that compensates for the aberration in individuals with astigmatism. The method models astigmatism by convolving the transformed image with a simulated point spread function (PSF) based on second-order Zernike polynomials, specifically Z_2^2 and Z_2^{-2} . A convolutional neural network (CNN) is then trained to minimize the error between a well-defined reference image and the image generated by convolving the transformed image, thereby learning the optimal transformation for individuals with astigmatism. The method allows for evaluating correction performance across varying severities of both horizontal and vertical astigmatism, offering an alternative when physical optical correction is not feasible. Additionally, upsampling techniques are employed to simulate zoom effects, enhancing the clarity of the reconstructed image. Quantitative results demonstrate that bicubic scaling with a factor of 2 significantly improves the perceptual quality, increasing SSIM from 0.86 to 0.98, PSNR from 25.8 dB to 37.7 dB, and reducing MSE from 2.6 to 0.2 under moderate distortion (severity 1.0). The proposed method is validated on the KITTI dataset using physically realistic degradations, highlighting the effectiveness of combining physics-based modeling with deep learning for the correction of eye aberrations.

II. MODELING OPTICAL ABERRATIONS USING ZERNIKE POLYNOMIALS

Zernike polynomials form a set of orthogonal basis function defined over the unit disk in polar coordinates (ρ, θ) [10]. Each mode $Z_n^m(\rho, \theta)$ consists of a radial polynomial $R_n^{|m|}(\rho)$ and an angular part expressed as $\cos(m\theta)$ or $\sin(|m|\theta)$, depending on the sign of m [11]. These polynomials are orthogonal with respect to the inner product weighted by ρ , and are commonly used in optical systems with circular pupils due to their completeness and ease of interpretation in terms of classical aberrations [11]. The general form of a Zernike polynomial is mathematically described as:

$$Z_n^m(\rho, \theta) = \begin{cases} R_n^{|m|}(\rho) \cos(m\theta), & m \geq 0 \\ R_n^{|m|}(\rho) \sin(|m|\theta), & m < 0 \end{cases} \quad (1)$$

where the radial polynomial is defined as:

$$R_n^{|m|}(\rho) = \sum_{k=0}^{(n-|m|)/2} (-1)^k \frac{(n-k)!}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!}$$

Second-order Zernike modes are especially important for representing defocus and astigmatism. Specifically, $Z_2^0(\rho, \theta)$ corresponds to defocus, while $Z_2^2(\rho, \theta)$ and $Z_2^{-2}(\rho, \theta)$ represent orthogonal components of astigmatism [11, 12].

A general wavefront $W(\rho, \theta)$ can be expressed as a linear combination of Zernike modes:

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_{n,m} Z_n^m(\rho, \theta), \quad (2)$$

where the coefficients $c_{n,m}$ are determined by projecting the wavefront onto each basis function using the orthonormal property [11]. Once the wavefront is known, it can be used to

model the complex pupil function, which describes how the optical system modulates the wavefront:

$$P(\rho, \theta) = A(\rho, \theta) \exp\left(i \frac{2\pi}{\lambda} W(\rho, \theta)\right), \quad (3)$$

where $A(\rho, \theta)$ represents the amplitude modulation of the light passing through the optical system, and λ is the wavelength of light [12]. This pupil function is crucial for understanding how aberrations affect the imaging process and is the key input for simulating optical systems.

To propagate the wavefront from the pupil to the image plane, Fourier optics is used [13]. Specifically, the Fourier transform of the pupil function gives the point spread function (PSF) of the system. The change from the polar coordinate system (ρ, θ) to Cartesian coordinates (x, y) is necessary for representing the spatial field in the image plane. For simplicity, in the paraxial approximation, the transformation is given by:

$$x = \rho \cos(\theta), \quad y = \rho \sin(\theta),$$

which converts the radial and angular components of the pupil function to Cartesian coordinates. This transformation is critical when modeling the propagation of light through the optical system.

The resulting PSF can be obtained by propagating the wavefront through the optical system using the Fourier transform of the complex pupil function, mathematically expressed as:

$$H(u, v) = \mathcal{F}\{P(\rho, \theta)\}, \quad (4)$$

where $H(u, v)$ represents the spatial frequency components of the PSF, and \mathcal{F} denotes the Fourier transform. To obtain the intensity distribution at the image plane, the PSF is typically normalized and expressed in Cartesian coordinates:

$$h(x, y) = |\mathcal{F}\{P(\rho, \theta)\}|^2. \quad (5)$$

The PSF is given by the squared magnitude of the Fourier transform, which describes how each point in the object space is imaged onto the image plane under the influence of the optical aberrations represented by the Zernike polynomials [14, 15].

Assuming spatially variant properties of the PSF, the observed image is modeled as the convolution of the input image with the PSF. For an input image $I(x, y)$, the final image $I_{\text{out}}(x, y)$ in the presence of optical aberrations is modeled as:

$$I_{\text{out}}(x, y) = h(x, y) * I(x, y), \quad (6)$$

where $*$ denotes the convolution operation. This convolution represents the distortion that the input image undergoes due to the optical aberrations present in the system.

The PSF can now be used to simulate the imaging of an object, allowing the assessment of image quality and the effect of various aberrations on the final image. By incorporating the optical aberrations described by the Zernike polynomials into this framework, one can model and correct various distortions in imaging systems such as the human eye or optical instruments.

III. COMPUTATIONAL IMAGE CORRECTION USING DEEP LEARNING

In the human eye, the captured image is a degraded version of the real scene due to optical aberrations, which can be modeled with the PSF. To simplify the notation, let I represent the original, and h denote the PSF modeling the aberrations. The goal of the proposed method is to find a new image, I_{new} , such that when observed by a person with astigmatism, it appears as I , i.e.,

$$I = h * I_{\text{new}}, \quad (7)$$

To achieve this, a U-Net architecture [16] is adapted, denoted as f_θ , and trained to generate a transformed image $I_{\text{new}} = f_\theta(I)$ from the original scene I , such that, when I_{new} is convolved with the eye's point spread function h , the resulting image approximates the original, i.e., $I \approx h * I_{\text{new}}$. Specifically, given a training dataset, the method minimizes the following optimization problem:

$$\theta^* \in \arg \min_{\theta} \mathbb{E}_I \|h * f_\theta(I) - I\|_2^2, \quad (8)$$

where the objective function measures the error between the distorted image $h * f_\theta(I)$ and the original image I , ensuring consistency with the degradation model of the eye during the training process [17]. This approach allows the network to learn a transformation that corrects for the aberrations caused by astigmatism.

To analyze the effect of resolution, an upsampling operator $g(\cdot)$ is applied to the images to simulate the zoom version of the scene. This operator resizes the image I before degradation.

In this scenario, the learned optimization problem that directly incorporates the imaging model is obtained as

$$\theta^* = \arg \min_{\theta} \mathbb{E}_I \|h * f_\theta(g(I)) - g(I)\|_2^2. \quad (9)$$

Here, $f_\theta(g(I))$ is the transformed image, and $h * f_\theta(g(I))$ simulates the forward model of the eye. Notice that Equation 9 generalizes Equation 8, where $g(\cdot)$ functions as the identity operator. It is important to emphasize that the effect of $g(\cdot)$ is to perform an upsampling of the image, effectively increasing the resolution of the scene. This can be interpreted as a zoomed version or, when displayed on a higher-resolution screen, as a more detailed scene. Finally, the training process is summarized

Algorithm 1: Training Procedure for PSF-Aware Image Restoration

```

Input : Training dataset  $\mathcal{D} = \{I_i\}_{i=1}^M$ , resolution  $s$ ,  

         Zernike coefficients  $c_{n,m}$ , initial parameters  $\theta$   

1 foreach batch of images  $I_i \in \mathcal{D}$  do  

2    $\tilde{I}_i \leftarrow g(I_i; s)$ ; // Upsampling  

3    $W \leftarrow \sum c_{n,m} Z_n^m$ ; // (Eq. 2)  

4    $P \leftarrow A \exp\left(i \frac{2\pi}{\lambda} W\right)$ ; // (Eq. 4)  

5    $h \leftarrow |\mathcal{F}\{P\}|^2$ ; // (Eq. 5)  

6    $\hat{I}_{\text{new}i} \leftarrow f_\theta(\tilde{I}_i)$ ; // (Eq. 7)  

7    $\mathcal{L}_i \leftarrow \|h * \hat{I}_{\text{new}i} - \tilde{I}_i\|_2^2$ ; // (Eq. 8)  

8    $\theta_i \leftarrow \theta_i - \eta \nabla_{\theta} \mathcal{L}_i$ ; // Gradient update  

Output : Optimized parameters  $\theta^*$ 

```

in Algorithm 1, where gradient updates can be performed using state-of-the-art optimization algorithms such as Adam [20].

IV. SIMULATIONS AND RESULTS

The proposed method is evaluated under four levels of horizontal astigmatism severity, defined as (0.5, 1.0, 2.0, 3.0). To this, A U-Net architecture was trained independently for each level of severity. The training process employed images at three resolution levels: the original size (256×256), upsampled using the operator $g(\cdot)$ to 512×512 via bicubic interpolation, and further upsampled to 1024×1024 . This scaling simulates different levels of visual zoom as part of the correction strategy.

Table I: Quantitative evaluation of restoration quality (SSIM, MSE, PSNR) under increasing astigmatism severity for KITTI and DIV2K datasets.

Dataset	Severity	$g(I): 256 \times 256$						$g(I): 512 \times 512$						$g(I): 1024 \times 1024$					
		Astigmatism			Corrected			Astigmatism			Corrected			Astigmatism			Corrected		
		SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR
KITTI [18]	0.5	0.79	4.93	23.16	0.98	0.42	34.51	0.93	1.32	29.0	0.99	0.01	45.2	0.98	0.42	34.3	0.99	0.05	48.1
	1.0	0.74	5.71	22.4	0.86	2.61	25.85	0.90	1.63	28.0	0.98	0.22	37.7	0.97	0.41	33.5	0.99	0.06	46.0
	2.0	0.63	8.72	20.6	0.77	5.52	22.63	0.82	3.25	25.0	0.92	1.32	28.8	0.93	0.99	30.2	0.99	0.07	39.0
	3.0	0.59	11.3	19.5	0.72	7.81	21.12	0.78	4.76	23.3	0.86	3.01	25.2	0.89	1.82	27.4	0.96	0.09	32.0
DIV2K [19]	0.5	0.82	3.73	24.9	0.92	1.12	29.18	0.95	0.44	36.2	0.97	0.33	37.8	0.98	0.22	39.0	0.99	0.13	41.3
	1.0	0.75	5.10	23.2	0.87	1.92	27.82	0.91	0.91	31.2	0.94	0.54	34.12	0.96	0.33	36.0	0.98	0.15	38.2
	2.0	0.68	7.23	21.0	0.78	3.74	24.34	0.84	2.02	27.1	0.91	1.23	29.53	0.95	0.62	31.7	0.97	0.32	34.0
	3.0	0.61	10.4	19.5	0.70	6.32	21.62	0.76	4.13	23.5	0.85	2.66	25.8	0.91	1.51	28.2	0.95	0.72	31.0



Figure 3: Visual comparison of restored images under varying astigmatism severity. The outputs of the U-Net model trained on original and bicubic-interpolated inputs are shown alongside the corresponding degraded images, highlighting progressive improvements in structural detail.

Datasets. The model was trained on the KITTI dataset [18], which contains 7481 RGB images (1242×375 pixels) of real-world urban environments captured under natural conditions. These images include diverse structural elements such as roads, vehicles, and buildings. To evaluate the generalization capacity of the trained model, inference was conducted on the DIV2K dataset [19], which comprises 1000 high-resolution natural images with a broad range of textures, illumination levels, and spatial frequencies. Unlike KITTI, DIV2K is commonly used for image super-resolution and restoration, offering a complementary distribution for out-of-domain validation.

Metrics. The performance of image restoration was evaluated using three full-reference metrics that quantify similarity between the restored image and the corresponding ground truth. The Structural Similarity Index (SSIM) measures structural consistency by comparing spatial, luminance, and contrast patterns between both images [21]. The Mean Squared Error (MSE) computes the average of squared pixel-wise differences, serving as a direct measure of reconstruction error [22]. The Peak Signal-to-Noise Ratio (PSNR), derived from MSE, expresses reconstruction quality in decibels [23]. All experiments were conducted on a workstation equipped with an AMD Ryzen 7 5700X processor at 3.40 GHz, 64 GB of RAM, and an NVIDIA GeForce RTX 4070 GPU with 12 GB of VRAM.

Table I presents a comparative evaluation of the correction framework applied to images affected by horizontal astigmatism across three resolution levels. The results indicate that resolution upsampling improves reconstruction accuracy. Scaling to 512×512 enhances metric values, and further interpolation to 1024×1024 yields additional gains. A performance drop is observed as astigmatism severity increases. These results are consistent with the growing difficulty of correcting spatial distortions introduced by higher-order wavefront components.

To assess generalization, inference was conducted on the DIV2K dataset [19]. The framework maintained consistent metric behavior across datasets and severity levels. For example, SSIM increased from 0.75 to 0.98 and PSNR from 23.2 dB to 38.2 dB at severity 1.0. These results support the transferability of the model trained on simulated optical degradations to new image distributions.

Figure 3 presents a qualitative comparison of correction performance at severity level 3.0 using samples from the DIV2K dataset. Each column shows an individual sample under different stages: aberrated input ($h * g(I)$), transformed output ($f_\theta(g(I))$), final corrected result $h * f_\theta(g(I))$, and ground-truth image ($g(I)$). The examples illustrate how the model progressively mitigates structural distortions introduced by astigmatism, with improved results observed when the input is processed through the upsampling operator $g(\cdot)$.

V. CONCLUSIONS

A computational image correction framework for eye aberrations, based on physical modeling with second-order Zernike polynomials and a deep learning strategy using a U-Net architecture, is presented. The method enables the simulation and correction of astigmatic distortions by integrating a physics-based point spread function (PSF) into the training loss, aligning the learning objective with the optical properties of the human visual system. The performance of the framework was validated across multiple image resolutions and datasets, demonstrating consistent improvements in SSIM, MSE, and PSNR metrics. For instance, under moderate astigmatism (severity 1.0), SSIM improved from 0.74 to 0.98, PSNR increased from 22.4 dB to 37.7 dB, and MSE decreased from 5.71 to 0.22 on the KITTI dataset. Bicubic upsampling of a 4-factor enhanced reconstruction quality, particularly under higher distortion levels. The approach also generalized well to unseen datasets with different visual characteristics, achieving SSIM scores up to 0.99 and PSNR values above 41 dB on DIV2K.

REFERENCES

- [1] B. Monroy, K. Contreras, and J. Bacca, “Autoregressive high-order finite difference modulo imaging: High-dynamic range for computer vision applications,” *arXiv preprint arXiv:2504.04228*, 2025.
- [2] M. Demir, S. A. Kurna, and et al., “Assessment of aberrations and visual quality differences between myopic and astigmatic eyes before and after contact lens application,” *Northern Clinics of Istanbul*, vol. 2, no. 1, pp. 1–6, 2015.
- [3] S. V. Bădescu, R. Barac, S. Schmitzer, and C. Tătaru, “The influence of optical aberrations in refractive surgery,” *Romanian Journal of Ophthalmology*, vol. 59, no. 4, pp. 217–222, 2015.
- [4] K. Contreras, B. Monroy, and J. Bacca, “High dynamic range modulo imaging for robust object detection in autonomous driving,” *arXiv preprint arXiv:2504.11472*, 2025.
- [5] M. Sundy, D. McKnight, C. Eck, and F. Rieger, “Visual acuity outcomes of toric lens implantation in patients undergoing cataract surgery at a residency training program,” *Missouri Medicine*, vol. 113, no. 1, pp. 40–43, 2016.
- [6] M. A. Jr., A. B. Barreto, and M. Adjouadi, “Digital image inverse filtering for improving visual acuity for computer users with visual aberrations,” *Inverse Problems in Science and Engineering*, vol. 16, no. 8, pp. 957–966, 2008.
- [7] F.-C. Huang and R. Raskar, “Eyeglasses-free display: towards correcting visual aberrations with computational light field displays,” *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 33, no. 4, pp. 59:1–59:12, 2014.
- [8] P. J. Schubert, R. Saxena, and J. Kornfeld, “Deepfocus: fast focus and astigmatism correction for electron microscopy,” *Nature Communications*, vol. 15, p. 948, 2024.
- [9] L. Zhang and H. Shroff, “Deep learning-based aberration compensation improves contrast and resolution in fluorescence microscopy,” *Nature Communications*, vol. 16, p. 313, 2025.
- [10] R. J. Noll, “Zernike polynomials and atmospheric turbulence,” *Journal of the Optical Society of America*, vol. 66, no. 3, pp. 207–211, 1976.
- [11] V. N. Mahajan, “Zernike circle polynomials and optical aberrations of systems with circular pupils,” *Journal of the Optical Society of America*, vol. 71, no. 1, pp. 75–85, 1981.
- [12] A. B. Watson, J. D. Yellott, and D. M. E. Cooper, “Computing human optical point spread functions,” *Journal of Vision*, vol. 15, no. 2, p. 26, 2015.
- [13] H. Arguello, S. Pinilla, Y. Peng, H. Ikoma, J. Bacca, and G. Wetzstein, “Shift-variant color-coded diffractive spectral imaging system,” *Optica*, vol. 8, no. 11, pp. 1424–1434, 2021.
- [14] A. Silva-Lora and R. Torres, “Superconical aplanatic ovoid singlet lenses,” *Journal of the Optical Society of America A*, vol. 37, no. 7, pp. 1155–1165, 2020.
- [15] A. Silva-Lora and R. A. Torres, “Explicit cartesian oval as a superconic surface for stigmatic imaging optical systems with real or virtual source or image,” *Proceedings of the Royal Society A*, vol. 476, no. 2235, p. 20190894, 2020.
- [16] S. Guan, A. A. Khan, S. Sikdar, and P. V. Chitnis, “Fully dense unet for 2-d sparse photoacoustic tomography artifact removal,” *IEEE journal of biomedical and health informatics*, vol. 24, no. 2, pp. 568–576, 2019.
- [17] J. Bacca, “Projection-based correction for enhancing deep inverse networks,” *arXiv preprint arXiv:2505.15777*, 2025.
- [18] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, pp. 3354–3361.
- [19] E. Agustsson and R. Timofte, “Ntire 2017 challenge on single image super-resolution: Dataset and study,” in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 126–135.
- [20] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [22] Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it? a new look at signal fidelity measures,” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.
- [23] Q. Huynh-Thu and M. Ghanbari, “Scope of validity of psnr in image/video quality assessment,” *Electronics Letters*, vol. 44, no. 13, pp. 800–801, 2008.