VISION CORRECTION FOR THE VISUALLY IMPAIRED VIA DIGITAL IMAGE PROCESSING

by

Onur Keleş

B.S., Electrical and Electronics Engineering, Boğaziçi University, 2016

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University
2019

ACKNOWLEDGEMENTS

I would like to thank to Prof. Dr. Emin ANARIM for providing me with a comfortable working atmosphere.

I am thankful to Op. Dr. Efekan COŞKUNSEVEN and, via him, Dünyagöz Hospitals, without whom the necessary knowledge and data about the human eye and its working principles would not be collected, correctly understood and mathematically modeled.

I am grateful to Buğra YILMAZ and Cem KARAGÖZLÜ for their effective and worthy contributions to this work.

I am thankful to Boğaziçi Zeybek Team for that I always feel their supportive presence during my undergraduate and master of science programs.

I am eternally grateful to my family for their invaluable and inexpressible support without which I cannot complete the master thesis.

The author is supported by "*TÜBİTAK-BİDEB Eğitim Burs ve Destekleri Grubu 2210A-Yurt İçi Yüksek Lisans Burs Programı*" between 2016 and 2018 during the master of science program.

This research is supported by Boğaziçi University BAP under the name of "*Sayısal Ekranların Görüntü İşleme Yardımıyla Göz Kırılma Kusurlarını Telafi Edecek Biçimde Ayarlanması*" with project number 15221P and project code 19A02P1 in 2019.

ABSTRACT

VISION CORRECTION FOR THE VISUALLY IMPAIRED VIA DIGITAL IMAGE PROCESSING

In this study, an adjustment based on the mathematical modeling on digital screens for people with refractive visual impairments is proposed. According to the laws of refraction and diffraction, the human eye can be considered as a linear and space-invariant system. In such a case, if the impulse response of the eye (point spread function) is known, the convolution law can determine how the light entering the eye (input) is shaped according to the response of the eye. When the studies using point spread function for this purpose are examined, it is seen that the point spread function of the eye is found only on the basis of the crystalline lens. In this study, the incoming distance of light, the cornea structure, accommodation, age of the person and gradient-index structure of the crystalline lens are included in the human eye model. The idea in this study is that, based on the inverse convolution (deconvolution) method, if the light entering the eye is shaped according to the inverse of the impulse response of the eye, theoretically, the patient clearly sees the object whose light comes into the eye. The biggest problem with this approach, which can easily be adapted to personal computers, portable devices and smart glasses that are expected to be developed in the future, is that the deconvolved digital image has a much higher dynamic range than the dynamic range of the standard screens, hence, the disappearance of the necessary contrast. In order to solve this problem, methods that are different from studies in the literature are proposed. In the simulations, it is seen that the newly presented methods give better visual results than the existing ones. Although the existing methods are improved with this study, it is suggested as the future work that dynamic range compression and contrast enhancement methods should be further investigated and more complex methods should be proposed.

ÖZET

GÖRME KUSURU OLANLAR İÇİN SAYISAL İMGE İŞLEME ARACILIĞIYLA GÖRÜNTÜ DÜZELTME

Bu çalışmada; kırılmaya bağlı görme kusurları olan kişiler için, sayısal ekranlarda matematiksel modellemeye dayalı uyarlama önerilmiştir. Kırılma ve kırınım yasalarına göre, insan gözü, doğrusal ve uzayda değişmez bir sistem olarak ele alınabilir. Böyle bir durumda, gözün dürtü yanıtı (nokta yayılım fonksiyonu) bilinirse, göze giren ışığın (girdinin) gözün yanıtına göre nasıl biçimleneceği evrişim yasasıyla bulunabilir. Bu amaç için nokta yayılım fonksiyonu kullanan çalışmalar incelendiğinde, gözün nokta yayılım fonksiyonunun yalnızca göz merceği esas alınarak yapıldığı görülmüştür. Bu çalışmadaysa ışığın geliş mesafesi, gözün boynuzsu tabakası, göz uyumu, kişinin yaşı ve göz merceğinin değişken indeksli yapısı da modellemelere dahil edilmiştir. Bu çalışmadaki fikir, ters evrişim yöntemine dayanarak, göze girecek ışığın gözün dürtü tepkisinin tersine göre biçimleneceği durumda, matematiksel olarak hastanın, gözüne ışığı gelen cismi net göreceğidir. Kişisel bilgisayarlar, taşınabilir aygıtlar ve ileride geliştirilmesi beklenen akıllı gözlüklere kolaylıkla uyarlanabilecek bu yaklaşımdaki en büyük sorun, ters evrişim uygulanan sayısal imgenin, standart ekranların dinamik aralığından çok daha yüksek bir dinamik aralığa sahip olması ve sonucunda asıl görüntüde elde edilen karşıtlığın yok olmasıdır. Bunun çözümü için, alanyazında önerilen yöntemlerden farklı olan yöntemler öne sürülmüştür. Yapılan benzetimlerde, yeni sunulan yöntemlerin, var olanlardan görsel olarak daha iyi sonuç verdiği görülmüştür. Bu çalışmayla var olan yöntemler her ne kadar geliştirilmiş olsa da ileride yapılması gereken iş olarak da dinamik aralık sıkıştırma ve karşıtlık geliştirme yöntemlerinin daha çok incelenmesi ve daha karmaşık yöntemlerin ileri sürülmesi önerilmiştir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES xii
LIST OF SYMBOLS xiii
LIST OF ACRONYMS/ABBREVIATIONS xxiii
1. INTRODUCTION 1
2. WAVE THEORY FOR IMAGE FORMATION SYSTEMS 4
2.1. Huygens and Fermat's Principles 4
2.2. Reflection 6
2.3. Refraction 9
2.3.1. Index of Refraction 13
2.3.2. Dispersion 14
2.3.3. Refraction on Spherical Surfaces 17
2.3.4. Thick and Thin Lenses 19
2.4. Diffraction 23
2.4.1. Huygens–Fresnel Principle 24
2.4.2. Diffraction on an Aperture 28
2.4.3. Fresnel Approximation 29
2.4.4. Fraunhofer Approximation 34
2.5. Light Passing Through an Aperture with Lens 41
2.6. Image Formation Passing Through an Aperture with a Lens 46
2.6.1. Coherent and Incoherent Illumination 51
2.6.2. Optical Transfer Function and Modulation Transfer Function . . 53
3. HUMAN VISION SYSTEM 54
3.1. Human Eye Anatomy 55
3.2. Image Formation in Human Eye 59
4. RELATED WORK 65

4.1. Detail Enhancement Methods 65
4.1.1. Adaptive Filtering 65
4.1.2. Wideband Enhancement 68
4.1.3. Histogram Hyperbolization 70
4.1.4. Unsharp Masking and Extremum Sharpening 71
4.1.5. Compression-based Enhancement 73
4.1.6. Object or Edge Stressing 80
4.2. Deconvolutional Methods 89
4.2.1. Using Visual Degradation Transfer Function 90
4.2.2. Using Wavefront Aberration Function 91
5. METHODOLOGY 94
5.1. Zernike Polynomials 94
5.2. Wavefront Aberration Function 100
5.3. Ray Tracing 101
5.4. Point Spread Function 103
5.5. Mathematical Model of Emmetropic and Ametropic Eye 106
5.6. The Factors That Changes Point Spread Function 109
5.6.1. Different Pupil Diameter 110
5.6.2. Different Wavelength 118
5.6.3. Effect of Age and Accommodation 121
5.7. Deconvolution 123
5.8. Contrast Preservation 124
6. COMPARISON OF RESULTS 137
7. CONCLUSION 143
REFERENCES 152

LIST OF FIGURES

Figure 2.1. Huygens Principle. 5

Figure 2.2. Derivation of law of reflection. 7

Figure 2.3. Refraction of waves. 9

Figure 2.4. Derivation of law of refraction. 10

Figure 2.5. Wavelength–Refractive index relationship for silicon. 16

Figure 2.6. The geometry for refraction on spherical surfaces. 17

Figure 2.7. Derivation of lens equation. 20

Figure 2.8. Huygens–Fresnel Principle. 24

Figure 2.9. The geometry for Fresnel's derivation. 26

Figure 2.10. Diffraction effect on a point behind the aperture on an opaque screen. 29

Figure 2.11. The Fresnel integrals. 33

Figure 2.12. Fraunhofer diffraction of an electric field from a rectangular aperture. 37

Figure 2.13. The geometry for Fraunhofer diffraction on a circular aperture. . . 37

Figure 2.14. The Bessel functions of the first kind of order $m = 0, 1, ..., 4$. . . . 39

Figure 2.15. Fraunhofer diffraction of an electric field from a circular aperture. 40

Figure 2.16. The lens thickness geometry. 42

Figure 2.17. The geometry for image formation via an aperture with a lens. . . 47

Figure 3.1. Human eye anatomy. 56

Figure 3.2. Human photoreceptor cell absorbency according to the wavelengths. 63

Figure 4.1. Example of adaptive filtering. 67

Figure 4.2. Flowchart of the wideband enhancement. 70

Figure 4.3. Example of histogram hyperbolization: 71

Figure 4.4. Examples of unsharp masking and extremum sharpening. 73

Figure 4.5. JPEG DCT basis functions and JPEG zigzag ordering. 75

Figure 4.6. The block diagram of JPEG encoding and decoding. 76

Figure 4.7. Example of edge enhancement using scene simplification. 83

Figure 4.8. Example of the TRON algorithm. 85

Figure 4.9. Examples of cartoonization algorithm. 88

Figure 4.10. A representation of sine wave grating pattern. 90

Figure 4.11. Result of edge tapering. 92

Figure 4.12. Single-sided contrast mapping. 93

Figure 5.1. Standard reference axis and positive direction for reporting eye aberrations. 96

Figure 5.2. First 45 Zernike polynomials. 99

Figure 5.3. Optical path difference. 100

Figure 5.4. Ray tracing working principle. 102

Figure 5.5. Refraction in gradient-index (GRIN) lens. 104

Figure 5.6. Mathematical example of an emmetropic eye. 107

Figure 5.7. Mathematical example of an ametropic eye. 108

Figure 5.8. Representation difference of the same aperture for different aperture radii. 111

Figure 5.9. Ordering schematic of polynomials and their parameters in calculations of new coefficients for different pupil size. 112

Figure 5.10. Effect of the different wavelength on point spread function. 119

Figure 5.11. Wavefront aberration function and the corresponding PSF graphs for different wavelength values. 120

Figure 5.12. Age versus diopters plot. 122

Figure 5.13. Result of deconvolution operation. 125

Figure 5.14. Result of dynamic range scaling. 126

Figure 5.15. Results of histogram equalization, histogram hyperbolization and brightness preserving bi-histogram equalization (BBHE) using edge information. 127

Figure 5.16. Result of single-sided contrast enhancement. 128

Figure 5.17. Profile and result of double-sided contrast enhancement. 130

Figure 5.18. Sigmoid function. 131

Figure 5.19. Result of sigmoid mapping. 132

Figure 5.20. Result of gradient domain HDR compression. 134

Figure 5.21. iCAM06 method flowchart. 134

Figure 5.22. Result of iCAM06-based HDR compression. 135

Figure 6.1. More results of the methods. 141

Figure 7.1. PSF graphs for different focused distances. 145

Figure 7.2. GRIN change with age. 145

Figure 7.3. Different point spread functions. 147

Figure 7.4. Cornea and lens result versus lens only result. 147

LIST OF TABLES

Table 6.1. The results for contrast measurement for each method. 139

Table 6.2. The results for contrast measurement for each method for Figure 6.1.142

Table 7.1. The sum of Gaussians fit parameters for each cell type. 150

LIST OF SYMBOLS

a	Person's age
A	Source point
A_o	Wave amplitude
$\text{asf}(y_s, z_s;\, y_o, z_o)$	Amplitude spread function
B	Image point
B_c	Average band amplitude with total c cycles/block in matrix $\mathbf{D}$
$\hat{B}_c$	Enhanced band with total c cycles/block
c	Constant for histogram hyperbolization *or* middle value of sigmoid function
$\mathbf{c}$	Zernike polynomial coefficients column vector
$\mathbf{c'}$	Zernike polynomial coefficients column vector for different pupil diameter
c_k	Coefficient of k^{th} Zernike polynomial
C	Center of curvature of medium boundary
$\mathbf{C}$	Transformation matrix for new coefficients given the old coefficients and radius
C_1	Center of curvature of lens front surface
C_2	Center of curvature of lens back surface
C_c	Contrast measure in band with total c cycles/block
$\hat{C}_c$	Contrast measure in enhanced band with total c cycles/block
$C_i(u, v)$	Bandpass filter of i^{th} scale in frequency domain
$C(u)$	Normalization constant for frequency coordinate u
$C(v)$	Normalization constant for frequency coordinate v
$C(x)$	Fresnel cosine integral
cmn	Contrast measure number
d	Lens thickness
D	Diopter
$\mathbf{D}$	DCT coefficient matrix

d_{acc}	Minimum accommodation distance
d_c	Thickness of cornea
d_{mn}	Coefficient for m cycles/block horizontal frequency and n cycles/block vertical frequency in matrix $\mathbf{D}$
$\hat{d}_{mn}$	Enhanced coefficient for m cycles/block horizontal frequency and n cycles/block vertical frequency in matrix $\mathbf{D}$
d_0	Maximum lens thickness
d_1	Maximum thickness of the first curved surface of a lens
d_2	Thickness of the flat part of the lens
d_3	Maximum thickness of the second curved surface of a lens
$\mathbf{D}_d$	Directional difference along direction d
$d(y_a, z_a)$	Coordinate-dependent lens thickness
$\mathrm{div}(\cdot)$	Divergence operator
$\mathrm{DoG}(x, y)$	Difference of Gaussians image
$DSC(l)$	Double-sided contrast enhancement
E	Energy of a photon
f	Frequency of the photon *or* focal distance
$f_{ssc}(l)$	Non-linear part of single-sided contrast enhancement
f_{y_a}	Fourier transform evaluation frequency
f_{z_a}	Fourier transform evaluation frequency
f_1	The focal distance of the first lens
f_2	The focal distance of the second lens
f_k	Strength factor of each frequency
$f(x; \alpha, c)$	Sigmoid function
$f(x)$	Converter function from CIEXYZ to CIELAB
$f^{-1}(x)$	Converter function from CIELAB to CIEXYZ
F	Focal point
$\mathbf{F}$	Filtering matrix for contrast enhancement
$[\mathbf{F}]_{mn}$	m^{th} row and n^{th} column of matrix $\mathbf{F}$
$\mathcal{F}\{\cdot\}$	Fourier transform operation
$\mathcal{F}^{-1}\{\cdot\}$	Inverse Fourier transform operation
FI	Feature image

$g(\cdot)$	Function for diffusion (conduction) coefficient
$\mathbf{G}$	Gradient magnitude matrix
$G(k,l)$	Gaussian window
$G_c(x,y)$	Central Gaussian in cmn calculation
$G_s(x,y)$	Surrounding Gaussian in cmn calculation
h	Total vertical distance from source to the image point *or* Planck constant
h_1	Vertical distance from light source to surface touch point
h_2	Vertical distance from surface touch point to image point
h_z	Rectangular aperture height
H_1	q-recursive method coefficient
H_2	q-recursive method coefficient
H_3	q-recursive method coefficient
$H(I_L)$	High-pass amplifier function of I_L
i	Scale number *or* single index for Zernike polynomial
I	Image
I_1	Incident light ray
I_2	Reflected light ray
$I_{dc}(x,y)$	Deconvolved image
I_H	High frequency content of image I
I_i	Image I in scale i
$I_i(y_s,z_s)$	Image intensity
I_L	Low frequency content of image I
$I_o(y_o,z_o)$	Object intensity
I_x	First derivative of image I along direction x
I_{xx}	Second derivative of image I along direction x
I_{xxx}	Third derivative of image I along direction x
I'_H	Modified high frequency content
I'_L	Modified low frequency content
$I(u,v)$	Image in discrete cosine transform domain
$I(x,y;t)$	Pixel value of image I at (x,y) in time t
$I(x,y)$	Pixel value at (x,y) of image I

$I^{(t)}(x, y)$	Pixel value of image I at (x, y) in time t
j	Imaginary number
$J(I)$	Transformation for histogram hyperbolization of image I
$J_m(x)$	Bessel function of the first kind of order m
k	Wave number
K	Threshold value in TRON algorithm *or* noise canceler variable in deconvolution
$K(\theta)$	Obliquity factor
l	Pixel value
L	Total path length
$L(I_L)$	Lowpass amplifier function of I_L
m	Changed variable *or* Bessel function order *or* minimum value of the new image after single-sided enhancement *or* azimuthal order of Zernike polynomial
m_e	Mass of an electron
m'	Dummy variable in Fresnel integrals
M	System magnification
$\mathbf{M}$	Zernike polynomial azimuthal matrix
$[\mathbf{M}(m)]$	Block matrix of $\mathbf{M}$
$M^m(\theta)$	Azimuthal component of Zernike polynomial with azimuthal order m
n	Changed variable *or* radial order of Zernike polynomial
N	Surface normal *or* number of charges per unit volume *or* number of bins used when determining the quantization step size Δq
$\mathbf{N}$	Zernike polynomial normalization matrix
n_1	Refractive index of medium 1
n_2	Refractive index of medium 2
n_c	Refractive index of cornea
n_d	Refractive index difference between center and periphery of the eye lens
n_l	Refractive index of lens *or* number of levels in cmn calculation

n_m	Refractive index of lens surrounding medium
N_n^m	Normalization component of Zernike polynomial with radial order n and azimuthal order m
n_o	Refractive index of middle of eye lens
n_{pl}	Number of pixels in level l
n'	Dummy variable in Fresnel integrals
O	Object point
O'	Image point
$\overleftrightarrow{OO'}$	Optical axis
$\mathrm{OPD}(y, z)$	Optical path difference
$\mathrm{OPL}(y_a, z_a)$	Optical path length a ray takes until the lens
$\mathrm{OTF}(Y_s, Z_s)$	Optical transfer function
p	Distance from object to the medium boundary *or* power determining the exponential decay in GRIN lens profile equation
P	Surface touch point
$\mathbf{P}$	Permutation matrix for ordering the Zernike polynomial parameters
$\mathrm{P}(y_a, z_a)$	Pupil window function
$\mathcal{P}(y_a, z_a)$	Generalized pupil function
p_1	Distance between object and the lens front surface
p_2	Distance between the lens back surface and image created by the lens front surface
$p_a(a)$	Age-dependent axial power value of GRIN lens profile
$P_a(y_a, z_a)$	Point on the aperture
p_c	Center pixel
$\mathbf{P}_d$	Prewitt edge operator along direction d
$p_e(a)$	Age-dependent equatorial power value of GRIN lens profile
P_o	Object point
$P_s(y_s, z_s)$	Image point
p_s	Surrounding pixel of a center pixel p_c
$p(I)$	Empirical probability density function of pixel brightness
$\mathrm{psf}(y_s, z_s)$	Point spread function

$\mathrm{PSF}(Y_s, Z_s)$	Fourier transform of point spread function
q	Distance from image to the medium boundary
$\mathbf{Q}$	Quantization matrix in DCT
$\hat{\mathbf{Q}}$	Quantization matrix for contrast enhancement
q_1	Distance between lens front surface and image created by the lens front surface
q_2	Distance between lens back surface and image created by the lens back surface
q_e	Charge on an electron
Q_n	Nearest quantized channel value
QL^*	Quantized L^* channel
r	Radius of curvature
$\mathbf{R}$	Zernike polynomial radial matrix
$[\mathbf{R}(m)]$	Block matrix of $\mathbf{R}$
r_1	Radius of curvature of lens front surface
r_2	Radius of curvature of lens back surface
r_a	Circular aperture radius
r_{as}	Distance from image point P_s to the aperture point P_a
r_c	Size for central Gaussian in cmn calculation
r_f	Radial component of cylindrical coordinates in frequency domain
$R_n^m(\rho)$	Radial component of Zernike polynomial with radial order n and azimuthal order m
r_o	Distance between point wave source and primary wave *or* perpendicular distance from object to the aperture
r_{oa}	Distance from object to the aperture point P_a
r_s	Distance between the differential surface element on primary wave and the point on which the superposition of waves is calculated *or* perpendicular distance from aperture to the image point *or* size for surrounding Gaussian in cmn calculation
r'_a	New pupil radius
s	Horizontal distance from light source to image point
S	Point on which the superposition of waves is calculated

s_1	Horizontal distance from light source to image point
s_2	Horizontal distance from surface touch point to image point
S_{A}	Aperture area
$\mathbf{S}_d$	Sobel edge operator along direction d
S_{w}	Surface area of the primary wave
$S(x)$	Fresnel sine integral
$SSC(l)$	Single-sided contrast enhancement
t	Threshold value for single-sided and double-sided contrast enhancement
$\mathbf{T_S}$	Sharpness transition matrix for cartoonization algorithm
t_t	Total travel time
th_i	Threshold value in scale i
TI_i	Thresholded image in scale i
TrI	Resulting image of TRON algorithm
ts_m	Minimum value for sharpness transition matrix
ts_m	Maximum value for sharpness transition matrix
u	Spatial frequency Cartesian coordinate
U_{A}	Constant illumination on lens in an aperture
$U_{\mathrm{A}}(y_a, z_a)$	Illumination on lens in an aperture
$U_{\mathrm{A}_0}(y_a, z_a)$	Illumination amplitude on lens in an aperture
$U_{al}(y_a, z_a)$	Wave expression after lens
$U_{bl}(y_a, z_a)$	Wave expression before lens
$U_i(y_s, z_s)$	Image
$U_o(y_o, z_o)$	Object
U_{P_s}	Wave contribution on point P_s
U_S	Total wave contribution on point S
$u(r_o, t)$	Wave expression in space–time
$U(r_o)$	Wave expression in space
v	Spatial frequency Cartesian coordinate
v_1	Average speed of light ray in medium 1
v_2	Average speed of light ray in medium 2
w	Weight fit for eye cell response

$\mathbf{W}$	Image weighting matrix
$W_A(y,z)$	Aberrated wavefront
W_D	Denominator of lowpass filter
$W_N(y,z)$	Undisturbed wavefront
W_N	Numerator of lowpass filter
w_y	Rectangular aperture width
x	Spatial Cartesian coordinate
$\mathbf{X}$	Horizontal gradient matrix
y	Spatial Cartesian coordinate
$\mathbf{Y}$	Vertical gradient matrix
(y_a, z_a)	Aperture point coordinates
$(\tilde{y_a}, \tilde{z_a})$	Normalized aperture coordinates
(y_o, z_o)	Object coordinates
$(\tilde{y_o}, \tilde{z_o})$	Normalized object coordinates
(y_s, z_s)	Image point coordinates
(Y_s, Z_s)	Spatial frequency coordinates
$\mathbf{Z}$	Zernike polynomial row vector
$Z_n^m(\rho,\theta)$	Zernike polynomial with radial order n and azimuthal order m
α	Angle between optical axis and object-touch point line *or* aperture angular coordinate *or* contrast enhancement factor *or* steepness parameter of sigmoid function
β	Angle between optical axis and center-touch point line *or* image point angular coordinate *or* steepness parameter *or* contrast enhancement factor
γ	Dissipation constant *or* angle between optical axis and image-touch point line *or* diffusion (conduction) coefficient
γ_d	Diffusion (conduction) coefficient along direction d
$\gamma_d^{(t)}(x,y)$	Diffusion (conduction) coefficients in direction d changing with time t
$\gamma(x,y;t)$	Position- and time-dependent diffusion (conduction) constant
$\delta(\cdot)$	Dirac delta function

Δ	Laplacian operator
Δq	Step size of quantization
$\Delta W(y, z)$	Wavefront aberration function
$\Delta W_c(y_a, z_a)$	Wavefront aberration function of cornea
$\Delta W_l(y_a, z_a)$	Wavefront aberration function of eye lens
ϵ_0	Permittivity of vacuum
ε_A	Source strength per unit area
η	Ratio between old and new pupil diameter
$\boldsymbol{\eta}$	Matrix of ratio between old and new pupil diameter
θ	Angle between primary wave surface normal and point on which the superposition of waves is calculated *or* angle between r_{oa} and r_{as} *or* angle of polar coordinate for Zernike polynomial
θ_i	Angle between incident light ray and surface normal
θ_{rl}	Angle between reflected light ray and surface normal
θ_{rr}	Angle between refracted light and surface normal
κ	Complex constant *or* constant for anisotropic diffusion filter
λ	Wavelength of the photon
ρ	Normalized polar coordinate for Zernike polynomials
$\boldsymbol{\rho}$	Zernike polynomial radial component row vector
$\boldsymbol{\rho'}$	Zernike polynomial radial component row vector for different pupil diameter
ρ_a	Aperture radial coordinate
ρ_s	Image point radial coordinate
ς	Stabilizing constant
τ_m	Lower threshold for cartoonization algorithm
τ_M	Higher threshold for cartoonization algorithm
$\phi(y_a, z_a)$	Illumination phase on lens in an aperture
$\phi(I_x)$	Function of I_x, introduced for ease of notation
$\phi'(I_x)$	First derivative of $\phi(I_x)$
$\phi''(I_x)$	Second derivative of $\phi(I_x)$
ω	Angular frequency of the light

ω_0	Resonant frequency of an atom
∇	Gradient operator
$\bigtriangledown_d I$	Directional pixel difference of image I along direction d

LIST OF ACRONYMS/ABBREVIATIONS

asf	Amplitude Spread Function
BBHE	Brightness preserving Bi-Histogram Equalization
DCT	Discrete Cosine Transform
DoG	Difference of Gaussians
GRIN	Gradient indexed
HDR	High Dynamic Range
iCAM	Image Color Appearance Model
JPEG	Joint Photographic Experts Group
MPEG	Moving Picture Experts Group
OPD	Optical Path Difference
OPL	Optical Path Length
otf	Optical Transfer Function
psf	Point Spread Function
PSF	Point Spread Function
RGB	Red-Green-Blue color space
TRON	Tinted Reduced Outlined Nature
varDoG	Variance of difference of Gaussians image
YCbCr	Luma-Blue difference-Red difference color space

1. INTRODUCTION

Human visual system is the strongest and most important connection with outside world. Due to the importance and the fact that vision is the our primary source for information, its sensory organ, eye, is the most complicated one among other sensor organs. This vitality requires special care for eye, and a problem occurring in the eye causes negative impact on our daily lives.

Generally, visual deficiencies happen in the form of blurred vision, decrease in incoming light, color blindness, field of vision alteration and loss of vision. Among them, the blurred vision is the most common. It is encountered so common that the people tried, and achieved, to correct the vision for centuries [1]. Correction of blur requires re-direction of the light, which is known thanks to the knowledge of refraction of light. For this purpose, light-refracting lenses are traditionally used. As the time passes, the techniques for production of lenses become more developed, too. In time, it is understood that the causes and types of blurred vision may be different, and lenses are started to be made according to those differences. As the technology advanced, the treatment for visual impairments diversified. Together with eye glasses, contact lenses are started to be used. Moreover, for some certain type of visual deficiencies, medical operations become a device-free solution for patients.

Although the techniques to correct blurred vision advances together with technology, it is understood that some visual impairments do not positively react to the treatment. For instance, the lenses can be manufactured for low-degree impairments like nearsightedness, farsightedness, astigmatism and presbyopia. But the current technology does not allow to produce lenses for higher-ordered deficiencies such as trefoil and secondary coma. Although they can be treated by medical surgeries, it is common that the operations either cannot fully eliminate the impairment or introduce new ones. Also, some impairments, such as keratoconus, occur in cornea instead of crystalline lens and their degree, by nature, can be very high to produce a lens. Therefore, where medical methods are insufficient, technology can offer solution.

Digital image processing is a widely used area from security to aviation, from film industry to medicine. As a solution for the medical problem of vision correction, digital image processing was first proposed by Peli and Lim [2]. They suggested that by measuring and obtaining the contrast loss or degradation of people suffering a refractive error, the images can be pre-compensated according to that loss. This measurement was done by a sine wave grating pattern. For each frequency and contrast degree, there is a point where the pattern cannot be distinguished. By plotting the contrast and frequency data, a function is obtained. This sensitivity function, then, can be used to enhance images on television or on newspapers for people with visual deficiencies.

After that first study, the refractive error correction field via digital image processing started to draw attention. The following studies began to look new methods to enhance the visual detail since what the refractive error in eye does is actually decrease the received detail. When the study [2] was published, the method to determine the visual acuity loss was the sine wave grating. But this method does not specify the visual impairment type since two people may have the same sensitivity but with different types of impairments, hence it is not deficiency-specific method. Following the development of Hartmann-Shack wavefront sensor, however, the vision deficiencies started to be correctly and specifically found. After that, the image processing became more responsive to compensate visual impairments on images. The first work was [3] and then the study became the main approach to the problem.

For the refractive error compensation on digital screens, the same approach in [3] is used. Basically, the transfer function of eye is found via a device giving the necessary data to calculate it, and its inverse is applied to the images or digital screens. Theoretically and mathematically, the person with visual deficiency should be able to get a clear image which is normally seen by a healthy eye. However, in reality and practically, there are some limitations for a clear vision. The main problem is the dynamic range of a standard digital screens and commercial printers. Being this dynamic range very small compared to the dynamic range of images that the method produces causes the contrast loss in the image.

In this thesis, new solutions that have not been suggested before for this problem are proposed. Moreover, an update to the modeling of human eye that the other calculations do not take into account for this task is presented. The thesis is organized as following: Chapter 2 gives the theory that is necessary to understand and develop the method. Chapter 3 includes the anatomy of human vision system as well as the "seeing"and the necessary information for eye modeling. Chapter 4 widely summarizes the methods that have been proposed and tried for solving the problem in literature. Chapter 5 explains the method that is followed in this work as well as the new solutions and the new modeling of human eye for the problem. In Chapter 6, the results obtained in this work are evaluated and interpreted. Chapter 7 concludes the thesis and draws a frame for future work.

2. WAVE THEORY FOR IMAGE FORMATION SYSTEMS

Light is a type of electromagnetic radiation which is the fundamental component to carry, transmit and receive the visual information. Many daily and optical phenomena are the consequences of light and its nature. Different colors that we encounter in every day are the consequences of light with different frequencies. The different colors of sky during a day are observed thanks to *scattering* property of the light. We see ourselves in a mirror or on the surface of a reflective material due to *reflection* of light. Rainbows cannot be seen without the *dispersion* property of light. The eye glasses or lenses that we use to clarify our distorted vision rely on *refraction* of the light, and the vision in our eyes cannot be formed without *diffraction* property of the light. All those phenomena are explained by wave nature of the light which has a wave-matter dual nature.

For explaining the propagation behaviors of the light, ray optics, which is also referred as geometric optics, is used. In this approach, the light rays are assumed to move in a straight line while passing through an optically uniform medium and they do not change direction unless either they enter a different (non)uniform medium or a nonuniformity in the optical properties of the medium occurs in space or time. Also, this approach assumes that those light rays are perpendicular to the wavefront of the light wave.

To be able to understand the image formation and its dynamics in human eye, phenomena that play role in image formation such as reflection, refraction, dispersion and diffraction should be analyzed and modeled.

2.1. Huygens and Fermat's Principles

In order to explain the aforementioned phenomena, two important principles, namely, Huygens Principle and Fermat's Principle, are used.

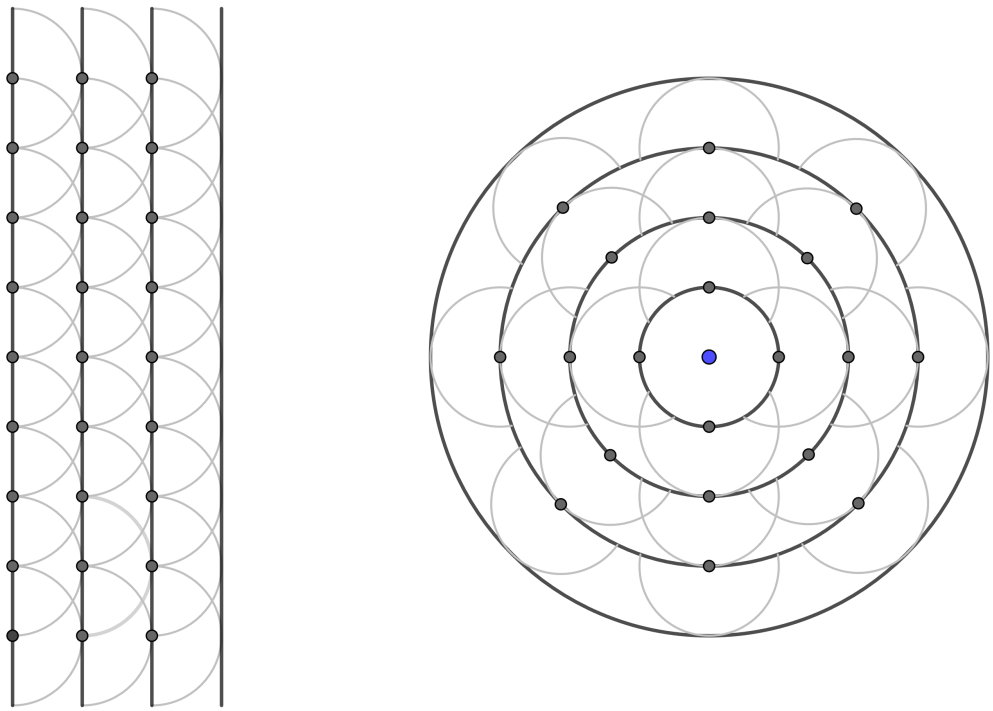


Figure 2.1. Huygens Principle. A plane wave (left) and a spherical wave (right) source are shown. Note that the wavelets (light gray waves) do not spread behind the waves that their sources are on to.

In 1678, the Dutch physicist Christian Huygens wrote a treatise. In that treatise, he stated that every point on a wavefront behaves as a point source emitting spherical wavelets in the same frequency with the wavefront, and sum of these secondary wavelets form the next wavefront with the same frequency in an optically uniform medium. He assumed that those wavelets always move forward; in other words, they do not spread behind the wave on which their point sources sit; but he did not provide a mathematical basis for this statement. A pictorial representation of Huygens Principle is shown in Figure 2.1. Note that on a wavefront, there are infinitely many point sources that emits spherical waves, but only some of them are shown in figure for visualization purposes.

Another simple yet important principle is the Fermat's Principle. Also known as the principle of least time, it states that a light ray traveling between two points

chooses the path that it goes fastest.

Although this principle is stated by Euclid, Ptolemy and Ibn al-Haytham long before Pierre de Fermat [4], Fermat explained the generalized principle in modern form. Moreover, it can be derived from other principles such as Huygens Principle. Suppose that a light ray on a wavefront goes from one point to another via all possible routes without the limitations of geometric optics. Since the path lengths exceed one wavelength, the arriving rays have a wide range of different phases. This high variety in phases tends to cause the interference to be destructive. However, if the shortest path is followed, then all the rays traveling on this route and near to this route have no or very little phase difference, enabling the interference to be constructive. Hence, the light is seen to travel this fastest path.

These two important principles are the keys to explain the image formation phenomena such as reflection, refraction and diffraction. By using these, image formation and its dynamics can be modeled properly.

2.2. Reflection

Reflection is one of the basic yet most important features of the light. It is the fundamental property of light for vision acquisition. When the light encounters with an object that is not a light source itself, it reflects back from the object and the reflected light rays come to our eyes. The same principle is valid for mirrors. In mirror, we see the objects, the reflected light of which also reflects from the mirror. That means we see objects via the light that is reflected from the objects. Without reflection, we would not be able to see the objects that do not emit light by itself.

When light rays traveling in a medium encounter the boundary of another medium, part of the parallel incident light rays is reflected. If the boundary, or the reflective surface, is a smooth surface; i.e. the variations on the surface are very small, then all the reflected light rays are parallel to each other. This type of reflection is called *specular reflection*. When the variations on the surface is not small, then the reflected

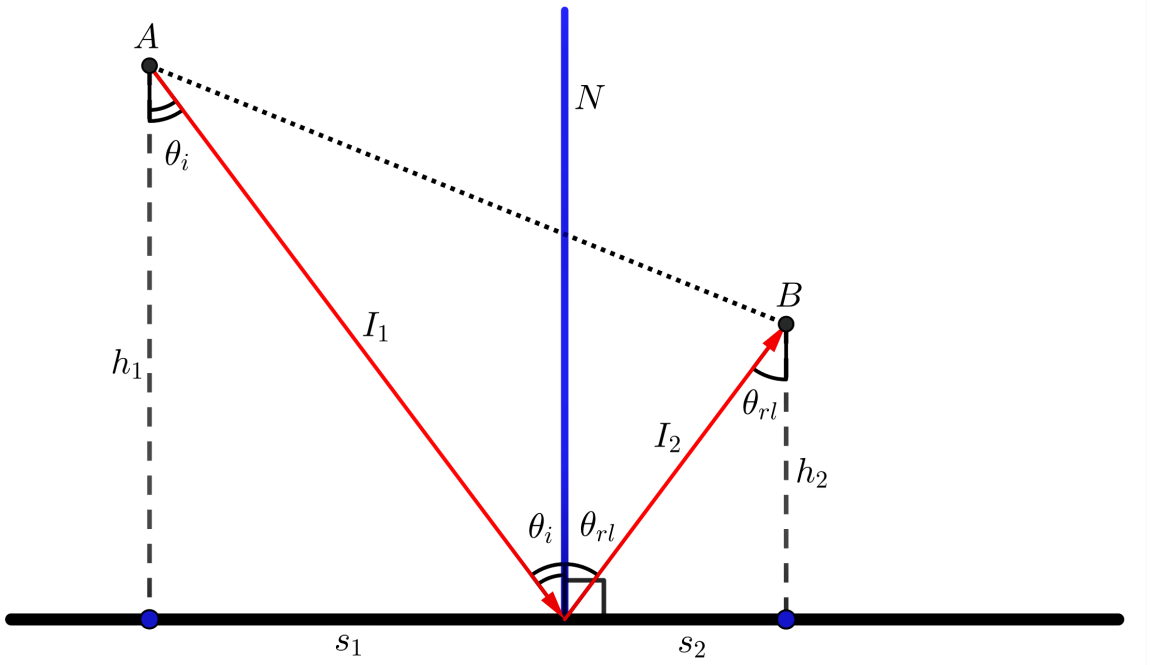


Figure 2.2. Derivation of law of reflection.

light rays are not parallel to each other. This reflection type is called *diffuse reflection*. After here in this work, only the specular reflection is considered and meant when the term "reflection"is used for its simplicity and generality.

Fermat's Principle can be applied to reflection to find a general rule. Suppose that a light ray travels from point A to point B in an optically uniform medium as in Figure 2.2. The shortest path between those two points are the direct path indicated with the dashed line connecting A and B. But suppose that the light ray makes single reflection and by that reflection does it arrive to point B. Suppose that the incident light I_1 makes an angle θ_i with the surface normal N and the reflected light I_2 makes θ_{rl} angle. The total length of path taken by rays I_1 and I_2, L, is

$$L = \sqrt{h_1^2 + s_1^2} + \sqrt{h_2^2 + s_2^2}$$

Let $s = s_1 + s_2$. Then

$$L = \sqrt{h_1^2 + s_1^2} + \sqrt{h_2^2 + (s - s_1)^2}$$

Since the medium is optically uniform, the light ray does not change its speed. Therefore, the fastest path is the shortest path. To find an expression for the shortest path, if we take the derivative of L with respect to s_1 and make it equal to 0,

$$\begin{aligned} \frac{dL}{ds_1} &= \frac{2s_1}{2\sqrt{h_1^2 + s_1^2}} + \frac{-2(s - s_1)}{2\sqrt{h_2^2 + (s - s_1)^2}} \\ &= \frac{s_1}{\sqrt{h_1^2 + s_1^2}} + \frac{-(s - s_1)}{\sqrt{h_2^2 + (s - s_1)^2}} \\ &= \frac{s_1}{\sqrt{h_1^2 + s_1^2}} + \frac{-s_2}{\sqrt{h_2^2 + s_2^2}} = 0 \end{aligned} \tag{2.1}$$

From Eq. (2.1), it can be written that

$$\frac{s_1}{\sqrt{h_1^2 + s_1^2}} = \frac{s_2}{\sqrt{h_2^2 + s_2^2}} \tag{2.2}$$

When looked at the Figure 2.2, it is seen that the left hand side of Eq. (2.2) is $\sin\theta_i$. Similarly, the right hand side of Eq. (2.2) is $\sin\theta_{rl}$. There are two cases in which sine values of two angles are equal. First one is $\theta_i = \pi - \theta_{rl}$. But if one of the angles is bigger than $\pi/2$, then it means that one of the light rays has changed medium, which is not the case. Then, the second option becomes the solution, which is

$$\boxed{\theta_i = \theta_{rl}} \tag{2.3}$$

The Eq. (2.3) is known as *the law of reflection*, which simply states that any incident ray coming to a surface with an angle with respect to the surface normal is reflected with the same angle to the surface normal.

2.3. Refraction

Refraction can be described as bowing of waves (light, in this case) when they are passing from one medium to another in which their speed is different. When light encounters with a medium which is not opaque and has different density than the medium it leaves, it tries to continue its way. If the light does not enter the new medium with a 90° angle between itself and the surface, then some portion of the wavefront experiences new medium before other parts. That portion changes speed and this change in speed causes the wavefront to alter direction. The effect can be visualized as in Figure 2.3.

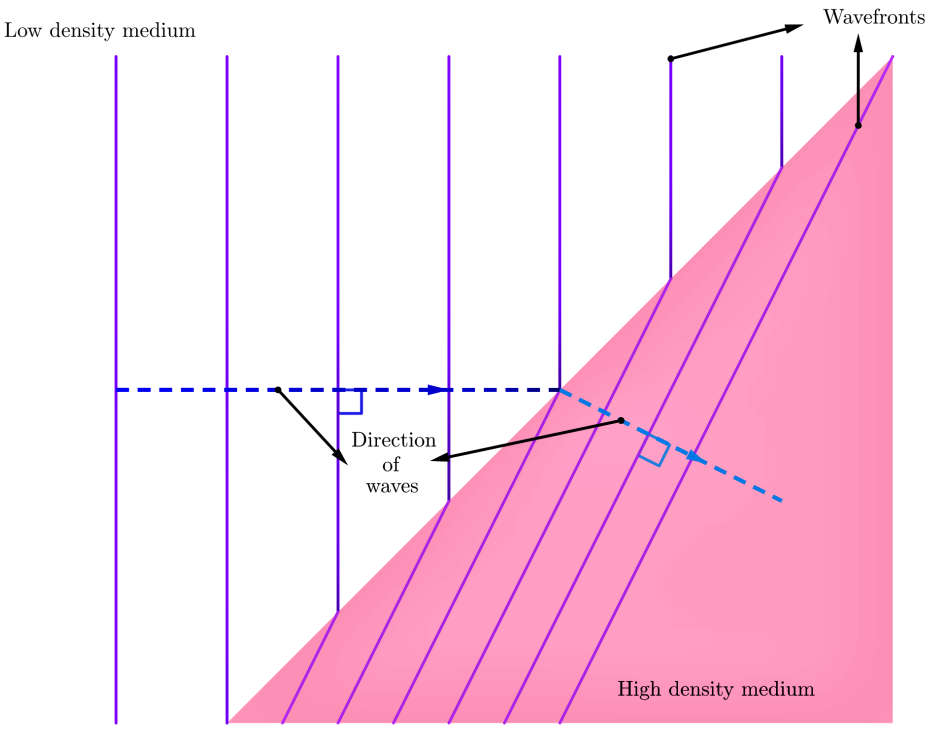


Figure 2.3. Refraction of waves.

Refraction occurs when the light enter a new medium in which its speed is different. Actually, light speed is not different, it always has the same constant value, denoted by c, which is exactly equal to $2.99\,792\,458 \times 10^8\,\mathrm{m/s}$. However, if the light enters a denser medium, the probability and the number of encounters with light and

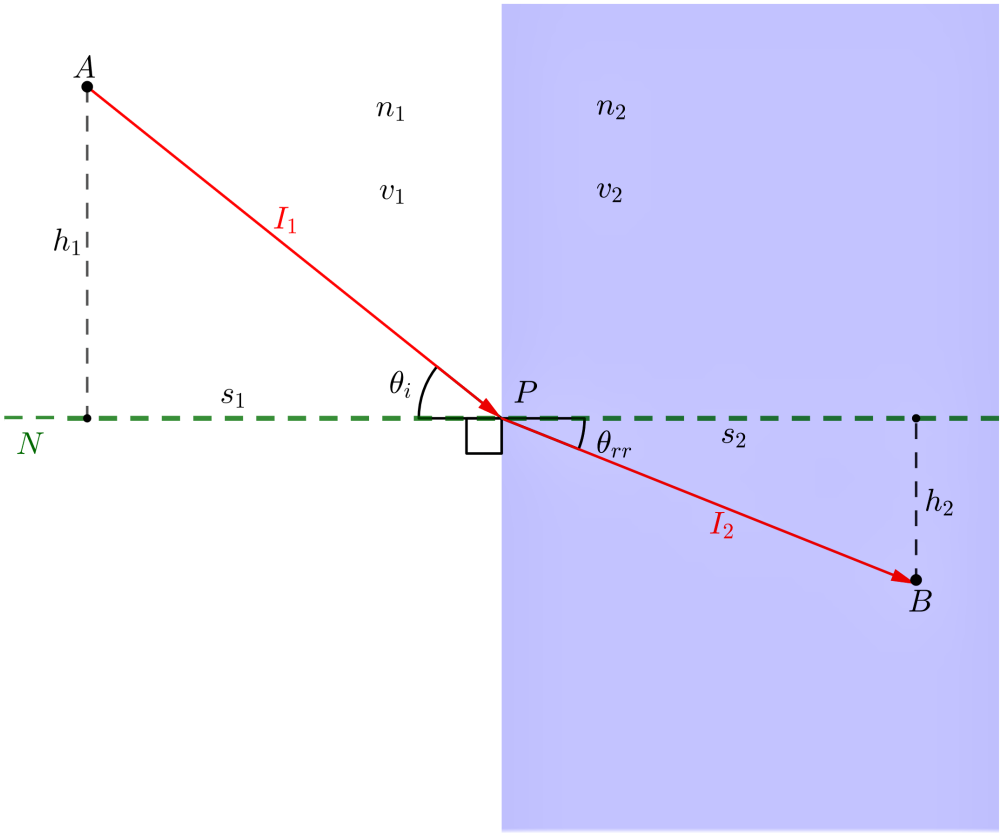


Figure 2.4. Derivation of law of refraction.

matter increase. When the light interacts with matter, it is completely absorbed by the electron which is then in an excited state. To be able to stable again, the electron releases a photon whose energy is equal to the one that it absorbed before. When the photon is released, it moves with the speed c just to interact with another electron. Due to these encounters, the average speed of the light has a lower value than c. Hence, when it is said that light has a lower speed, it means the average phase speed, not the actual speed of light [5].

Fermat's Principle, again, can be applied to refraction to find an expression about it. Suppose that the incident monochromatic light ray I_1 in a medium in which its speed is v_1 interacts with a boundary of a medium in which light speed is v_2, and passes to that medium as the light ray I_2 as in Figure 2.4. The incident ray I_1 makes θ_i angle with the surface normal and the refracted ray I_2 makes the angle θ_{rr}. According to the Fermat's Principle, the light should go from A to B in the shortest time. The total

travel time, t_t, is

$$t_t = \frac{\sqrt{h_1^2 + s_1^2}}{v_1} + \frac{\sqrt{h_2^2 + s_2^2}}{v_2} \tag{2.4}$$

Let $h = h_1 + h_2$. Then

$$t_t = \frac{\sqrt{h_1^2 + s_1^2}}{v_1} + \frac{\sqrt{(h - h_1)^2 + s_2^2}}{v_2}$$

The least time for light to travel from A to B can be found by taking derivative of t_t with respect to h_1 and making it equal to 0. In that case,

$$\begin{aligned} \frac{dt_t}{dh_1} &= \frac{2h_1}{2v_1\sqrt{h_1^2 + s_1^2}} + \frac{-2(h - h_1)}{2v_2\sqrt{(h - h_1)^2 + s_2^2}} \\ &= \frac{h_1}{v_1\sqrt{h_1^2 + s_1^2}} + \frac{-(h - h_1)}{v_2\sqrt{(h - h_1)^2 + s_2^2}} \\ &= \frac{1}{v_1}\frac{h_1}{\sqrt{h_1^2 + s_1^2}} + \frac{1}{v_2}\frac{-h_2}{\sqrt{h_2^2 + s_2^2}} = 0 \end{aligned} \tag{2.5}$$

From Eq. (2.5), it can be written that

$$\frac{1}{v_1}\frac{h_1}{\sqrt{h_1^2 + s_1^2}} = \frac{1}{v_2}\frac{h_2}{\sqrt{h_2^2 + s_2^2}} \tag{2.6}$$

From Figure 2.4, it can be seen that the fraction term with square root on the left hand side of Eq. (2.6) is $\sin\theta_i$, and the fractional term with square root on the right hand side is $\sin\theta_{rr}$. Hence, Eq. (2.6) becomes

$$\boxed{\frac{\sin\theta_i}{v_1} = \frac{\sin\theta_{rr}}{v_2}} \tag{2.7}$$

The Eq. (2.7) is knows as *the law of refraction*. It's also known as *the Snell's Law*, named after the astronomer Willebrord Snellius, although the relationship was known

by Ibn Sahl long before Snellius [6].

The Eq. (2.7) creates some useful results. First of all, the process is reversible; that is, the light would follow the same path if the light rays go from B to A instead of A to B. Second, this equation reveals the density relationship between materials. Without loss of generality, assume that the light moves in the first medium (on average) faster than it moves in the second medium. According to the law of refraction, the incident angle θ_i is bigger than the refraction angle θ_{rr} in this case. So, when light passes to a "denser"medium (the intended meaning is explained in Section 2.3.1), it bends towards the surface normal; the denser the matter is, the more the light bends.

As another result, assume that the incident ray comes perpendicular to the medium surface. This means that θ_i is 0, which makes $\sin\theta_i$ equal to 0. To satisfy the Eq. (2.7), the other side should be 0. Assuming that the new material is a transparent material, meaning that it allows the light to pass through itself, the light has a nonzero velocity inside of it. In this case, $\sin\theta_{rr}$ must be 0, meaning that θ_{rr} is 0. As a result, it follows that a light ray perpendicular to the surface of the boundary continues its way without being refracted, although the speeds are different in each medium.

Assume that a light ray goes from B to A (from denser material to a less dense one) in Figure 2.4. If Eq. (2.7) is reordered to give an expression of θ_i, [1]

$$\sin\theta_i = \frac{v_1}{v_2}\sin\theta_{rr} \tag{2.8}$$

Since $v_1 > v_2$ and both speeds are positive, the expression v_1/v_2 is bigger than 1. In this case, an angle θ_{rr} can be found such that $0 < \theta_{rr} < \pi/2$ making the right hand side of Eq. (2.8) equal to 1; i.e., $\sin\theta_{rr} = v_2/v_1$. Thus, θ_i becomes $\pi/2$, meaning that the refracted ray travels on the boundary between the two media. The angle θ_{rr} satisfying this condition is called *critical angle*. When the light ray makes the critical angle with

[1] Since the process is reversible, θ_i and θ_{rr} are used as in Figure 2.4, although by θ_i the incident ray angle is meant and θ_{rr} stands for refraction angle.

the normal moving from a dense medium to a less dense medium, it moves on the surface of the dense medium. If the angle continues to grow, the right hand side of Eq. (2.8) gets bigger than 1, which is impossible as the left hand side of the equation is a sine function. Therefore, the law of refraction is no longer valid in this case, and the light makes total internal reflection, obeying the law of reflection in Eq. (2.3).

2.3.1. Index of Refraction

As mentioned before, the (average) speed of light decreases when it enters a denser medium. That means light moves at the maximum speed c when there is no matter, in other words, in vacuum. Let a number n be defined as the ratio of its maximum speed and its speed in a medium, v;

$$\boxed{n = \frac{c}{v}} \tag{2.9}$$

n is called *index of refraction*. As can be seen from Eq. (2.9), its minimum value is 1, which is the index of refraction of vacuum, since the speed of light is the highest speed possible in current understanding of physics. For different matters, n is different since it depends on the molecular structure, temperature, density of the material, and so on. For vacuum, n is 1; for air at 1 atm pressure and 0℃ , it's 1.000 293; for water at 20℃ it's 1.333; for diamond, n is equal to 2.419 [7].

If Eq. (2.9) is put into Eq. (2.7), another form of law of refraction can be obtained as follows:

$$\boxed{n_1 \sin\theta_i = n_2 \sin\theta_{rr}} \tag{2.10}$$

In this equation, it can be concluded that the bigger the index of refraction of a material, the smaller the angle of refraction. By the description "denser" material in this chapter, it is meant that the material has bigger index of refraction. For instance,

density of diamond is 3.51 g/cm^3 and density of lanthanum fluoride is 5.936 g/cm^3. However, the index of refraction of diamond is 2.4166, and for lanthanum fluoride, it is 1.6040 [8]. That means the molecular density does not imply proportional index of refraction, and diamond is "denser" than lanthanum fluoride in this work.

For Figure 2.4, if the Eq. (2.9) is inserted into with Eq. (2.4), the result is

$$\begin{aligned} t_t &= \frac{n_1\sqrt{h_1^2+s_1^2}}{c} + \frac{n_2\sqrt{h_2^2+s_2^2}}{c} \\ &= \frac{1}{c}\left(n_1\sqrt{h_1^2+s_1^2} + n_2\sqrt{h_2^2+s_2^2}\right) \\ &= \frac{1}{c}\left(n_1\overline{AP} + n_2\overline{PB}\right) \end{aligned} \tag{2.11}$$

The term inside the parentheses in Eq. (2.11) is called *optical path length* (OPL), which is the traveled distance times the refractive index of the medium that the path takes place. In its modern form, the Fermat's Principle also dictates that a light ray must go from a path whose length is stationary with respect to the variations of the path.

2.3.2. Dispersion

When a photon encounters a boundary of a different medium, it does not gain or lose energy since there is not a mechanism to do that. Photon may be absorbed by an electron, but is released again by the same electron with the same energy. Hence, the energy of a photon does not change after experiencing a new medium. According to the Planck-Einstein Relation, energy of a photon, E, is

$$E = hf$$

where f is the frequency of the photon, and h is the Planck constant which is equal to $6.62\,607\,015 \times 10^{-34}$ J.s [9]. Energy and h are constant here, so is the frequency f on the boundary of the new medium.

Speed of a traveling wave, v, can be found by

$$v = \lambda f \tag{2.12}$$

where λ is the wavelength of the wave. If Eq. (2.9) is put into Eq. (2.12) and the obtained expression is rearranged, the result is

$$\boxed{\frac{c}{f} = n\lambda} \tag{2.13}$$

Eq. (2.13) tells that for the same frequency (the same light ray), the wavelength and the index of refraction are inversely proportional. The fact that refractive index depends on the wavelength of the incident ray is called *dispersion*. All the index of refraction values given in Section 2.3.1 and other sources in literature are conventionally obtained for the yellow doublet D-line of sodium, which nearly corresponds to 589.29 nm wavelength [7], to prevent confusion.

n value is strongly dependent on the frequency of the incident light ray. In fact, its effect is more dominant than, for instance, that of molecular structure or temperature of material. For gases, the index of refraction, n, can be calculated from

$$n^2 = 1 + \frac{Nq_e^2}{m_e\epsilon_0}\frac{1}{(-\omega^2 + j\gamma\omega + \omega_0^2)} \tag{2.14}$$

and for dense materials

$$3\left(\frac{n^2-1}{n^2+2}\right) = \frac{Nq_e^2}{m_e\epsilon_0}\sum_k \frac{f_k}{-\omega^2 + j\gamma_k\omega + \omega_{0k}^2} \tag{2.15}$$

where all ω's are the angular frequency of the incoming light. These equations are given to show the incident light ray frequency dependence of the refractive index. For further detail, please check [10]. An example for the effect of the wavelength on refractive index is shown in Figure 2.5.

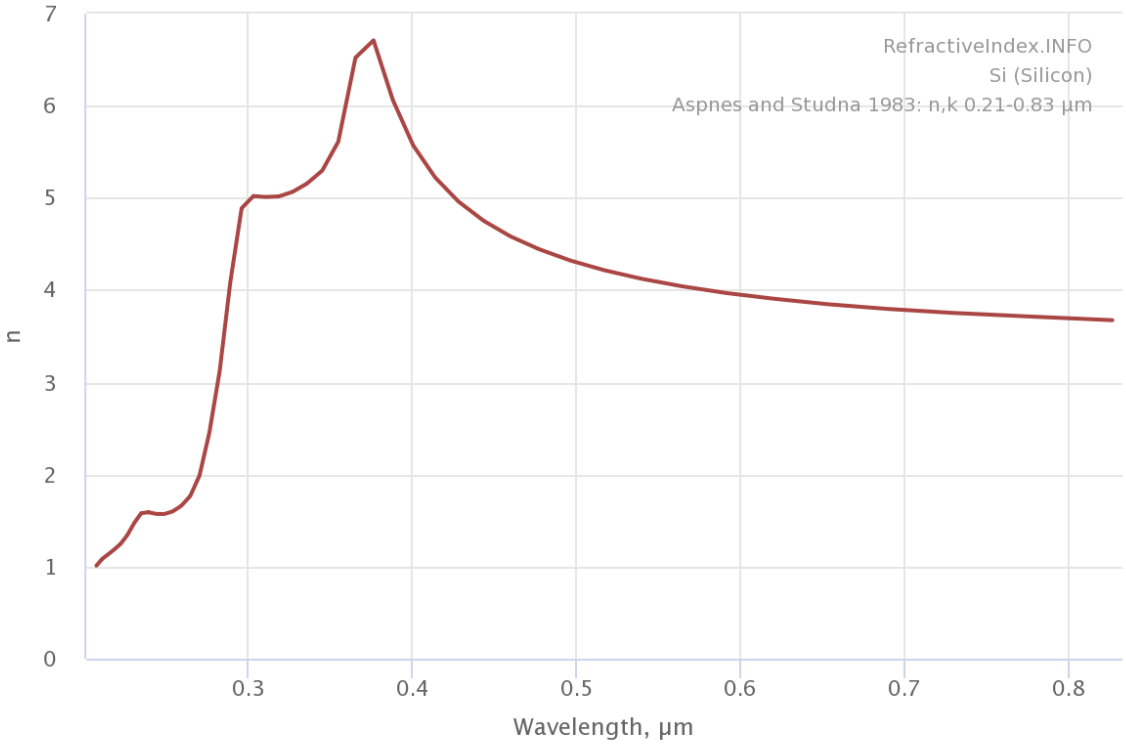


Figure 2.5. Wavelength–Refractive index relationship for silicon. Obtained from [11].

Recall that the refraction in Section 2.3 is explained with monochromatic light. Polychromatic light would create different light rays with different refraction angle due to dispersion. An example of dispersion phenomenon is rainbow. The sun light is a polychromatic light source, meaning that it contains many different wavelength values in visible range which is 390 nm to 780 nm. From Eq. (2.13), it is obvious that light rays with different wavelength values have different indices of refraction, and as a result of Eq. (2.10) they refract with different refraction angles. In this case, different colors are exposed to angular separation leading them to different directions when they pass through water droplets in air during rainy weather. The highest frequency (the shortest wavelength) experiences the biggest refractive index. Thus, it bends more towards to surface normal of the boundary. The longest wavelength in visual spectrum belongs to red, followed by orange, yellow, green, blue and violet, the shortest wavelength. This ordering is also the same seen in a rainbow.

2.3.3. Refraction on Spherical Surfaces

A flat transparent medium boundary always bends the planar wave of light with the same angle because the coming angle is the same for every point on the wavefront. However, if the medium boundary is a curved one, then the surface normal would be different for all points on planar wave since the surface normal for a curved plane is the line passing from the center of curvature and the contact point. In this case, the refracted wavefront will no longer be planar but curved. Depending on the surface's having convex or concave geometry, the planar wavefront either converges to a point or diverges from a point, respectively.

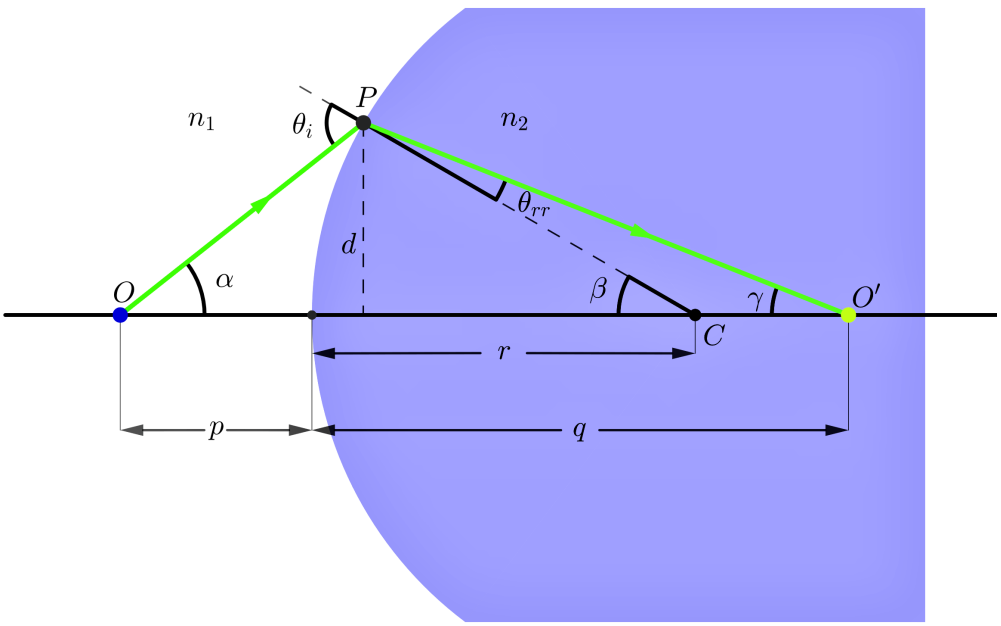


Figure 2.6. The geometry for refraction on spherical surfaces.

Assume that a light ray leaves from a point O inside a medium whose refractive index is n_1, touches an on–boundary point P, and is refracted by a medium with index of refraction n_2 such that the light ray forms an image on point O', as in Figure 2.6.

Here, it is assumed that $n_2 > n_1$. From the figure and the geometry, it is obvious that

$$\theta_i = \alpha + \beta \quad \text{and} \quad \theta_{rr} = \beta - \gamma$$

Note that β is the common variable for both angles. For Figure 2.6, the optical path length is

$$OPL = n_1\overline{OP} + n_2\overline{PO'} \tag{2.16}$$

where

$$\overline{OP} = \sqrt{r^2 + (p+r)^2 - 2r(p+r)\cos\beta}$$
$$\overline{PO'} = \sqrt{r^2 + (q-r)^2 + 2r(q-r)\cos\beta}$$

If the Fermat's Principle in its modern form is applied to Eq. (2.16) with the consideration that the optical path length varies with β,

$$\frac{d}{d\beta}OPL = 0 = \frac{n_1 r(p+r)\sin\beta}{2\overline{OP}} + \frac{-n_2 r(q-r)\sin\beta}{2\overline{PO'}}$$

which, after some simplifications, results to

$$\boxed{\frac{n_1}{\overline{OP}} + \frac{n_2}{\overline{PO'}} = \frac{1}{r}\left(\frac{n_2 q}{\overline{PO'}} - \frac{n_1 p}{\overline{OP}}\right)} \tag{2.17}$$

Although Eq. (2.17) is exact, the equation is not simple to apply since it depends on the location of the contact point P. For simplification, assume that the distance between contact point P and the optical axis $\overleftrightarrow{OO'}$ is not a comparable length with object-to-boundary distance p, boundary-to-image distance q or radius of curvature r. Hence, the angles θ_i and θ_{rr} are small, as well as α, β and γ. Since $\overline{CP}$ is close to

$\overleftrightarrow{OO'}$, so is the incident light ray. In this case, the incident ray is "close to the axis". Such rays are called *paraxial* rays; *para-* means "beside, close, side-by-side" in Latin, and *axial* means "about the axis". For paraxial ray approximation, $\beta \to 0$, resulting in $\overline{OP} \to p$ and $\overline{PO'} \to q$. These approximations lead to the final result

$$\boxed{\frac{n_1}{p} + \frac{n_2}{q} = \frac{n_2 - n_1}{r}} \tag{2.18}$$

Eq. (2.18) indicates that for a fixed object distance p, the paraxial rays are focused on the same point regardless of the incident angle. In this equation, all terms are positive. However, it does not have to be the case always. Eq. (2.18) is obtained with the assumption that $n_2 > n_1$. In this case, the sign convention for Figure 2.6 is as follows:

- r is positive if the center of curvature and the light source are at the opposite side of the refracting surface.
- q is positive if the light source and its image are at the opposite side of the refracting surface.
- p is positive if the incident light rays and the light source are at the same side of the refracting surface.

Note that although this equation is obtained with the assumption that $n_2 > n_1$, the Eq. (2.18) is valid regardless of which index of refraction is bigger. For $n_1 > n_2$, the sign convention is the opposite of the signs explained above.

2.3.4. Thick and Thin Lenses

Assume that two spherical surfaces with radii r_1 and r_2 are merged together to create a lens with thickness d. The lens has a refractive index n_l and the medium in which the lens is put has n_m with a relationship $n_l > n_m$. Assume that an object O is put on the optical axis of the lens p_1 units away in front of the lens. The setup is

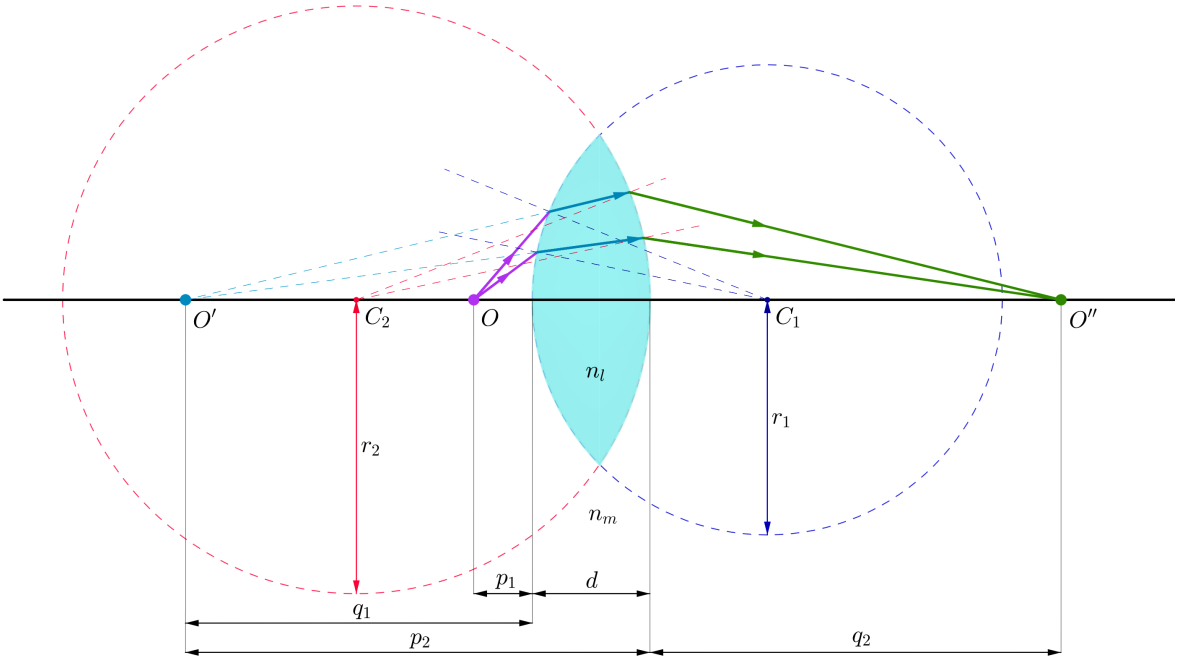


Figure 2.7. Derivation of lens equation.

pictured in Figure 2.7.

If Eq. (2.18) is written for the first contact of light and lens on the surface with radius of curvature r_1,

$$\frac{n_m}{p_1} + \frac{n_l}{q_1} = \frac{n_l - n_m}{r_1} \tag{2.19}$$

Assume that the object O is at such a position that the refracted light rays do not converge. In this case, extensions of the refracted light rays intersect to form a virtual image O' at q_1 units away from the lens surface. Since the image is not real, by the sign convention, q_1 is negative in Eq. (2.19); that is, $|q_1| = -q_1$. p_1 is positive since the incident light rays and the object are at the same side of the refracting surface. Center of the curvature C_1 is at the opposite side with the object, thus, r_1 is positive.

The second surface "sees" the light rays as if they are coming from O' and from a medium whose index of refraction is n_l. The object distance is $|p_2| = |q_1| + d$. Since

the incident light rays and the source are at the same side with respect to the surface, p_2 is positive, hence, $p_2 = d - q_1$. If Eq. (2.18) is written for the second surface,

$$\frac{n_l}{d - q_1} + \frac{n_m}{q_2} = \frac{n_m - n_l}{r_2} \tag{2.20}$$

In Eq. (2.20), q_2 is positive, and r_2 is negative since the center of curvature C_2 and the light source are at the same side. If Eq. (2.19) and Eq. (2.20) are summed and the obtained expression is rearranged,

$$\frac{1}{p_1} + \frac{1}{q_2} = \left(\frac{n_l}{n_m} - 1\right)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) - \frac{n_l d}{n_m q_1 (d - q_1)} \tag{2.21}$$

Eq. (2.21) is not convenient to use since it includes a term which requires the knowledge of the position of first image created by the first surface of the lens. Instead, [12] derived a more useful equation via small angle approximation ($\sin\theta \approx \tan\theta \approx \theta$ for $\theta < \pi/18$ radians) and lots of algebraic manipulation for thick lenses, which is

$$\frac{1}{p_1} + \frac{1}{q_2} = \left(\frac{n_l}{n_m} - 1\right)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n_l - n_m)d}{n_l r_1 r_2}\right] \tag{2.22}$$

Eq. (2.22) presents some interesting features. Firstly, the equation does not include a term containing the first image and second object locations. Hence, it is safe to drop the subscripts from p_1 and q_2. Second, the right hand side of the equation depends on the lens parameters and the medium, which do not have any relationship with the light source parameters or position. This makes that side a constant value. By calling this constant value as $1/f$, the Eq. (2.22) can be written as

$$\boxed{\frac{1}{p} + \frac{1}{q} = \frac{1}{f} = \left(\frac{n_l}{n_m} - 1\right)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n_l - n_m)d}{n_l r_1 r_2}\right]} \tag{2.23}$$

Eq. (2.23) is known as the *thick lens equation*. The distance f is called the *focal length* of the lens because when the object is at infinity; in other words, $p = \infty$, all the incident rays are focused on the point that f units away from the second surface of the lens. The value $1/f$ is called the *optical power* of the lens. It is so common in equations and expressions in optics that although the SI unit for $1/f$ is m^{-1}, it has another name in optics, *diopters*, shown by D. The lenses with positive optical power are called *convex* or *positive* lenses, and the ones with negative optical power are called *concave* or *negative* lenses.

For a lens whose thickness can be neglected, $d \to 0$, which makes the rightmost term 0 in Eq. (2.23), which results in

$$\boxed{\frac{1}{p} + \frac{1}{q} = \frac{1}{f} = \left(\frac{n_l}{n_m} - 1\right)\left[\frac{1}{r_1} - \frac{1}{r_2}\right]} \tag{2.24}$$

Eq. (2.24) is called the *lens law*, *thin lens equation* or *lens maker's equation* since it is used in lens making process to determine r_1 and r_2 when the focal point f and the index of refraction of the lens n_l are known. Although Eq. (2.23) and Eq. (2.24) are obtained with the geometry in Figure 2.7 under the condition $n_l > n_m$, these equations are valid in all cases with the proper application of sign convention.

Suppose that two thin lenses with focal distances f_1 and f_2 are in contact and an image is to be created by those lenses. If the thin lens maker's equation is written for separate lenses,

$$\frac{1}{p_1} + \frac{1}{q_1} = \frac{1}{f_1} \quad \text{and} \quad \frac{1}{p_2} + \frac{1}{q_2} = \frac{1}{f_2}$$

According to Figure 2.7, the image created by the first lens falls behind the lens and that image serves as the object for the second lens. However, since the lenses are considered thin, the image created by the first lens, which is the object for the second

lens, falls behind the second lens. Hence, by the sign convention, p_1 and q_1 are positive, as well as q_2, but p_2 should be negative and equal to $-q_1$. In this case,

$$\frac{1}{p_1}+\frac{1}{q_1}=\frac{1}{f_1} \quad \text{and} \quad \frac{1}{-q_1}+\frac{1}{q_2}=\frac{1}{f_2}$$

If those two equations are added and the subscripts of p_1 and q_2 are dropped as in Eq. (2.22), the result is

$$\boxed{\frac{1}{p}+\frac{1}{q}=\frac{1}{f}=\frac{1}{f_1}+\frac{1}{f_2}} \tag{2.25}$$

Eq. (2.25) tells that when the lenses are cascaded, the optical power of resultant system is the summation of the optical power of all lenses in contact.

2.4. Diffraction

In the 17$^{\text{th}}$ century, Francesco Maria Grimaldi published a work stating that in the experiment he conducted, he concluded that the passage of the light cannot be explained by rectilinear propagation of light. He called this deviation as *diffraction* [7], from the Latin word *diffringere* meaning "to break into pieces". In a similar wording, Arnold Sommerfeld explained diffraction as "any deviation of light rays from rectilinear paths which cannot be interpreted as reflection or refraction" [13].

Diffraction is a general characteristic of waves, either sound, matter or electromagnetic, happening when a part of the wavefront is prevented by an either transparent or opaque obstacle causing it to change amplitude or phase [7]. When there is no block for wavefront to move forward or change in the optical characteristics of medium, the Huygens' Principle holds and the superposition of the secondary wavelets creates the next wavefront generally in the same shape with the previous wavefront. However,

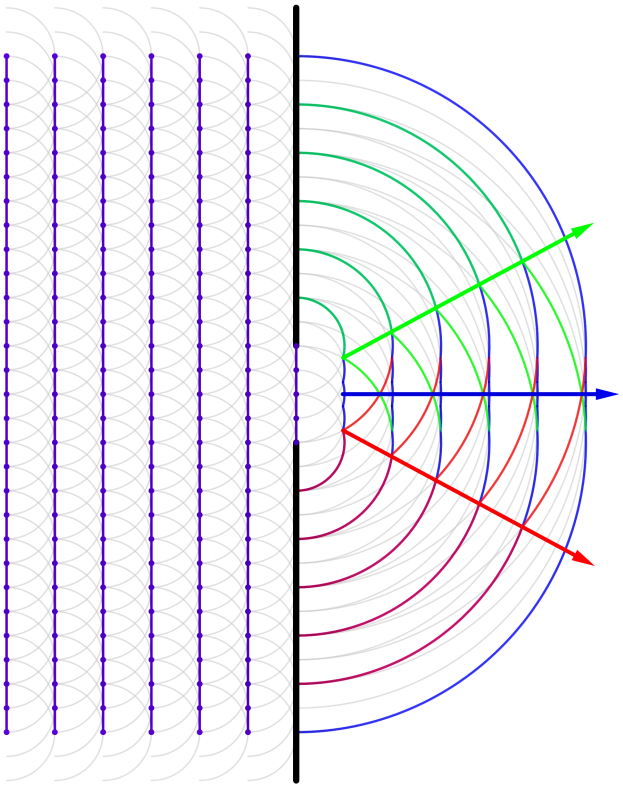


Figure 2.8. Huygens–Fresnel Principle.

when some portion of the wave meets an obstacle, some part of the wavefront is lost to create the next wavefront in the same shape. Hence, the shape of next wavefront is not a perfect copy of the previous one. The effect is shown in Figure 2.8 for planar waves. Since the wavefront meets an opaque screen with a small aperture, not all the wavefront is able to pass it, and beyond the screen it starts to be curved.

2.4.1. Huygens–Fresnel Principle

Note that Huygens' Principle in Section 2.1 does not specify the wave type or the wavelength. Suppose that in a sunny day you tell something to a friend, between you and who is a tree and the sun is behind of you. When you say something loud enough, your friend can hear you, meaning that the sound waves can make it behind the tree. However, there is also a shadow of the tree on the ground, meaning that the light waves could not reach behind the tree. Since both sound and light are waves, although with different frequencies, they should obey the Huygens' Principle, but according to that

the shadow should not exist; light should be able to fall behind the tree. This was the inadequacy of Huygens' Principle.

Huygens' Principle also cannot explain the diffraction phenomenon. When waves pass from a slit as in Figure 2.8, it could predict that the waves drawn by blue is formed, but with this formation the pattern on a viewing screen behind the aperture cannot be formed, which are called diffraction pattern. Moreover, as mentioned before, Huygens stated that the secondary spherical wavelets propagate in the direction of the primary wave but he did not give a mathematical explanation for why those "spherical" wavelets do not propagate behind the wavefront.

To mathematically correct these unexplained behaviors of the waves, Augustine-Jean Fresnel corrected the Huygens' Principle with the addition of *interference*. He corrected that as "every point on an unobstructed wavefront behaves as a point source of spherical wavelets at the same frequency, and the amplitude of the optical field beyond the wavefront is the superposition of all of those wavelets" [7]. By the superposition principle, he could explain the patterns qualitatively. In the Figure 2.8, the green parts of the secondary wavelets are in phase although they start their propagation at different times. Hence, those waves add up to produce a wavefront whose normal is shown with a green arrow. The same is valid for the waves shown in red. Although there are other in-phase wavelets and other point sources, only some portion of them are shown for illustrative purposes.

The second contribution of Fresnel is the explanation of why the spherical wavelets do not go beyond the primary wave that Huygens did not explain mathematically. The wave expression in space–time is

$$u(r_o, t) = \frac{A_o}{r_o} e^{j(kr_o - \omega t)} \tag{2.26}$$

where A_o is the wave amplitude, r_o is the traveled distance by the wave as in Figure 2.9, ω is the angular wave frequency and $k = 2\pi/\lambda$ is the wave number with λ being

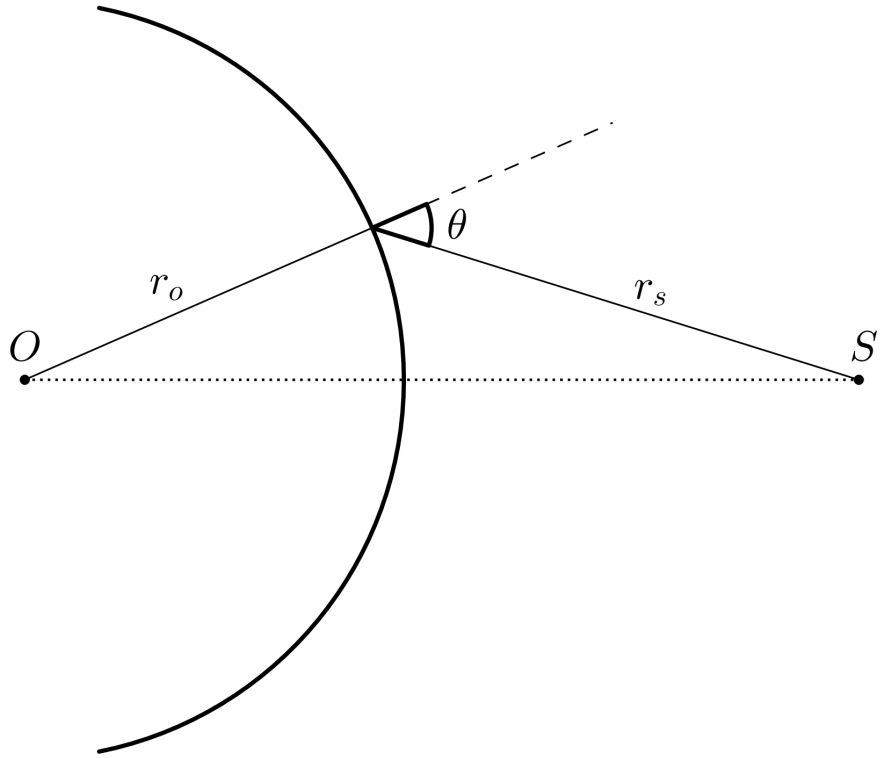


Figure 2.9. The geometry for Fresnel's derivation.

the wavelength. The complex amplitude of this wave in space is

$$U(r_o) = \frac{A_o}{r_o} e^{jkr_o} \tag{2.27}$$

This expression is easy to interpret: The amplitude of the wave decreases proportionally as the traveled distance r_o increases, and the phase of the wave is kr_o.

From his experiments, Fresnel *assumed* that the secondary wavelets should have $\pi/2$ out of phase with respect to primary waves, and since they propagate with the same frequency, hence, the same wavelength in an optically uniform medium, their amplitude ratio should be $1 : \lambda$. Therefore, he added $e^{-j\pi/2}/\lambda = -j/\lambda = 1/j\lambda$ term. Moreover, he *assumed* that an obliquity factor $K(\theta)$ should be added to explain that waves do not spread backwards. Together with the superposition principle, he expressed the wave contribution on point S, which is r_s units away from the differential surface element

dS_{w} on the primary wave, with the time dependence omitted, as

$$dU_S = \frac{A_o}{r_o} e^{jkr_o} \frac{1}{j\lambda} \frac{e^{jkr_s}}{r_s} K(\theta) dS_{\mathrm{w}}$$

The equation above can be explained as follows: The first two terms are the complex (primary) wave amplitude equation in space originated from r_o units away. The $1/j\lambda$ term is the assumed constant phase difference and $1 : \lambda$ amplitude ratio between the secondary waves and the primary wave. e^{jkr_s}/r_s is the equation of a wave originated on a point on the primary wave. $K(\theta)$ is the obliquity factor preventing the secondary waves from propagating the back of the primary wave. Finally, this contribution comes from a differential surface area dS_{w} on the primary wave; as the area increases, the contribution increases.

As a result, all the contribution on the point S from the primary wave is [14]

$$\boxed{U_S = \frac{A_o}{j\lambda} \iint_{S_{\mathrm{w}}} \frac{e^{jk(r_o+r_s)}}{r_o r_s} K(\theta) dS_{\mathrm{w}}} \tag{2.28}$$

where S_{w} is the surface area of the primary wave, θ is the angle between the point S and the surface normal of the primary wave as shown in Figure 2.9.

Although Fresnel assumed all of $\pi/2$ phase difference, $1 : \lambda$ amplitude ratio, and the obliquity factor $K(\theta)$, Gustav Kirchhoff showed via his integral theorem, Helmholtz wave equation and Green's theorem [13] that the term $1/j\lambda$ indeed comes from the nature of waves. Moreover, he showed that the obliquity factor is equal to

$$K(\theta) = \frac{1 + \cos\theta}{2}$$

Note that it is 1 when $\theta = 0$ and 0 for $\theta = \pi$. Eq. (2.28) is known as the *Fresnel–Kirchhoff diffraction formula*, and is the exact mathematical expression for the Huygens–Fresnel Principle [7].

To correct the Huygens's Principle, Fresnel used the concept interference. Although they have different names, diffraction and interference work in a quite similar way. In fact, the distinction between interference and diffraction is not physically clear. However, generally, the term interference is used when the number of waves in superposition is small, and the diffraction is used when there are large number of waves [7].

2.4.2. Diffraction on an Aperture

As mentioned before, diffraction occurs when some portion of the wavefront is blocked by an obstacle, either opaque or transparent, changing the amplitude and/or the phase of the wave. Assume that a light source sits on point P_o which is r_o units in front of an opaque surface. Behind the obstacle is the point $P_s(y_s, z_s)$ sitting r_s units away from the obstacle, which is illuminated by the light coming from the small aperture. The described geometry is shown in Figure 2.10. Note that the aperture is drawn as a rectangular for ease of illustration. However, this does not have to be case.

If the disturbance on point P_s due to a differential area dS located at $P_a(y_a, z_a)$ on the aperture is written by using Eq. (2.28), the result is

$$dU_{P_s} = \frac{1}{j\lambda} \frac{A_o e^{jk(r_{oa}+r_{as})}}{r_{oa} r_{as}} K(\theta) dS_{\mathrm{A}} \tag{2.29}$$

where S_{A} is the aperture area. r_{oa} and r_{as} can be written as

$$r_{oa} = \sqrt{r_o^2 + y_a^2 + z_a^2}$$
$$r_{as} = \sqrt{r_s^2 + (y_a - y_s)^2 + (z_a - z_s)^2}$$

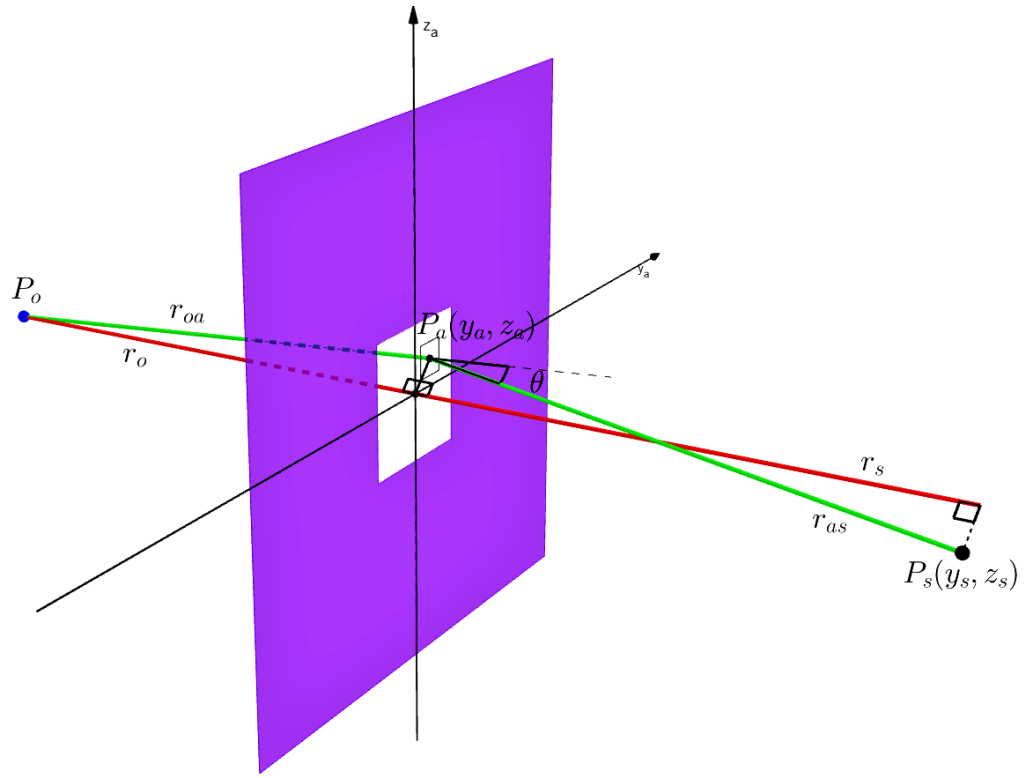


Figure 2.10. Diffraction effect on a point behind the aperture on an opaque screen.

If the expressions are rewritten by factoring out r_o and r_s from the square roots,

$$\begin{aligned} r_{oa} &= r_o\sqrt{1+\frac{y_a^2}{r_o^2}+\frac{z_a^2}{r_o^2}} \\ r_{as} &= r_s\sqrt{1+\frac{(y_a-y_s)^2}{r_s^2}+\frac{(z_a-z_s)^2}{r_s^2}} \end{aligned} \tag{2.30}$$

For the Eqs. (2.30) inserted in Eq. (2.29), there are two approximations. One is done by Augustine–Jean Fresnel and the other is made by Joseph Ritter von Fraunhofer.

2.4.3. Fresnel Approximation

For Figure 2.10, Fresnel assumed that the aperture sizes are very small compared to the distances r_o and r_s. This assumption also makes the deviation of the light from

the aperture very small. Hence, for the amplitude terms, r_{oa} and r_{as} are nearly equal to r_o and r_s, respectively. But for the phase term, the same approximation cannot be done since $k = 2\pi/\lambda$ is a huge number for visible light. Consider the binomial expansion of the square root

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots$$

which converges for $|x| < 1$. If Eq. (2.30) are expanded using the binomial expansion of square root up to the second term,

$$\begin{aligned} r_{oa} &\approx r_o \left[1 + \frac{y_a^2 + z_a^2}{2r_o^2}\right] \\ r_{as} &\approx r_s \left[1 + \frac{(y_a - y_s)^2}{2r_s^2} + \frac{(z_a - z_s)^2}{2r_s^2}\right] \end{aligned} \tag{2.31}$$

So, the total disturbance on point P_s is

$$U_{P_s} = \frac{A_o}{j\lambda r_o r_s} \iint_{S_\mathrm{A}} e^{jk(r_{oa}+r_{as})} K(\theta) dS_\mathrm{A} \tag{2.32}$$

where the terms r_{oa} and r_{as} are as in Eqs. (2.31).

Eq. (2.32) is interesting to analyze. If r_{oa} and r_{as} are written into the equation, the phase term becomes

$$\exp\left\{jk\left(r_o + r_s + \frac{y_s^2 + z_s^2}{2r_s}\right)\right\} \cdot \exp\left\{jk\frac{(r_o + r_s)(y_a^2 + z_a^2)}{2r_o r_s}\right\} \cdot \exp\left\{\frac{-jk}{r_s}\left(y_a y_s + z_a z_s\right)\right\}$$

The first term is independent of the integration variables that are y_a and z_a, since S_A depends on the coordinates y_a and z_a. Hence, the first term can be taken out of the

integral. If the Eq. (2.32) is rearranged,

$$\boxed{\begin{aligned} U_{P_s} &= \frac{K(\theta)}{j\lambda}\exp\left\{jk\left(\frac{y_s^2+z_s^2}{2r_s}\right)\right\} \\ &\times \iint_{S_\mathrm{A}}\left[\frac{A_o e^{jk(r_o+r_s)}}{r_o r_s}\exp\left\{jk\frac{(r_o+r_s)(y_a^2+z_a^2)}{2r_o r_s}\right\}\right] \\ &\times \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(y_a y_s + z_a z_s\right)\right\} dS_\mathrm{A} \end{aligned}} \tag{2.33}$$

where $K(\theta)$ is taken out of the integral as it does not depend on the integration variables since r_s is so large compared to the aperture sizes that they are neglected in the calculation of $\cos\theta$. A close look at the Eq. (2.33) reveals that the term in square brackets are the disturbance just in front of the aperture and a quadratic phase term. Other than the multiplicative phase and amplitude terms, the complete equation in Eq. (2.33) is the Fourier transform of the wave coming from the aperture and a quadratic phase [14]. It can be thought as that because when the limits of the integral go from $-\infty$ to ∞ as the Fourier transform requires, the wave amplitude would become 0 where the aperture coordinates exceed the opening size.

If the center illumination is examined; in other words, if $y_s = 0$ and $z_s = 0$ and the aperture is assumed a rectangle, then the phase terms outside of the square brackets are 1, meaning that there is no phase change by those terms. If the Eq. (2.33) is rearranged in this case,

$$U_{P_s} = \frac{A_o e^{jk(r_o+r_s)}}{j}K(\theta)\iint_{S_\mathrm{A}}\frac{1}{r_o r_s \lambda}\exp\left\{j\frac{2\pi}{\lambda}\left(\frac{(r_o+r_s)(y_a^2+z_a^2)}{2r_o r_s}\right)\right\} dy_a dz_a$$

For the phase term inside the integral,

$$\exp\left\{j\frac{2\pi}{\lambda}\left(\frac{(r_o+r_s)(y_a^2+z_a^2)}{2r_o r_s}\right)\right\} = \exp\left\{j\frac{\pi(r_o+r_s)y_a^2}{\lambda r_o r_s}\right\}\cdot\exp\left\{j\frac{\pi(r_o+r_s)z_a^2}{\lambda r_o r_s}\right\}$$

If the change of variables is applied,

$$m = y_a\sqrt{\frac{2(r_o + r_s)}{\lambda r_o r_s}} \quad \text{and} \quad n = z_a\sqrt{\frac{2(r_o + r_s)}{\lambda r_o r_s}}$$

then the resulting integral is

$$U_{P_s} = \frac{K(\theta)}{j}\frac{A_o e^{jk(r_o+r_s)}}{2(r_o + r_s)}\int_m e^{j\pi m'^2/2}dm' \int_n e^{j\pi n'^2/2}dn' \tag{2.34}$$

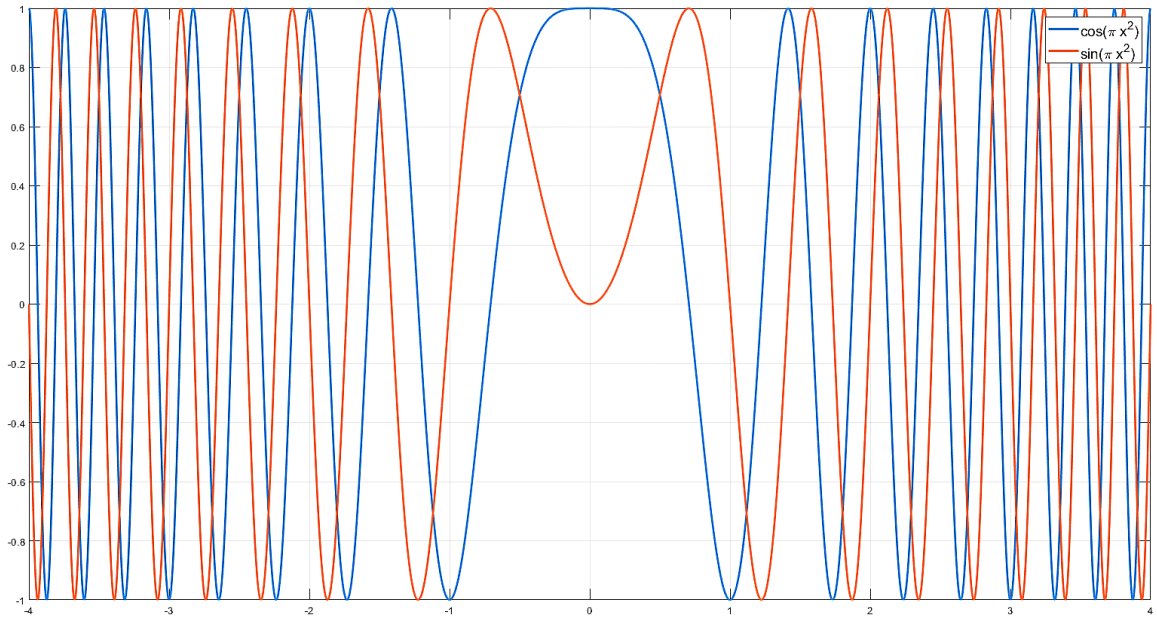
The term in front of the integral is the half of the unobstructed wave with a $\pi/2$ phase delay.[2] The integrals can be calculated using *Fresnel integrals*, which are

$$C(x) = \int_0^x \cos(\pi u^2)du \quad \text{and} \quad S(x) = \int_0^x \sin(\pi u^2)du$$

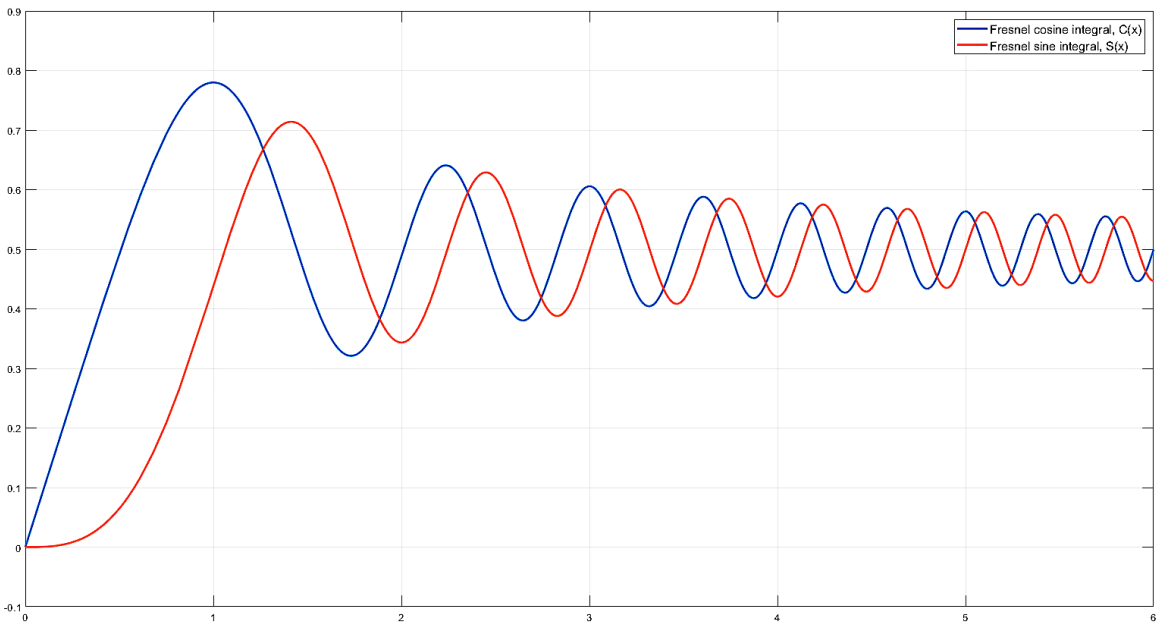
The functions $\cos(\pi x^2)$ and $\sin(\pi x^2)$, together with $C(x)$ and $S(x)$ are plotted in Figure 2.11. The graph 2.11(b), the Fresnel integrals, shows the radiance distribution pattern on a screen with light coming from a rectangular aperture illuminated by a point source. That is, if a photodiode is placed on the screen on which the pattern is expected to occur, the detected distribution has a shape related to those functions [7].

The Fresnel approximation is also called *near-field* approximation. Because in the calculation, although the distances r_o and r_s are considered big and thus eliminate some of the terms, they are assumed not big enough to neglect, for instance, quadratic terms in Eqs.(2.30) and (2.31), so the source and the screen are still close to aperture.

[2] $1/j = e^{-j\pi/2}$. Moreover, the term $K(\theta)$ is accepted as 1 since for $y_s = 0$ and $z_s = 0$, and with the assumption that r_s is much bigger than y_a and z_a, $\cos\theta$ is nearly 1 making $K(\theta)$ equal to 1.



(a)



(b)

Figure 2.11. The Fresnel integrals. (a) is $\cos(\pi x^2)$ and $\sin(\pi x^2)$, (b) shows the Fresnel integrals $C(x)$ and $S(x)$.

2.4.4. Fraunhofer Approximation

Consider the Eq. (2.30) for Figure 2.10,

$$r_{oa} = r_o\sqrt{1 + \frac{y_a^2}{r_o^2} + \frac{z_a^2}{r_o^2}}$$
$$r_{as} = r_s\sqrt{1 + \frac{(y_a - y_s)^2}{r_s^2} + \frac{(z_a - z_s)^2}{r_s^2}}$$

For the amplitude-related terms in Fresnel approximation, those quantities are taken equal to r_o and r_s, respectively. In Fraunhofer, or *far-field*, approximation, r_o and r_s are considered bigger than those in the Fresnel approximation. In this case, for amplitudes, $r_{oa} = r_o$ and $r_{as} = r_s$. For phase terms, if the binomial expansion of square root is used for r_{oa}, the result is as in Eq. (2.31). r_o is accepted so larger than the aperture sizes y_a and z_a that the fractional terms are considered insignificant even after the multiplication with $k = 2\pi/\lambda$, which is a very big number for light waves. Hence, in the phase terms also, $r_{oa} = r_o$. For r_{as}, if the binomial expansion is applied,

$$\begin{aligned} r_{as} &= r_s\sqrt{1 + \frac{(y_a - y_s)^2}{r_s^2} + \frac{(z_a - z_s)^2}{r_s^2}} \\ &\approx r_s\left[1 + \frac{(y_a - y_s)^2}{2r_s^2} + \frac{(z_a - z_s)^2}{2r - s^2}\right] \\ &= r_s\left[1 + \frac{y_a^2 + z_a^2}{2r_s^2} + \frac{y_s^2 + z_s^2}{2r_s^2} - \frac{2(y_a y_s + z_a z_a)}{2r_s^2}\right] \\ &\approx r_s + \frac{y_s^2 + z_s^2}{2r_s} - \frac{y_a y_s + z_a z_s}{r_s} \end{aligned}$$

in which at the last step, r_s is considered so bigger than the aperture coordinates that the terms containing only the aperture coordinates can be ignored even if they are contributing the phase. It is because the term is so small that it actually makes the exponent 0, making the phase contribution is equal to 1. Moreover, the obliquity factor $K(\theta)$ can be taken as 1. Since the point P_s very far from the aperture, θ can be kept small, although y_s and z_s are bigger than the aperture, making $\cos\theta$ nearly 1, as well as $K(\theta)$. If the values for r_{oa}, r_{as} and $K(\theta)$ are put into Eq. (2.29) and obtained

expression is integrated over the aperture, the result is

$$\boxed{U_{P_s} = \frac{\exp\left\{jk\left(\frac{y_s^2+z_s^2}{2r_s}\right)\right\}}{j\lambda} \iint\limits_{S_\mathrm{A}} \frac{A_o e^{jk(r_o+r_s)}}{r_o r_s} \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(y_a y_s + z_a z_s\right)\right\} dS_\mathrm{A}} \tag{2.35}$$

The Eq. (2.35) is called the *Fraunhofer diffraction formula*. It says that apart from the quadratic phase term depending on the screen (observer) coordinates, the Fraunhofer diffraction is the Fourier transform of the incident wave coming from the aperture evaluated at frequencies $f_{y_\mathrm{a}} = y_s/\lambda r_s$ and $f_{z_\mathrm{a}} = z_s/\lambda r_s$. Eq. (2.33) and Eq. (2.35) look similar, but are not completely the same. The first difference is the obliquity factor seen in the Fresnel diffraction. It can be ignored in Fraunhofer version because of the assumption that the aperture is very far from both the source and the observation point. The second difference is the quadratic phase in the integral. In the Fresnel formula, a quadratic phase term written in the aperture coordinates does not disappear since the distances are not big enough to let the aperture sizes be eliminated. This term also emphasizes the biggest and the most important difference between Fresnel and Fraunhofer approximations: The linearity in the aperture coordinates in amplitude and phase terms is the necessity for the Fraunhofer diffraction, whereas for Fresnel approximation such a condition is not sought [7, 14].

An important assumption in Fraunhofer diffraction is the distance between the source and the aperture going to infinity [13]. This assumption ensures that the arriving incident waves are planar waves and the illumination at every point on the aperture is equal due to the absence of the obliquity factor.

Suppose the setting in Figure 2.10 with rectangular aperture whose width is w_y and height is h_z. In that case, the Eq. (2.35) becomes

$$U_{P_s} = \varepsilon_A \int_{-w_y/2}^{w_y/2} \int_{-h_z/2}^{h_z/2} \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(y_a y_s + z_a z_s\right)\right\} dy_a dz_a \tag{2.36}$$

where

$$\varepsilon_A = \frac{A_o \exp\left\{jk\left(r_o + r_s + \frac{y_s^2+z_s^2}{2r_s}\right)\right\}}{j\lambda r_o r_s} \tag{2.37}$$

This variable ε_A can be thought as the source strength per unit area, which can also be understood by looking at the units. The Eq. (2.36) can also be written as

$$U_{P_s} = \varepsilon_A \iint_{-\infty}^{\infty} \text{rect}\left(\frac{y_a}{w_y}\right) \text{rect}\left(\frac{z_a}{h_z}\right) \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(y_a y_s + z_a z_s\right)\right\} dy_a dz_a \tag{2.38}$$

where $\text{rect}(x)$ is defined as

$$\text{rect}(x) = \begin{cases} 1 & , |x| < 1/2 \\ 0.5 & , |x| = 1/2 \\ 0 & , \text{otherwise} \end{cases}$$

By using a very well known fact that Fourier transform of the rectangular function $\text{rect}(x/a)$ is $a\,\text{sinc}(af) = a\sin(\pi af)/\pi af$, Eq. (2.38) results to

$$U_{P_s} = \varepsilon_A (w_y h_z)\,\text{sinc}\left(\frac{w_y y_s}{\lambda r_s}\right) \text{sinc}\left(\frac{h_z z_s}{\lambda r_s}\right) \tag{2.39}$$

As told before, ε_A is the source strength per unit area, and $(w_y h_z)$ is the aperture area, which gives a consistent outcome. The resulting field distribution is given in Figure 2.12. Here, the aperture is wider than it is tall. This results to more frequent peaks on horizontal axis than the vertical axis.

As another, and more related-with-this-work, setup, consider a circular aperture as in Figure 2.13. In this case, the coordinates in Eq. (2.35) should be modified for spherical coordinates as follows:

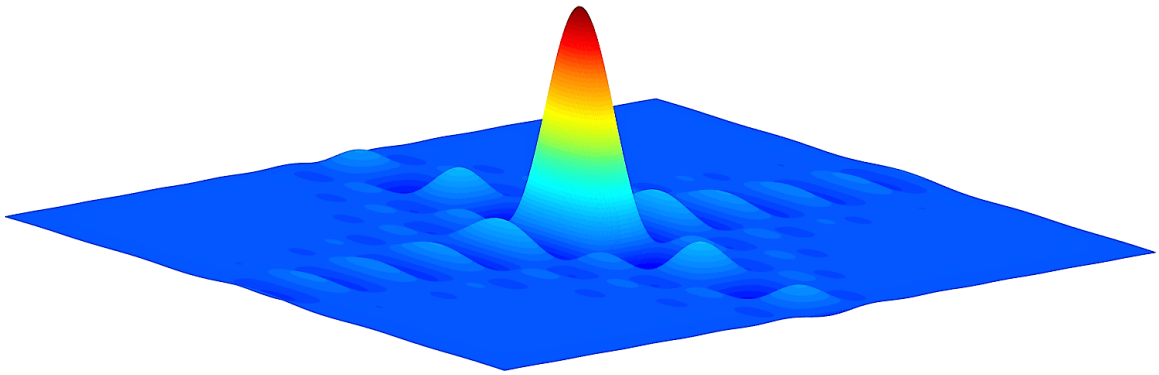


Figure 2.12. Fraunhofer diffraction of an electric field from a rectangular aperture.

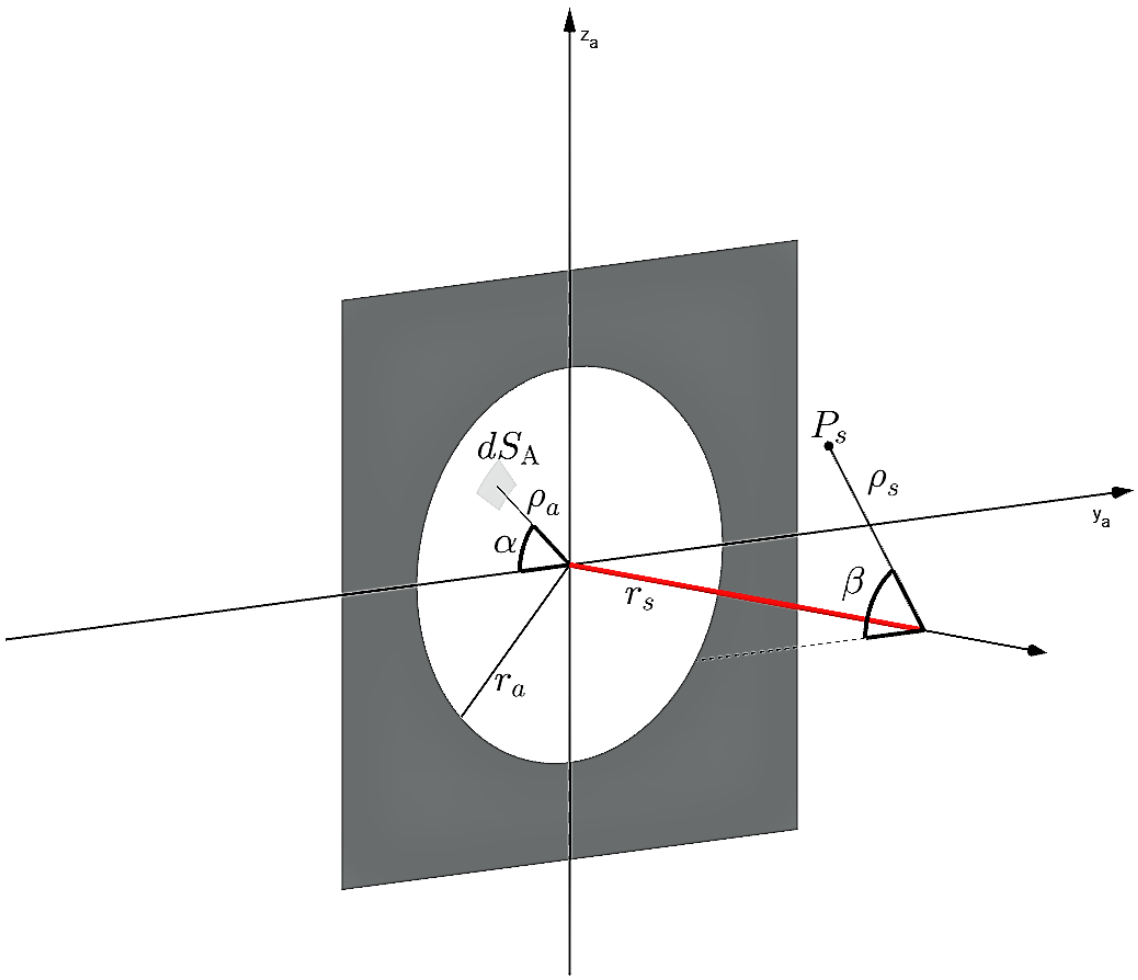


Figure 2.13. The geometry for Fraunhofer diffraction on a circular aperture.

$$\begin{aligned} y_a &= \rho_a \sin\alpha \qquad z_a = \rho_a \cos\alpha \qquad dS_{\mathrm{A}} = \rho_a d\rho_a d\alpha \\ y_s &= \rho_s \sin\beta \qquad z_s = \rho_s \cos\beta \end{aligned} \tag{2.40}$$

If Eqs. (2.40) are put into Eq. (2.35), the result is

$$\begin{aligned} U_{P_s} &= \varepsilon_A \iint \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(\rho_a\rho_s \sin\alpha\sin\beta + \rho_a\rho_s\cos\alpha\cos\beta\right)\right\}\rho_a d\rho_a d\alpha \\ &= \varepsilon_A \int_0^{r_a}\int_0^{2\pi} \rho_a \exp\left\{\frac{-j2\pi}{\lambda r_s}\rho_a\rho_s\cos(\alpha-\beta)\right\} d\alpha d\rho_a \\ &= \varepsilon_A \int_0^{r_a} \rho_a \left[\int_0^{2\pi} \exp\left\{\frac{-j2\pi}{\lambda r_s}\rho_a\rho_s\cos\alpha\right\} d\alpha\right] d\rho_a \end{aligned} \tag{2.41}$$

The first step in Eq. (2.41) is reached via putting the expressions in Eq. (2.40) into Eq. (2.35). The second step is obtained by the trigonometric identity

$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$$

The last step uses the linearity of integral operation to separate ρ_a from the integral depending on α, and the axial symmetry in the screen geometry to eliminate the β dependence of the integral. In every steps, ε_A is defined as in Eq. (2.37).

Consider the term in the square brackets in Eq. (2.41). If the expression is rearranged,

$$\begin{aligned} &\int_0^{2\pi} \cos\left(\frac{2\pi\rho_a\rho_s\cos\alpha}{\lambda r_s}\right) d\alpha - j\int_0^{2\pi} \sin\left(\frac{2\pi\rho_a\rho_s\cos\alpha}{\lambda r_s}\right) d\alpha \\ =& \int_0^{2\pi} \cos(x\cos\alpha) d\alpha - j\int_0^{2\pi} \sin(x\cos\alpha) d\alpha \\ =& \int_0^{2\pi} \cos(x\cos\alpha) d\alpha \end{aligned}$$

where $x = 2\pi\rho_a\rho_s/\lambda r_s$. If $\cos(x\cos\alpha)$ and $\sin(x\cos\alpha)$ are plotted, it can be understood that the integral with sin term actually equals to 0, leaving only the integral on the left. This integral cannot be reduced more, and is a specific function named after the German astronomer Friedrich Wilhelm Bessel.

Denoted by $J_m(x)$, the Bessel function of the first kind of order m is defined as

$$J_m(x) = \frac{j^{-m}}{2\pi} \int_0^{2\pi} e^{j(m\alpha + x\cos\alpha)} d\alpha$$

The graph of the Bessel functions of the first kind of order $m = 0, 1, ..., 4$ are shown in Figure 2.14.

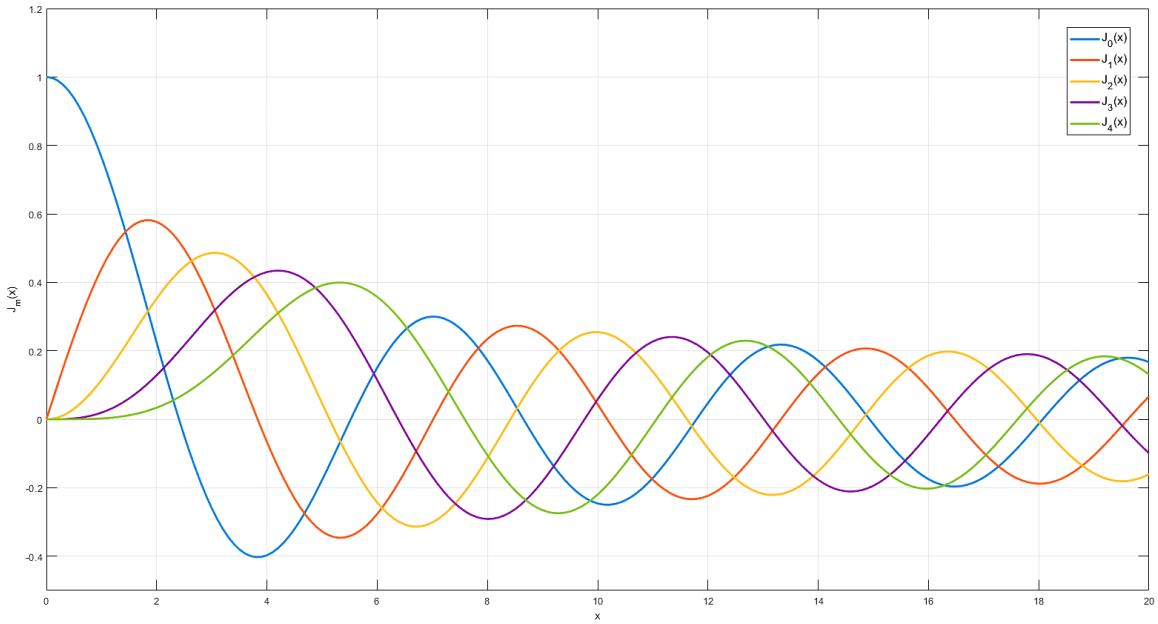


Figure 2.14. The Bessel functions of the first kind of order $m = 0, 1, ..., 4$.

The Bessel functions have many representations, properties and applications, most of which are out of scope of this work. But an interesting and useful property is the recurrence relations, which states that

$$\left(\frac{1}{x}\frac{d}{dx}\right)^n [x^m J_m(x)] = x^{m-n} J_{m-n}(x)$$
$$\left(\frac{1}{x}\frac{d}{dx}\right)^n \left[\frac{J_m(x)}{x^m}\right] = (-1)^n \frac{J_{m+n}(x)}{x^{m+n}}$$

Going back to Eq. (2.41), observe that, except the normalization constant, the term in square brackets is exactly equal to $J_0(x)$. In this case, Eq (2.41) becomes

$$U_{P_s} = 2\pi\varepsilon_A \int_0^{r_a} \rho_a J_0\left(\frac{2\pi\rho_a\rho_s}{\lambda r_s}\right) d\rho_a \tag{2.42}$$

If the variables are changed in this integral as

$$x = \frac{2\pi\rho_a\rho_s}{\lambda r_s}, \qquad dx = d\rho_a \frac{2\pi\rho_s}{\lambda r_s}$$

the integral becomes

$$U_{P_s} = 2\pi\varepsilon_A \left(\frac{\lambda r_s}{2\pi\rho_s}\right)^2 \int_0^{2\pi r_a\rho_s/\lambda r_s} x J_0(x) dx \tag{2.43}$$

As the last step, by using the first recurrence relation with $n = 1$ and $m = 1$, the Fraunhofer diffraction from a circular aperture becomes

$$U_{P_s} = \varepsilon_A \left(\frac{\lambda r_s r_a}{\rho_s}\right) J_1\left(\frac{2\pi r_a\rho_s}{\lambda r_s}\right) \tag{2.44}$$

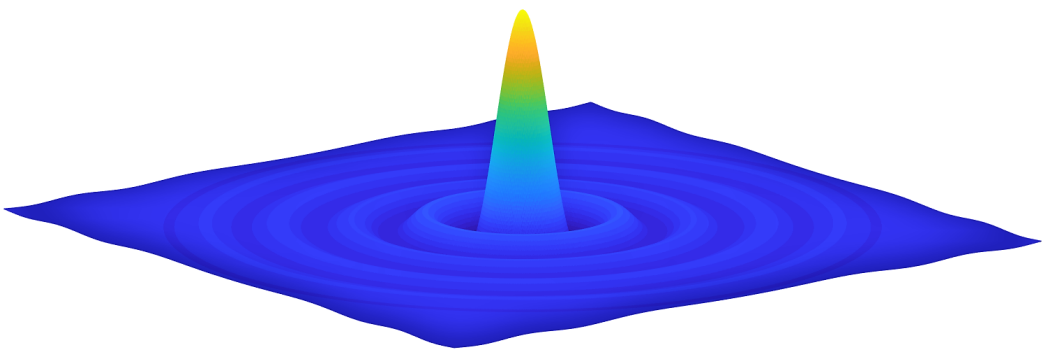


Figure 2.15. Fraunhofer diffraction of an electric field from a circular aperture.

The resulting pattern is shown in Figure 2.15. This pattern is called *Airy pattern*, named after George Biddell Airy. The disk at the middle is defined as the best-focused location of a light that a perfect lens can make with a circular aperture which is diffraction limited.

Fraunhofer, or far-field, approximation for optical frequencies is not actually very practical one. For instance, the patterns in Figure 2.12 and Figure 2.15 are obtained with 4 mm-wide and 2 mm-tall aperture and a circular aperture with 2 mm radius, respectively, using 589.29 nm wavelength. But the simulated observations are done on $1.5\,\mathrm{m} \times 1.5\,\mathrm{m}$ plane placed 1000 m away from the apertures. However, Fraunhofer diffraction can be observed much closer than these distances under the right circumstances. It can be shown that converging spherical waves from the aperture to the observation plane create Fraunhofer diffraction pattern although in the calculation the given distance requires Fresnel approximation [14]. Together with the plane wave requirement entering the aperture, a system with one converging lens in front of the aperture and one just behind obeys the Fraunhofer diffraction [7]. The first lens ensures that the entering waves are planar, and the second one provides the converging spherical waves to the observer.

2.5. Light Passing Through an Aperture with Lens

Consider the Fresnel integral in Eq. (2.33) with rearranged terms

$$\begin{aligned} U_{P_s} = {} & \frac{K(\theta)}{j\lambda r_s} \exp\left\{jk\left(\frac{y_s^2 + z_s^2}{2r_s}\right)\right\} \times \\ & \iint_{S_\mathrm{A}} \left[U_\mathrm{A} \exp\left\{jk\left(\frac{y_a^2 + z_a^2}{2}\right)\left(\frac{r_o + r_s}{r_o r_s}\right)\right\}\right] \exp\left\{\frac{-j2\pi}{\lambda r_s}\left(y_a y_s + z_a z_s\right)\right\} dS_\mathrm{A} \end{aligned} \tag{2.45}$$

where

$$U_\mathrm{A} = \frac{A_o e^{jk(r_o + r_s)}}{r_o}$$

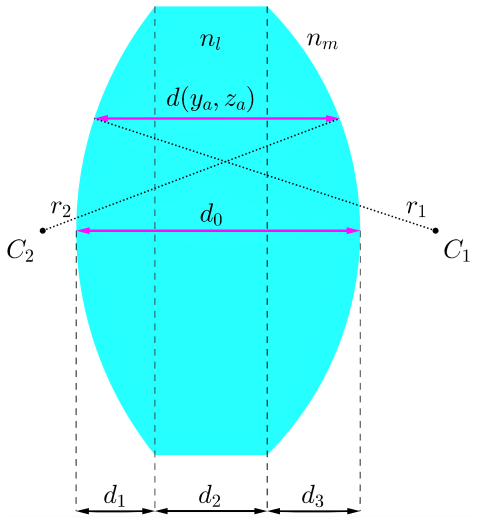
For this equation, assume that an aperture is illuminated by a plane wave. For incident planar wave, the disturbance seems to come from the infinity; in other words, $r_o \to \infty$. In that case, U_A becomes a planar wave with constant amplitude, and for the limiting case the phase term inside the square brackets becomes

$$\exp\left\{jk\left(\frac{y_a^2 + z_a^2}{2r_s}\right)\right\} \tag{2.46}$$

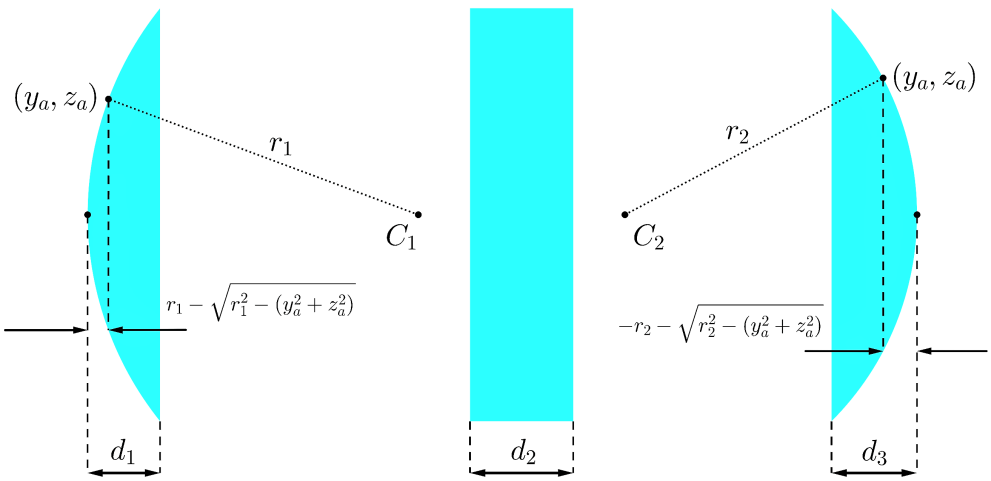
Suppose that there is a converging lens in front of the aperture. In that case, the illumination on the aperture is not constant but a function of the aperture coordinates in phase and the amplitude; that is,

$$U_{\mathrm{A}}(y_a, z_a) = U_{\mathrm{A}_0}(y_a, z_a) e^{j\phi(y_a, z_a)} \tag{2.47}$$

Although the amplitude $U_{\mathrm{A}_0}(y_a, z_a)$ is also important, the bigger care for the phase term $\phi(y_a, z_a)$ must be given.



(a)



(b)

Figure 2.16. The lens thickness geometry. (a) is the lens, (b) is the assumed lens components.

A lens which has one surface with radius of curvature r_1 and another surface having r_2 is put in front of the aperture. The thickest part of the lens has d_0 thickness. The lens having a refractive index n_l is inside a medium with n_m index of refraction. The lens is in such a shape that it can be thought as it is the combination of two curved lens and one rectangular material. The explained geometry is shown in Figure 2.16.

Assume that the lens is clean and lets the light pass through without loss in amplitude. So, the lens only changes the phase of the incident light, meaning that $U_{\mathrm{A}_0}(y_a, z_a)$ is constant. The phase change, however, depends on the optical path that light travels.

The optical path length (OPL) traveled by the incident light coming from the left (first touching the surface with radius of curvature r_1) is

$$\mathrm{OPL}(y_a, z_a) = n_l d(y_a, z_a) + n_m(d_0 - d(y_a, z_a))$$

where the first term on the right hand side is the optical path length traveled in the lens and the second part is the optical path length traveled in the medium. If this expression is rearranged,

$$\mathrm{OPL}(y_a, z_a) = n_m d_0 + (n_l - n_m) d(y_a, z_a) \tag{2.48}$$

where $d(y_a, z_a)$ is the thickness function of the lens. This function can be found by dividing the lens into three components as in Figure 2.16(b). The first part (the leftmost part in Figure 2.16(b)) has the thickness equal to

$$d_1 - \left(r_1 - \sqrt{r_1^2 - (y_a^2 + z_a^2)}\right) = d_1 - r_1\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_1^2}}\right) \tag{2.49}$$

The second part, obviously, has a thickness d_2. The thickness of the last portion of the lens is

$$d_3 - \left(-r_2 - \sqrt{r_2^2 - (y_a^2 + z_a^2)}\right) = d_3 + r_2\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_2^2}}\right) \tag{2.50}$$

where the r_2 comes out from the square root as $-r_2$ and the first r_2 has a negative sign in front of it due to the sign convention. So, the total thickness of the lens is

$$\begin{aligned} d(y_a, z_a) &= d_1 - r_1\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_1^2}}\right) + d_2 + d_3 + r_2\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_2^2}}\right) \\ &= d_0 - r_1\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_1^2}}\right) + r_2\left(1 - \sqrt{1 - \frac{y_a^2 + z_a^2}{r_2^2}}\right) \end{aligned} \tag{2.51}$$

where $d_0 = d_1 + d_2 + d_3$ as in the Figure 2.16(a).

For a small aperture with the center of the lens placed at the center of it, the incoming light rays can be considered as paraxial rays. With the use of paraxial ray approximation, the square root terms can be expanded using first two terms of its binomial expansions. Hence, Eq. (2.51) becomes

$$\begin{aligned} d(y_a, z_a) &= d_0 - r_1\left[1 - \left(1 - \frac{y_a^2 + z_a^2}{2r_1^2}\right)\right] + r_2\left[1 - \left(1 - \frac{y_a^2 + z_a^2}{2r_2^2}\right)\right] \\ &= d_0 - \frac{y_a^2 + z_a^2}{2r_1} + \frac{y_a^2 + z_a^2}{2r_2} \\ &= d_0 - \frac{y_a^2 + z_a^2}{2}\left[\frac{1}{r_1} - \frac{1}{r_2}\right] \end{aligned} \tag{2.52}$$

If Eq. (2.52) is put into Eq. (2.48),

$$\begin{aligned} \mathrm{OPL}(y_a, z_a) &= n_m d_0 + (n_l - n_m)\left(d_0 - \frac{y_a^2 + z_a^2}{2}\left[\frac{1}{r_1} - \frac{1}{r_2}\right]\right) \\ &= n_l d_0 - (n_l - n_m)\frac{y_a^2 + z_a^2}{2}\left[\frac{1}{r_1} - \frac{1}{r_2}\right] \\ &= n_l d_0 - n_m \frac{y_a^2 + z_a^2}{2}\left\{\left(\frac{n_l}{n_m} - 1\right)\left[\frac{1}{r_1} - \frac{1}{r_2}\right]\right\} \end{aligned}$$

The term in the curly brackets is exactly the same with Eq. (2.24). Hence, this term can be changed with $1/f$, making the OPL equal to

$$\mathrm{OPL}(y_a, z_a) = n_l d_0 - n_m \frac{y_a^2 + z_a^2}{2f}$$

Optical path length, thereby the index of refraction, should be taken into account in light propagation. This is also valid for Fresnel and Fraunhofer diffraction. However, in these calculations, the medium is taken as vacuum; in other words, the refractive index of the medium is considered as 1. If a different number is used, for instance n_m, then every r_{oa} and r_{as} term has to have a multiplicative constant n_m inside of them, which affects both the amplitude and the phase of the propagated light. Keeping this in mind, if it is assumed that $n_m = 1$, then

$$\mathrm{OPL}(y_a, z_a) = n_l d_0 - \frac{y_a^2 + z_a^2}{2f} \tag{2.53}$$

The complete phase term for Eq. (2.47) is then

$$\phi(y_a, z_a) = k\mathrm{OPL}(y_a, z_a)$$

making the Eq.(2.47) as

$$\begin{aligned} U_{\mathrm{A}}(y_a, z_a) &= U_{\mathrm{A}_0}(y_a, z_a) e^{j\phi(y_a, z_a)} \\ &= U_{\mathrm{A}_0}(y_a, z_a) e^{jk\mathrm{OPL}(y_a, z_a)} \\ &= U_{\mathrm{A}_0}(y_a, z_a) e^{jkn_l d_0} \exp\left\{-jk\frac{y_a^2 + z_a^2}{2f}\right\} \end{aligned} \tag{2.54}$$

where $k = 2\pi/\lambda$. Finally, in Eq. (2.45), the r_s the distance where the complex electric field is calculated. When a converging lens is used, however, the light falls onto the focal point, In this case, all the r_s terms are equal to f, the focal point. Approximating

the obliquity factor to 1 and inserting Eq. (2.54) and Eq. (2.46), the Eq. (2.45) becomes

$$\begin{aligned} U_{P_s} = & \frac{1}{j\lambda f} \exp\left\{jk\left(\frac{y_s^2+z_s^2}{2f}\right)\right\} \\ & \times \iint_{S_\mathrm{A}} \left[U_{\mathrm{A}_0}(y_a,z_a) e^{jkn_l d_0} \exp\left\{-jk\frac{y_a^2+z_a^2}{2f}\right\} \exp\left\{jk\left(\frac{y_a^2+z_a^2}{2f}\right)\right\}\right] \\ & \times \exp\left\{\frac{-j2\pi}{\lambda f}\left(y_a y_s + z_a z_s\right)\right\} dS_\mathrm{A} \end{aligned}$$

The quadratic phases inside the square brackets exactly cancel each other, leaving

$$U_{P_s} = e^{jkn_l d_0} \frac{\exp\left\{jk\left(\frac{y_s^2+z_s^2}{2f}\right)\right\}}{j\lambda f} \iint_{S_\mathrm{A}} U_{\mathrm{A}_0}(y_a,z_a) \exp\left\{\frac{-j2\pi}{\lambda f}\left(y_a y_s + z_a z_s\right)\right\} dS_\mathrm{A} \quad (2.55)$$

If Eq. (2.35) and Eq. (2.55) are compared, it is easily seen that the two expressions are the same except a constant phase delay and the notation difference. Although the calculations are made from the Fresnel approximation, the resulting electric field exhibits a Fraunhofer diffraction pattern. This result proves the claim at the end of Section 2.4.4.

2.6. Image Formation Passing Through an Aperture with a Lens

The basic property of a lens is the ability of image formation. When an object is illuminated, the lens creates a vision, real or virtual, resembling the object under appropriate circumstances. This illumination is called an image of the object and is created by the lens.

Suppose that an object sitting at a distance r_o from aperture spreads monochromatic plane waves and those waves create an image at r_s distance away from the aperture. Let the plane wave emitted by the object be $U_o(y_o, z_o)$ and the image be denoted as $U_i(y_s, z_s)$. By the linearity of the wave propagation [14], the image can be

expressed by the superposition integral as

$$U_i(y_s, z_s) = \iint_{-\infty}^{\infty} \text{asf}(y_s, z_s;\, y_o, z_o) U_o(y_o, z_o) dy_o dz_o \tag{2.56}$$

where $\text{asf}(y_s, z_s;\, y_o, z_o)$ is the impulse response of the system at (y_s, z_s) for a wave sitting on (y_o, z_o). If the imaging system is perfect, then

$$\text{asf}(y_s, z_s;\, y_o, z_o) = \kappa\delta(y_s \pm My_o, z_s \pm Mz_o)$$

where κ is a constant, and $\delta(\cdot)$ is Dirac delta function in which M is the system magnification which makes the sign positive if the image inversion does not occur and negative if the image is inverted.

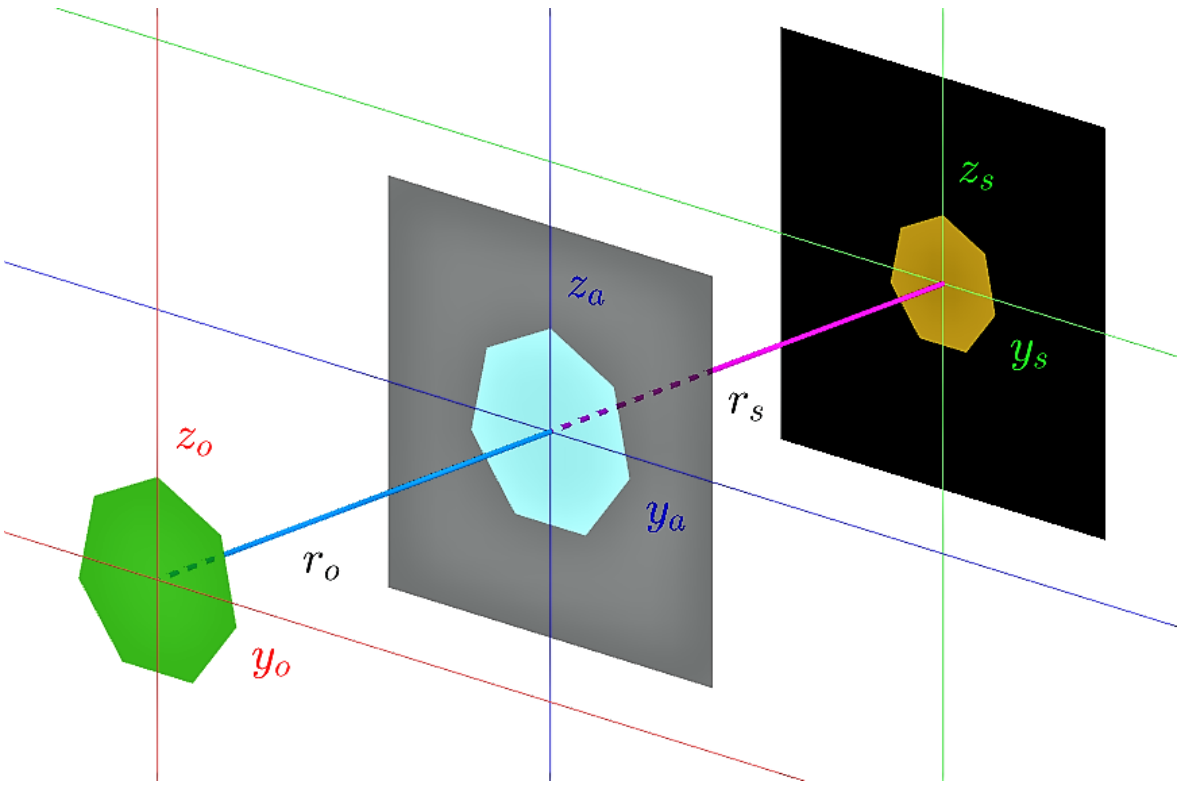


Figure 2.17. The geometry for image formation via an aperture with a lens.

For the general case, the impulse response should be found. Suppose that an object lies on $y_o z_o$ plane at a distance r_o from the aperture. The aperture, having a

positive lens with refractive index n_l, focal distance f and central thickness d_0, sits on $y_a z_a$ plane. The $y_s z_s$ plane is the observing screen which takes place at r_s units away from the aperture. The explained geometry is shown in Figure 2.17.

To find the impulse response, all the system components should be examined under the unit amplitude wave. For the spherical wave before the lens, $U_{bl}(y_a, z_a)$, the paraxial approximation gives [14]

$$U_{bl}(y_a, z_a) = \frac{1}{j\lambda r_o} \exp\left\{ jk \frac{(y_a - y_o)^2 + (z_a - z_o)^2}{2r_o} \right\} \tag{2.57}$$

After passing through the lens, the wave becomes, as in Eq. (2.54)

$$U_{al}(y_a, z_a) = e^{jkn_l d_0} U_{bl}(y_a, z_a) \mathrm{P}(y_a, z_a) \exp\left\{ -jk \frac{(y_a^2 + z_a^2)}{2f} \right\} \tag{2.58}$$

where

$$\mathrm{P}(y_a, z_a) = \begin{cases} 1 & , (y_a, z_a) \in \text{Aperture} \\ 0 & , \text{otherwise} \end{cases}$$

is the pupil window function. At the end, due to the limited aperture and the observing distance, the wave exhibits Fresnel diffraction. Remembering that the all terms containing r_o in Eq. (2.33) comes from the incident wave, the result after Eqs. (2.57) and (2.58) are put into the Fresnel diffraction is

$$\begin{aligned} \mathrm{asf}(y_s, z_s;\, y_o, z_o) = {} & \frac{-K(\theta) e^{jk(r_s + n_l d_0)}}{\lambda^2 r_o r_s} \exp\left\{ jk \frac{(y_s^2 + z_s^2)}{2r_s} \right\} \exp\left\{ jk \frac{(y_o^2 + z_o^2)}{2r_o} \right\} \\ & \times \iint_{-\infty}^{\infty} \mathrm{P}(y_a, z_a) \exp\left\{ jk \left[\frac{1}{r_o} + \frac{1}{r_s} - \frac{1}{f} \right] \frac{(y_a^2 + z_a^2)}{2} \right\} \\ & \times \exp\left\{ -jk \left[y_a \left(\frac{y_o}{r_o} + \frac{y_s}{r_s} \right) + z_a \left(\frac{z_o}{r_o} + \frac{z_s}{r_s} \right) \right] \right\} dy_a dz_a \end{aligned} \tag{2.59}$$

For an easy-to-use expression, some simplifications should be made over Eq. (2.59).

Consider the Eq. (2.24), the lens law. It says that when light passes through a lens, the object location and image location have a relationship. If the object position is r_o and the image location is r_s, the relation is

$$\frac{1}{r_o}+\frac{1}{r_s}=\frac{1}{f}\Longrightarrow\frac{1}{r_o}+\frac{1}{r_s}-\frac{1}{f}=0$$

which is the case in Eq. (2.59). Hence, if the lens law is satisfied, which should be the case for image formation, the first phase term in the integral vanishes.

Consider the Eq. (2.26). Until this point, the expressions are evaluated by dropping the time dependence and considering the complex amplitude of this expression. However, the complex amplitude of a field is not a measurable quantity. Due to that, the irradiance is of more of concern for observations and calculations. Irradiance is the flux or the radiant energy per unit area. It is the time average of square of the real part of the electric field; in other words,

$$I(r_o)=\left\langle(\Re\{u(r_o,t)\})^2\right\rangle_T=\frac{1}{2}U(r_o)U^*(r_o)$$

where $U^*(r_o)$ is the complex conjugate of $U(r_o)$. By the complex conjugation, the imaginary parts and the phase terms vanish. In Eq. (2.59), the phase term containing only y_s, z_s and r_s terms can be neglected if the only concerning and measurable result is the irradiance as this terms will be disappear after the calculation of irradiance.

In this specific case where the lens is a converging lens, for the image to be acceptable as an image, the blur should be very minimum or does not exist at all. That means an object point (y_o, z_o) should be spread very little around the corresponding image point (y_s, z_s). In this case, the phase factor $k(y_o^2+z_o^2)/2r_o$ should change so little that the each point can be considered as falling onto the corresponding image point exactly. This situation must satisfy $(y_o, z_o) = (y_s/M, z_s/M)$ where M is the system magnification expressed as $M = -r_s/r_o$ in which the minus sign comes due to the image inversion. After this conversion is applied to the phase factor depending only on

y_o, z_o and r_o, the new phase term can be neglected via the reasoning used to eliminate the phase term containing only y_s, z_s and r_s [14].

Again with an obliquity factor reasonably set to 1 and taking the pure phase term out of the integral which will be added at the end, Eq. (2.59) becomes

$$\begin{aligned} \mathrm{asf}(y_s, z_s;\, y_o, z_o) \approx \frac{1}{\lambda^2 r_o r_s} \iint_{-\infty}^{\infty} \mathrm{P}(y_a, z_a) \\ \times \exp\left\{-jk\frac{\left[y_a\left(y_s - My_o\right) + z_a\left(z_s - Mz_o\right)\right]}{r_s}\right\} dy_a dz_a \end{aligned} \tag{2.60}$$

which is, apart from the extra multiplier $1/\lambda r_o$, the Fraunhofer diffraction pattern centered on the image coordinates $(y_s, z_s) = (My_o, Mz_o)$.

The coordinates can be normalized for a simpler expression. If the variables are changed as

$$\tilde{y_a} = \frac{y_a}{\lambda r_s}, \quad \tilde{z_a} = \frac{z_a}{\lambda r_s}, \quad \tilde{y_o} = My_o, \quad \tilde{z_o} = Mz_o$$

the result becomes

$$\begin{aligned} \mathrm{asf}(y_s, z_s;\, \tilde{y_o}, \tilde{z_o}) = -M \iint_{-\infty}^{\infty} \mathrm{P}(\lambda r_s \tilde{y_a}, \lambda r_s \tilde{z_a}) \\ \times \exp\left\{-j2\pi\left[\tilde{y_a}\left(y_s - \tilde{y_o}\right) + \tilde{z_a}\left(z_s - \tilde{z_o}\right)\right]\right\} d\tilde{y_a} d\tilde{z_a} \end{aligned} \tag{2.61}$$

Eq. (2.61) does depend only on the difference of the coordinates. Hence

$$\mathrm{asf}(y_s, z_s;\, \tilde{y_o}, \tilde{z_o}) = \mathrm{asf}(y_s - \tilde{y_o}, z_s - \tilde{z_o})$$

As the last modification, consider the object and the image. If the imaging system is perfect, then the image should be magnified version of the object on the

corresponding image coordinates; in other words [14],

$$U_i(y_s, z_s) = \frac{1}{M} U_o\left(\frac{y_s}{M}, \frac{z_s}{M}\right) = \frac{1}{M} U_o(y_o, z_o) = \frac{1}{M} U_o\left(\frac{\tilde{y_o}}{M}, \frac{\tilde{z_o}}{M}\right)$$

Together with this result and Eq. (2.61), Eq. (2.56) becomes

$$\boxed{\begin{aligned} U_i(y_s, z_s) &= \frac{-e^{jk(r_s + n_l d_0)}}{M^2} \iint_{-\infty}^{\infty} \mathrm{asf}(y_s - \tilde{y_o}, z_s - \tilde{z_o}) \left[\frac{1}{M} U_o\left(\frac{\tilde{y_o}}{M}, \frac{\tilde{z_o}}{M}\right)\right] d\tilde{y_o} d\tilde{z_o} \\ &= \left(\frac{-e^{jk(r_s + n_l d_0)}}{M^2}\right) \mathrm{asf}(y_s, z_s) * \left[\frac{1}{M} U_o\left(\frac{y_s}{M}, \frac{z_s}{M}\right)\right] \end{aligned}} \tag{2.62}$$

where $*$ is the convolution operation and

$$\boxed{\mathrm{asf}(y_s, z_s) = -M \iint_{-\infty}^{\infty} \mathrm{P}(\lambda r_s \tilde{y_a}, \lambda r_s \tilde{z_a}) \exp\left\{-j2\pi \left(\tilde{y_a} y_s + \tilde{z_a} z_s\right)\right\} d\tilde{y_a} d\tilde{z_a}} \tag{2.63}$$

is the *amplitude spread function* [7] (asf) affected from diffraction. This very important result says the effect of the diffraction is the convolution of the ideal image with Fraunhofer diffraction pattern of the pupil window function or with the Fourier transform of the pupil window function.

2.6.1. Coherent and Incoherent Illumination

Until now, it is assumed that the object is illuminated via monochromatic light source. However, in reality it is nearly impossible to obtain a monochromatic light source. Even the lasers emit light in narrowband around a center frequency. Hence the discussion should be extended to the polychromatic light. For this, the concepts *coherent* and *incoherent* should be introduced.

Consider the illumination where the complex space amplitudes at all object points vary in harmony. In this case, although the actual phases are different, they change

with time in a correlated way; in other words, the phase difference between the object points remains constant. Having also the same wave shape and the frequency, this illumination is said *spatially coherent*, or shortly *coherent*. Such illumination occurs when the light seems to coming from a single point. This can be, for instance, a laser beam or a powerful bulb light whose light rays are first passed from a pinhole. Another type of illumination is the one that the complex amplitudes and phases of the waves change independent from each other. Such illumination is called *spatially incoherent*, or simply *incoherent*. When the illumination is incoherent, the light waves have neither the same frequency nor the phase and amplitude. The example for this illumination can be the sun light.

When the illumination is coherent, the amplitudes vary in unison. Hence, the resulting wave amplitude should be obtained by the addition of the complex amplitudes. The field is linear in complex amplitudes [14]. Hence, under the coherent illumination,

$$U_i(y_s, z_s) = \iint_{-\infty}^{\infty} \text{asf}(y_s - \tilde{y_o}, z_s - \tilde{z_o}) U_o(\tilde{y_o}, \tilde{z_o}) d\tilde{y_o} d\tilde{z_o} \tag{2.64}$$

in which the constant amplitude and phase terms are not written. Under the incoherent illumination, since the incident light rays do not have a common component, their addition should be done in terms of power intensity [14]. Thus,

$$\boxed{I_i(y_s, z_s) = \iint_{-\infty}^{\infty} \left|\text{asf}(y_s - \tilde{y_o}, z_s - \tilde{z_o})\right|^2 I_o(\tilde{y_o}, \tilde{z_o}) d\tilde{y_o} d\tilde{z_o}} \tag{2.65}$$

where $I_i(y_s, z_s)$ is the image intensity, $I_o(y_o, z_o)$ is the object intensity, and $|\text{asf}(\cdot)|^2$ is the *point spread function* (psf). Again, the constant term in front of the integral is omitted.

2.6.2. Optical Transfer Function and Modulation Transfer Function

The presence of convolution operation brings more convenient ways to calculate the resulting image of an object. Convolution theorem says that the convolution of two functions results in a function whose (inverse) Fourier transform is equal to the multiplication of (inverse) Fourier transforms of two convolved functions. Consider the Eq. (2.65),

$$I_i(y_s, z_s) = \iint_{-\infty}^{\infty} \mathrm{psf}(y_s - \tilde{y_o}, z_s - \tilde{z_o}) I_o(\tilde{y_o}, \tilde{z_o}) d\tilde{y_o} d\tilde{z_o} = \mathrm{psf}(y_s, z_s) * I_o(y_s, z_s)$$

If the convolution theorem is applied,

$$I_i(y_s, z_s) = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\mathrm{psf}(y_s, z_s)\right\} \mathcal{F}\left\{I_o(y_s, z_s)\right\}\right\} = \mathcal{F}^{-1}\left\{\mathrm{PSF}(Y_s, Z_s) \mathcal{F}\left\{I_o(y_s, z_s)\right\}\right\}$$

where (Y_s, Z_s) is the spatial frequency coordinates, $\mathcal{F}\{\cdot\}$ is the Fourier transform operator, $\mathcal{F}^{-1}\{\cdot\}$ is the inverse Fourier transform, and $\mathrm{PSF}(Y_s, Z_s)$ is the Fourier transform of the point spread function;

$$\mathrm{PSF}(Y_s, Z_s) = \iint_{-\infty}^{\infty} \mathrm{psf}(y_s, z_s) \exp\left\{-j2\pi(y_s Y_s + z_s Z_s)\right\} dy_s dz_s$$

If the $\mathrm{PSF}(Y_s, Z_s)$ is normalized by the value at the frequency $(0, 0)$,

$$\mathrm{OTF}(Y_s, Z_s) = \frac{\displaystyle\iint_{-\infty}^{\infty} \mathrm{psf}(y_s, z_s) \exp\left\{-j2\pi(y_s Y_s + z_s Z_s)\right\} dy_s dz_s}{\displaystyle\iint_{-\infty}^{\infty} \mathrm{psf}(y_s, z_s) dy_s dz_s}$$

the result is known as the *optical transfer function* (otf), and the modulus of is, $|\mathrm{OTF}(Y_s, Z_s)|$ is called the *modulation transfer function* (mtf).

3. HUMAN VISION SYSTEM

Our visual system is the primary information source of the outside world for us. The importance of this system can be understood from the fact that the sensor organ of visual system is the most advanced one and is the only sensory organ directly connected to the brain. Among our five sensory systems; namely visual, auditory, olfactory, tactual and gustatory systems, the visual system is the most complex one in terms of the way of processing the sensor organ signals of brain and sensory organs of the systems.

In the simplest meaning, eye, the visual system sensory organ, is a light-sensitive structure providing the vision, receiving and processing visual details, and performing some other biological tasks that are not related with the vision, such as adjustment of circadian rhythms [15, 16], among the livings. Since the need for eye and its capabilities are different for every species, its structure, mission and working way vary from species to species. Scientists classify eyes into two different categories, simple eye and compound eye [17]. Simple eyes have one chamber for light to pass and reach the sensor cells, whereas compound eyes can be thought as comprised of "many simple eyes". Each category consists of five types of eye, making totally 10 different eye types among living beings. Human eye is in the type that can perform corneal refraction in simple eye category [17].

This sensory organ is actually very powerful. Its dynamic range to notice different brightness values is reported as 10 000 000:1 [18].[3] A perfect eye can distinguish 0.35 mm-wide white and black lines from 1 m distance [19] and even can detect a single photon under right circumstances [20]. To be able to understand the power of eye and where this power comes from, its anatomy should be examined and understood.

[3]The actual range is bigger because in the correspondent study, the intensities higher than the full moon are not considered.

3.1. Human Eye Anatomy

As it can be seen from Figure 3.1, human eye consists of many layers, muscles, fibers, veins and neural cells, all of which have many different structure, working principle and responsibilities so much detailed and complicated that covering all requires a separate study. Since the main concept of this study is vision, only the parts that are related with vision formation are introduced here.

The eye ball is nearly a sphere. It sits within a pliable but strong shell called *sclera*. Sclera is white and opaque (it's shown as a grayish shell around the eye in Figure 3.1) except in front of the eye. That part is transparent and called *cornea*. It is the first part interacting with the incoming light, and the strongest convex component of the eye [7]. It is equipped with high number of nerve cells perceiving touch and pain signals. Hence, any disturbance, even the smallest one, caused by external agents such as dust triggers the reflexes of blinking and producing more tears [22]. Cornea has a thickness about 0.449 mm, but this varies due to the fact that it has different inner and outer curvatures: The front radius of cornea is about 7.259 mm and the back radius is about 5.585 mm [23].

At the back of the cornea lies the *anterior chamber*. This chamber is filled with water-like liquid called *aqueous humor*, and has a depth around 2.794 mm [23]. This liquid has two important responsibilities: to nourish the anterior part of the eye and to redirect the light coming from the cornea a little bit more [7]. Deeply buried into the aqueous humor is the *iris*. The word comes from the Latin word for "rainbow"and the messenger of Gods *Iris* in Greek mythology. This part of eye carries pigments, giving the colors of our eyes. Iris is the aperture controller of the eye to limit the amount of entering light via *pupil*. By the radial and circular muscles, respectively, iris can expand the pupil up to 8 mm diameter and can close it until the diameter of pupil becomes 2 mm [7, 22].

Crystalline lens, suspended by *zonular fibers*, is placed just behind of the pupil. Its diameter is around 9 mm and its thickness is about 4–5 mm. Its front part (the

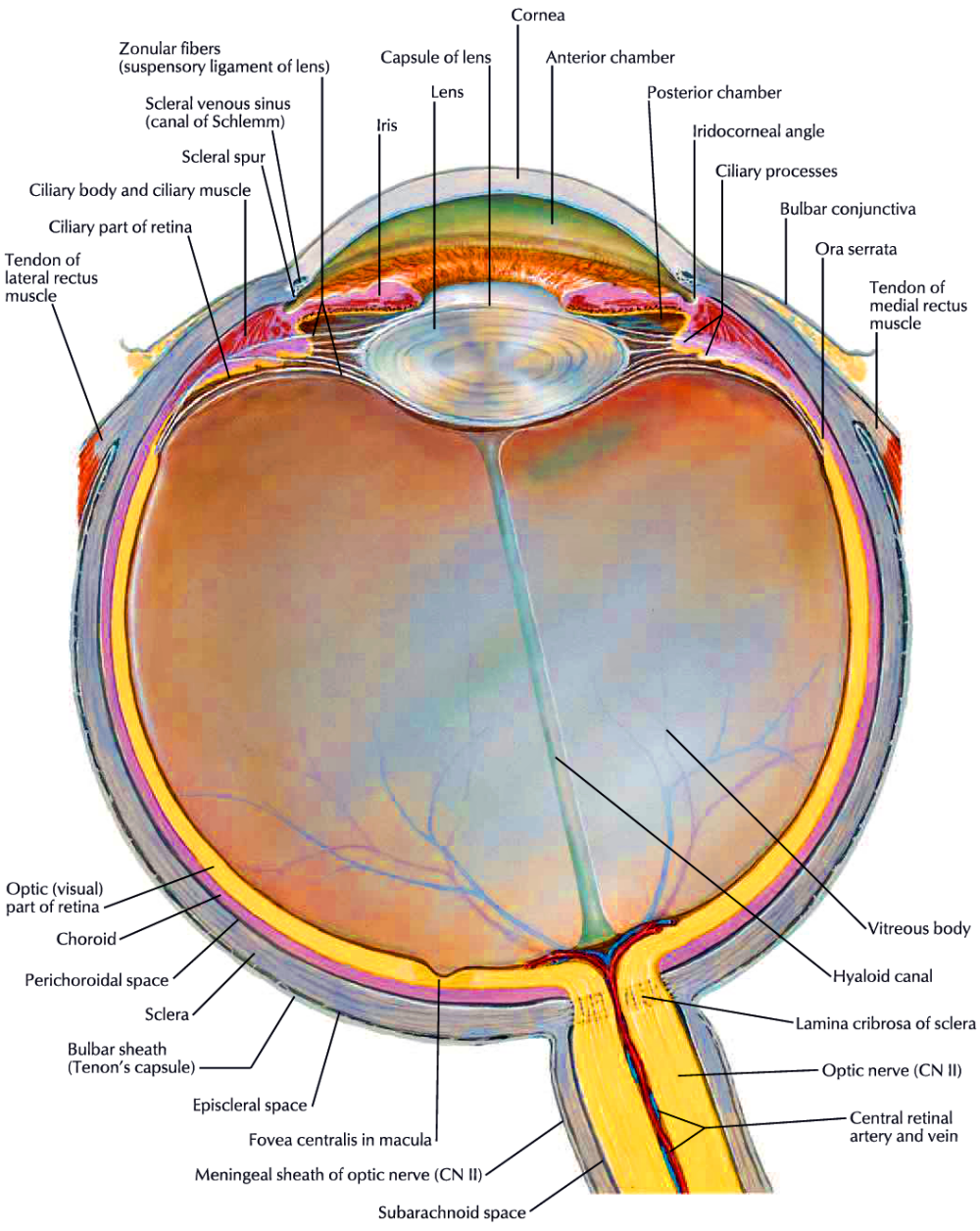


Figure 3.1. Human eye anatomy. Adopted from [21].

part at the exit of the pupil) is flatter than the its back part. Lens has a multi-layered structure and elastic shell. The crystalline lens has some distinctive features: Its growth in size is continuous. The lens remains flexible for 40–50 years, after which it causes inevitable sight problems. Its structure gives it the ability to bend the light inside of it; i.e., the crystalline lens has gradient-index system. Moreover, by *ciliary muscles*, it can change its shape, hence the focal length of it, making easier the accommodation [7].

Behind the lens is another chamber called *posterior chamber*, filled with *vitreous humor*. This gel-like matter contains proteins and cellular remains, and supports the eye ball to cover its shape. The walls of posterior chamber has two more layers other than sclera. The first layer inside of sclera is *choroid*. It is rich of blood vessels and a pigment called *melanin* which helps eye with absorbing the light that cannot be focused properly and scatters inside of the eye, being able to create irrelevant and non-sharp images [7]. The second layer which is responsible for absorbing light and converting it into the electrochemical signals to create image is *retina*.

Retina, which itself has three different cell layers, has a thickness about 0.1–0.5 mm [7]. The first layer consists of *ganglion cells* which directs the electrical signals to the brain by optic nerves. The second layer is filled with *bipolar cells*, which sets connection between ganglion cells and light sensitive cells, *amacrine cells* connecting bipolar cells to ganglion cells, and *horizontal cells* which enables transmission between bipolar cells and the sensor cells. The last layer contains *rod cells* and *cone cells*, the structure which are responsible to sense the incoming light.

Around the center of the retina lies a yellow-colored region with 2.5–3 mm diameter, *macula*. Here, the densities of the sensor cells rods and cones are twice of the densities at the other regions of retina. At the center of the macula, a region called *fovea centralis* (or *fovea*, in short) lies. This region has 0.3 mm diameter, does not contain any rod cell and is the region where the formed image is the sharpest. In fovea, the cones are thinner than the cones placed in other parts and the most densely packed compared to other regions of retina. Fovea is so important that each cone in fovea is connected to one nerve cell, a feature that is not seen anywhere in retina, and without

fovea, the eye loses its visual capability around 90–95% [7]. Due to this importance, eye ball tries to fall the incoming light onto the fovea by eye movements to create the sharpest image possible.

Rods and cones are different in many aspects such as mission, structure, size, shape, sensitivity, number in different regions, number of connected cells and so on. These differences can be listed shortly as below:

- There are about 125 million rod and cone cells totally in each eye which are unequally distributed over the retina. Rods outnumber the cones; there are only 6–7 million of cone cells in retina [7].
- Rods are 2 μm in diameter; whereas cones are 1.5–3 μm wide in fovea and 6 μm wide outside of the fovea [7].
- Rods are longer and more slender compared to cones that are shorter, wider and tapered.
- Rods are very sensitive, hence they are responsive to dim light and creates the same response in bright light after some threshold value for brightness. Cones give reactions to bright light and responsible for color vision and resolving fine details [22].
- All rods contain the same and one type of pigment responsive to the light. Cones, however, are divided into three types according to the pigment they carry differentiated by the wavelength they respond.

On the retina, one region is receptor cell free and unresponsive to the light; that is, this particular region does not contain any rod or cone cell. Consequently, when the light falls onto this region no chemical reaction is triggered and no image is formed. This area from which blood veins and optic nerves leave the eye ball is called *blind spot*.

3.2. Image Formation in Human Eye

As mentioned before, eye can be considered as a system with two positive lenses. The first lens of this system interacting with light is cornea. Although it is a part of the nearly-spherical structure sclera, it is a little bit flattened to prevent the possible spherical aberrations [7]. Cornea has an index of refraction about 1.376. Consequently, the most of the optical power of eye comes from the air-cornea interface. Since the difference between indexes of refraction of two successive media in human vision system is the biggest on this interface, most of the bending of light occurs here. The reason that we cannot see sharply under water also comes from the difference between two media's indexes of refraction. Since it is 1.333 for water, the difference between indexes of refraction of water and cornea is smaller than that of between air and cornea. Hence, the light cannot be bent adequately [7].

Optical power of cornea on air-cornea interface is about +51–+52 D. However, the back face of the cornea forms a concave lens and the next refractive part after cornea has a lower index of refraction. Hence the optical power at the back of cornea is negative, specifically around -7 D [23]. So totally, the cornea has a optical power around +44–+45 D.

The incoming light is bent in cornea and then passes through aqueous humor. The index of refraction of this liquid is around 1.336 [7]. Its index is close to that of water's, but not the same because it also carries some nutritious matter for front part of eye. Aqueous humor directs the light towards pupil that can change size, and after pupil, the light interacts with crystalline lens.

Crystalline lens comes just after the iris. Incoming light enters through the lens and is bent to be further focused. It has some distinctive features for it to focus the light properly. First, the lens has a layered structure and can be likened as "an onion formed of around 22 000 very fine layers" [7]. Due to this layered structure, the light seems to be bent discretely and step by step instead of being bowed just once and necessary enough. Also, as mentioned before, the lens has gradient-indexed (GRIN)

system. Index of refraction is different for different points on the crystalline lens: On center, it is around 1.406, it decreases polynomially towards the end, and at the very edges it becomes 1.386 [7]. This feature of lens allows light to follow an arc-shaped path inside the lens and to occur a part of focusing the light in the lens.

Optical power of the lens is about +19 D when it is surrounded by air [7]. This optical power is distributed for front part and back part of the lens because, as mentioned before, the front part and back part have different radii of curvature. For front part of the lens, the optical power is about +8.1 D, and for the backside of the lens, it is about +10.9 D [23]. However, since it is enclosed by aqueous humor and vitreous humor, the focal length, hence the optical power, is different [7].

Another property of crystalline lens is accommodation. It is basically the fine focusing duty of the lens. Lens is suspended behind the iris in an aligned position with pupil. This suspension is done via zonular fibers attached to the ciliary muscles. When these muscles are completely relaxed, the fibers are in stretched phase to hold the lens relatively flat shape. In this case, the light coming from an object at infinity is focused on the retina, which is accepted as 5–6 m and more for all practical purposes [7]. For an unaccommodated eye, the object position whose light can be focused on retina is called *far point*. As the object gets closer than the far point, ciliary muscles start to contract, zonular fibers lose their tension. In this case, the crystalline lens starts to expand towards the posterior chamber via its elastic forces and its surface becomes rounder. By doing so, lens decreases its focal point to make the light fall onto retina. The closer the object comes, the more the crystalline lens accommodates. After a point, called *near point*, the lens cannot change shape to focus light onto retina. The near point and age are proportional: For normal eye, near point lies on 7 cm distance for a teenager, about 12 cm for a young adult, 28–40 cm for a middle-aged person, and around 1 m for a 60-year-old [7].

Crystalline lens also affects our vision by its blocking nature. Human vision systems is sensitive to light whose wavelength lies on the range from 390 nm to 780 nm. This range is called *visual spectrum*. Below of this range is called *ultraviolet* and

the spectrum above the largest wavelength of the visual spectrum is called *infrared*. Although, generally, the visual spectrum is defined as the range from 390 nm to 780 nm, some studies extend this range from 310 nm to 1050 nm. For the ultraviolet spectrum, our visual range is limited by crystalline lens since it absorbs the ultraviolet light. Some patients whose crystalline lens had been surgically removed reported that they can "see" X-rays or ultraviolet light, indicating that the removal operation increased their ultraviolet light sensitivity [7].

After leaving the crystalline lens, the light interacts with vitreous humor. This gelatinous matter consists of collagen and hyaluronic acid, both of which have protein-based origins. Supporting the eye ball, this liquid has an index of refraction about 1.337, close to that of aqueous humor's [7].

At the last stage of the optical par of vision, the light falls onto retina. The retina, as mentioned before, consists of three different cell layers; ganglion cells, bipolar cells and sensor cells, in the order that light interacts with. Ganglion cells are responsible for the signals produced by sensor cells to be transmitted to brain. Bipolar cells carry information between sensor cells and ganglion cells. Sensory cells detect the light and turn this information into the electrochemical signals according to the features of the light.

In human vision system, there are two types of sensor cells, rod cells and cone cells. Rods are responsible for vision in dim light. Cone cells are active in bright light conditions. Cone cells have the duty to distinguish color and fine details. Rods, however, do not have those abilities. This is why our night vision is not both sharp and colorful under dim light conditions.

Rod cells contain one type of pigment whose peak excitement occurs with 498 nm light. While rod cells are of one type, cone cells are divided into three different types according to the pigments they carry and the wavelengths they respond to and absorb most: "red", "green", and "blue" cone cells. This naming is, however, is not truly descriptive for the cells due to the following reasons [22]:

- When a color name is pronounced for an object, it is the color of the light that it reflects. However, in case of the receptor cells, the mentioned colors are the colors of absorbed light by the receptor cells.
- The wavelengths at which the peak sensitivities of the cone cells occur are about 420–430 nm, 530–535 nm, and 560–565 nm. Those wavelengths actually correspond to violet, blue-green and yellow-green colors, respectively.
- If each type of cone cell could be excited separately, the perceived colors would be violet, green and yellowish-green.

The sensitivity graphic obtained from the study [24] is shown in Figure 3.2. Note that the peak sensitivity of rods correspond to green light. Overall, the green light is the most stimulant light for human vision system. This fact causes humans to perceive the objects "greener" than the actual color of the light that the objects reflect, and enables human visual system to become more sensitive to green light and distinguish the luminance change better at green color. This is the reason for that the night vision cameras give output as green images, and for that the most widespread color filter array, Bayer color filter array, to filter the light according to their wavelengths for pixel sensors uses the ratio 1:2:1 for colors red, green and blue, respectively.

The area on retina which the light falls onto is called *receptive field*. Since the light only stimulates the cells on the receptive field, only the ganglion cells on this area are active. However, this activity is not activation of all cells lying on the receptive field. When a small circular area on the center of this field is excited by light, the center ganglion cells produce electrical signals. When the light on the center is gone, however, the cells on area around the center produce electrical signals while the cells on the center do not. Those cells are named *on-center* and *off-center cells* by the discoverer of those cells, Stephen Kuffler [22].

Although there are three and only three types of cone cells in human retina, eye does not work completely restricted to those three colors. Ewald Hering theorized that the human vision system relies on three opponent processes; one for red-green sensation, one for yellow-blue sensation and one for black-white sensation [22]. This

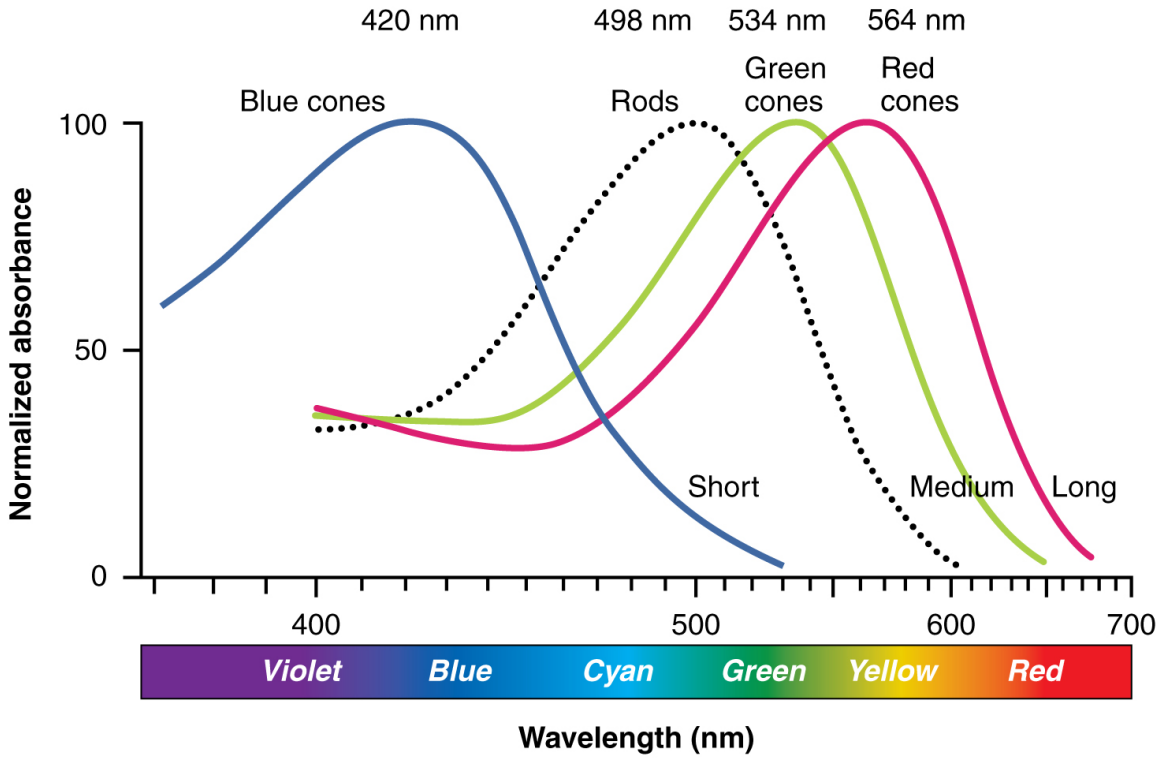


Figure 3.2. Human photoreceptor cell absorbency according to the wavelengths [24].

theory also brings explanation for some perceived colors like brown which cannot be produced by mixing red, green and blue lights on a surface. He suggested that the color perception depends not only the incoming light from an object but also brightness of incoming light from the surrounding of object. This explains, for example, that the color brown is not a result of linear addition of red, green and blue lights but actually the yellow light whose perceived brightness is less that its surroundings.

This theory was somehow supported by Gunnar Svaetichin's findings in 1956 [22]. He found that the horizontal cells in human retina are three types, and work in a similar on-center and off-center manner of ganglion cells: One type, that he called *L type*, is stimulated by light regardless of its wavelength. The *r-g type* cells are excited for red-green light with maximum hyperpolarization by green light and maximum depolarization by red light. The last type, called *y-b type*, are activated with yellow-blue light with maximum hyperpolarization by blue and maximum depolarization by yellow light [22]. These findings supported the idea of Hering, and led the creation of L*a*b*

color space representation.

After the light falls onto retina and resulting electrical signals are produced, the ganglion cells transmit the those signals to brain and the brain processes them to obtain what we see. Since both the brain part of the vision is not completely understood and is not relevant with this work, it will not be explained here.

4. RELATED WORK

Together with the understanding and mathematical modeling of human eye, people have tried to develop computational and model-based solutions for problems related with the human vision. Being a developing and solution-offering area, image processing and its means have been used to correct or recover visual acuity. In general, those image processing methods to compensate vision deficiency for visually handicapped people can be divided into two categories: detail enhancement methods and deconvolutional methods.

4.1. Detail Enhancement Methods

Detail enhancement methods refer to the techniques that enhance the image features desirable for visual acuity patients or that help such people to realize and distinguish what they see. Those methods do not try to fully correct the corrupted vision; instead, they try to increase the recognition rate of some features or objects in the image such as faces, texts, main object edges and so on. This is generally done by enhancing the high frequency content of image because of that visual acuity depends on the eye's ability to resolve object boundaries, or remapping the image to carry the information-rich areas to more clear line of sight of the person. Hence, this category of methods better serves people with impairments such as low light entry like cataract, reduced information region disorders like scotoma, macular degeneration and visual acuity decrease like mild myopia, astigmatism and so on.

4.1.1. Adaptive Filtering

In the signal processing area, the first work studying vision correction via image processing was proposed by Peli and Peli [25]. They locally separated image into two components; low-frequency image and high-frequency image. Since the work [25] dealt with the low vision impairments like cataract, they enhanced high-frequency content according to the low-frequency part of the local image, and then combined the processed

images into one image.

To perform the image enhancement, they applied a method called spatial adaptive filtering [2]. To apply that, first, authors of [25] extract the low frequency content of the local image, $I_L(x, y)$, by averaging it with a $(2N_1 + 1) \times (2N_2 + 1)$ mask; in other words,

$$I_L(x, y) = \frac{1}{W_D} \sum_{k=x-N_1}^{x+N_1} \sum_{l=y-N_2}^{y+N_2} I(k, l) W_N \tag{4.1}$$

where $I(x, y)$ is the original image. W_D and W_N depend on the applied averaging type. If the average is simple mean [2], then

$$W_D = (2N_1 + 1)(2N_2 + 1) \quad \text{and} \quad W_N = 1$$

If weighted averaging is applied, then

$$W_D = \sum_{k=-N_1}^{N_1} \sum_{l=-N_2}^{N_2} G(k, l) \quad \text{and} \quad W_N = G(k - x, l - y)$$

where $G(k, l)$ is a Gaussian window with controllable variance [25]. After obtaining the low frequency content, the high frequency part, $I_H(x, y)$, is obtained by subtracting the low frequency image from the original image; in other words,

$$I_H(x, y) = I(x, y) - I_L(x, y) \tag{4.2}$$

Being separated, $I_L(x, y)$ and $I_H(x, y)$ are processed in different ways for different purposes. Highpass content is handled in a way that both it is amplified to stress details and it is prevented from being clipped due to the amplification. It is done by amplifying the high frequency content with a scalar-valued lowpass-dependent function. That is,

$$I'_H(x, y) = I_H(x, y) H(I_L) \tag{4.3}$$

where $I'_H(x, y)$ is the high frequency content of the enhanced image and $H(I_L) > 1$ is the scalar amplifier function of local luminance mean $I_L(x, y)$.

Together with highpass content, low frequency image is also modified for enhancement. It is processed in a nonlinear way that the local mean of the lowpass image is shifted towards the midrange of the dynamic range of the gray image which is [0 255]. This is achieved by

$$I'_L(x, y) = [I_L(x, y) - 128]\, L(I_L) + 128 \tag{4.4}$$

where 128 is the midrange of the standard display range, and $L(I_L)$ is scalar-valued function of $I_L(x, y)$ within the range [0 1]. As the last step, two resultant components $I'_L(x, y)$ and $I'_H(x, y)$ are summed to obtain the enhanced image. An example of the result of this method is given in Figure 4.1. Note that although some areas like the hair of the child becomes indistinguishable from the background, some features can be seen in more details, the bird figure on the coat, for instance.

(a) (b)

Figure 4.1. Example of adaptive filtering. (a) is the original image, (b) is the adaptive filtered image with weighted averaged by 5×5 Gaussian with $\sigma = 2$, lowpass gain 0.9 and highpass gain 7.

To evaluate the effects of spatially adaptive filter on reading, people employed it to several image types. Authors in [26] applied the method to text images. However, they found that the increase rate in reading speed is not substantial in a sample containing people with visual acuity between 20/200 and 20/800, and central field loss (central scotoma). Moreover, since the text images are white letters on a black background, they came to conclusion that the increase rate may be due to the increased luminance in the images instead of contrast increase in the text. In another study looking at the still and motion images, authors found that the visually impaired people prefer the moderately enhanced images helping them to see more clear object boundaries [27].

4.1.2. Wideband Enhancement

Alternative to the narrowband enhancement in [25], [28] offered a wideband enhancement for the people with vision defect. In this article, authors search for a specific feature, edges in this case, in different scales by bandpass filtering the image in frequency domain [29]. The filters in different scales, having one octave bandwidth and separated by one octave in frequency domain, are applied to four different channels; red color channel, green color channel, blue color channel and luminance channel. The bandpass filters, $C_i(u, v)$, are given by

$$C_i(u,v) = \begin{cases} 0.5\left[1 + \cos(\pi \log_2 r_{\mathrm{f}} - \pi i)\right] & , 2^{i-1} \leq r_{\mathrm{f}} \leq 2^{i+1} \\ 0 & , \text{elsewhere} \end{cases} \tag{4.5}$$

where i is the scale number (order of the filter), u and v are spatial frequency Cartesian coordinates, and r_{f} is the radial frequency coordinate with $r_{\mathrm{f}} = \sqrt{u^2 + v^2}$. Images from this filtering are converted into spatial domain from the frequency domain without resizing. Then, obtained different-scale images, $I_i(x, y)$, are compared with a threshold. Thresholding can be either hard-thresholding to binarize the image or soft-thresholding according to the human visual model and the patient's vision acuity. That is, for hard-

thresholding,

$$TI_i(x,y) = \begin{cases} +1 & ,\text{if } I_i(x,y) \geq 0 \\ -1 & ,\text{if } I_i(x,y) < 0 \end{cases} \tag{4.6}$$

and for soft-thresholding,

$$TI_i(x,y) = \begin{cases} +1 & ,\text{if } I_i(x,y) \geq +|th_i| \\ 0 & ,\text{if } -|th_i| < I_i(x,y) < +|th_i| \\ -1 & ,\text{if } I_i(x,y) \leq -|th_i| \end{cases} \tag{4.7}$$

where $TI_i(x,y)$ is the thresholded image in scale i, th_i is the threshold value in scale i which can be derived according to [30].

After thresholding, the scales are compared with each other in a XOR-like process, namely phase congruency test. If a feature in one scale coincides with the same feature in a different scale, then that feature is kept for enhancing the original image. Mathematically speaking,

$$FI(x,y) = \begin{cases} +1 & ,TI_i(x,y) = 1,\ \forall i \\ 0 & ,\text{otherwise} \\ -1 & ,TI_i(x,y) = -1,\ \forall i \end{cases} \tag{4.8}$$

where $FI(x,y)$ is the feature image. Obtained the features in whole image, the algorithm then adds those features to or subtracts those features from the corresponding pixels to enhance the edges, an operation done according to the edges being "dark" edges or "bright" edges. The enhancement level can be controlled by the patient. Flowchart of the algorithm is shown in Figure 4.2.

In [28], it is found that the wideband enhancement method is preferred by the subjects over narrowband enhancement. However, it is reported that if the level of

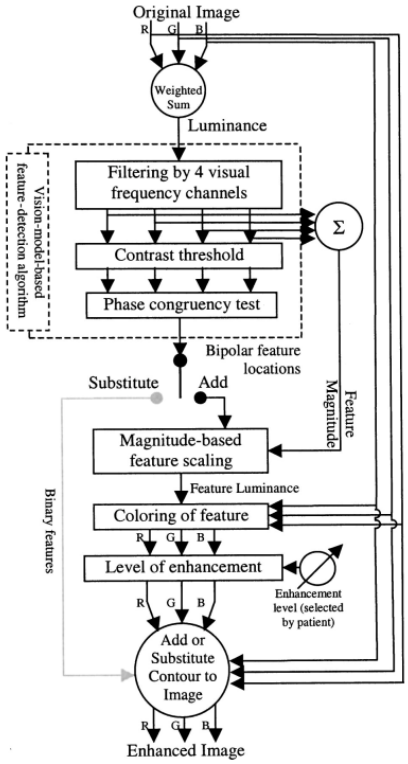


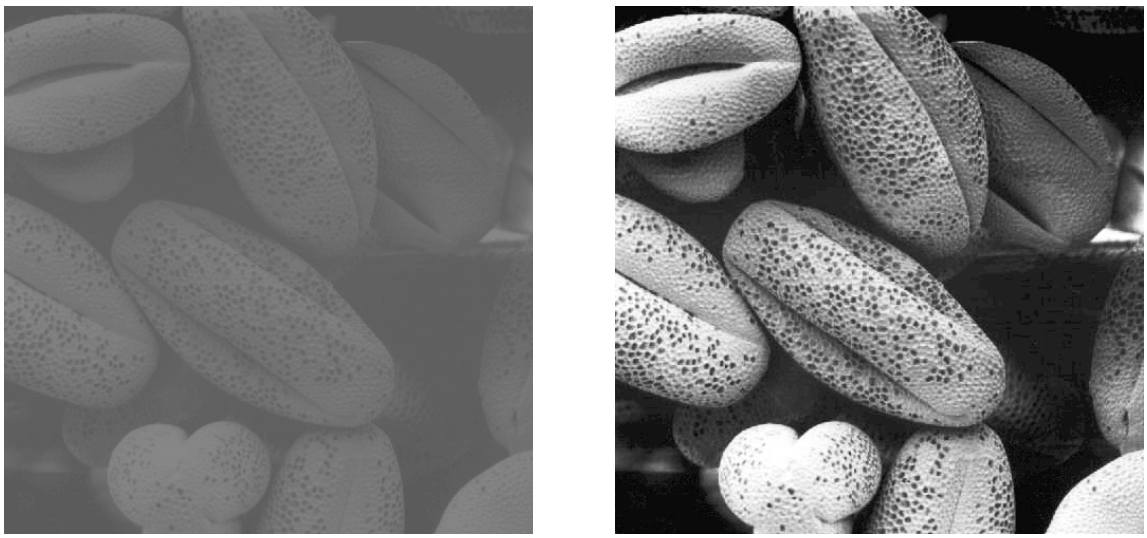
Figure 4.2. Flowchart of the wideband enhancement. Adopted from [28].

enhancement is too high, then patients do not prefer wideband enhancement. Because the method enhances edges and boundaries together with the noise associated with the still or motion picture, no matter whether it is hard-thresholded or soft-thresholded.

4.1.3. Histogram Hyperbolization

The work [31] studied different approaches and methods to enhance images for visually impaired people. As the first approach, contrast manipulating methods were reviewed. Those methods are popular and well-known methods in image processing; namely gamma changing, gray level redistribution, thresholding, histogram equalization and histogram hyperbolization. Among them, the histogram hyperbolization is reported as the best method to enhance image for vision degradation patients [31].

Proposed by Frei, histogram hyperbolization benefits from the fact that pleasing pictures with well-adjusted contrast generally have histograms shaped similar to



(a) (b)

Figure 4.3. Example of histogram hyperbolization. (a) is the original image, (b) is the histogram hyperbolized image with $c = 0.5$.

Rayleigh, exponential or hyperbolic functions [32]. Hence, histogram hyperbolization tries to make the shape of the histogram of the picture like a hyperbolic function. It is done for an image $I(x, y)$ by the transformation $J(I)$ as

$$\boxed{J(I) = c\left[\exp\left(\ln(1 + 1/c)\int_0^I p(I)dI\right) - 1\right]} \tag{4.9}$$

where $p(I)$ is the empirical probability density function of pixel brightness values and c is an experimentally-determined constant. For discrete-natured digital images, the integration becomes summation, hence, the histogram of the resulting image does not have a perfect hyperbolic shape [32]. However, the image does still have better contrast than the original image. Figure 4.3 shows an example result of histogram hyperbolization.

4.1.4. Unsharp Masking and Extremum Sharpening

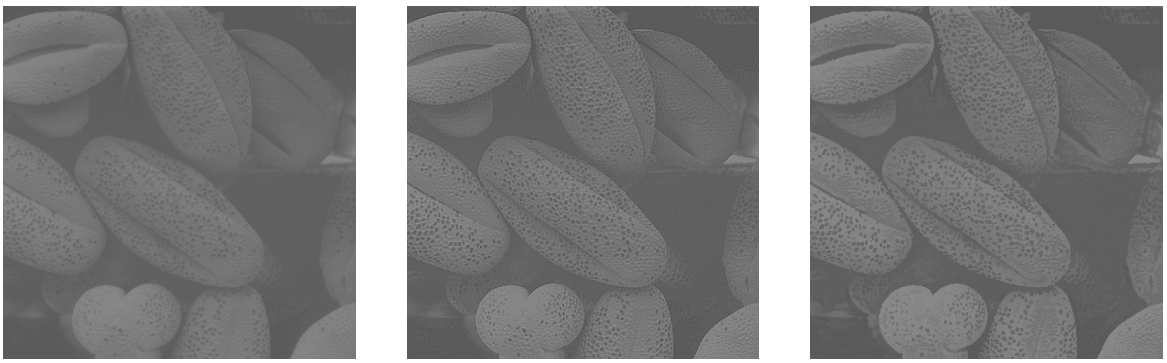
Another approach that [31] investigated is spatial filtering. Together with the aforementioned spatial adaptive filtering [2], unsharp masking and extremum sharpen-

ing were also examined to help the visually impaired.

Unsharp masking is a relatively simple but effective way to make the images sharper. In this method, the image is firstly lowpass filtered. This filter can be averaging filter, weighted average filter, Gaussian filter and so on. Then, the obtained lowpass image is subtracted from the original image, giving the highpass component. As the last step, the highpass content is added to the original image, and this addition can be done several times. As a result, the image becomes sharper than the original image.

Extremum sharpening is another simple approach for image enhancement [33]. To enhance the image, a pixel is compared to the minimum and the maximum of its neighboring pixels in a window. If it is out of the range of those extremum values; i.e. smaller than the smallest or larger than the largest, then its original value is conserved. If it is in the range, then the pixel value becomes the closest one of those extremum points.

The outputs of unsharp masking and extremum sharpening are given in Figure 4.4. Note that these methods do not increase contrast, luminance or visibility of the objects in the image. Instead, they try to sharpen the object boundaries and edges in the image. Although such techniques give good enhancement results for normal vision people, contrast enhancement may not be the ideal solution for the visually impaired. According to study [34], contrast enhancement in images has some limitations and challenges. The first limitation comes from the high frequency nature of the noise in images. Most of the visual defects like opacities (for example, cataract), visual field loss (for example, scotoma) and visual disorders (for example, astigmatism) bring the reduction of high frequency sensitivity. To compensate this vision loss in images, they have to be highpass filtered, enhancing the details. However, the inevitable noise in pictures is also enhanced compared to lower frequencies after the filter is applied. This noise amplification can cause disruption to user more than the enjoy that stressed details give to the visually impaired. On the other hand, lowpass filtering the image to get rid of the noise can sweep the highpass content that needs to be embellished off



(a) (b) (c)

Figure 4.4. Examples of unsharp masking and extremum sharpening. (a) is the original image, (b) is the unsharp masked image, (c) is the extremum sharpened image with window size 5×5.

the image. The second limitation is the range of standard digital displays. If an image occupies most of the dynamic range and is exposed to high-frequency amplification to some degree, then the amplified values may exceed the available dynamic range of the screens. In this case, the enhanced pixel values will be clipped. If the pixel values are to be rescaled to [0 255] after enhancement, the contrast difference between pixels may decrease, reducing the effect of the enhancement.

4.1.5. Compression-based Enhancement

Together with image and video compression methods such as JPEG, JPEG2000, MPEG and so on, spatial frequency filtering started to be applied in cooperation with those standards. Since they are similar in terms of the methodology, only JPEG standard will be explained briefly before the explanation of enhancement using those standards .

JPEG (Joint Photographic Experts Group) picture standard was developed and proposed in 1991 as a picture compression standard for digital devices [35]. Basically, this method reduces or destroys the contribution of the high frequency components in an image to compress the image data.

In the beginning,[4] encoding part divides the picture into 8×8 image blocks. Then, each channel of each block is shifted towards the dynamic range from [0 255] to [−128 127] by subtracting 128 from each pixel value, and goes through discrete cosine transform (DCT);

$$I(u,v) = \frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7} I(x,y)\cos\left(\frac{(2x+1)u\pi}{16}\right)\cos\left(\frac{(2y+1)v\pi}{16}\right) \tag{4.10}$$

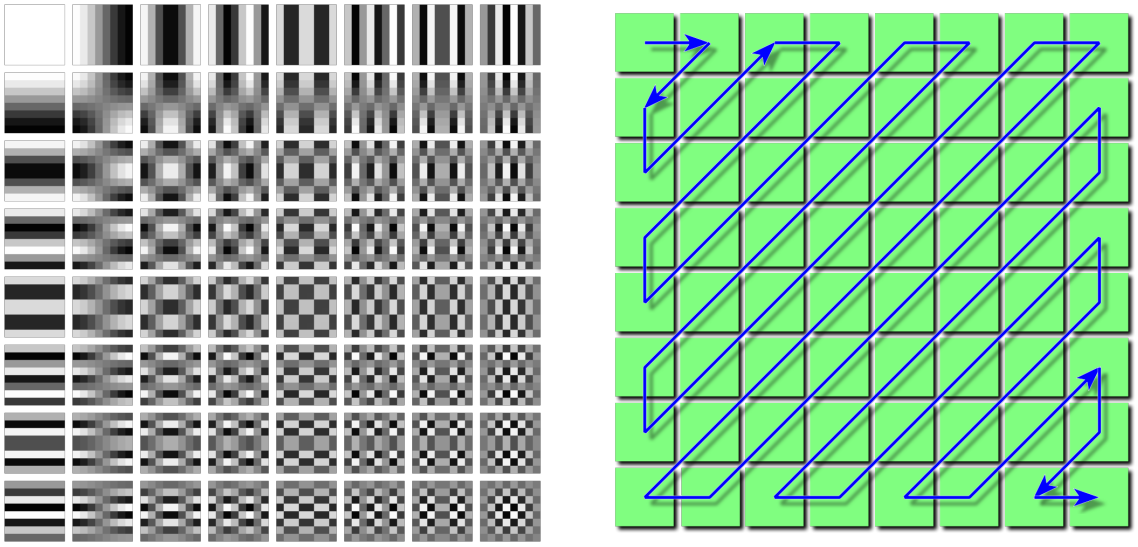
for

$$C(u) = \begin{cases} 1/\sqrt{2} & , u = 0 \\ 1 & , \text{otherwise} \end{cases} \quad \text{and} \quad C(v) = \begin{cases} 1/\sqrt{2} & , v = 0 \\ 1 & , \text{otherwise} \end{cases} \tag{4.11}$$

where $I(u,v)$ is the pixel value of image I at DCT domain coordinates (u,v), $I(x,y)$ is the pixel value of image I at spatial domain coordinates (x,y), $C(u)$ and $C(v)$ are the normalization constants given in Eqs. (4.11). After obtained, the frequency components are quantized according to a quantization table dependent on the quality factor of compression. Obtained DCT image block is element-wise divided by the quantization matrix and the results are rounded to the nearest integer. This part is where the loss in image information and quality occur. The rounding to the nearest integer operation is irreversible, and thanks to the reduction of information in this part, the image is compressed. An example of 8×8 quantization matrix $\mathbf{Q}$ with quality factor 50 specified by JPEG community is given in Eq. (4.12).

As the last step, each 8×8 DCT block is ordered and grouped in a manner that the similar frequencies lie on the same group. This ordering is called zigzag ordering and it has the same pattern as the DCT basis function ordering from low frequency to high frequency in each axis. DCT basis functions and JPEG zigzag ordering are given in Figure 4.5. After this ordering, a lossless compression method like Huffman coding

[4]The conversion from RGB to YCbCr and chroma subsampling of CbCr are omitted here.



(a) (b)

Figure 4.5. JPEG DCT basis functions and JPEG zigzag ordering. (a) is the JPEG DCT basis functions, adopted from [36]. (b) is the JPEG zigzag ordering, adopted from [37].

is applied to the image and it can be either stored or transferred as this.

$$\mathbf{Q} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \tag{4.12}$$

Decoding part of JPEG is the reverse-ordered version of encoding. First, entropy decoding is applied and the image blocks are reordered as the normal image is supposed to be. Second, dequantization is applied where each element of the dequantizing matrix, which is same as the quantization matrix, is multiplied element-by-element with the

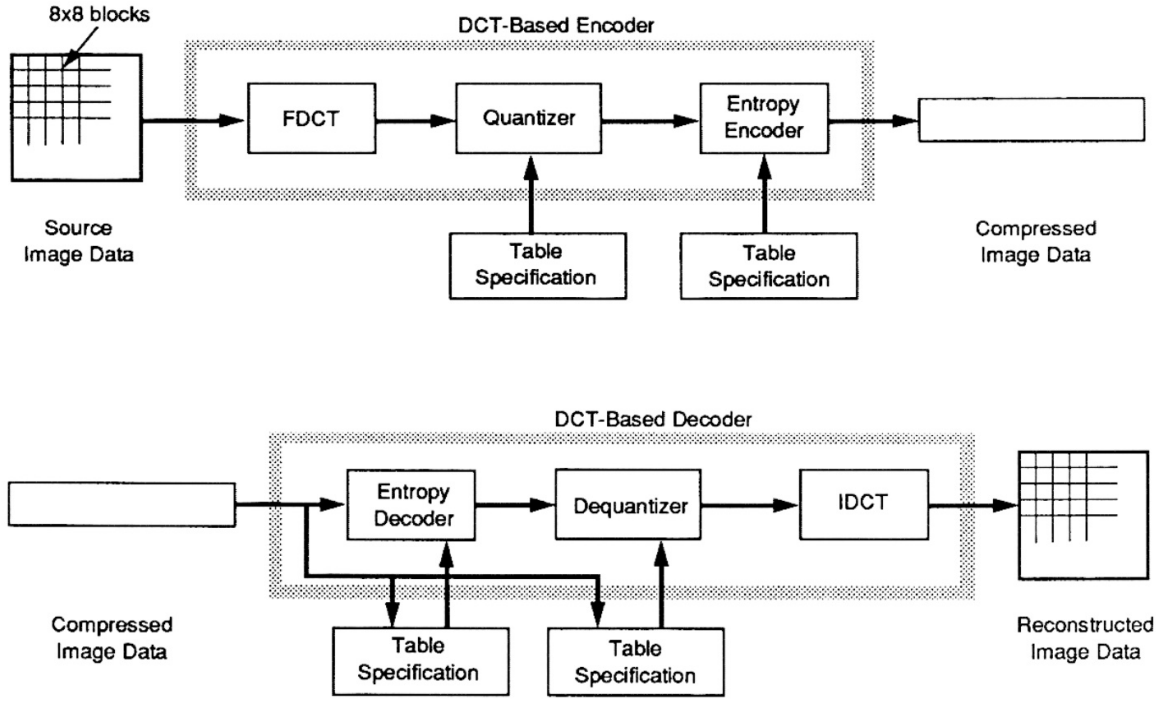


Figure 4.6. The block diagram of JPEG encoding and decoding. Adopted from [35].

blocks. Then, 8×8 inverse discrete cosine transform, IDCT, is applied as in Eq. (4.13),

$$I(x,y) = \frac{1}{4} \sum_{u=0}^{7} \sum_{v=0}^{7} C(u)C(v)I(u,v) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2y+1)v\pi}{16}\right) \qquad (4.13)$$

where $C(u)$ and $C(v)$ are defined in Eq. (4.11). After inversion, the values are rounded to the nearest integers. As the last step, the image range is shifted from $[-128\ 127]$ to $[0\ 255]$ by adding 128 to all pixels. The block diagram of JPEG compression is shown in Fig 4.6.

The difference in video encoding, such as in Moving Picture Experts Group (MPEG) standard, is inter-frame coding additional to intra-frame coding which is the same as JPEG encoding. In inter-frame coding, the difference between two blocks placed in same positions in consecutive images is encoded and decoded. Also, each block has its own encoded position vector and movement vector indicating the current position in the reference image and the next position in the consecutive frame. The next frame is reconstructed from movement vector for each block and the difference

between the image blocks.

The JPEG standard is a little bit manipulated to enhance the high frequency content of the images in real time for the visually impaired people in studies [38, 39]. Authors take advantage of the quantization table and its zigzag ordering from the lowest frequency to the highest. They modify the dequantization table in decoding part to boost the high frequencies in the image. To do that, they multiply the dequantization matrix with a filtering matrix.

To apply the enhancement technique in [38, 39], suppose that $\mathbf{D}$ is an 8×8 DCT coefficient matrix of an image block. In this block, each column represents vertical spatial frequency and each row represents the horizontal spatial frequency. Each entity, d_{mn}, of matrix $\mathbf{D}$ shows the magnitude of the block with m cycles/block in horizontal direction and n cycles/block in vertical direction. For instance d_{32} component involves 3 cycles/block horizontal spatial frequency and 2 cycles/block vertical spatial frequency, and d_{00} is the DC component of the image. Suppose that B_c indicates the average band amplitude with total c cycles/block frequency in matrix $\mathbf{D}$. For example, B_1 band is shown with cerulean, B_5 band with red and B_{11} band with green in Eq. (4.14).

$$\mathbf{D} = \begin{bmatrix} d_{00} & d_{01} & d_{02} & d_{03} & d_{04} & d_{05} & d_{06} & d_{07} \\ d_{10} & d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ d_{20} & d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} & d_{27} \\ d_{30} & d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} & d_{37} \\ d_{40} & d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} & d_{47} \\ d_{50} & d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} & d_{57} \\ d_{60} & d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} & d_{67} \\ d_{70} & d_{71} & d_{72} & d_{73} & d_{74} & d_{75} & d_{76} & d_{77} \end{bmatrix} \tag{4.14}$$

A contrast measure for each band in band-limited DCT domain can be defined as

$$C_c = \frac{B_c}{B_{c-1}} \quad \text{for} \quad 1 \leq c \leq 14,\ n \in \mathbb{Z} \tag{4.15}$$

where aforementioned B_c is mathematically expressed as

$$B_c = \frac{\sum_{m+n=c} |d_{mn}|}{\sum_{m+n=c} 1} \tag{4.16}$$

Suppose that the contrast in band c, except the DC component $c = 0$, is to be enhanced by a factor, say, α. The new contrast, $\hat{C}_c$, is then

$$\frac{\hat{B}_c}{\hat{B}_{c-1}} = \hat{C}_c = \alpha C_c = \frac{\alpha B_c}{B_{c-1}} \tag{4.17}$$

If we reorder and expand the Eq. (4.17), then we have

$$\begin{aligned}
\hat{B}_c &= \frac{\alpha B_c}{B_{c-1}} \hat{B}_{c-1} \\
&= \frac{\alpha B_c}{B_{c-1}} \frac{\alpha B_{c-1}}{B_{c-2}} \hat{B}_{c-2} = \alpha^2 \frac{B_c}{B_{c-2}} \hat{B}_{c-2} \\
&= \alpha^2 \frac{B_c}{B_{c-2}} \frac{\alpha B_{c-2}}{B_{c-3}} \hat{B}_{c-3} = \alpha^3 \frac{B_c}{B_{c-3}} \hat{B}_{c-3} \\
&= \dots \\
&= \alpha^c \frac{B_c}{B_0} \hat{B}_0 = \alpha^c B_c
\end{aligned} \tag{4.18}$$

where $B_0 = \hat{B}_0$ since DC component is not to be enhanced. If Eqs. (4.16) and (4.17) are combined, then the enhancement can be found as

$$\hat{d}_{mn} = \alpha^{m+n} d_{mn} \tag{4.19}$$

where $\hat{d}_{mn}$ is the new DCT coefficient for horizontal frequency m cycles/block and vertical frequency n cycles/block. This enhancement can be done via the quantization

matrix, $\mathbf{Q}$. If $\mathbf{Q}$ is modified to $\hat{\mathbf{Q}}$ such that

$$\hat{\mathbf{Q}} = \mathbf{Q} \circ \mathbf{F}$$

where $[\mathbf{F}]_{mn} = \alpha^{m+n}$, $0 \leq m, n \leq 7$, and $\circ$ is the element-wise (Hadamard) product, the enhancement can be realized. The frequency improvement can also be done directionally. For instance, if only the vertical frequencies are to be bolstered, $\mathbf{F}$ can be modified as

$$[\mathbf{F}]_{mn} = \begin{cases} \alpha^{m+n} & , m \leq n \\ 1 & , m > n \end{cases}$$

For video enhancement, the process is similar with the image enhancement. Authors of [40,41] multiply the quantization matrix for both intra- and inter-frame coding with a filtering matrix, $\mathbf{F}$;

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & \alpha & \alpha & \alpha & 1 & 1 \\ 1 & 1 & \alpha & \alpha & \alpha\beta & \alpha & 1 & 1 \\ 1 & \alpha & \alpha & \alpha & \alpha\beta & \alpha & 1 & 1 \\ \alpha & \alpha & \alpha & \alpha\beta & \alpha & 1 & 1 & 1 \\ \alpha & \alpha & \alpha & \alpha & 1 & 1 & 1 & 1 \\ \alpha & \alpha & \alpha & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{4.20}$$

Here, the matrix has non-identity elements for only the frequencies that are found to be useful to enhance in their experiments [40]. By eliminating low frequency bands B_1, B_2 and B_3, low and high ends of B_6, and two ends of B_7, the blocking effect from the enhancement is canceled. Also, by the asymmetry that is applied to enhance the vertical edges a little bit more by variable β, flickering effect stemming from interlaced images on television is reduced [40].

4.1.6. Object or Edge Stressing

Detail enhancement for vision degradation patients does not necessarily have to be done via either contrast enhancement or spatial filtering. Some studies also worked on the enhancement of the foreground or salient objects in images with a motivation that most of time the important content of image is not hidden in the background. Another motivation is that since the patients mostly lose their visual sensitivity for object boundaries or edges, they can be stressed by some different techniques.

Atabany and Degenaar added edges detected from a simplified image [42]. To obtain clear and connected edges, they first simplify the image via anisotropic diffusion filter. Proposed firstly by Perona and Malik [43], anisotropic filtering reduces the variations on a region while enhancing the boundaries of those regions. Authors of [43] look at the heat diffusion equation to use in image processing;

$$\frac{\partial I(x,y)}{\partial t} = \mathrm{div}(\gamma \nabla I(x,y)) \tag{4.21}$$

where $I(x,y)$ is the image pixel value at (x,y) coordinates, t is time, $\mathrm{div}(\cdot)$ is divergence operator, ∇ is the gradient operator and γ is the positive diffusion (conduction) constant. In order edges to be enhanced, the conduction coefficient must be 1 within regions and 0 on the boundaries on the regions, which means that the diffusion coefficient should not be a constant. They claim that there is no reason for γ to be constant, and change it to $\gamma(x,y;t)$, a position- and time-dependent variable. So, the Eq. (4.21) becomes

$$\begin{aligned} \frac{\partial I(x,y;t)}{\partial t} &= \mathrm{div}(\gamma(x,y;t)\nabla I(x,y;t)) \\ &= \nabla\gamma(x,y;t)\cdot\nabla I(x,y;t) + \gamma(x,y;t)\Delta I(x,y;t) \end{aligned} \tag{4.22}$$

where Δ is the Laplacian operator. To solve this equation, they let $\gamma(x,y;t)$ to be a function of the gradient of the image brightness; in other words, $\gamma = g(\nabla I)$, where the spatial coordinates x, y and time indicator t are omitted for simpler and cleaner

notation. Assume that, without loss of generality, the smooth edge is aligned with y axis, hence the edges are one dimensional. Then $\nabla I = I_x$ and $\gamma = g(I_x)$, where I_x is the change in edge slope. In this case, Eq. (4.22) becomes

$$\begin{aligned}\frac{\partial I}{\partial t} &= \operatorname{div}(g(I_x) \cdot I_x) \\ &= g'(I_x) \cdot I_{xx} \cdot I_x + I_{xx} \cdot g(I_x) \\ &= \phi'(I_x) \cdot I_{xx} \end{aligned} \tag{4.23}$$

where $\phi(I_x) = g(I_x) \cdot I_x$ is introduced for notational simplicity. Looking at the variation of the slope of edge in time, if $\gamma > 0$, then the differentiation in $\partial I_x / \partial t$ may change order, and becomes

$$\frac{\partial I_x}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\partial I}{\partial t}\right) = \phi''(I_x) \cdot I_{xx}^2 + \phi'(I_x) \cdot I_{xxx} \tag{4.24}$$

Suppose that the edge is oriented such that $I_x > 0$. So, $I_{xx} = 0$ where the slope of edge is maximum, and $I_{xxx} \ll 0$. So, signs of $\partial I_x / \partial t$ and $\phi'(I_x)$ are opposite; if $\phi'(I_x) < 0$, the slope of edge increases with time, enhancing the edge.

The discretization of Eq. (4.22) can be done as [43]

$$I^{(t+1)}(x,y) = I^{(t)}(x,y) + \varsigma \left[\gamma_U \bigtriangledown_U I + \gamma_D \bigtriangledown_D I + \gamma_L \bigtriangledown_L I + \gamma_R \bigtriangledown_R I\right]^{(t)}(x,y) \tag{4.25}$$

where (x, y) is the pixel coordinate, (t) is the time (iteration number), ς is a stabilizer constant with $0 \leq \varsigma \leq 1/4$, $\bigtriangledown_{(\cdot)} I$s are defined as directional differences; in other words,

$$\begin{aligned}\bigtriangledown_U I^{(t)}(x,y) &= I^{(t)}(x-1,y) - I^{(t)}(x,y) \\ \bigtriangledown_D I^{(t)}(x,y) &= I^{(t)}(x+1,y) - I^{(t)}(x,y) \\ \bigtriangledown_L I^{(t)}(x,y) &= I^{(t)}(x,y-1) - I^{(t)}(x,y) \\ \bigtriangledown_R I^{(t)}(x,y) &= I^{(t)}(x,y+1) - I^{(t)}(x,y)\end{aligned}$$

γ_U, γ_D, γ_L and γ_R are the diffusion coefficients for directions up, down, left and right, respectively, defined as a function of the absolute value of directional differences; that is,

$$\gamma_U^{(t)}(x,y) = g\left(|\bigtriangledown_U I^{(t)}(x,y)|\right)$$
$$\gamma_D^{(t)}(x,y) = g\left(|\bigtriangledown_D I^{(t)}(x,y)|\right)$$
$$\gamma_L^{(t)}(x,y) = g\left(|\bigtriangledown_L I^{(t)}(x,y)|\right)$$
$$\gamma_R^{(t)}(x,y) = g\left(|\bigtriangledown_R I^{(t)}(x,y)|\right)$$

Perona and Malik proposed two possible choices for function $g(\cdot)$: To prefer high contrast edges over low contrast ones;

$$g\left(\nabla I\right) = e^{-\left(|\nabla I|/\kappa\right)^2} \tag{4.26}$$

and for privileging the wide regions over the smaller ones, the used formula is

$$g\left(\nabla I\right) = \frac{1}{1+\left(\dfrac{|\nabla I|}{\kappa}\right)^2} \tag{4.27}$$

where κ is either determined experimentally or found from noise estimation [44].

Atabany and Degenaar applied anisotropic diffusion operations with little modifications [42]: They choose

$$g\left(\nabla I\right) = \frac{1}{1+\sqrt{\sum_d \left(\nabla_d I\right)^2}} \tag{4.28}$$

where d denotes the eight directions; up, down, left, right, up-right, up-left, down-right and down-left, and modify Eq. (4.25) for those eight directions.

(a) (b)

Figure 4.7. Example of edge enhancement using scene simplification. (a) is the original image, (b) is the resultant image with $\varsigma = 1/7$, $\kappa = 30$, γ is as in Eq. (4.27), and iteration number is 2.

To detect edges over the simplified image with anisotropic filter, they use Sobel edge detectors. The detector filters are shown in Eq. (4.29),

$$\mathbf{S}_h = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \qquad \mathbf{S}_v = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$
$$\mathbf{S}_{md} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \mathbf{S}_{ad} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \tag{4.29}$$

where $\mathbf{S}_h$ is used for horizontally aligned edges, $\mathbf{S}_v$ is for vertical ones, $\mathbf{S}_{md}$ is for edges along with main diagonal (i.e. through upleft-downright), and $\mathbf{S}_{ad}$ is for antidiagonal edges.

After the edges are obtained, they are added to the original image to stress the object boundaries or edges. An example of this enhancement is shown in Figure 4.7.

Similar to the study [42], [45] also detects edges with gradient operators and adds them to the original image to emphasize the edges. Their differences are that the work [42] simplifies the image with diffusion filtering while study [45] does not do that, and they use different edge detector operators. Atabany and Degenaar use Sobel edge detectors given in Eq. (4.29), while Wolffshon et al. use Sobel and Prewitt operators together, in which the latter is defined as

$$\mathbf{P}_h = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \qquad \mathbf{P}_v = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \tag{4.30}$$

where $\mathbf{P}_h$ is for horizontal edges and $\mathbf{P}_v$ is for vertical edges. The reason that they use both edge detectors is to preserve every edge detail; Sobel operator extract square contours while Prewitt operator sketches curved lines [45]. The last difference between the studies is that [42] applies the filter to gray-scale image that is either the original image or converted from RGB image. On the other hand, [45] applies the edge detectors to each color channels separately.

To increase the details and contrast difference that low vision people receive, [46] applies a similar approach with [42]. Referred as tinted reduced outlined nature (TRON) algorithm, this work first simplifies the image via anisotropic diffusion filter described in Eqs. (4.25) and (4.28). Then, edge detection is done with edge detector in [47], the modified version of Canny's [44]: Instead of $[-1\ 1]$ masks, they use $[-1\ 0\ 1]$ masks to compute first derivatives in four main directions; horizontal, vertical and two diagonals. Named as $\mathbf{D}_h$, $\mathbf{D}_v$, $\mathbf{D}_{d_1}$ and $\mathbf{D}_{d_2}$, those four directional differences are used to calculate the $\mathbf{X}$ and $\mathbf{Y}$ gradients, as

$$\mathbf{X} = \mathbf{D}_h + \frac{\mathbf{D}_{d_1} + \mathbf{D}_{d_2}}{2} \quad \text{and} \quad \mathbf{Y} = \mathbf{D}_v + \frac{\mathbf{D}_{d_1} - \mathbf{D}_{d_2}}{2} \tag{4.31}$$

After the calculation of gradients $\mathbf{X}$ and $\mathbf{Y}$, the last stage of TRON algorithm is to multiply the image by a weighting matrix $\mathbf{W}$. It is obtained from the gradient

(a) (b)

Figure 4.8. Example of the TRON algorithm. (a) is the original image, (b) is the resultant image with edge detectors in Eq. (4.29); $\varsigma = 1/7$, $\kappa = 30$, $g(\nabla I)$ is as in Eq. (4.27), iteration number 1 for Eq. (4.25), and threshold $K = 0.2$.

magnitude matrix, $\mathbf{G}$, calculated as

$$\mathbf{G} = \left(\mathbf{X}^{\textcircled{2}} + \mathbf{Y}^{\textcircled{2}}\right)^{\textcircled{.5}} \tag{4.32}$$

where $\textcircled{2}$ is the element-wise square of a matrix and $\textcircled{.5}$ is point-wise square root of a matrix. Then $\mathbf{G}$ is normalized to [0 1] and thresholded with K, values below of which are raised to K to obtain $\mathbf{W}$. Then the multiplication is done in an element-by-element manner; that is,

$$TrI = I \circ \mathbf{W} \tag{4.33}$$

where TrI is the TRON image. The result of the algorithm can be seen in Figure 4.8. Note that the procedure tries to stress edges by reducing the color information of the image, thus makes it darker.

In this study [46], another image enhancement technique to be used for helping the visually impaired people is cartoonization algorithm. The motivation of this algorithm is that in cartoon images edges and boundaries are well-drawn and not blended. Also, objects in cartoon images do not have much of the visual details such as texture and smooth color transition. Therefore, cartoonization reduces the details in applied images

and tries to keep main object boundaries and edges.

The first step in cartoonization algorithm is to reduce or erase the textures and details. To do this, first, an RGB (red/green/blue) image is converted to an YCbCr (luminance/blue-difference/red-difference) image. The conversion from an 8-bit RGB image to YCbCr image is done according to the formulas

$$\begin{aligned} Y &= 16 + (65.481 \times R + 128.553 \times G + 24.966 \times B) \\ Cb &= 128 + (-37.797 \times R + -74.203 \times G + 112 \times B) \\ Cr &= 128 + (112 \times R + -93.786 \times G + -18.214 \times B) \end{aligned} \tag{4.34}$$

In this conversion, luminance component (Y) lies in the range [16 235], and chrominance components (Cb and Cr) lie in the range [16 240]. To eliminate texture, only Y component is filtered with anisotropic diffusion filter. Then the new YCbCr image is converted back to RGB color space with the formulas

$$\begin{aligned} R &= 0.00457 \times (Y - 16) + 0.00000 \times (Cb - 128) + 0.00626 \times (Cr - 128) \\ G &= 0.00457 \times (Y - 16) - 0.00154 \times (Cb - 128) - 0.00319 \times (Cr - 128) \\ B &= 0.00457 \times (Y - 16) + 0.00791 \times (Cb - 128) + 0.00000 \times (Cr - 128) \end{aligned} \tag{4.35}$$

The second step in cartoonization is the calculation and normalization of the gradient image. Again, the calculation of the gradient image is done as in Eqs. (4.31) and (4.32). After normalizing the gradient image to [0 1], two threshold values are determined, τ_m and τ_M. The values of gradient image below τ_m are set to 0, and the values above τ_M become 1.

To create a cartoon image effect, the colors in original image are quantized because, generally, cartoon images do not have a smooth color transition. This quantization is done in CIELAB color space. To do that, first the RGB image is converted

into CIEXYZ color space image via the formula [48]

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49000 & 0.31000 & 0.20000 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00000 & 0.01000 & 0.99000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.36}$$

Then the CIEXYZ (XYZ) image is converted into CIELAB ($L^*a^*b^*$) color space image by the formulas [49]

$$\begin{aligned} L^* &= 116 f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* &= 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\ b^* &= 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{aligned} \tag{4.37}$$

where

$$f(x) = \begin{cases} x^{1/3} & , x > (24/116)^3 \\ (841/108)\, x + (16/116) & , x \leq (24/116)^3 \end{cases} \tag{4.38}$$

and X_n, Y_n and Z_n are the reference white color values in CIEXYZ space. The reference lightness for white is taken as $[X_n\ Y_n\ Z_n] = [100\ 100\ 100]$. The conversion from RGB to $L^*a^*b^*$ cannot be done directly because $L^*a^*b^*$ color space is defined with respect to XYZ. The quantization is done on the L^* channel of the converted image, as

$$QL^* = Q_n + \frac{\Delta q}{2} \tanh\left(\mathbf{T_S} \circ (L^* - Q_n)\right) \tag{4.39}$$

Here, QL^* is the quantized L^* channel; $\Delta q = 100/(N-1)$ in which N is the number of bins and 100 is the maximum value that L^* can take; $\mathbf{T_S}$ is the sharpness transition matrix, effectively the gradient image rescaled to the range between two empirically determined values ts_m and ts_M, and Q_n is the nearest quantized value of the channel; that is, $Q_n = \Delta q \times \lfloor L^*/\Delta q \rceil$, where $\lfloor \cdot \rceil$ is the nearest integer function.

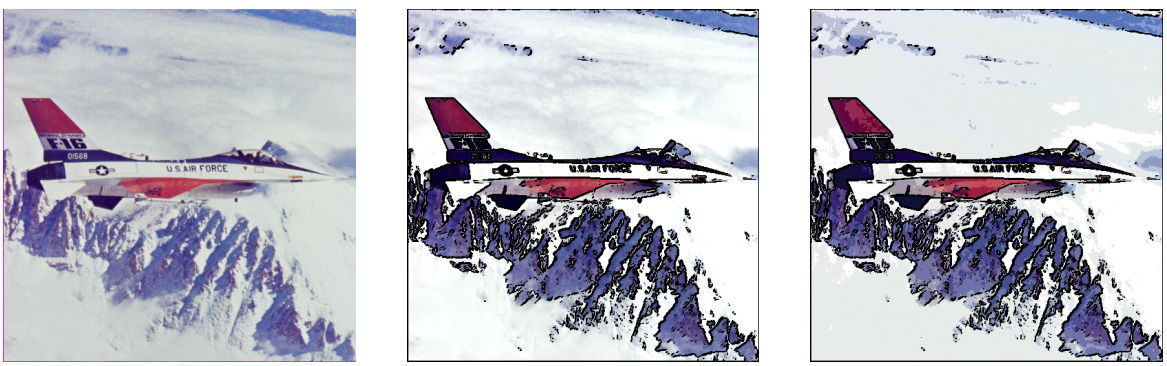
As the last step, the image is converted back to the RGB image with the quantized L* channel, and is combined with the negative of the gradient map to emphasize the edges. Conversion from L*a*b* to RGB is done indirectly. First, L*a*b* image is converted to XYZ image via

$$\begin{aligned} X &= X_n f^{-1}\left(\frac{L^*+16}{116}+\frac{a^*}{500}\right) \\ Y &= Y_n f^{-1}\left(\frac{L^*+16}{116}\right) \\ Z &= Z_n f^{-1}\left(\frac{L^*+16}{116}-\frac{b^*}{200}\right) \end{aligned} \tag{4.40}$$

where

$$f^{-1}(x) = \begin{cases} x^3 & , x > (24/116) \\ \frac{108}{841}\left(x-\frac{16}{116}\right) & , x \le (24/116) \end{cases} \tag{4.41}$$

Then, XYZ image is converted back to RGB image via the transformation



(a) (b) (c)

Figure 4.9. Examples of cartoonization algorithm. (a) is the original image, (b) is the cartoonized image without color quantization, (c) is the cartoonized image to which color quantization is applied with $N = 8$, $ts_m = 3$, $ts_M = 25$. In both of the results, $\tau_m = 0.1$ and $\tau_M = 0.4$. The parameters for anisotropic diffusion filter are the same as the results in Figure 4.8.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0.4184657 & -0.15866088 & -0.0828349 \\ -0.0911690 & 0.2524314 & 0.0157075 \\ 0.0009209 & -0.0025498 & 0.1785989 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4.42}$$

In Figure 4.9, the results of the cartoonization algorithm are shown. Note that the difference between the result with and without color quantization applied: In color quantized version, the color transition is not smooth but has some value jumps while this is not the case the result without the quantization.

4.2. Deconvolutional Methods

Deconvolution is a term used for the process to invert the effect of a linear, time- or space-invariant system on the output to get the original input. In the deconvolutional approach to vision correction, human eye is modeled as a linear, space-invariant system with an impulse response. When the light comes into the eye, it is processed with the impulse response of the eye and perceived as that processed version of the actual input. If the impulse response is corrupted and steers away from the ideal response, then the perceived image is corrupted. To recover the actual image, which can also be called desired image, the output of eye should be convolved with inverse of (i.e. deconvolved with) the transfer function of eye. Since the output is the retinal image and in this case deconvolution must be done after rod and cone cells send the image signal but before it reaches the image processing part of the brain, this is currently not possible. But, thanks to the commutativity property of convolution process, deconvolution can be performed in a different stage. For a screen, if image is deconvolved with the impulse response of the human eye and then displayed to the person, then the perceived image, in theory, must be very close to the image that a healthy eye perceives. This idea is the basis for deconvolutional vision correction methods.

4.2.1. Using Visual Degradation Transfer Function

A deconvolutional approach for vision correction was firstly discussed in [25]. In this work, the transfer function of the eye is obtained via sine wave grating patterns. An example of this pattern is shown in Figure 4.10. Normally in this graph, frequency increases from left to right and contrast decreases in upward direction in logarithmic fashion. Note that the given figure is not a standard measurement chart. This image is created without following the necessary standards and is for the explanatory reasons.

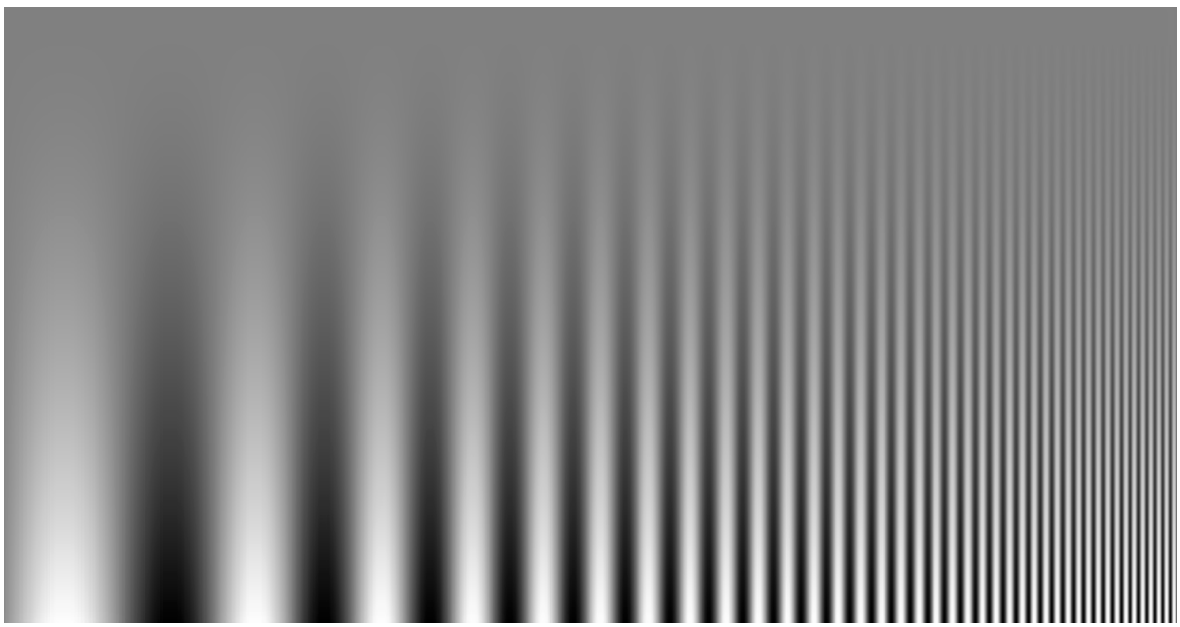


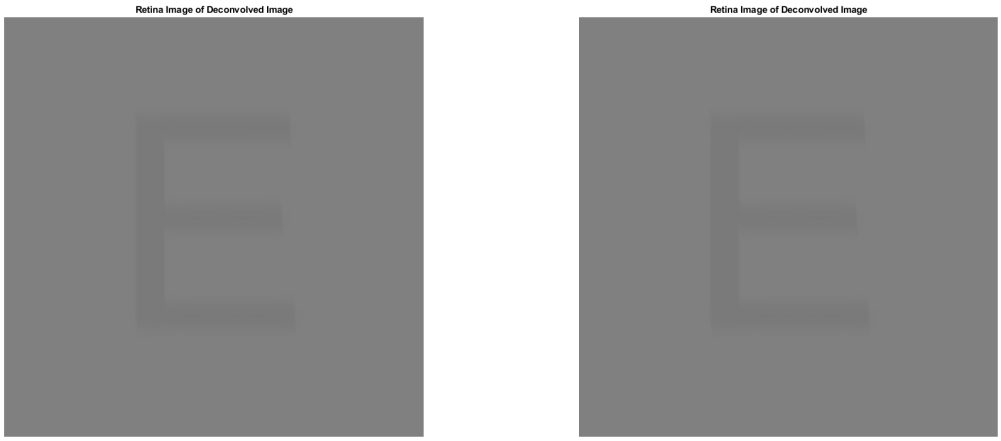
Figure 4.10. A representation of sine wave grating pattern.

For each contrast threshold value, the ability of patient to distinguish the grating pattern is different, which is a situation specifying the characteristics of the visual system. For every threshold, there is a special frequency value (cycles/mm or cycles/deg in units) for which the patient cannot separate the gratings. The reciprocal of this value is the sensitivity for the specific frequency, and when the sensitivity values are plotted as a function of the spatial frequency, the contrast sensitivity function of the visual system is obtained. According to the authors of [25], this function can be used as the modulation transfer function of the system, which can be related with the complete characteristics of the visual system as explained in Section 2.6.2.

After obtaining the necessary transfer function, the authors apply the deconvolution operation to the image. But the result is not satisfactory. The first, and the most disturbing, reason is the amplified noise. The obtained transfer function has low-pass nature. Consequently, the inverse of this function exhibits high-pass characteristics. When it is applied to an image, it enhances details as well as the noise effective in the high-pass spectrum. Hence, it disturbs the patient more than it presents a solution for visual degradation. The second reason is the lost contrast. The inverse of the function has very big and small values exceeding the dynamic range of the standard screens. Thus, the necessary contrast difference for a proper correction cannot be displayed. As a result, clipping and the loss of information are inevitable. The last reason for unsatisfactory results is the way of obtaining the transfer function. It is found by the answers that are given by the patient to the change in contrast and spatial frequency. But the examination is neither the patient- nor the impairment-specific. For instance, both an astigmatic and a myopic patient suffer from blurred vision in a distance. Both patients may respond to the same contrast threshold value and the spatial frequency, giving the same contrast sensitivity function, thus, the same transfer function. However, those impairments have different characteristics and cannot have the same transfer function.

4.2.2. Using Wavefront Aberration Function

The last problem is overcome via the Hartmann-Shack wavefront sensor [50]. The sensor is composed of an array of small lenses, called "lenslets". The aperture array is developed by Johannes Franz Hartmann in 1900 [51], and the array is modified by Roland Shack as an array of lenslets [52]. Each lenslet has the same focal length and a priori known focus point which hosts a sensor for light detection. When a distorted wave meets with the lens array, the lenslets cannot focus the part of the wave to the assumed focal point. Difference between the normal focal point on which the normal wave is focused and the actual focal point on which the distorted wave is focused can be calculated for each lenslet. Since any phase distortion can be approximated with discrete tilts, all the measured differences are used to approximate the wavefront aberration function. This approach led to better presentation of point spread function



(a) (b)

Figure 4.11. Result of edge tapering. (a) is the retina image without edge tapering, (b) is the with edge tapering.

than the visual degradation function, and was used in various studies [3, 53].

The lost contrast is the biggest problem since the standard dynamic range of 8-bit digital screen are not even close to be adequate for deconvolution results. As shown in Figure 5.13 in Section 5.8, the dynamic range of the result may typically lie between around -6000 and about +7000, changing with different image and transfer function. After shifting and scaling the resultant dynamic range linearly to the necessary one, the vast differences between the actual pixel values are also scaled, decreasing the actual intensity difference, hence the contrast. The main idea in deconvolutional methods using wavefront aberration function is to increase the lost contrast.

One of the methods proposed to increase the perceived contrast is the edge tapering method [54]. In this method, deconvolved image is edge-tapered with a Gaussian filter whose size is 20×20 pixels and standard deviation (σ) is 5, after which it is displayed on the screen. The outcomes in study [54] seem very promising. However, those results are obtained with the assumption that the images can be displayed on a high dynamic range screen instead of a standard one. The results with and without edge tapering on an 8-bit standard display are shown in Figure 4.11.

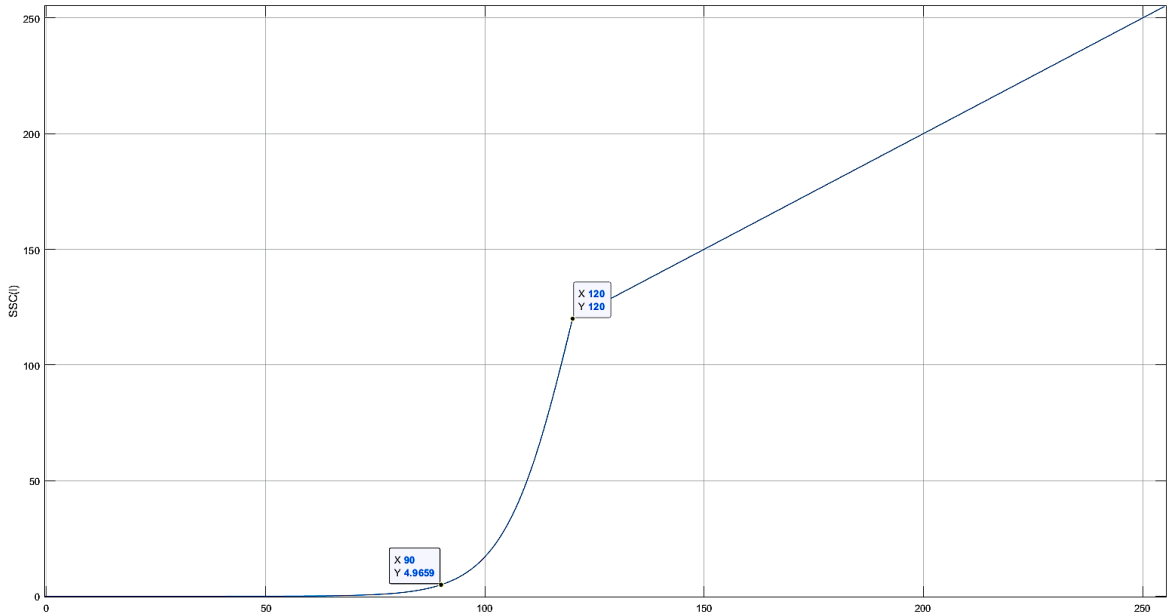


Figure 4.12. Single-sided contrast mapping for $\beta = 9$, $t = 120$, and $m = 50$.

Another method encountered in the literature to increase the visibility of image is the single-sided contrast enhancement [55]. This post-processing method relies on the fact that shifting and scaling the dynamic range of an image do not have an impact on non-zero frequencies. In single-sided enhancement, deconvolved image is shifted and scaled to the dynamic range of an ordinary digital screen. Then, the transformation function in Eq. (4.43) is applied;

$$SSC(l) = \begin{cases} t \cdot \dfrac{1 - \exp\left\{-\beta\left(\frac{l-m}{t-m} - 1\right)\right\}}{1 + \exp\left\{-\beta\left(\frac{l-m}{t-m} - 1\right)\right\}} + t & , l \leq t \\ l & , l > t \end{cases} \tag{4.43}$$

where $SSC(l)$ is single-sided contrast enhancement, l is the luminance value of a pixel, β is the steepness parameter for the enhancement curve, t is the threshold value, and m is the minimum value of the new image. This method is called single-sided because it transforms the values under the threshold according to the Eq. (4.43), and the values above it are linearly mapped to [0 255]. The example mapping is in Figure 4.12 for values $\beta = 9$, $t = 120$, and $m = 50$.

5. METHODOLOGY

For vision correction for the people with visual impairment, the followed approach in this work is the deconvolutional approach that uses the point spread function of human eye. In this work, the eye data that help calculation of the point spread function are collected from a device that uses the ray tracing method to determine the aberration map. From those data, the wavefront aberration function is calculated by employing the Zernike polynomials in the calculation of wavefront aberration function. As the author knows, this work is the first which calculates the wavefront aberration function of human eye cornea as well as the crystalline lens. The wavefront aberration functions of two refractive components of human eye are included into the point spread function calculations. The obtained function can be called "raw"because after the main calculation, the effects that can change the point spread function are taken into account. Those effects are the changing pupil diameter, different wavelength value that changes the defocus term in Zernike polynomials, accommodation distance and the person's age. After the final point spread function is calculated, the deconvolution operation is applied to the images via the Wiener filtering. As the last step, the mapping and contrast enhancement methods are applied to display the deconvolved images whose dynamic range originally expands out of the standard screen dynamic range with as much contrast as possible for clear and sharp retina image.

5.1. Zernike Polynomials

Zernike polynomials are a set of continuous polynomials that are orthogonal on the unit circle. They were first used by Frits Zernike, the winner of the 1953 Nobel Prize in Physics due to the invention of phase-contrast microscopy [56], in his phase-contrast method [57], and standardized by OSA in 2002 for reporting optical aberrations in human eye [58].

Zernike polynomials are one of the many polynomial sets that can be represented with two real variables, and orthogonal in a continuous manner over the interior of

the unit circle. Although they are orthogonal when they are continuous and generally not orthogonal in discrete domain [59], they have several mathematical and practical advantages over the other polynomial sets to be used in ophthalmology. First of all, Zernike polynomials form a complete set, meaning that they can represent any circular surface given enough terms. They have rotational symmetry that provides them with being able to be written in the form of polynomial product of a radial polynomial, written with radius ρ, and a periodic azimuthal polynomial, written in θ. Second, n^{th} degree radial polynomial does not contain any term whose degree is less than m, the degree of the azimuthal polynomial. Also, if degree of azimuthal polynomial, m, is even, the radial polynomial is also even, and if m is odd, so is the radial function [59]. Moreover, some of the Zernike polynomials have exactly the same shape with the low order aberrations, which makes the aberrations easier to be represented. When expressing the wavefront aberration function, the coefficients of used Zernike polynomials are independent of the number of polynomials used, meaning that any number of additional Zernike polynomials will not affect the previously computed or used ones. Lastly, the balanced representation of classical aberrations with Zernike polynomials shows minimum variance over a circular pupil [60], meaning that the representation of wavefront aberration with Zernike polynomials over a circular pupil has the minimum change compared to representations with other basis functions.

Zernike polynomials are generally defined in polar coordinates (ρ, θ) with ρ being the radial component ranging between 0 and 1, and θ being azimuthal coordinate with values between 0 and 2π. They can be written as product of polynomials, and a Zernike polynomial have three components: radial component, azimuthal component and normalization component. Together with n being the highest radial order and m being the azimuthal frequency, a Zernike polynomial in double indexing scheme, $Z_n^m(\rho, \theta)$, can be written as

$$\boxed{Z_n^m(\rho, \theta) = N_n^m R_n^m(\rho) M^m(\theta)} \tag{5.1}$$

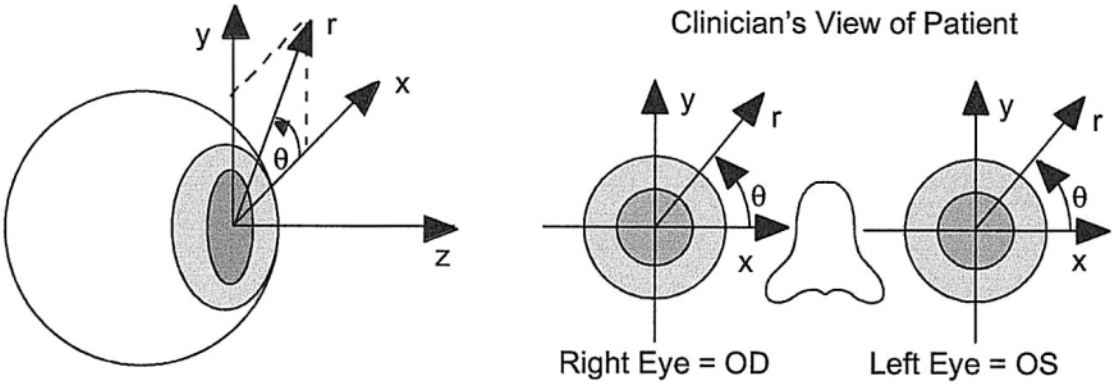


Figure 5.1. Standard reference axis and positive direction for reporting eye aberrations. Adopted from [58].

In Eq. (5.1), the polynomial $M^m(\theta)$ is the azimuthal, or meridional, component with m being the azimuthal frequency and θ being the azimuthal coordinate. Although the reference axis is taken as positive y axis and the positive direction is accepted as clockwise in some studies [60,61], the right handed coordinate system is standardized for representation of aberrations as shown in Figure 5.1 [58], hence the standard approach is used here for coordinate θ. The azimuthal component is defined as

$$\boxed{M^m(\theta) = \begin{cases} \cos(m\theta) & , m \geq 0 \\ -\sin(m\theta) & , m < 0 \end{cases}} \tag{5.2}$$

The radial polynomial $R_n^m(\rho)$ is described as

$$\boxed{R_n^m(\rho) = \sum_{k=0}^{(n-|m|)/2} \frac{(-1)^k (n-k)!}{k! \left[\frac{n+|m|}{2} - k\right]! \left[\frac{n-|m|}{2} - k\right]!} \rho^{n-2k}} \tag{5.3}$$

where n is the highest degree of Zernike polynomial, m is the azimuthal frequency, and ρ is the radial coordinate. Note that $R_n^m(\rho) = R_n^{-m}(\rho)$.

Although the definition of the radial polynomial $R_n^m(\rho)$ is as in Eq. (5.3), it is not the best way to calculate it because of the fact that factorial operation spends the most of the time during computation of the radial polynomial. Moreover, multiplication and division operations may cause numerical instabilities such as overflow, underflow or floating point precision during the calculation of higher order polynomials. To overcome these problems, some alternative methods to direct method (Eq. (5.3)) are proposed, such as Kintner's algorithm [62], modified Kintner's algorithm [63], Prata's algorithm [64], Belkasim's method [65], coefficient method, p-recursive method [66] and q-recursive method [63]. Among them, q-recursive method is the best algorithm in terms of computation time [63] and numerical stability robustness [67].

In q-recursive method, the radial polynomial $R_n^m(\rho)$ is calculated as following:

$$R_n^m(\rho) = \begin{cases} \rho^n & , n = m \\ nR_n^n(\rho) - (n-1)R_m^m(\rho) & , n = m+2 \\ H_1 R_n^{m+4}(\rho) + \left(H_2 + \dfrac{H_3}{\rho^2}\right) R_n^{m+2}(\rho) & , \text{otherwise} \end{cases} \tag{5.4}$$

where

$$H_1 = \frac{(m+4)(m+3)}{2} - (m+4)H_2 + \frac{H_3(n+m+6)(n-m-4)}{8}$$

$$H_2 = \frac{H_3(n+m+4)(n-m-2)}{4(m+3)} + m + 2$$

$$H_3 = \frac{-4(m+2)(m+1)}{(n+m+2)(n-m)}$$

The method calculates $R_n^m(\rho)$ recursively over m. It is called q-recursive because authors of [63] use $R_p^q(\rho)$ notation instead of $R_n^m(\rho)$. Also, this algorithm uses the fact that $R_n^m(\rho) = R_n^{-m}(\rho)$ by definition in Eq. (5.3), because, for instance, H_2 would be undefined for azimuthal frequency $m = -3$ without it.

The normalization constant is chosen from $\int_0^1 \rho R_n^m(\rho) R_{n'}^m(\rho) d\rho = \frac{R_n^m(1)\delta_{nn'}}{2(n+1)}$, so

$$\boxed{N_n^m = \sqrt{\frac{2(n+1)}{1+\delta(m)}}} \tag{5.5}$$

where $\delta(m)$ is Kronecker delta function which is 0 other than $m = 0$. In some studies, a separate normalization term as in Eq. (5.5) is not used [68]. Instead, its components are distributed; $\sqrt{n+1}$ is embedded into $R_n^m(\rho)$, and $\sqrt{2}$ is written inside of $M^m(\theta)$ with the constraint $M^0(\theta) = 1$, yielding the same result.

In double index scheme, radial order n and azimuthal frequency m should satisfy some conditions, all of which can be interpreted from Eq. (5.3). These conditions are

- $n \geq 0$,
- $|m| \leq n$,
- $n + m \equiv 0 \mod 2$.

Together with double index scheme, a single index representation is also recommended by OSA to be used only in graphical display of the coefficients of Zernike polynomials. In this representation,

$$Z_i(\rho, \theta) = Z_n^m(\rho, \theta)$$

where the conversion between these two schemes can be done by

$$i = \frac{n(n+2)+m}{2}; \quad n = \left\lceil \frac{-3+\sqrt{9+8i}}{2} \right\rceil, \quad m = 2i - n(n+2) \tag{5.6}$$

provided that $\lceil \cdot \rceil$ is the ceil (round-up) function [58].

The first 45 Zernike polynomials, obtained by Eq. (5.1) with both double indexed and single indexed names, are given in Figure 5.2. In the figure, yellow areas show "higher" points and blue areas indicate "lower" values. Note that radial order increases

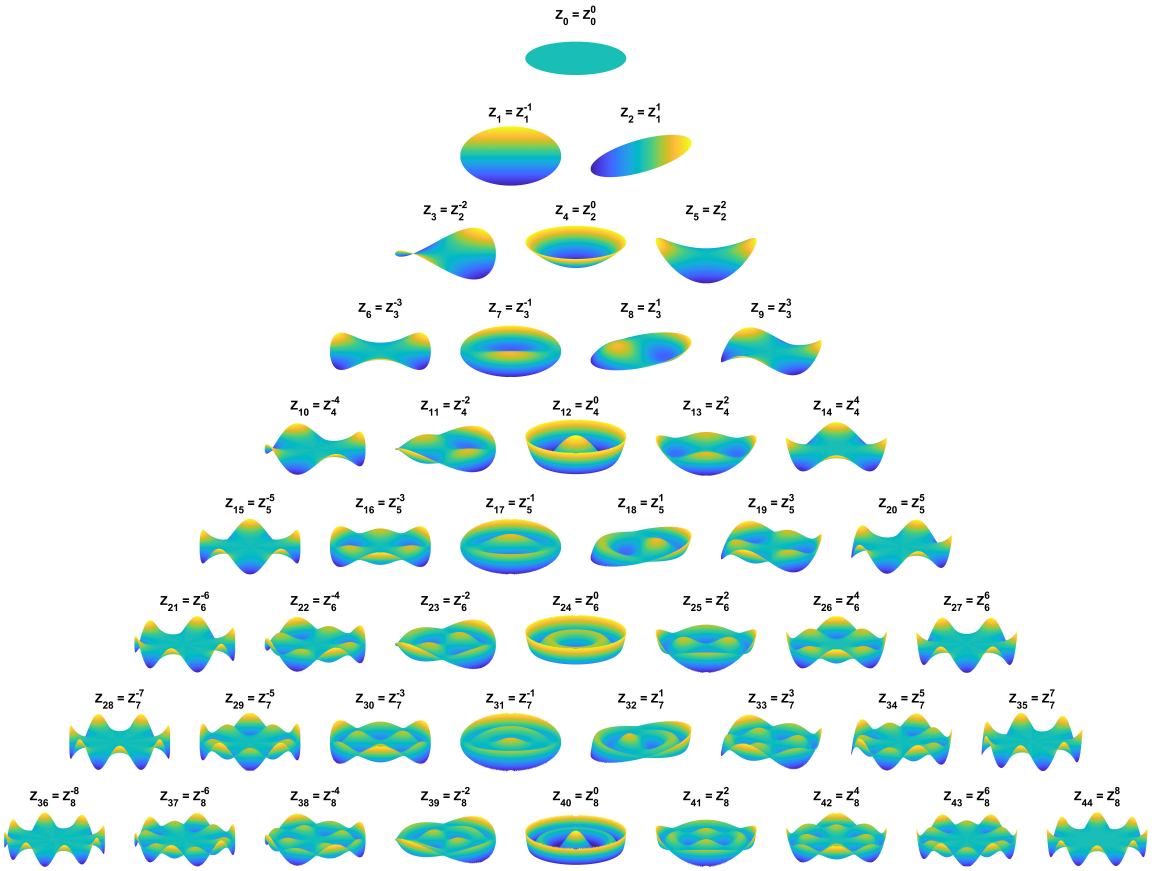


Figure 5.2. First 45 Zernike polynomials. Yellow colored areas indicate bigger/higher values than the blue ones, smaller/lower values. The double indexed and single indexed representations are given on top of the each polynomial.

as moving down in the pyramid, and azimuthal frequency increases from left to right. Note also that the numbering of single indexed scheme is done first from left-to-right and then top-to-bottom.

5.2. Wavefront Aberration Function

From the discussions in Chapter 2, it is obvious that a perfect converging lens focuses a planar wave not extending beyond the limits of the paraxial approximation onto its focal point F. However, if the lens is not perfect in some ways, for instance it is dirty, damaged, deformed or not homogeneous, then the refracted wave cannot have the same phase at every point on itself. In this case, the wavefront is not a perfectly spherical wave and is said aberrated.

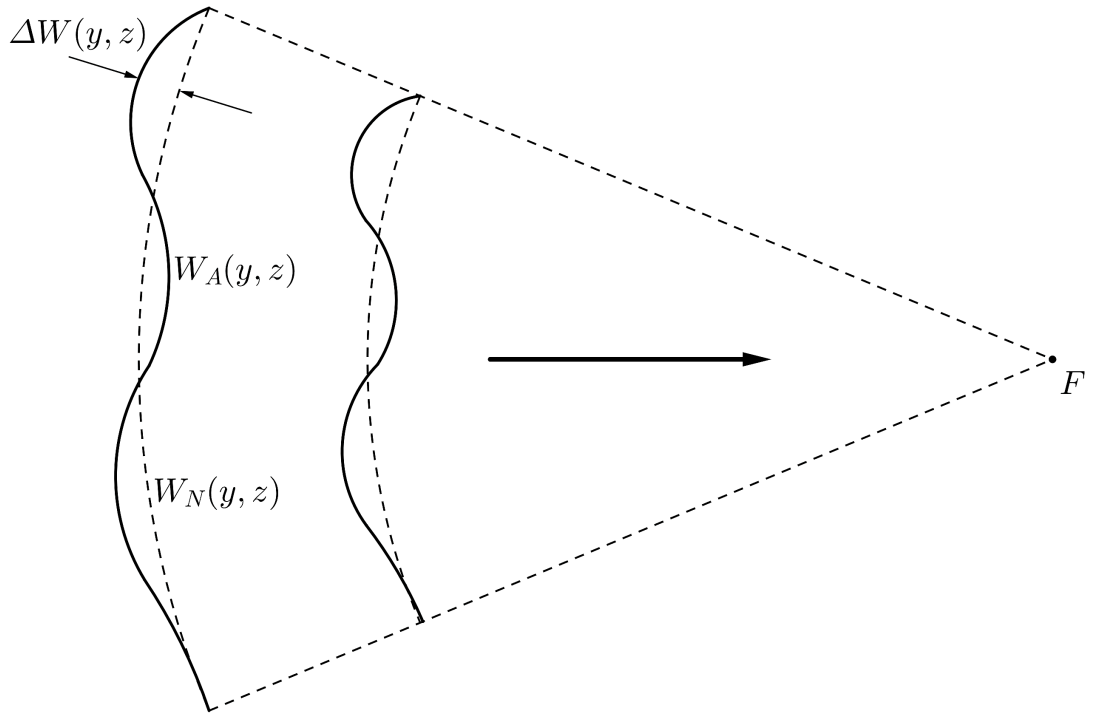


Figure 5.3. Optical path difference.

Consider the Figure 5.3. Here, the point F is the focal point for a normal and undisturbed spherical wave $W_N(y,z)$. $W_A(y,z)$ is the aberrated wavefront due to an imperfect optical device. For this setup, optical path difference (OPD) is defined as

the difference between aberrated wavefront and the normal wavefront [59]; that is,

$$\mathrm{OPD}(y,z) = \Delta W(y,z) = W_A(y,z) - W_N(y,z) \tag{5.7}$$

$\Delta W(y,z)$ is called *wavefront aberration function* because it shows how much the resulting wavefront is aberrated from the perfect spherical wave. Conventionally, the difference is taken as positive if the aberrated wave leads the spherical wave, and negative for a following distorted wave. In general, the wavefront aberration function is measured either in micrometers (μm) or in wavelengths.

The wavefront aberration function is used as the optical path difference. Hence, in the calculations for diffraction, it is used in the phase term as in Section 2.6. Since the waves considered in this work are the light waves with very small wavelength, the optical path in micrometer or several wavelength scale makes a very important impact due to the wave number term $k = 2\pi/\lambda$ in the phase. To measure such small length differences, special methods must be used.

5.3. Ray Tracing

There are several methods to determine the wavefront aberration function such as automatic retinoscope [69], Tscherning [70], Hartmann-Shack wavefront sensor [50] and ray tracing [71, 72]. The most common method is to use Hartmann-Shack wavefront sensor [73], the principals of which are explained in Section 4.2.2, and this work uses ray tracing method since the available device uses it.

The ray tracing methodology is pictured in Figure 5.4. In this method, a laser beam is sent to the eye parallel to its line of sight. After refracted onto the retina, the light is reflected back. The scattered light then passes through the eye lens affected from the aberrations of it. The reflected light is directed via a beam splitter and a lens onto, and captured by, a linear array of photodetectors. The "healthy" falling position on the photodetector and the actual one are compared for the specific position on the

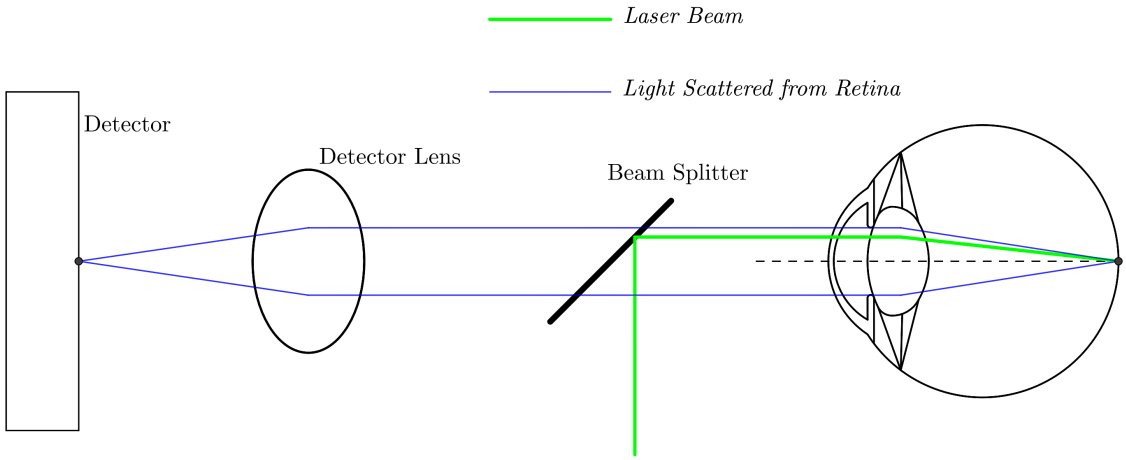


Figure 5.4. Ray tracing working principle.

pupil the laser beam is sent to. After that, the beam is shifted to another position until the the whole pupil is scanned according to a scanning geometry.

When the scanning is complete, a set of data showing the deviation from the "healthy" position of the focal point is processed. This processing includes the calculation of the necessary surfaces to obtain the actual surface deviating each laser beam as much as measured deviation. The required surface is tried to be reconstructed from the Zernike polynomials that the device can utilize. Reconstruction is done by using the least squares method [73]. As the result, the coefficients of the Zernike polynomials, c_k, that help them to be linearly added to find the necessary surface in the formula

$$\boxed{\Delta W(\rho,\theta) = \sum_k c_k Z_k(\rho,\theta)}$$

are obtained.

The ray tracing has several advantages. First of all, the method uses forward projection. That means the method projects the focal shift onto the retina and measures it via detector, whereas the backward projection requires a source on the retina and the focal shift is projected onto a camera, which itself can have aberrations. The second advantage is that it is a parallel method. That means the measurements are taken

point by point on the aperture. In series methods, for instance the in the Hartmann-Shack sensors, the all data are taken in one measurement. That is, the reflected light is evaluated as a point source back of the eye optics and the aberrations from the point source are measured in one take. By this way, the cross-over effect, which is the confusion of which aberration stems from where, is eliminated. The flexibility of the method allows different scanning geometry other than rectilinear or circular. Finally, the simplicity of the method makes it very robust to higher order aberrations.

In this work, the ray-tracing method is applied by iTrace System developed by Tracey™ Technologies. The device uses 785 nm infrared light on the order of 100 μm width [74]. iTrace projects this beam parallel to eye's line of sight to 64 points in the concentric rings geometry, and for each point it happens 4 times (totally 256 points), completed in 250 milliseconds. For an emmetropic eye, those beams are focused on one point on the retina due to the eye optics. For ametropic eye, they are spread to different points. These points are then interpolated and a surface is obtained. It is then tried to be approximated by least squares fit with the first 45 Zernike polynomials, and wavefront aberration function is obtained in micrometers.

5.4. Point Spread Function

Consider the Eq. (2.58);

$$U_{al}(y_a, z_a) = e^{jkn_l d_0} U_{bl}(y_a, z_a) \mathrm{P}(y_a, z_a) \exp\left\{-jk\frac{(y_a^2 + z_a^2)}{2f}\right\}$$

In this equation, $U_{al}(y_a, z_a)$ and $U_{bl}(y_a, z_a)$ are the wave-related terms, and the phase term containing a fraction is related with the focal distance of the lens (see the derivation in Section 2.5) which is constant for crystalline lens of the eye. Hence the pupil- and lens-related terms are $\mathrm{P}(y_a, z_a)$ and $e^{jkn_l d_0}$.

In Eq. (2.63), only the pupil window function is put into the integral. If, however, the lens is not perfect, as said in Section 5.2, then the traveled distance by the wave inside the lens, hence the phase of each point, is different. In this case, an additional

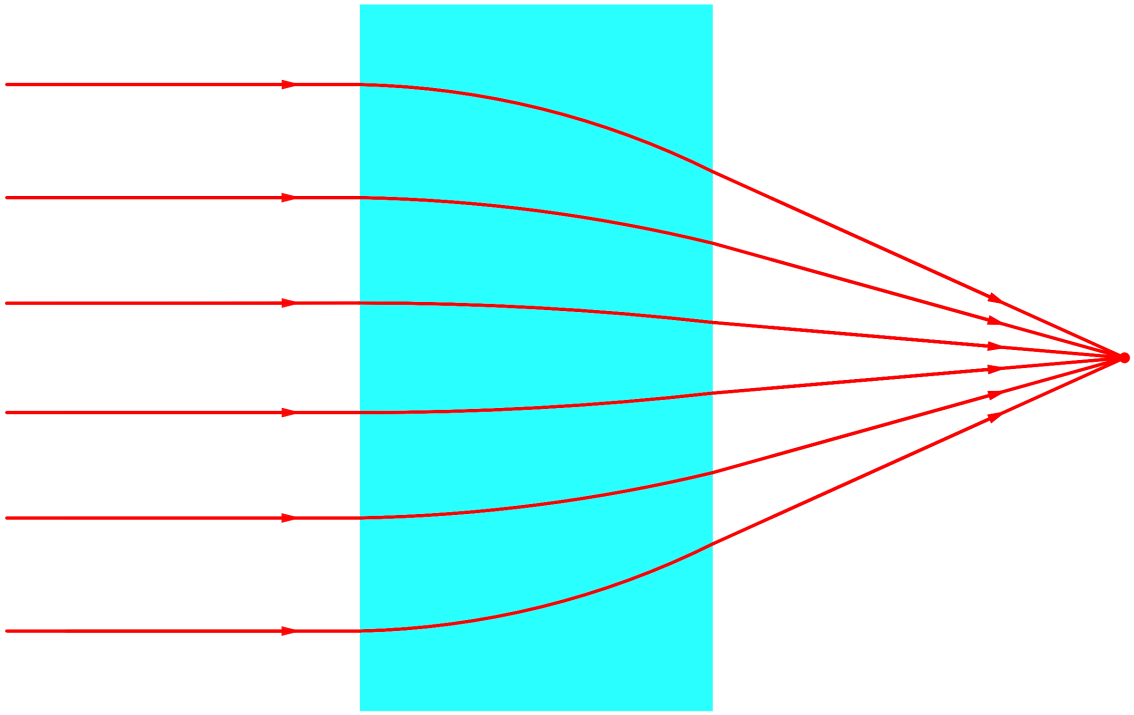


Figure 5.5. Refraction in gradient-index (GRIN) lens.

phase term should be included into the Eq. (2.63). The generalized pupil function, $\mathcal{P}(y_a, z_a)$, is described as

$$\mathcal{P}(y_a, z_a) = \mathrm{P}(y_a, z_a) \exp\{jkn_l(d_0 + \Delta W(y_a, z_a))\}$$

where $\mathrm{P}(y_a, z_a)$ is the pupil window function, n_l is the refractive index of lens, d_0 is the maximum thickness of the lens and $\Delta W(y_a, z_a)$ is the wavefront aberration function. $\Delta W(y_a, z_a)$ is multiplied by the lens' index of refraction because it is the optical path difference inside a medium whose refractive index is n_l.

As mentioned before, the crystalline lens in the eye has gradient-index (GRIN) nature, whose effect is represented in Figure 5.5. This lens type has varying refractive index inside of itself to focus the light a point better. Actually, the eye lens has the same feature; its center refractive index is 1.406 and at the edges, it's 1.386. GRIN lens of the eye has a polynomial change in refractive index with changing radial distance

from the centers, according to the equation [75]

$$\boxed{n_l(r) = n_o + n_d r^p} \tag{5.8}$$

where n_l is the refractive index of lens, n_o is the index of refraction at the center, n_d is the index difference between periphery and the center of the lens, r is the normalized radial distance on the lens, and p is the power determining the exponential decay. The power p is different for equatorial and axial part of the lens. This equation is a function of the radial distance on the lens, meaning that it also depends on the aperture coordinates. So, the equation should be modified as

$$\mathcal{P}(y_a, z_a) = \mathrm{P}(y_a, z_a) \exp\left\{jkn_l(y_a, z_a) \times (d_0 + \Delta W(y_a, z_a))\right\} \tag{5.9}$$

As the last step, the strongest refractive part of the eye, cornea, should be included in the equations. Normally, its refractive index is constant, and together with the crystalline lens, they create a system whose focal point is 17.1 mm away from the exit of the lens [7]. However, the visual impairments can stem from the imperfect crystalline lens as well as a deformation in the cornea. Consequently, including the contribution of cornea, the generalized pupil function for human eye is

$$\boxed{\mathcal{P}(y_a, z_a) = \mathrm{P}(y_a, z_a) \times e^{jkn_c[d_c + \Delta W_c(y_a, z_a)]} \times e^{jkn_l(y_a, z_a)[d_0 + \Delta W_l(y_a, z_a)]}} \tag{5.10}$$

where $\mathrm{P}(y_a, z_a)$ is the pupil window function, $k = 2\pi/\lambda$ is the wave number, n_c is the refractive index of cornea, d_c is the cornea thickness, $\Delta W_c(y_a, z_a)$ is the wavefront aberration function coming from the cornea, $n_l(y_a, z_a)$ is coordinate dependent refractive index of crystalline lens, d_0 is the maximum lens thickness and $\Delta W_l(y_a, z_a)$ is the wavefront aberration function coming from the crystalline lens. If Eq. (5.10) is inserted

into Eq. (2.63), the point spread function becomes

$$\boxed{\text{psf}(y_s, z_s) = \left| -M \iint_{-\infty}^{\infty} \mathcal{P}(\lambda r_s \tilde{y_a}, \lambda r_s \tilde{z_a}) \exp\left\{-j2\pi \left(\tilde{y_a} y_s + \tilde{z_a} z_s\right)\right\} d\tilde{y_a} d\tilde{z_a} \right|^2} \tag{5.11}$$

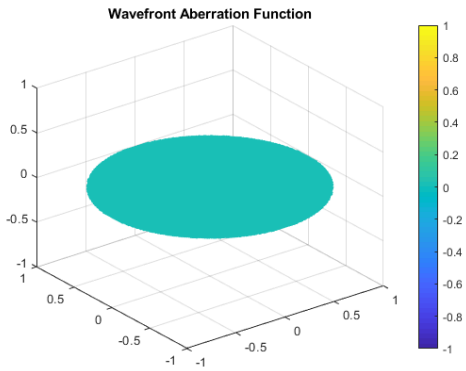
after which the $\text{psf}(y_s, z_s)$ is normalized such that $\iint_{-\infty}^{\infty} \text{psf}(y_s, z_s) dy_s dz_s = 1$. Note that the phase term depending on the cornea thickness in Eq. (5.10) is independent from the pupil coordinates. Hence, it can be taken out of the Fourier transform. In this case, it can be ignored since absolute value of any complex exponential is equal to 1.

5.5. Mathematical Models of Emmetropic and Ametropic Eye

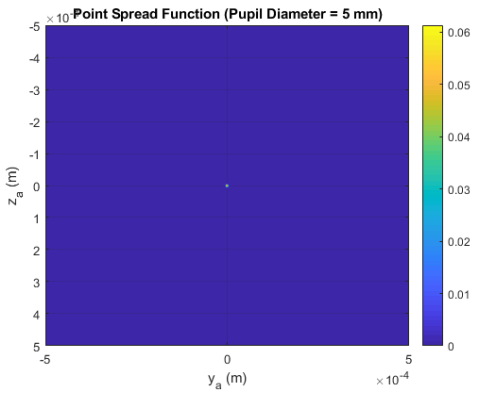
Emmetropia is the condition in which an object at infinity[5] is sharply focused by an relaxed eye on the retina exactly. It comes from the Greek words *en* meaning "in", *metron* meaning "measure", and *ops* for "sight"; hence, literally, it means "well-adjusted sight". Emmetropic eye is the eye, cornea and lens of which have no deformation or problem. Ametropia is, on the other hand, the condition where the vision is not sharply focused on retina. The Greek prefix *a* means "to take away". Ametropic eyes have imperfections in cornea and/or lens structure.

Since emmetropic eye does not have deficiency either in cornea or crystalline lens, the waves passing through them will not experience an additional phase difference stemming from wavefront aberration. Hence, in Eq. (5.10), $\Delta W_c(y_a, z_a)$ and $\Delta W_l(y_a, z_a)$ are both zero for every (y_a, z_a) pair. The wavefront aberration functions for both cornea and the crystalline lens appear as $Z_0^0(\rho, \theta)$ in Figure 5.2, and the Fourier transform of pupil function becomes Airy pattern as in Figure 2.15. But since the distance from the pupil to the retina is not as big as the necessary distance to observe the pattern clearly, the point spread function will appear as a point on the retina. Hence, the impulse response of eye can be considered as a Dirac delta function in two dimensional space, and the image will be sharply focused.

[5]For all practical purposes, 6 m and further away is accepted as infinity in human vision.



(a)

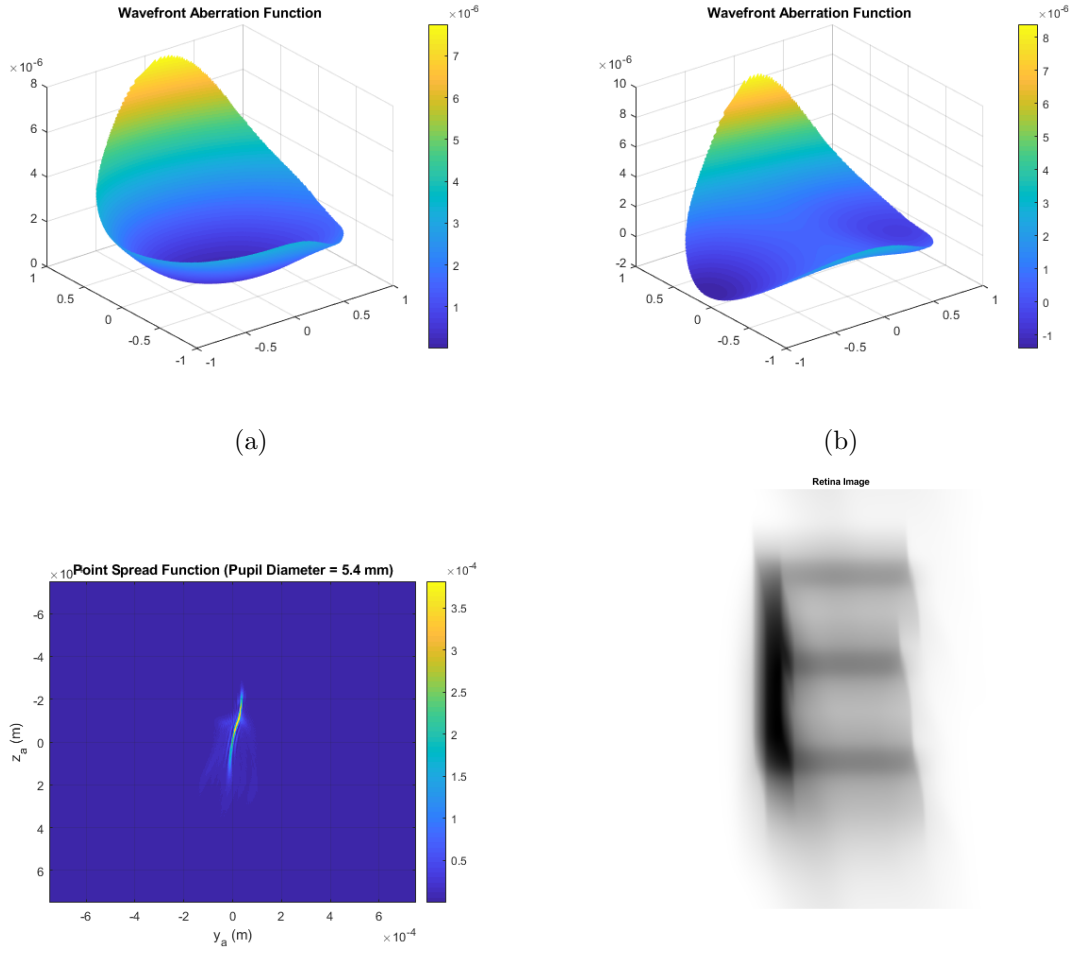


(b)

(c)

(d)

Figure 5.6. Mathematical example of an emmetropic eye. (a) is the wavefront aberration function, (b) is the total point spread function, (c) is the original image, (d) is the image falling onto the retina. Here, the lens is taken as a GRIN lens with n_o is 1.406, n_d is -0.02, p for equatorial part is 6.30 and p for axial part of the lens is 4.90 [76].



(a) (b)

(c) (d)

Figure 5.7. Mathematical example of an ametropic eye. (a) is the wavefront aberration function of lens only, (b) is the wavefront aberration function of cornea alone, (c) is the total point spread function, (d) is the image falling onto the retina. Here, the lens is taken as a GRIN lens with the same parameters as in Figure 5.6. Cornea has a constant refractive index of 1.376.

The resulting wavefront aberration function, point spread function, original image and retina image are shown in Figure 5.6. Note that the retina image is a little blurred. This stems from the diffraction-limited nature of eye's imaging system. Moreover, the brain has its compensation for that much blur. The image is what falls onto the retina, not what the brain sees. The nervous system processes the image further in the brain to merge the incoming signals from different rod and cone cells, to correct the upside-down image and to eliminate the blur that is inevitable for the eye's imaging system.

For an ametropic eye, the cornea and/or the crystalline lens introduce non-zero wavefront aberration function which is represented by the linear combination of Zernike polynomials. In this case, the impulse response of the eye will not be a Dirac delta function on the retina but a spread point. Hence the image becomes blurred according to the resulting point spread function. In Figure 5.7, the wavefront aberration functions for cornea and crystalline lens are shown together with the resulting point spread function and the fallen image onto retina.

Please note that for both Figure 5.6 and Figure 5.7, the simulations do not consider the image distance from the eye or the image size covered in the visual area. The results are obtained via direct convolution of point spread function matrix and the original image. This is the reason for black areas appeared in Figure 5.7(d) and border-like part in Figure 5.6(d).

5.6. The Factors That Changes Point Spread Function

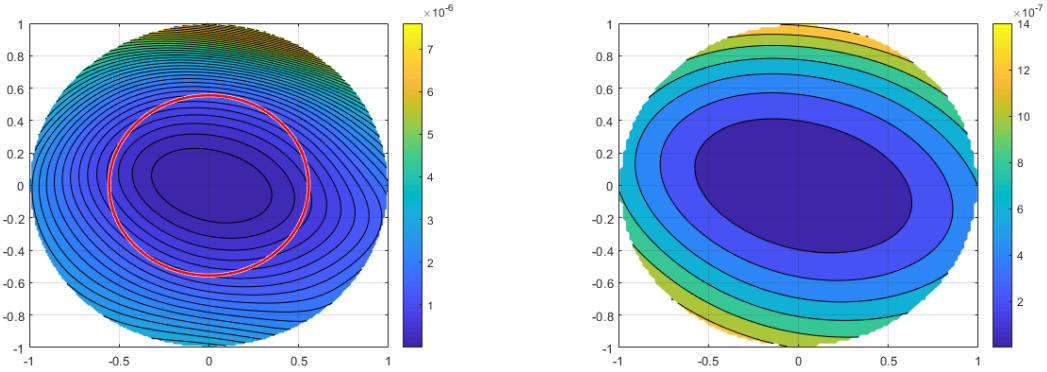
As can be seen in Eq. (5.11), the point spread function depends on several factors. They can be listed as the distance between the pupil and the retina (exit pupil distance), the wavelength of the light, refractive index of the eye lens and cornea, the effect of which are obvious since it is mathematically changes the point spread function, and the pupil diameter. Indirectly, the object distance may also be effective on the point spread function since the accommodation of the lens changes the focal point of the crystalline lens and the exit pupil distance.

5.6.1. Different Pupil Diameter

As mentioned before, the Zernike polynomials are orthogonal over the unit circle; that is, $0 \leq \rho \leq 1$. To benefit this feature, the circular aperture radius to be represented by Zernike polynomials is scaled to [0 1]. Although representation of the system is done on the same unit circle, the different aperture radius of the same system means different Zernike polynomial combination. Consider the Figure 5.8 for instance. In Figure 5.8(a), the phase distortion is given introduced by a system (an eye) having 5.4 mm radius. When the aperture gets smaller, the phase distortion changes since the path that light can take differs. In this case, the phase distortion is introduced by the part which the change in radius does not cover. For Figure 5.8(b), the aperture has 3 mm diameter. Note that the effect is the same as limiting the wavefront radius from 5.4 mm to 3 mm (red circle in the first contour map). In both figures, the radius is scaled to a range in which the maximum value is 1. The coefficients for the first five Zernike polynomials for Figure 5.8(a) and Figure 5.8(b), respectively, are

$$[1.7543,\ 0.4927,\ -0.0858,\ 0.2223,\ 1.0749]$$
$$[0.4872,\ 0.0174,\ -0.0348,\ 0.0814,\ 0.2924]$$

Since a larger pupil contains every information for a smaller pupil, the Zernike polynomial coefficients can, and should, be calculated for the small pupil from the larger one due to the fact that the Zernike polynomials for a large pupil are not orthogonal for a smaller one with the same coefficients. The first study to scale the pupil for new set of Zernike polynomials was done by Kenneth Goldberg and Kevin Geary [77]. They thought that the since the orthogonality must hold, the changing part must be the radial part for the new fit according to the ratio of new and old radii. Later, Jim Schwiegerling introduced a new method [78] to find the new Zernike polynomial weights from the old ones and the new found coefficients recursively, which was later corrected in a minor way and simplified by GuangMing Dai [79]. Apart from those



(a) (b)

Figure 5.8. Representation difference of the same aperture for different aperture radii. (a) is the sum of first 45 Zernike polynomials for a system having an aperture with radius 5.4 mm, (b) is the sum of first 45 Zernike polynomials with different coefficients for the same system with 3 mm aperture. Note that the red circle in (a) is the same wavefront as (b). The contour line step is 0.2 μm.

methods, this work uses the matrix-based method developed by Charles Campbell [80].

Wavefront aberration function can be represented as the linear combination of the Zernike polynomials as

$$\Delta W(\rho,\theta) = \sum c_n^m Z_n^m(\rho,\theta)$$

This equation can be written in terms of matrix product;

$$\Delta W(\rho,\theta) = \mathbf{Zc} \tag{5.12}$$

where $\mathbf{c}$ is the column vector of the polynomial coefficients, and $\mathbf{Z}$ is the row vector of the Zernike polynomials. The ordering in both $\mathbf{c}$ and $\mathbf{Z}$ is a little bit different, which can be understood best visually in the Zernike polynomial pyramid.

The ordering is shown in Figure 5.9. It starts from the bottom row of the pyramid from the leftmost element, then continues to right checking the "higher floors". If there

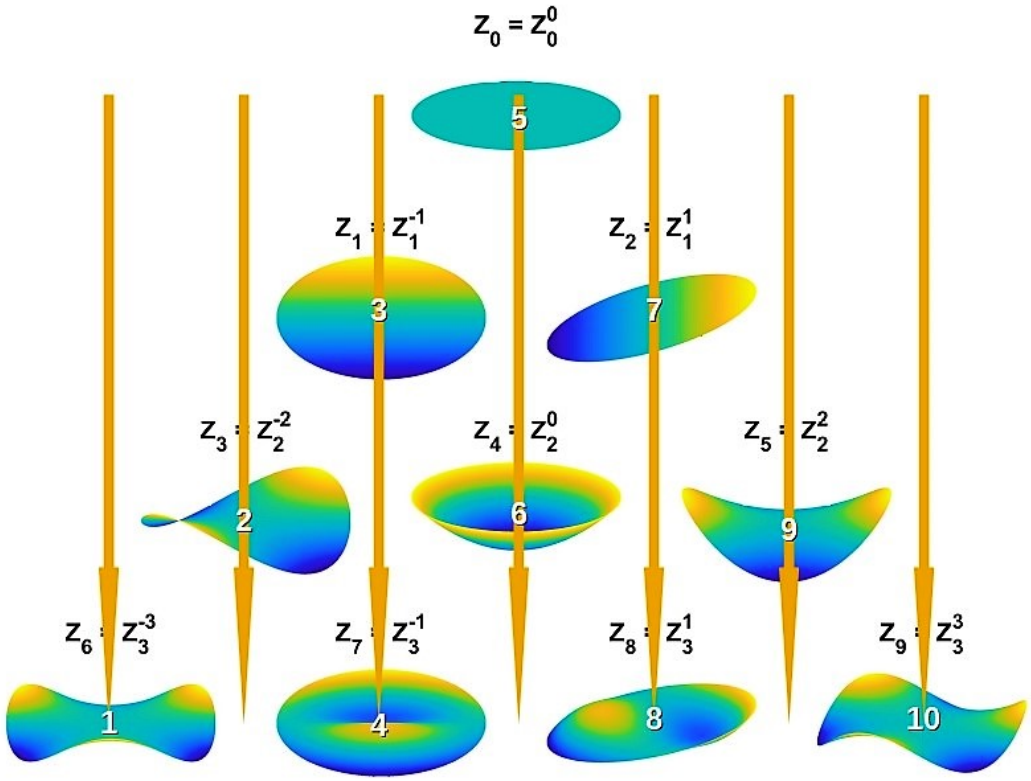


Figure 5.9. Ordering schematic of polynomials and their parameters in calculations of new coefficients for different pupil size.

is any element on the upper floor, they are put into the vector starting from top to bottom. Finishing the procedure, it moves to right. This placing procedure continues until the pyramid is finished. The coefficient vector **c** is a column vector with this ordering from top to bottom, and **Z** is a row vector using the same placing from left to right. For the Figure 5.9, the Zernike polynomial row vector **Z** and coefficient column vector **c** are

$$\mathbf{Z} = \begin{bmatrix} Z_3^{-3} & Z_2^{-2} & Z_1^{-1} & Z_3^{-1} & Z_0^0 & Z_2^0 & Z_1^1 & Z_3^1 & Z_2^2 & Z_3^3 \end{bmatrix}$$
$$\mathbf{c} = \begin{bmatrix} c_3^{-3} & c_2^{-2} & c_1^{-1} & c_3^{-1} & c_0^0 & c_2^0 & c_1^1 & c_3^1 & c_2^2 & c_3^3 \end{bmatrix}^T$$

Since the Zernike polynomials are put into o matrix form as **Z**, the Eq. (5.1) should be put into the same form. Beginning with the normalization constant N_n^m, its

matrix form is

$$\mathbf{N} = \begin{bmatrix} N_{n_{max}}^{-n_{max}} & 0 & 0 & \cdots & 0 \\ 0 & N_{n_{max}-1}^{-(n_{max}-1)} & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & N_{n_{max}-1}^{n_{max}-1} & 0 \\ 0 & 0 & \cdots & 0 & N_{n_{max}}^{n_{max}} \end{bmatrix} \tag{5.13}$$

Note that the ordering is done in such a way that the correct normalization constant is multiplied by the corresponding polynomial. Again, the example $\mathbf{N}$ is

$$\mathbf{N} = \begin{bmatrix} N_3^{-3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & N_2^{-2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & N_1^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & N_3^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_0^0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & N_2^0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & N_1^1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_3^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_3^3 \end{bmatrix}$$

The radial polynomial matrix should be matched with the ordering pattern, and it should be separated from the radial component. For the separation, Eq. (5.3) is rearranged with the change of variables $n - 2k = p$,

$$R_n^m(\rho) = \sum_{p=|m|;p\to p+2}^{n} \frac{(-1)^{\left(\frac{n-p}{2}\right)} \left[\frac{n+p}{2}\right]!}{\left[\frac{n-p}{2}\right]! \left[\frac{p+|m|}{2}\right]! \left[\frac{p-|m|}{2}\right]!} \rho^p$$

where $p \to p+2$ indicates that p increases 2 for the next step. By this arrangement, the radial component ρ becomes a variable weighted by the first fraction term, and the

matrix $\mathbf{R}$ can be created from those weights. Following the matrix formation pattern that the formation first considers the meridional index m, the matrix $\mathbf{R}$ should be a block diagonal, $(2n_{max} + 1) \times (2n_{max} + 1)$ matrix,

$$\mathbf{R} = \begin{bmatrix} [\mathbf{R}(-n_{max})] & 0 & 0 & \cdots & 0 \\ 0 & [\mathbf{R}(-(n_{max}-1))] & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & [\mathbf{R}(n_{max}-1)] & 0 \\ 0 & 0 & \cdots & 0 & [\mathbf{R}(n_{max})] \end{bmatrix} \tag{5.14}$$

where each $[\mathbf{R}(m)]$ is a matrix, each column of which corresponds to the polynomial order that is available in corresponding meridional index. The block matrix in the example can be formed as

$$\mathbf{R} = \begin{bmatrix} [R(-3)] & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ 0 & [R(-2)] & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{0} & [\mathbf{R}(-1)] & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{R}(0)] & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{R}(1)] & \mathbf{0} & \mathbf{0} \\ 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & [R(2)] & 0 \\ 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & [R(3)] \end{bmatrix}$$

where, for instance,

$$[\mathbf{R}(-1)] = \begin{bmatrix} R_1^{-1}(p=1) & R_3^{-1}(p=1) \\ 0 & R_3^{-1}(p=3) \end{bmatrix}$$

In the example, the bold number 0 indicates that they are 0 matrices in the proper sizes. After the separation, the row vector ρ for the radial component ρ becomes

$$\boldsymbol{\rho} = \begin{bmatrix} \rho^{n_{max}} & \rho^{n_{max}-1} & \rho^{n_{max}-2} & \rho^{n_{max}} & \cdots & \rho^{n_{max}-1} & \rho^{n_{max}} \end{bmatrix} \tag{5.15}$$

the equivalent of which for the example is

$$\boldsymbol{\rho} = \begin{bmatrix} \rho^3 & \rho^2 & \rho^1 & \rho^3 & 1 & \rho^2 & \rho^1 & \rho^3 & \rho^2 & \rho^3 \end{bmatrix}$$

As the last part, the azimuthal function $M^m(\theta)$ should be put into the matrix notation. The formation is similar with the radial matrix $\mathbf{R}$,

$$\mathbf{M} = \begin{bmatrix} [\mathbf{M}(-n_{max})] & 0 & 0 & \cdots & 0 \\ 0 & [\mathbf{M}(-(n_{max}-1))] & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & [\mathbf{M}(n_{max}-1)] & 0 \\ 0 & 0 & \cdots & 0 & [\mathbf{M}(n_{max})] \end{bmatrix} \tag{5.16}$$

The block matrices $[\mathbf{M}(m)]$ are formed in similar way with $[\mathbf{R}(m)]$. However, since for each meridional index, there is only one function according to Eq. (5.2), the block matrices are diagonal. For the example,

$$\mathbf{M} = \begin{bmatrix} [M(-3)] & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ 0 & [M(-2)] & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{0} & [\mathbf{M}(-1)] & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{M}(0)] & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{M}(1)] & \mathbf{0} & \mathbf{0} \\ 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & [M(2)] & 0 \\ 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & [M(3)] \end{bmatrix}$$

where

$$[\mathbf{M}(-1)] = \begin{bmatrix} -\sin(-\theta) & 0 \\ 0 & -\sin(-\theta) \end{bmatrix}$$

As a result, the Zernike polynomial row vector $\mathbf{Z}$ can be expressed as

$$\mathbf{Z} = \boldsymbol{\rho}\mathbf{RMN} \tag{5.17}$$

and the wavefront becomes

$$\Delta W(\rho, \theta) = \boldsymbol{\rho}\mathbf{RMNc} \tag{5.18}$$

Assume that the same surface is to be expressed in terms of different radius $\boldsymbol{\rho'}$ and different coefficients $\mathbf{c'}$, so that

$$\Delta W(\rho, \theta) = \boldsymbol{\rho}\mathbf{RMNc} = \boldsymbol{\rho'}\mathbf{RMNc'}$$

ρ is the scaled parameter for radial component r in the range [0 1]. The scaling is done as

$$\rho = \frac{r}{r_a}$$

where r_a is the aperture radius, and r is the radial component for polar coordinates. Consider a new radius r'_a. In this case,

$$\rho' = \frac{r}{r'_a} = \left(\frac{r}{r_a}\right)\left(\frac{r_a}{r'_a}\right) = \rho\frac{1}{\eta} \Rightarrow \rho = \eta\rho' \tag{5.19}$$

where

$$\eta = \frac{r'_a}{r_a}$$

Let the matrix $\boldsymbol{\eta}$ be a matrix whose elements are zeros and powers of η. To be able to obtain a row vector $\boldsymbol{\rho'}$ with the same size as $\boldsymbol{\rho}$, $\boldsymbol{\eta}$ must be a square diagonal matrix

post-multiplying $\boldsymbol{\rho'}$, each element of which must be powered with the same exponent as ρ in row vector $\boldsymbol{\rho}$. In this case,

$$\boldsymbol{\eta} = \begin{bmatrix} \eta^{n_{max}} & 0 & 0 & \cdots & 0 \\ 0 & \eta^{n_{max}-1} & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & \eta^{n_{max}-1} & 0 \\ 0 & 0 & \cdots & 0 & \eta^{n_{max}} \end{bmatrix} \tag{5.20}$$

Following the example,

$$\boldsymbol{\eta} = \begin{bmatrix} \eta^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta^3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \eta^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \eta^1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \eta^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \eta^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \eta^3 \end{bmatrix}$$

Assume that there is a transformation matrix $\mathbf{C}$ which converts the coefficient column vector $\mathbf{c}$ to $\mathbf{c'}$; that is,

$$\mathbf{c'} = \mathbf{Cc} \tag{5.21}$$

If the equality is written in terms of matrices,

$$\Delta W(\rho, \theta) = \boldsymbol{\rho'\eta}\mathbf{RMNc} = \boldsymbol{\rho'}\mathbf{RMNCc} \tag{5.22}$$

The aim is to find the transformation matrix $\mathbf{C}$. Note that the row vector $\boldsymbol{\rho'}$ and column vector $\mathbf{c}$ are pre- and post-multiplier, respectively, in the Eq. (5.22), hence they can be eliminated;

$$\boldsymbol{\eta}\mathbf{RMN} = \mathbf{RMNC}$$

From the Eqs. (5.13), (5.14), and (5.16), it can be seen that all of $\mathbf{R}$, $\mathbf{M}$ and $\mathbf{N}$ have inverses. Thus

$$\mathbf{C} = \mathbf{N}^{-1}\mathbf{M}^{-1}\mathbf{R}^{-1}\boldsymbol{\eta}\mathbf{RMN} \tag{5.23}$$

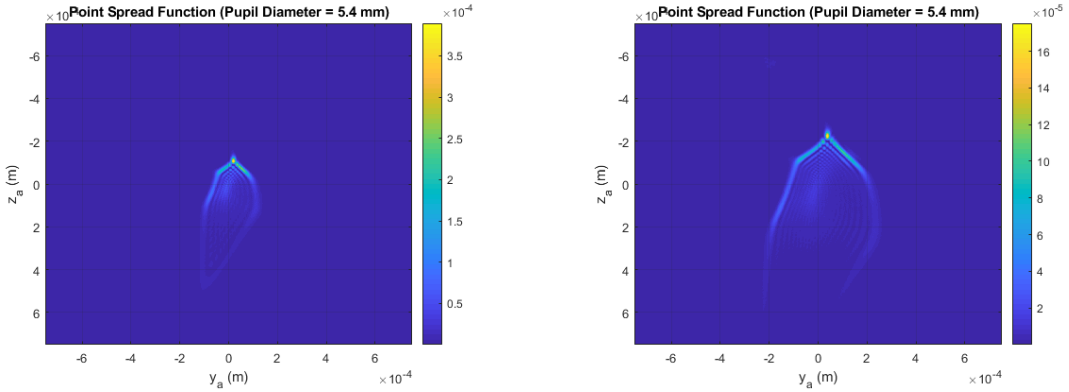
Eq. (5.23) gives the transformation matrix for the new coefficients of the new radius. It is noted in the study [80] that the meridional component matrix $\mathbf{M}$ and its inverse $\mathbf{M}^{-1}$ can be eliminated since their effect is only the scalar multiplication, and for the desired ordering of the polynomials, a permutation matrix $\mathbf{P}$ should be used, which results to

$$\boxed{\mathbf{C} = \mathbf{P}^{T}\mathbf{N}^{-1}\mathbf{R}^{-1}\boldsymbol{\eta}\mathbf{RNP}} \tag{5.24}$$

5.6.2. Different Wavelength

The wavelength of the incident light directly affects the point spread function. If Eqs. (5.10) and (5.11) are checked, it can be seen that the wavelength λ plays a role on the scaling of the coordinate system, and the phase of the pupil function via the wave number $k = 2\pi/\lambda$.

The mathematical direct effect of the changing wavelength can be seen in Figure 5.10. Note that as the wavelength gets smaller, the point spread function gets wider. This is consistent with the fact that light with smaller wavelength is refracted more, hence, spread more.



(a) (b)

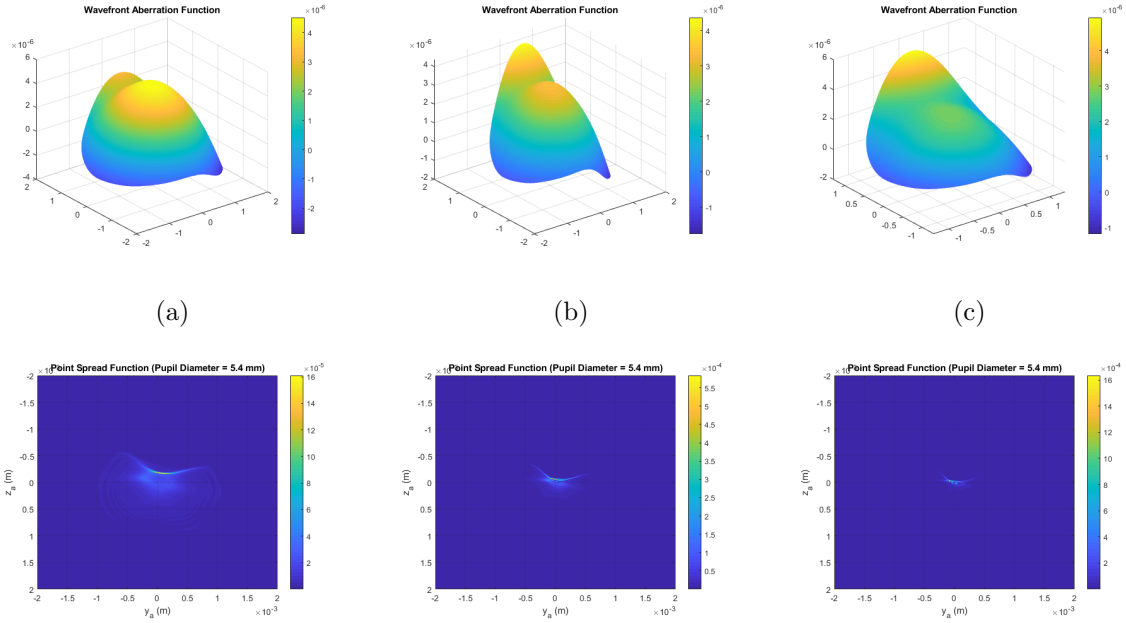
Figure 5.10. Effect of the different wavelength on point spread function. (a) is the PSF for 785 nm wavelength, (b) is for 400 nm. Here, only the aberrations on the crystalline lens are considered, the ones coming from cornea are ignored.

The biggest effect of the wavelength, however, comes from the dispersion phenomenon. As explained in Section 2.3.2, for different wavelength the index of refraction changes to another value. Because different wavelength for light means different frequency and color according to the formula $c = \lambda f$, and refractive index strongly depends on the frequency of the wave, which can be seen from Eqs. (2.14) and (2.15), the different colored lights are diffracted in separate ways.

Different diffraction pattern means different PSF, phase difference, and consequently, wavefront aberration function. In this case, the wavefront aberration representing Zernike polynomials must have coefficients other than the original ones which are obtained at different wavelength. That is; for different wavelength values, the Zernike polynomial coefficients should change. According to the studies [81, 82], the most notable and cannot-be-ignored change occurs at the coefficient of defocus term, Z_2^0. New coefficient can be calculated by the formula [82, 83]

$$\boxed{c_4 = c_2^0 = \frac{0.63346 \times r_a^2}{4\sqrt{3}} \left(\frac{1}{10^6 \lambda_0 - 0.2141} - \frac{1}{10^6 \lambda - 0.2141} \right)} \tag{5.25}$$

where $c_4 = c_2^0$ is the coefficient of defocus term, r_a is the pupil radius, λ_0 is the wavelength at which the eye is in focus, and λ is the desired wavelength, all of which are in meters. The Eq. (5.25) is obtained via polynomial fit of the data taken in diopters later converted into meters. In this work, since the measurements for crystalline lens are taken with 785 nm wavelength, λ_0 is calculated according to the mentioned value during new defocus term calculations.



(a) (b) (c)

(d) (e) (f)

Figure 5.11. Wavefront aberration function and the corresponding PSF graphs for different wavelength values. (a), (b) and (c) are the wavefront aberration functions; (d), (e) and (f) are the point spread functions. (a) and (d) are for 461 nm wavelength, (b) and (e) are for 539 nm, and (c) and (f) for 601 nm.

The effect of different wavelength on the PSF is shown in Figure 5.11. Note the change in the wavefront aberration function with the changing wavelength. As wavelength gets smaller, the effect of defocus term gets more apparent in the wavefront aberration function. Consequently, the PSF gets broader since blue-purple colored light is diffracted more than red light.

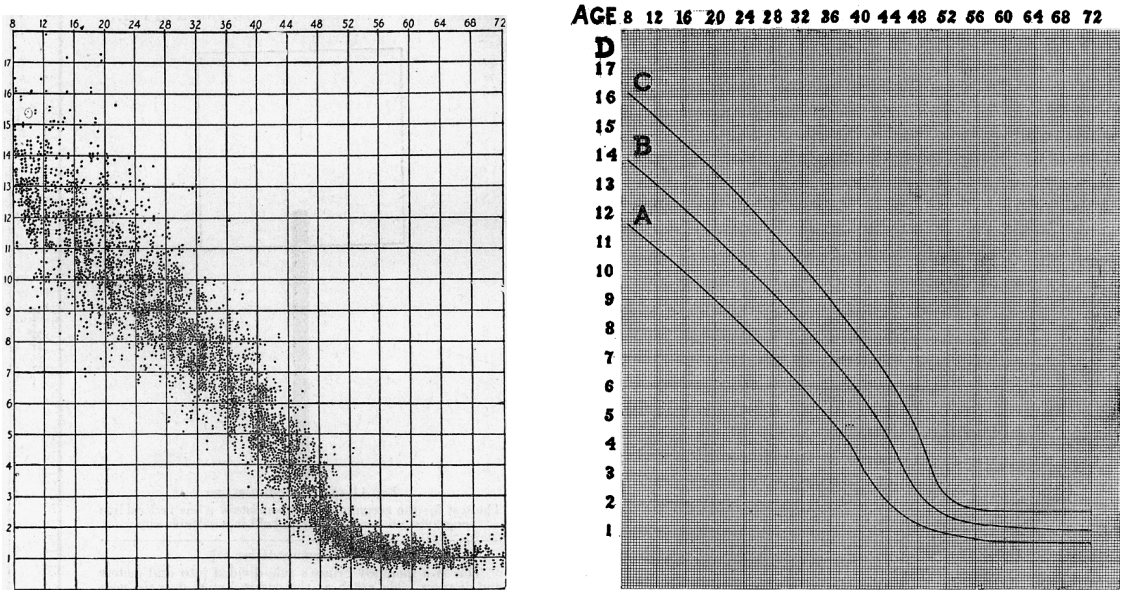
5.6.3. Effect of Age and Accommodation

As mentioned in Section 3.2, the crystalline lens of the eye is capable of accommodating the light. Accommodation property helps eye focus on the objects at different places and distances. This is done by changing the shape of crystalline lens: When the eye tries to focus on an object at a far distance, ciliary muscles are tightened and crystalline lens becomes flatter, longer and narrower. While the object is placed closer, ciliary muscles start to relax. In this case, the internal forces of the lens make it rounder, shorter and thicker. By this way, the focal point of the eye system is changed, hence, it is expected that the impulse response of eye also changes.

One of the first and most extensive studies about the accommodation was done by Alexander Duane [84]. In his study, he examined over 4200 eyes the age range of which lies between 8 and 72. Every eye exhibited a maximum accommodation value in diopters. By collecting al the data, Duane obtained an "age versus diopters" plot for accommodation. Then for every age, he extracted the maximum, minimum and average diopter values and drew separate curves for each one. Those graphs are given in Figure 5.12.

The data are extracted from the average curve by using MATLAB software. The image of the curves is given to, and via MATLAB, the data points on curve are marked. Then the coordinates of those points can be accepted as the data. For them to be useful, the diopter values are converted into the distance values by simply taking the reciprocal of the diopter value. Then the data become "age versus accommodation distance" data. Again, via MATLAB, a rational fit is found. The appropriate fit is chosen as rational fit because among the other fitting methods such as polynomial, sum of sines, Weibull, Fourier etc. the best fit presenting the ages both in [84] and ages that are not considered in the study is the rational fit. The found equation is

$$\boxed{d_{acc} = \frac{p_1 a^4 + p_2 a^3 + p_3 a^2 + p_4 a + p_5}{a^2 + q_1 a + q_2}} \tag{5.26}$$



(a) (b)

Figure 5.12. Graphs from the study [84]. (a) is the plot for age (x axis) versus diopters (y axis) of over 4200 eyes. (b) is the curves for maximum (top one), average (middle one) and minimum (bottom one) diopter curves for different human age.

where a is the person's age, and

$$p_1 = 0.000113027949$$
$$p_2 = -0.010661619305$$
$$p_3 = 0.380767402644$$
$$p_4 = -9.350307891314$$
$$p_5 = 189.669827703177$$
$$q_1 = -99.655450984002$$
$$q_2 = 2574.3649186179749$$

The focused distance and, implicitly, diopter value are important because the crystalline lens changes shape and dimension according to those values [76, 85]. When

focused to a near target, the eye lens becomes thicker. Ostrin et al. [86] gives the thickness change as 0.067 mm/D or 0.067 mm·m, a rate which is indirectly confirmed by the studies [76, 85] from their results.

Accommodation also alters the refractive index of the lens [76, 85]. Although it is accepted that accommodation changes the GRIN profile, there is no consistency in the way that change occurs. [76] claims that the axial change in refractive index becomes less steep in accommodated state whereas [85] finds that the p value in Eq. (5.8), consequently steepness, increases with accommodation.

p value also depends on the patient's age. In [76], the young people's age average is 22.82 and the olds' average is 64.26. The unaccommodated young eye has mean axial and equatorial p values 4.90 and 6.30, respectively. The same p values for old eye in the same state are 6.71 and 10.28, respectively. Assuming a linear change with age, the equations for axial and equatorial values, $p_a(a)$ and $p_e(a)$ respectively, according to an age, a, are

$$\boxed{p_a(a) = 0.04367761a + 3.90327703} \text{ and } \boxed{p_e(a) = 0.09604247a + 4.10831081} \quad (5.27)$$

5.7. Deconvolution

Deconvolution is most easily done in terms of Fourier transform since convolution in spatial domain is equivalent to multiplication in frequency domain. Suppose that the image is $I(x, y)$ and the point spread function is $\text{psf}(x, y)$. When the image is convolved with the point spread function, the resulting image, $I_c(x, y)$, is

$$\begin{aligned} I_c(x, y) &= \mathcal{F}^{-1}\left\{\mathcal{F}\left\{I(x, y)\right\}\mathcal{F}\left\{\text{psf}(x, y)\right\}\right\} \\ &= \mathcal{F}^{-1}\left\{\text{I}(X, Y)\text{PSF}(X, Y)\right\} \end{aligned} \quad (5.28)$$

where $\text{PSF}(X, Y)$ and $\text{I}(X, Y)$ are the Fourier transform of the point spread function and the image, respectively. The inverse of this operation, deconvolution of the image

with the point spread function, can be done mathematically by multiplying the image with the inverse of the point spread function in Fourier domain; that is,

$$I_{dc}(x,y) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{I(x,y)\}}{\mathcal{F}\{\mathrm{psf}(x,y)\}}\right\} = \mathcal{F}^{-1}\left\{\frac{\mathrm{I}(X,Y)}{\mathrm{PSF}(X,Y)}\right\}$$

Although this approach is correct, it has some problems. If the Fourier transforms of Figures 5.6(b) and 5.7(c) are checked, it can be seen that the point spread function behaves as a low-pass filter, having many small and zero values in its Fourier transform for large frequencies. In this case, mathematical division causes the pixel values to be extremely big or even indefinite. So, the operation has to be changed to eliminate this effect as much as possible without affecting the overall result. This can be achieved by the Wiener filter with a little modification, which is

$$\boxed{I_{dc}(x,y) = \mathcal{F}^{-1}\left\{\frac{\mathrm{I}(X,Y)}{\mathrm{PSF}(X,Y)}\frac{|\mathrm{PSF}(X,Y)|^2}{|\mathrm{PSF}(X,Y)|^2 + K}\right\}} \tag{5.29}$$

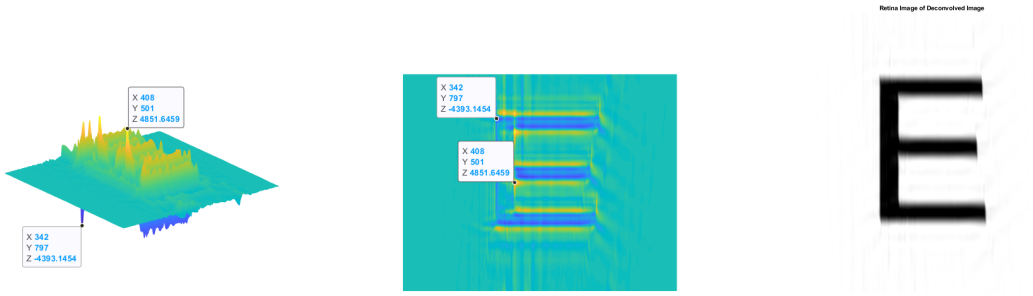
where K is a number controlled by three parameters, a, b and α, as following:

$$K = \begin{cases} 0 & , R = 0 \\ abe^{-\alpha R} + b & , \text{otherwise} \end{cases}$$

where $R = \sqrt{X^2 + Y^2}$ is the radial component in frequency domain. In this work, used values for a, b and α are 9, 0.0001 and 12, respectively, which are found empirically in study [55].

5.8. Contrast Preservation

The result of deconvolution operation should be the will-be-displayed image. If that image can be displayed, theoretically, the patient must have a sharp vision. The simulation of the deconvolved image and the image on retina are given in Figure 5.13.



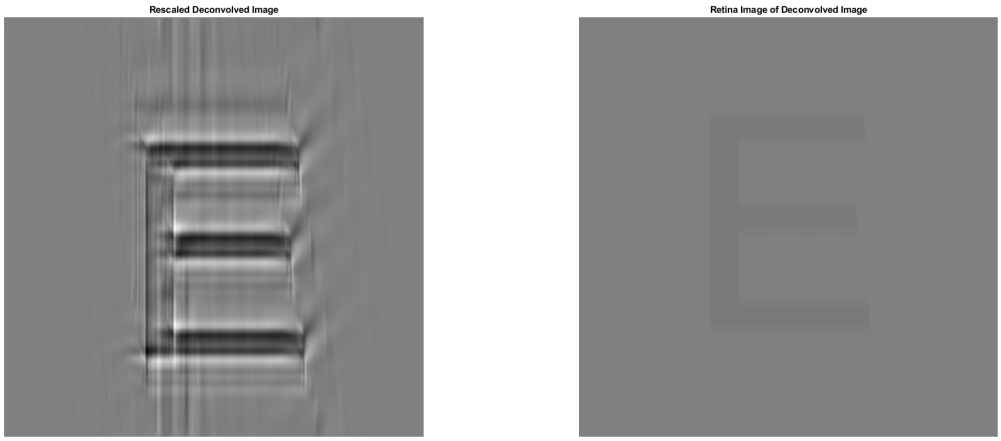
(a) (b) (c)

Figure 5.13. Result of deconvolution operation. (a) is the `surf` plot, (b) is the `surf` plot looked from over, (c) is the result of convolution with PSF and deconvolved image.

Although the outcome is promising, there is a practical limitation. As seen in Figure 5.13, the dynamic range of the deconvolved image is well beyond the regular digital screen dynamic range, [0 255]. To be able to display the deconvolved image, its dynamic range should be compressed to the necessary interval. This causes the contrast coming from the huge differences in pixel values to disappear. In this case, although the convolution with PSF gives clear image in terms of boundaries of the objects in image, there is very little contrast difference between the objects. This effect can be seen in Figure 5.14.

As a simple image processing task, the resultant image of convolution with PSF and deconvolved image can be exposed to contrast enhancement methods such as histogram equalization. However, in practice, this means modifying the image *after* it passes through the crystalline lens because the last effective part of eye shaping the point spread function is the lens, and current project is not interested in entering the eye to modify the retina image. Thus, all the contrast enhancement must be done after deconvolution but before displaying the image.

Three methods operating on the image histogram are chosen to enhance the contrast. The first method that can be tried for enhancement is histogram equalization. In this method, the overall histogram of the image is spread over the available dynamic



(a) (b)

Figure 5.14. Result of dynamic range scaling. (a) is the image that can be displayed on a common 8-bit digital screen, (b) is the simulation result of the retina image.

range, in this case [0 255], as much as possible. It is achieved by first scaling down the image to the necessary dynamic range and then creating a new image whose histogram is the cumulative distribution function of pixel values of the old image.

The second method for contrast enhancement is histogram hyperbolization. This approach is explained in detail in Section 4.1.3. The result of histogram hyperbolization together with equalization and simulated acquired vision are shown in Figure 5.15. The visual difference of those methods is that the histogram hyperbolization keeps the results "darker" because it uses the fact that human eye can detect contrast change better for low light levels than high intensity values; whereas histogram equalization tries to use all the available dynamic range for enhancement. This situation can be changed by playing with the parameter c in histogram hyperbolization.

Alternative to histogram equalization, brightness preserving bi-histogram equalization (BBHE) is proposed [87]. In BBHE, the image is divided into two sub-images by a threshold. Those images are then separately histogram-equalized. Equalization is done in a way that the image consisting of values lower than determined threshold is stretched from the minimum of the range to the threshold value, and the image

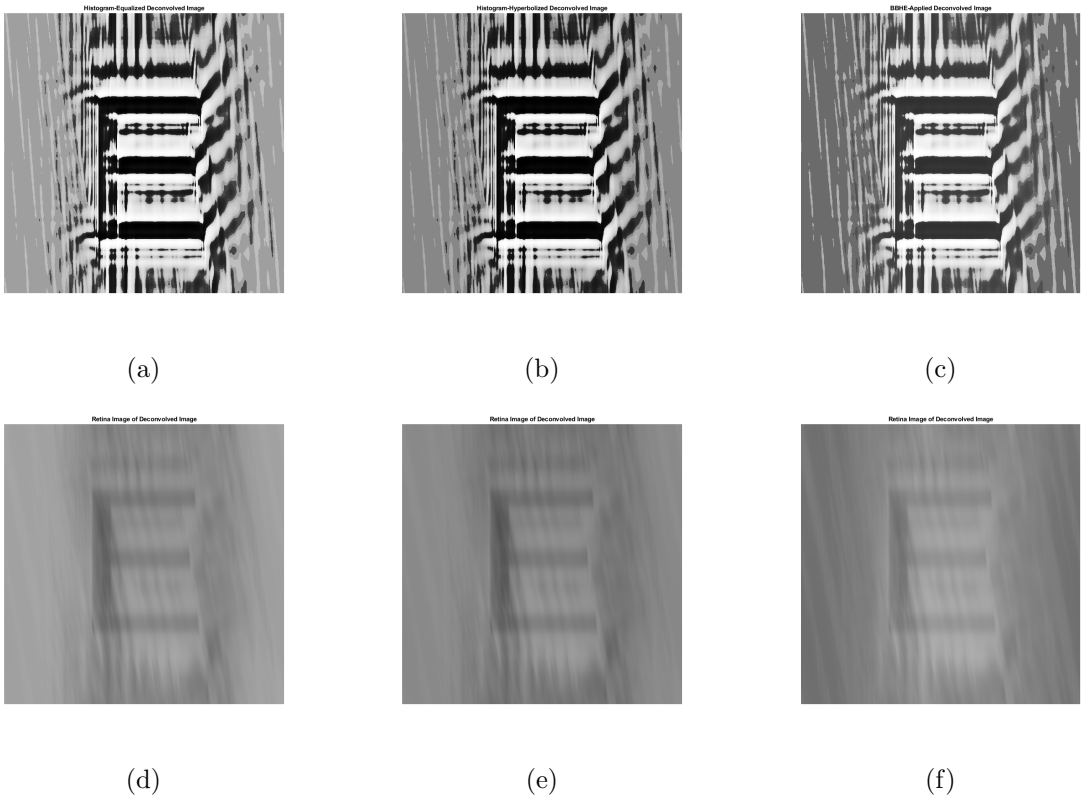
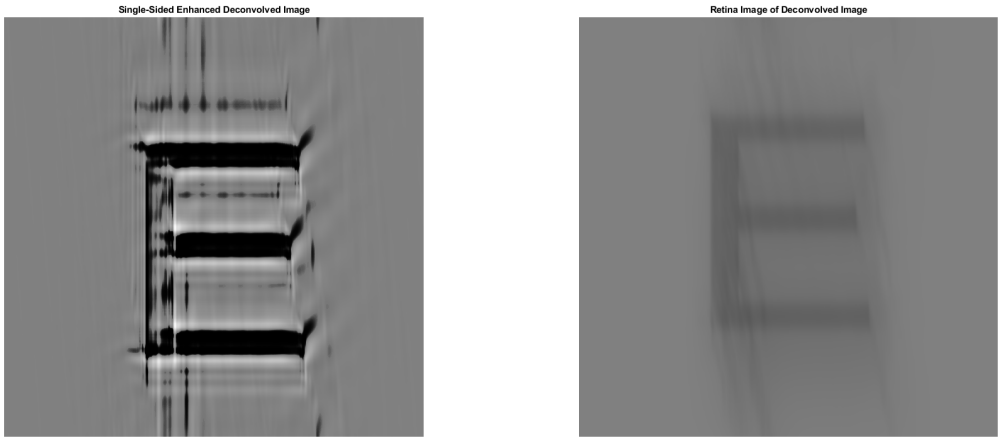


Figure 5.15. Results of histogram equalization, histogram hyperbolization and brightness preserving bi-histogram equalization (BBHE) using edge information. (a) is the histogram-equalized deconvolved image, (d) is the simulation result of the retina image for equalization. (b) is the histogram-hyperbolized deconvolved image with $c = 0.8$, (e) is the simulation result of the retina image for hyperbolization. (c) is the BBHE-applied deconvolved image with threshold determined by edge detection, (f) is the simulation result of the retina image for BBHE.

containing larger-than-threshold values is equalized from next value from threshold to the maximum of the range. Each equalization creates a mapping in their range, and the combined map is applied to the original image for total histogram equalization. In this way, the resulting image preserves its brightness. The threshold value can be chosen as mean or median value of the image, or can be determined by, for instance, edge detection [88]. In this work, threshold value is determined by edge detection, and the results are given in Figure 5.15.



(a) (b)

Figure 5.16. Result of single-sided contrast enhancement. (a) is the image that can be displayed on a common 8-bit digital screen with $\beta = 9$, $t = 120$, and $m = 50$, (b) is the simulation result of the retina image.

In literature, there is a method for enhancing the contrast of deconvolved images according to the point spread function, namely, single-sided contrast enhancement, which is described in Section 4.2.2. This method basically applies an exponential mapping on the values that are lower than a determined threshold value. Above those is applied the linear mapping; that is, all the values are mapped to themselves. The result of single-sided contrast is shown in Figure 5.16.

Single-sided contrast enhancement can be upgraded as double-sided contrast enhancement. It can be achieved by, visually, copying the behavior of profile below the threshold value to the linear part. Mathematically, this can be achieved by the equation

$$DSC(l) = \begin{cases} f_{ssc}(l) & , l \leq t \\ 2f_{ssc}(t) - f_{ssc}(2t-l) & , l > t \quad \text{and} \quad (2t-l) > 0 \\ 2f_{ssc}(t) & , l > t \quad \text{and} \quad (2t-l) \leq 0 \end{cases} \tag{5.30}$$

where

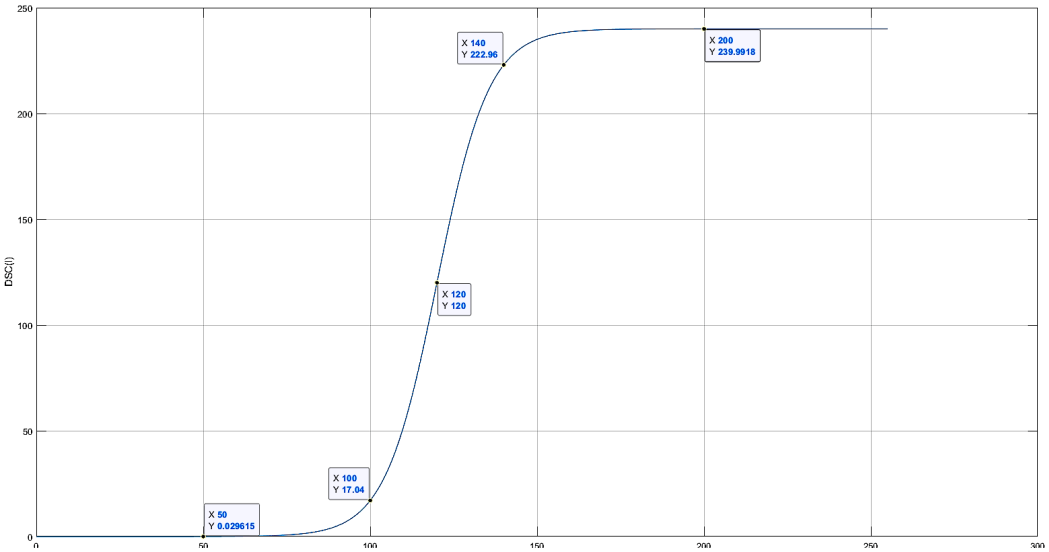
$$f_{ssc}(l) = t \cdot \frac{1 - \exp\left\{-\beta\left(\frac{l-m}{t-m} - 1\right)\right\}}{1 + \exp\left\{-\beta\left(\frac{l-m}{t-m} - 1\right)\right\}} + t$$

The profile and result of double-sided contrast enhancement are shown in Figure 5.17.

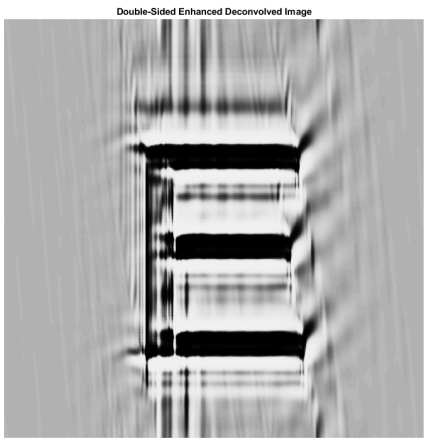
It is important to compare single-sided and double-sided contrast enhancement methods. First thing to note is that both techniques in Eqs. (4.43) and (5.30) are exactly the same in terms of mapping the pixel values that are lower than or equal to the threshold value. The difference starts for the values greater than the threshold values. Single-sided contrast method treats the bigger values linearly whereas double-sided technique takes the symmetry of the profile for $l \leq t$ around the threshold value for mapping. This difference creates some advantages and disadvantages for double-sided contrast enhancement. When the histogram of deconvolved image is examined, it can be seen that the majority of pixel values are accumulated in the middle of its dynamic range. This behavior does not change when pixel values are shifted and scaled to [0 255] range. That means there are significantly more pixels at the middle that should be clearly separated from each other for them to be visible; in other words, the contrast difference must be bigger for this interval. Double-sided enhancement allows it since its mapping is steeper than that of one-sided enhancement having a rising-slower mapping for those values. However, this property of single-sided allows it to use the available dynamic range correctly since it gives 255 as output for the highest intensity value. On the other hand, if t is chosen higher than 128, double-sided contrast enhancement starts to make pixel values exceed the highest intensity on the 8-bit screens, causing clipping of those values, hence, contrast loss again.

As a similar but different approach, this profile can be achieved by a sigmoid function;

$$\boxed{f(x; \alpha, c) = \frac{1}{1 + e^{-\alpha(x-c)}}}$$



(a)



(b) (c)

Figure 5.17. Profile and result of double-sided contrast enhancement. (a) is the mapping function of the enhancement, (b) is the image that can be displayed on a common 8-bit digital screen with $\beta = 9$, $t = 120$, and $m = 50$, (c) is the simulation result of the retina image.

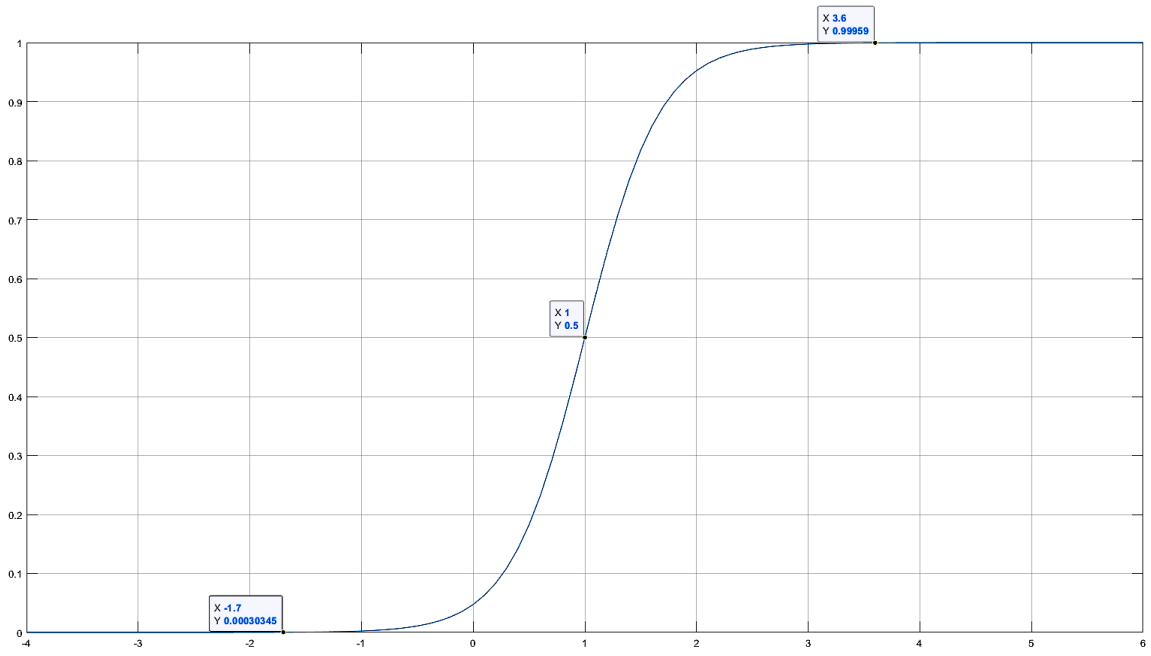
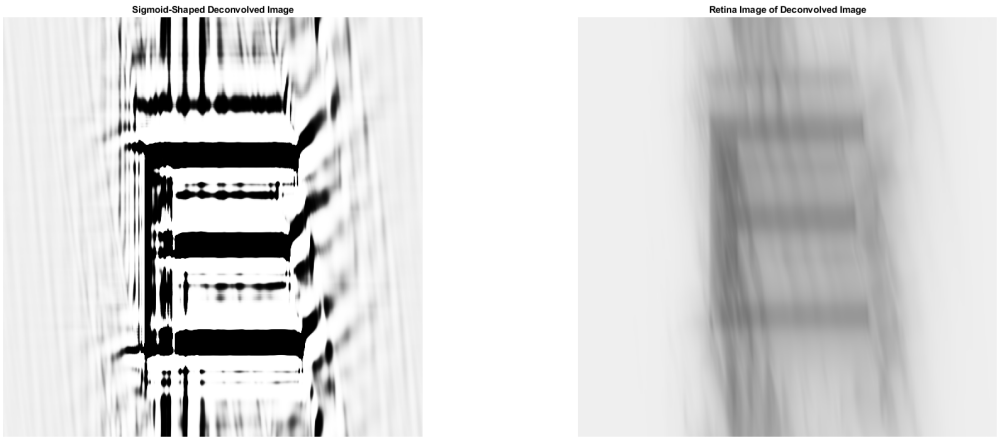


Figure 5.18. Sigmoid function from -4 to 6 for $\alpha = 3$ and $c = 1$.

where α is the steepness parameter and c is the value where $f(c) = 0.5$. An example graph of sigmoid function is shown in Figure 5.18.

Sigmoid function has some desirable properties. First of all, it has less parameters (α and c) to play with compared to singe-sided or double-sided enhancements (β, t, and m), which makes it easier to tune the function. Second, its output is always between 0 and 1 regardless of the chosen values of α and c. That means it is always bounded to an interval which is used for displayable images when the format of the image is `double` instead of `uint8`. That means this function does not introduce clipping when applied to an image. Lastly, this function does not need scaling and shifting of will-be-mapped image compared to single-sided and double-sided contrast enhancement methods requiring compressed dynamic range to [0 255] to operate. In that case, the contrast obtained in `double` format should be conserved better. Hence, the sigmoid mapping can be used directly on the deconvolved image.

The result of sigmoid mapping is shown in Figure 5.19. In double-sided enhancement, the result is darker (Figure 5.17(c)) compared to sigmoid mapping. This is



(a) (b)

Figure 5.19. Result of sigmoid mapping. (a) is the image that can be displayed on a common 8-bit digital screen with $\alpha = 0.01$ and $c = 120$, (b) is the simulation result of the retina image.

because the double-sided contrast method operates after dynamic range scaling and shifting, and has to have a threshold value chosen not to exceed the available dynamic range. On the other hand, sigmoid does not have these applications and restrictions. Moreover, the contrast in sigmoid mapping is better compared to single-sided (Figure 5.16(b)) and double-sided (Figure 5.17(c)) methods: The background is closer to white as in the original image whereas the letter is darker, again due to the restriction-free nature of method.

Note also that this method is also superior in terms of contrast of the image than the direct linear scaling (Figure 5.14(b)), histogram equalization (Figure 5.15(d)) and histogram hyperbolization (Figure 5.15(e)). Moreover, histogram equalization and hyperbolization methods create undesired artifacts in the image that are not present in the results obtained with sigmoid mapping.

As a more complicated way to increase the contrast, HDR (high dynamic range) compression methods can be used. Essentially, those methods try to compress the dynamic range of an image that does not fit to standard dynamic range without losing

the actual contrast. For this work, two methods are examined and applied: gradient domain high dynamic range compression and iCAM06.

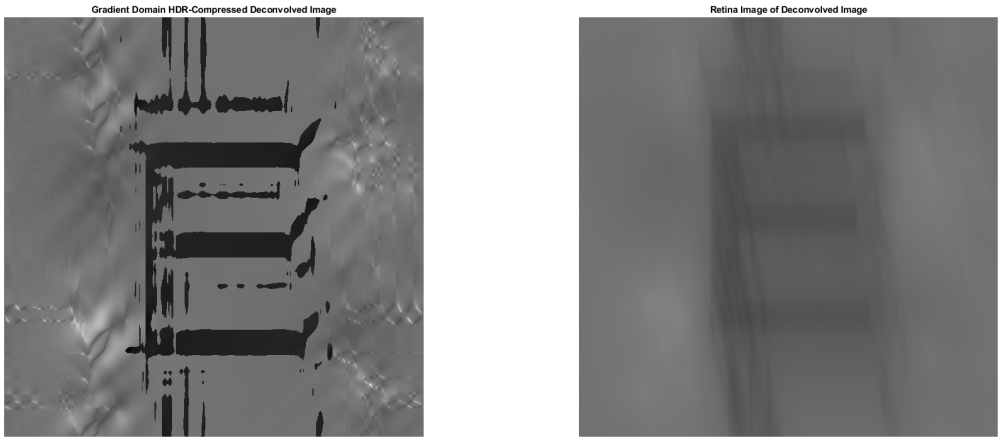
Gradient domain high dynamic range compression is proposed by Fattal et al. [89]. This method simply works on the gradient of the image because contrast is defined as the proportion of the difference between two neighbor pixels and the considered pixel. Due to the definition, the bigger the difference, the bigger the contrast. And since the change from one pixel to another can mathematically be defined by gradients, working on gradient is actually working on the contrast.

Basically, the gradient domain HDR compression method works as following: First, the luminance component of the image is extracted since color and luminance are different properties of an image and contrast is defined on the intensity values. This is done by converting an RGB image to $L^*a^*b^*$ image as described in Section 4.1.6, and taking only the L^* component. Then, to compress the higher intensities more than the lower pixel values, the logarithm of L^* component is used. Gradient is taken on the log scale of luminance component. After the compression is done via multiplication of a reduction factor, the steps are followed back.

This method is applied to deconvolved images since its dynamic range is also well beyond the standard dynamic ranges. Since the negative values in the deconvolved image cause the logarithm operation to give complex values, the image is shifted such that the minimum intensity value becomes 1. The result is shown in Figure 5.20.

Another applied method is iCAM06 method. The image color appearance model (iCAM) is first proposed in 2002 by Mark Fairchild and Garrett Johnson [90]. It is formulated "to simultaneously provide traditional color appearance capabilities, spatial vision attributes, and color difference metrics, in a model simple enough for practical applications" [90]. This model is also used for HDR image rendering [91].

The method used in this work is presented in [92], and its flowchart is given in Figure 5.21. It can be seen that gradient domain HDR compression and iCAM06



(a) (b)

Figure 5.20. Result of gradient domain HDR compression. (a) is the image that can be displayed on a common 8-bit digital screen, (b) is the simulation result of the retina image.

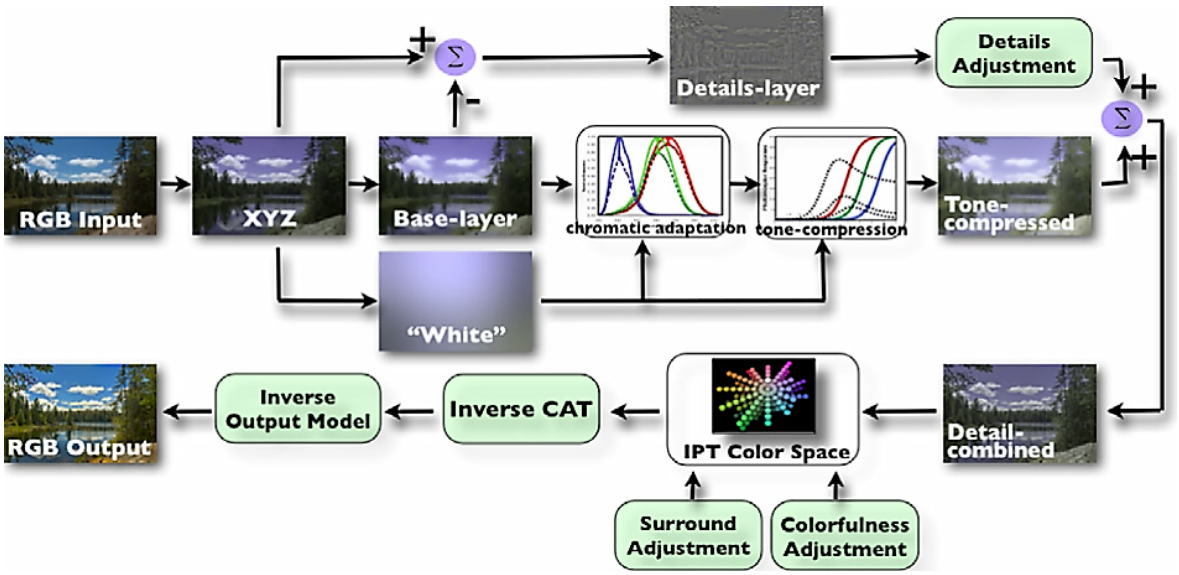
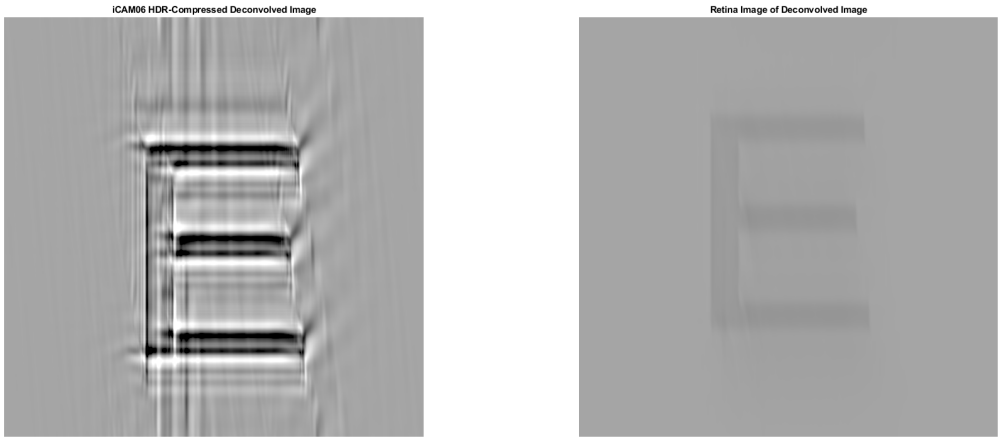


Figure 5.21. iCAM06 method flowchart. Adopted from [92].



(a) (b)

Figure 5.22. Result of iCAM06-based HDR compression. (a) is the image that can be displayed on a common 8-bit digital screen, (b) is the simulation result of the retina image.

methods have different approaches to the compression. The first difference is that Fattal et al. uses $L^*a^*b^*$ color space during processing whereas iCAM06 uses XYZ color space. Moreover, iCAM06 tries to preserve as much detail as possible by obtaining base layer with bilateral filtering. Also, iCAM06 applies tone compression according to the response of the eye to the colors. Gradient domain compression only aims to compress the dynamic range without specific detail or tone processing. The result of iCAM06 method is shown in Figure 5.22.

Gradient domain compression does not increase the contrast. The other methods create white or lighter colored areas around letter E in the images. However, gradient domain HDR compression makes those areas also gray, which is close to the color of the background and vertical artifacts around the letter. This may be stemming from the application of the method. Normally, HDR images do not contain negative intensity values. However, deconvolved images have those values. Therefore, when the method is implemented, negative values are made positive and the the whole image is processed. after which the pixels on the indexes where the negative values are placed are multiplied by -1. But this has to be done because when the method is applied after

shifting the whole image to the positive interval, the result gives gray area where there has to be a letter on the image, which actually destroys the image content.

iCAM06 method, on the other hand, gives satisfactory result in terms of the region color around the letter E. However, although it seems that it eliminates the vertical distortions, it decreases the contrast. It seems direct linear mapping of the dynamic range with just a little bit more contrast added to it.

6. COMPARISON OF RESULTS

Comparison of simulated retina images is not an easy and straightforward task since each method and approach has its own priorities, advantages and disadvantages. For instance, only scaling the dynamic range without further processing totally eliminates the artifacts resulting from the deconvolution operation but scaling and shifting the dynamic range collect all the pixel values around the middle point of the available dynamic range, hence, destroy the contrast. Histogram methods increase the details that are shown in the displayed and, thus, perceived image but they either stress artifacts around the main objects (histogram equalization and histogram hyperbolization) or make the main object "more clear" but "fainter"(BBHE). Single-sided contrast enhancement reduces the visibility of artifacts but also it makes the image darker. Double-sided contrast provides a good result but it is bounded to a limited range choice for the threshold value. Sigmoid function creates a result whose background and main object separation is the closest to the original image but it is distorted by the deconvolution side-effects. High dynamic range compressors do not perform as good as they are hoped to since the deconvolved image cannot be treated as a high dynamic range images due to that the pixel values of the image are not representing the intensities.

To be able to perform an objective assessment among the images, an evaluation approach based on an contrast measurement method is chosen [93,94]. In this approach, the original image is undersampled to a smaller image, by simply halving the image size. Then, each pixel is compared to its 8 neighbor pixels, and their differences are summed up and averaged. Then, each average obtained from each pixel in the image is summed up and averaged again. In this case, a number for the small image is obtained, and the operation continues to the next "level", another smaller image. After predetermined level number is reached, the obtained numbers are added and divided to the number of levels. As a result, the contrast measure number (cmn) is obtained. Mathematically

speaking;

$$\text{cmn} = \frac{1}{n_l} \sum_{\forall l} \left(\frac{1}{n_{pl}} \sum_{\forall p} \left(\frac{1}{8} \sum_{\forall p_s \in 8-\text{ngb}} |p_c - p_s| \right) \right)$$

where cmn is contrast measure number, n_l is number of levels, l is the current level, n_{pl} number of pixels in the level l, p_s is the surrounding pixel of the center pixel p_c and $8-\text{ngb}$ denotes the 8-neighborhood pixel.

This method is improved in study [94] such that instead of calculating differences with 8-neighborhood $|p_c - p_s|$, differences of Gaussians (DoG) are calculated for each central pixel as following:

$$\text{DoG}(x, y) = \frac{G_c(x, y) - G_s(x, y)}{G_c(x, y) + G_s(x, y)}$$

for

$$G_c(x, y) = \exp\left\{ -\left[\left(\frac{x}{r_c} \right)^2 + \left(\frac{y}{r_c} \right)^2 \right] \right\}$$
$$G_s(x, y) = 0.85 \left(\frac{r_c}{r_s} \right)^2 \exp\left\{ -\left[\left(\frac{x}{r_s} \right)^2 + \left(\frac{y}{r_s} \right)^2 \right] \right\}$$

where (x, y) is the pixel coordinate, r_c is the size for central Gaussian G_c, r_s is the size for surrounding Gaussian G_s, and $r_c < r_s$ always.

Applying this approach, every image has a single number indicating the contrast inside of it. The formulation of the method tells that the greater the number, the more the contrast. However, in the calculation it is seen that the method gives smaller numbers for images that are normally evaluated as having the best contrast by humans. Although the contrast perception is subjective, this comment is done since one of the smallest score for contrast is given to the unprocessed deconvolved image; that is, the image that gives the most clear result in simulations (Figure 5.13(c)) is given one of the smallest number. Due to this inconvenience, $\text{DoG}(x, y)$ image is not directly used.

Instead, its standard deviation is used since the contrast can be seen as the separability from the background or mean value of the image, and standard deviation is the measure for the average distance between each pixel and the mean value.

Table 6.1. The results for contrast measurement for each method.

	cmn (DoG)	**cmn (varDoG)**
Sigmoid	0.736268755689938	0.535002891663704
No Processing	0.738147786032337	0.375999941162451
Double-Sided	0.738273165176045	0.050272999265041
Single-Sided	0.738439659974214	0.049785224346678
Gradient-Domain HDR	0.743531776796119	0.020553913827253
H. Hyperbolization	0.743303433876251	0.016860989153078
H. Equalization	0.743598906777573	0.014887498081050
BBHE	0.744248986651603	0.008059302684952
iCAM06-Based HDR	0.744576348075489	0.007275782177001
Shift and Scale	0.744780087020244	0.005903970614623

The results for both approaches and for each method with 7 levels are given in Table 6.1. First of all, it should be noted that two measures give nearly the opposite results; the highest score in sum of DoG (soD) is the lowest score in variance of DoG (varDoG). The only difference occurs in the ordering of iCAM06-Based HDR and histogram hyperbolization. Since it is more convenient to think the higher the score the better the contrast, the analysis is continued on the score of varDoG approach.

Scaling and shifting only method gets the lowest score on the evaluation. The result is obvious when looking at the figures because it offers the least contrast difference among the pixel values in the whole results. It is also an expected result since scaling and shifting is the linear mapping only without considering the actual intensity differences between the pixel values. Moreover, the mapping is done to a very small range compared to the existing range, tending to collect most of the pixels around the middle of the available dynamic range.

The highest score belongs to sigmoid method despite of the ideal situation (no processing). Sigmoid gets higher score than no processing method because of the fact that both of them have nearly white background compared to the letter E together with that the result of sigmoid has more artifacts that are visible. Hence, the number of pixels that are deviated from the mean value of the image are higher, effecting the total standard deviation, thus, the contrast measure number.

The highest three methods are, excluding the no processing approach, sigmoid, double-sided and single-sided approaches. From these results, it is clear that the two new proposed methods, namely double-sided and sigmoid, are better than the state-of-the-art method for enhancing the contrast for visually impaired people, single-sided enhancement. It can be confirmed also visually by looking at the Figures 5.16, 5.17 and 5.19.

Histogram-based methods are ordered from one after another in Table 6.1. Among them, BBHE gives the worst contrast measure number. It may be stemming from the fact that BBHE does not allow to change the brightness after equalization; it tries to keep the overall brightness of processed image similar to the to-be-processed image. This feature prevents it from stretching the whole image to the values that have less occurrence in the original image, which may help the display contrast better. Ordinary histogram equalization and histogram hyperbolization perform closely to each other. However, hyperbolization seems to outperform the ordinary equalization. This is because histogram equalization tries to spread the histogram without changing the so called "histogram shape". However, since hyperbolization uses the fact that human eye can detect differences in low intensity values than high intensity values, it tries to distribute the histogram accordingly.

Gradient domain HDR compression performs better than histogram methods and scaling and shifting method but visually it seems to offer less contrast than histogram methods. The reason can be understood by looking at the Figure 5.20(a). The enhanced image by gradient domain compression includes no white or light area around the letter E opposite to the rest of the result, except the scale and shift only. This

outcome does not include a light-colored contrast area around the dark-colored letter E, meaning that a looking eye has less chance to distinguish the letter from the background. Although it is the best fourth method in the Table 6.1, it only seems so because it only adds some new pixel values that are not present in scale and shift only result, giving a bigger standard deviation. On the other hand, iCAM06-based compression seems to give the same outcome with scale and shift only with additional brightness.

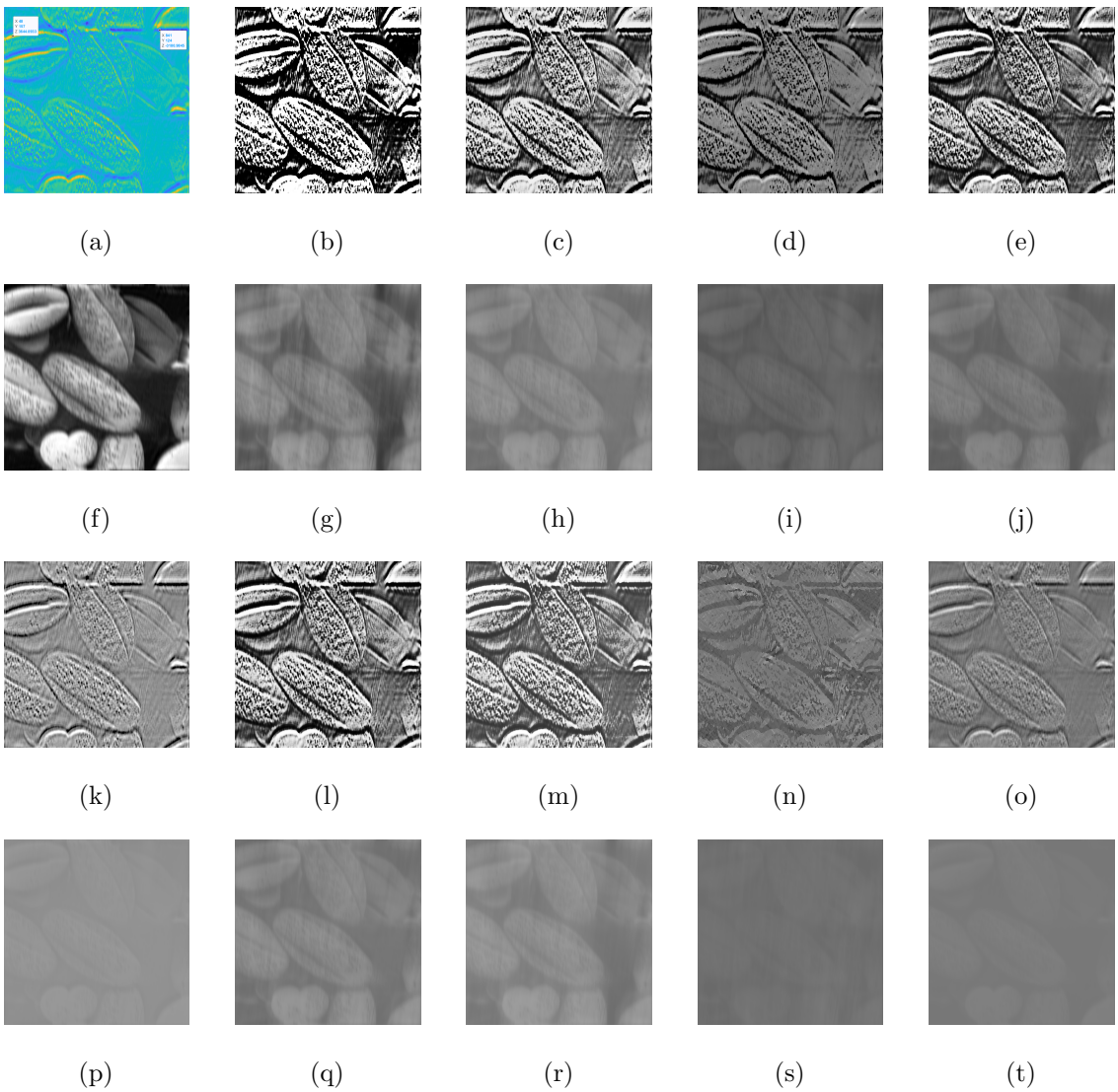


Figure 6.1. More results of the methods. The first and third rows are the deconvolved displayed images obtained from the methods. Second and fourth rows are the results of the first and third rows, respectively. The ordering of methods and the results are the same as in Table 6.2 from top to bottom.

More results with a different image are presented in Figure 6.1. The original image is the same as Figure 4.3. In these results, the highest varDoG score is 4.77597 belonging to the ideal solution, which is the natural result under normal circumstances. It is followed by sigmoid (varDoG = 0.22615), double-sided (varDoG = 0.03298) and single-sided (varDoG = 0.03197). Again, the worst approach is shifting and scaling only, with varDoG = 0.00616. The results of all methods for the new image is given in Table 6.2.

Table 6.2. The results for contrast measurement for each method for Figure 6.1.

	cmn (varDoG)
No Processing	4.775973600061374
Sigmoid	0.226152270699255
Double-Sided	0.032979943304808
Single-Sided	0.031968077758583
H. Hyperbolization	0.018706174278809
iCAM06-Based HDR	0.017714430348030
H. Equalization	0.015892556669131
BBHE	0.008435709075101
Gradient-Domain HDR	0.006177170852822
Shift and Scale	0.006164264654821

7. CONCLUSION

In this thesis, a signal processing-based approach to vision correction problem for people with refractive error is proposed. Visual impairments are generally, if they can be at all, corrected via goggles or contact lenses. If it is not an appropriate solution, medical surgery is considered. Both medical operations and corrective devices are costly. Lenses and goggles are not buy-once equipment since the impairments in the eye can change with time. Moreover, any damage to them may hamper people's daily life. Medical surgery, on the other hand, may be expensive as well as may not be an exact solution. Even if the doctors making operations are well qualified and trained, generally the surgery introduces another visual impairments to the visual system. The aim of this work is to offer people a scalable, flexible and easily reachable solution for vision problem.

In this work, the data for calculating the impulse response of eye are taken from a device working based on ray tracing principle instead of Hartmann-Shack principle. Hartmann-Shack wavefront sensor may not correctly represent and model the eye aberrations when they are large. Since ray tracing method does not have such a drawback, the data obtained by this method are more reliable, hence, its mathematical representation is more correct.

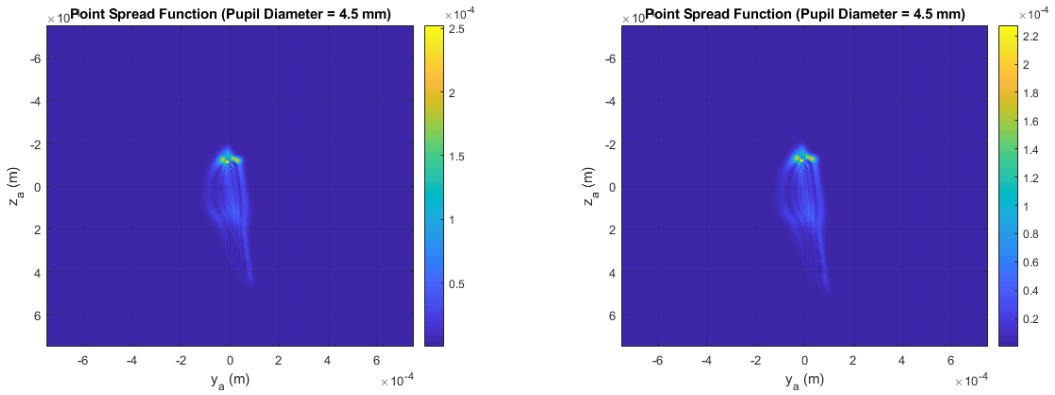
Determining the transfer characteristics of eye, point spread function calculations are mainly focused only on the crystalline lens because it is the main part of eye focusing the incident light onto the macula. However, the light first encounter with the cornea when traveling to the retina, and the biggest refraction occurs on the surface between weather and cornea. Thus, for more accurate and comprehensive modeling, contributions of cornea are also included to the calculations. iTrace device giving the lens aberration data can also determine the cornea topography. Since wavefront aberration functions are actually the distance differences between an aberrated and a normal wavefront, the cornea topography can be directly included to the wavefront aberration function calculations stemming from the cornea. Another inclusion of the

cornea is necessary for the calculation of incident wavelength coming to the crystalline lens. Due to the formula Eq. (2.13), the wavelength coming to the lens changes because it first passes through cornea and then aqueous humor, which have different indexes of refraction.

Together with the cornea contributions, lens model used in point spread function calculations in signal processing also needs to be updated. Crystalline lens has a dynamic structure. Its thickness alters with the object (or focused) distance. Also, its elasticity changes, more specifically decreases, as the eye gets older. When the elasticity changes, the ability of eye to focus different distances becomes weaker. Moreover, its refractive index is not same, for instance, at its center and its periphery. This dynamic structure also changes and affects the point spread function. For accurate modeling, object distance and effect of age are included to the model. During this work, it is seen that object distance indirectly affects the point spread function When the distance changes, the ciliary muscles reshape the lens to form a sharp image on fovea. Being thicker and/or thinner changes the focal point, hence the exit pupil distance, of the overall eye system. By this work, it is seen that the change in exit pupil distance only causes the point spread function becomes either "broader" or "narrower", in an unimportant scale, however. The result of focused distance can be seen in Figure 7.1. The calculation is done from Eq. (5.26) and by applying the fact that lens thickness change is 0.067 mm/D. Note that although the shapes and even the effects seem the same, the difference can be understood by looking at the color bar next to the graphs.

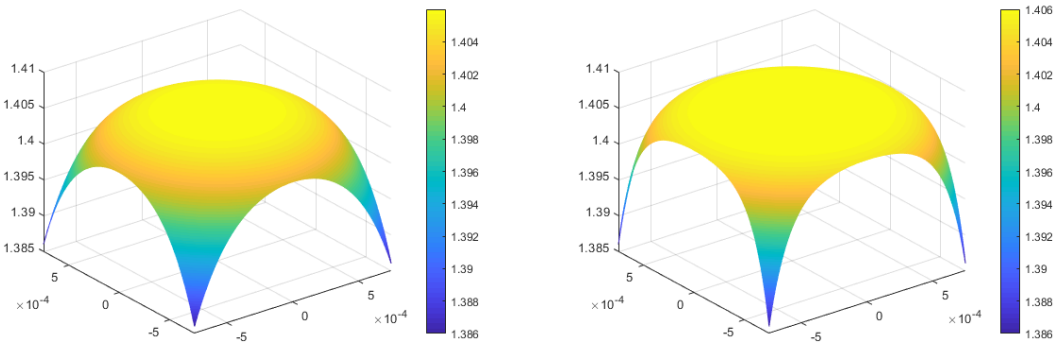
Another change in crystalline lens structure and eye model stems from the age of the eye. GRIN structure enables a lens to focus the light a point more easily since it additionally refracts light inside of it. Thus, GRIN structure of human lens helps for better focus. However, with age, the GRIN structure of eye also changes. The more an eye gets older, the more GRIN lens structure disappears. The effect of age on the GRIN lens structure is shown in Figure 7.2.

Two new approaches to the contrast enhancement for eye-based image processing are proposed. First one, double-sided contrast enhancement, can be thought as the



(a) (b)

Figure 7.1. PSF graphs for different focused distances. (a) is point spread function of eye focused on an object at 6 meters, (b) is for 25 cm object distance. For those, subject is thought to be 28 years old.



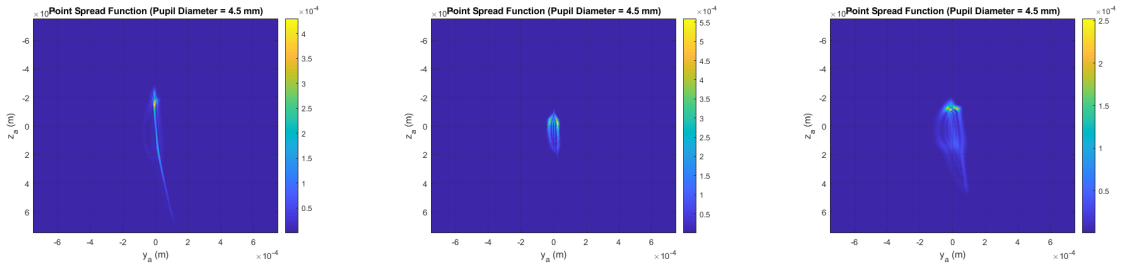
(a) (b)

Figure 7.2. GRIN change with age. (a) is GRIN lens structure calculated for a 7-year-old eye, (b) belongs to 70-year old eye. Note that the older eye has a "flatter" profile than a GRIN structure.

upgraded version of the single-sided mapping. Double-sided enhancement simply takes exponential profile of one-sided method, and shift and rotates it to the correct mapping region. While doing that, after a certain value it does not allow the exponential behavior to be dominant, and applies all-to-one mapping; that is, after a value in the image, all the values are map to one value. However, this approach has a limitation. After a certain threshold value, it exceeds the standard screen dynamic range and causes clipping in the mapping. This problem is surpassed by the second approach. Sigmoid function is a bounded function giving a result in the range [0 1] for any real value. This property of it enables the enhancement to be done directly on the deconvolved image without former application of shifting and scaling. By employing the sigmoid function, the values gathered around the middle of deconvolved image's dynamic range can be distributed such that they become placed further from each other without clipping or contrast loss thanks to the steep behavior of the function in its middle region.

To compare the methods that are not tried in literature about the eye-based contrast enhancement (except the gradient domain HDR compression method), a contrast evaluation method based on the looking at the differences of Gaussians in the image and its subsampled versions is applied. Since it is seen that the employed measure gives counter-intuitive results as "the smaller the contrast number the better the image", another method based on the standard deviation of the DoG image is tried. It is observed that the result of this method is easier to interpret and more accurate in a sense based on the user interpretation.

As the future work, there are a number of requirements and research areas for the project to be used in the daily life digital devices. The first area is the inclusion of cornea. In the literature, it is not encountered to include the cornea properties to the calculation of overall point spread function. This is interesting, because cornea is the first part to encounter with light and the strongest refractive component of the eye. One reason may be the notion that it is necessary to pass through cornea to reach the crystalline lens to be able to measure the aberrations on the lens, creating the thought that the measurements already measures the aberrations on the cornea. However, the iTrace device gives three different measurements: one for only the lens, one for only



(a) (b) (c)

Figure 7.3. Different point spread functions. (a) is the PSF both cornea and lens parameters included, (b) is cornea-only PSF, and (c) is the lens-only PSF.

the cornea topography, and one for the sum of them. All of them give different point spread functions shown in Figure 7.3.

(a) (b)

Figure 7.4. Cornea and lens result versus lens only result. (a) is the resulting image when PSF is calculated from only the lens parameters, (b) is from PSF both cornea and lens included. The chosen simulation is the ideal or requiring no processing situation.

Including the cornea to the calculation of PSF also creates deconvolved images whose convolution with the original point spread function does not yield a sharp image, even in the ideal solution case; whereas lens-only PSF gives more clear and sharper images. The situation is presented in Figure 7.4. Since taking cornea into account

makes it harder to obtain clear and sharp images, its inclusion should be handled more carefully.

This situation may be overcome by considering only the higher order aberrations for both cornea and the lens. Today, low order impairments such as myopia, hyperopia, astigmatism and presbyopia can be corrected via lenses. If they are excluded from the calculations, sharper and better images in terms of contrast can be obtained considering only the higher order aberrations.

In the borders of knowledge of this work's author, the accommodation of eye is not included in the eye modeling literature for signal processing purposes until this thesis. In this work, its inclusion is limited only with the change that it creates in the lens thickness and focal distance. However, aforementioned studies show that accommodation also changes the profile and refractive index of crystalline lens, although there is no consistency in the type and amount of change. This area is worth to further investigation.

In Chapter 6, the results for E letter image and the *seeds* image are presented. The two images have a vary basic and big difference: the E letter image is an unnatural image whereas the seeds image is a natural image. As you can see in Table 6.1 and Table 6.2, the quality of results for different image types are not the same. Generally, the unnatural images have a constant background behind them which also dominates the histogram and mean value of the image, and the necessary or relevant information takes place in foreground in such images. The enhancement methods applied to the unnatural images also enhances the background and the accompanying noise. As a promising research area, the enhancement methods concerning only the foreground or the area where the useful information takes place should be studied to prevent unnecessary enhancement of background.

Human eye is a dynamic organ. To focus on and see the objects at different places, people move their body and heads as well as their eyes. Movement of the eye changes the focus and creates a new image on the retina. In this case, the object

that was focused on before is seen different. There may be two reasons for that. The first reason is that by changing the visual axis that the light ray follows from target to retina, the focused object changes. In this case, the non-focused area creates a different response in rod and cone cells. As the second reason, since the visual axis is changed, the point spread function governing the image of surrounding of the focused object also changes. That point spread function can be calculated for not-focused objects and it can be used to reduce the computation burden of surrounding images because generally unfocused objects do not contain as much detail as the focused objects. For this purpose, an eye tracking system can be used, which is also helpful to determine the changes in human pupil diameter that affects the point spread function.

Although humans have two eyes, they use their "dominant eye" when looking at the distant objects or places. Dominant eye can be determined as follows: A person creates a hole with his/her hands by stretching his/her arms forward, and looks through the hole with two eyes open. Suppose that there is an object when looking through this hole and person can see it with two eyes open. Then (s)he closes one eye and looks the object through the hole with the open eye, and then changes the eye. With which eye the person sees the object through the hole is his/her dominant eye. That means, this person uses that eye when looking at the distant objects and for closer distances the other eye dominates the vision. This dominant eye concept can, and should, be investigated because it is nearly impossible the two eyes to have the same point spread function. The idea that using the dominant eye PSF for distant objects and the other eye's PSF for closer targets should be examined.

The biggest area that the research should be made on is the monochromatic and trichromatic approaches to the images. Until now, all the point spread calculations are done with only one wavelength value. However, in nature, there is no light containing only one specific wavelength. Hence, a group of wavelength value should be used during the calculations of point spread function. In this study, this approach is also applied. The notion is that since eye's response is known for each wavelength, the necessary ones can be calculated and added with some weight according to the eye's response. For this approach, weights are calculated from a fitted equation to the curves in Figure

Table 7.1. The sum of Gaussians fit parameters for each cell type.

	Rods	**Red Cones**	**Green Cones**	**Blue Cones**
$\mathbf{a_1}$	7.013	0	-5.362	17.48
$\boldsymbol{\mu_1}$	446.2	559.4	495.9	427.3
$\boldsymbol{\sigma_1}$	27.85	0.08667	160	30.52
$\mathbf{a_2}$	-1.434	-380.5	84.18	-0.9922
$\boldsymbol{\mu_2}$	629.3	533.9	539.1	405
$\boldsymbol{\sigma_2}$	11.85	63.97	60.26	11.49
$\mathbf{a_3}$	91.33	468	-9.146	56.51
$\boldsymbol{\mu_3}$	500.5	538	578.6	406.7
$\boldsymbol{\sigma_3}$	52.54	68.03	30.88	69.38
$\mathbf{a_4}$	4.153×10^{14}	35.26	39.01	-3.292
$\boldsymbol{\mu_4}$	-4384	393.5	370.8	380.7
$\boldsymbol{\sigma_4}$	869.9	48.14	172.5	18.12

3.2 obtained from study [24]. The fits are found by addition of four Gaussian. The parameters of Eq.(7.1) are given for each cell type in Table 7.1.

$$\boxed{w = a_1 e^{-\left(\frac{\lambda-\mu_1}{\sigma_1}\right)^2} + a_2 e^{-\left(\frac{\lambda-\mu_2}{\sigma_2}\right)^2} + a_3 e^{-\left(\frac{\lambda-\mu_3}{\sigma_3}\right)^2} + a_4 e^{-\left(\frac{\lambda-\mu_4}{\sigma_4}\right)^2}} \tag{7.1}$$

The problem with this approach is the small wavelength values. As it is shown in Figure 5.11, when the wavelength gets smaller, the coefficient of defocus term becomes bigger, hence, the point spread function spreads more. In this case neither image can be convolved properly nor the PSF behaves correctly since it exceeds the the dimensions and is wrapped on itself. This problem can easily be solved by making the sampling plane bigger. But this time either the sampling rate should be bigger, reducing the accuracy in calculation and representation of the point spread function, or it should be kept as is, making the images and PSF bigger and increasing the computation time exponentially.

As the last future work, the results of processed deconvolution images should be shown to and assessed by people suffering from visual impairment, instead of simulating the results on the computers. Since this project aims to ease the people's lives, it cannot be made more useful and perfect without human subject feedback.

REFERENCES

1. Ilardi, V., *Renaissance vision from spectacles to telescopes*, Vol. 259, American Philosophical Society, 2007.

2. Peli, T. and J. S. Lim, "Adaptive filtering for image enhancement", *Optical Engineering*, Vol. 21, No. 1, p. 211108, 1982.

3. Alonso, M., *On-screen pre-deblurring of digital images using the wavefront aberration function of the human eye to improve computer access for the visually impaired*, Master's Thesis, Florida International University, 2003.

4. Fitzpatrick, R., *Electromagnetism and optics: an introductory course*, 1999, `https://paginas.fe.up.pt/~fsal/EmagOptics.pdf`, accessed at December 2018.

5. Raymond A. Serway, J. W. J., *Physics for scientists and engineers with modern physics*, Cengage Learning, 9 edn., 2013.

6. Rashed, R., "A pioneer in anaclastics: Ibn Sahl on burning mirrors and lenses", *Isis*, Vol. 81, No. 3, pp. 464–491, 1990.

7. Hecht, E., *Optics*, Pearson, 5 edn., 2016.

8. Polyanskiy, M. N., *Refractive Index Database*, `https://refractiveindex.info/?shelf=main&book=LaF3&page=Laihoa-o`, accessed at December 2018.

9. Newell, D. B., F. Cabiati, J. Fischer, K. Fujii, S. G. Karshenboim, H. Margolis, E. de Mirandes, P. Mohr, F. Nez, K. Pachucki *et al.*, "The CODATA 2017 values of h, e, k, and Na for the revision of the SI", *Metrologia*, Vol. 55, No. 1, p. L13, 2018.

10. Feynman, R. P., R. B. Leighton and M. Sands, *The Feynman lectures on physics [vol 2]*, Vol. 2, A-W, 1964.

11. Polyanskiy, M. N., *Refractive Index Database*, `https://refractiveindex.info/?shelf=main&book=Si&page=Aspnes`, accessed at December 2018.

12. Morgan, J., "Introduction to geometrical and physical optics", *American Journal of Physics*, Vol. 21, No. 9, pp. 696–696, 1953.

13. Sommerfeld, A., "Optics: lectures on theoretical physics academic", *New York*, Vol. 4, 1953.

14. Goodman, J. W., *Introduction to Fourier Optics*, Roberts and Co., 3 edn., 2005.

15. Provencio, I., I. R. Rodriguez, G. Jiang, W. P. Hayes, E. F. Moreira and M. D. Rollag, "A novel human opsin in the inner retina", *Journal of Neuroscience*, Vol. 20, No. 2, pp. 600–605, 2000.

16. NIGMS, *Circadian Rhythms*, `https://www.nigms.nih.gov/education/Documents/CircadianRhythms.pdf`, accessed at October 2018.

17. Land, M. F. and R. D. Fernald, "The evolution of eyes", *Annual Review of Neuroscience*, Vol. 15, No. 1, pp. 1–29, 1992.

18. Blackwell, H. R., "Contrast thresholds of the human eye", *JOSA*, Vol. 36, No. 11, pp. 624–643, 1946.

19. Russ, J. C., *The image processing handbook*, CRC Press, 5 edn., 2006.

20. Tinsley, J. N., M. I. Molodtsov, R. Prevedel, D. Wartmann, J. Espigulé-Pons, M. Lauwers and A. Vaziri, "Direct detection of a single photon by humans", *Nature Communications*, Vol. 7, 7 2016.

21. Netter, F. H., *Atlas of human anatomy*, Netter Basic Science, Elsevier, 7 edn.,

2018.

22. Hubel, D. H., *Eye, brain, and vision*, Vol. 22 of *Scientific American Library Series*, New York, NY, US: Scientific American Library/Scientific American Books, 1995.

23. Nave, C. R., *Design of Scale-Model Eye*, `http://hyperphysics.phy-astr.gsu.edu/hbase/vision/eyescal2.html#c1`, accessed at October 2018.

24. Bowmaker, J. K. and H. Dartnall, "Visual pigments of rods and cones in a human retina", *The Journal Of Physiology*, Vol. 298, No. 1, pp. 501–511, 1980.

25. Peli, E. and T. Peli, "Image enhancement for the visually impaired", *Optical Engineering*, Vol. 23, No. 1, p. 230147, 1984.

26. Fine, E. M. and E. Peli, "Enhancement of text for the visually impaired", *JOSA A*, Vol. 12, No. 7, pp. 1439–1447, 1995.

27. Peli, E., "Perceived quality of video enhanced for the visually impaired", *Vision Science and Its Applications*, p. FC3, Optical Society of America, 1999.

28. Peli, E., J. Kim, Y. Yitzhaky, R. B. Goldstein and R. L. Woods, "Wideband enhancement of television images for people with visual impairments", *JOSA A*, Vol. 21, No. 6, pp. 937–950, 2004.

29. Peli, E., "Feature detection algorithm based on a visual system model", *Proceedings of the IEEE*, Vol. 90, No. 1, pp. 78–93, 2002.

30. Peli, E., L. E. Arend, G. M. Young and R. B. Goldstein, "Contrast sensitivity to patch stimuli: effects of spatial bandwidth and temporal presentation", *Spatial Vision*, Vol. 7, No. 1, pp. 1–14, 1993.

31. Peli, E., L. Arend Jr and G. Timberlake, "Computerized image enhancement for visually impaired persons: new technology, new possibilities", *J Vis Impairment Blind*, Vol. 80, pp. 849–854, 1986.

32. Frei, W., "Image enhancement by histogram hyperbolization", *Computer Graphics and Image Processing*, Vol. 6, No. 3, pp. 286–294, 1977.

33. Kramer, H. P. and J. B. Bruckner, "Iterations of a non-linear transformation for enhancement of digital images", *Pattern recognition*, Vol. 7, No. 1-2, pp. 53–58, 1975.

34. Peli, E., "Limitations of image enhancement for the visually impaired", *Optom Vis Sci*, Vol. 69, No. 1, pp. 15–24, 1992.

35. Wallace, G. K., "The JPEG still picture compression standard", *Communications of the ACM*, Vol. 34, pp. 30–44, 1991.

36. Wikipedia, *JPEG DCT-8x8*, `https://commons.wikimedia.org/wiki/File:DCT-8x8.png`, accessed at September 2018.

37. Wikipedia, *JPEG ZigZag*, `https://en.wikipedia.org/wiki/File:JPEG_ZigZag.svg`, accessed at September 2018.

38. Tang, J. and E. Peli, "33.2: Image enhancement in JPEG domain for low-vision patients", *SID Symposium Digest of Technical Papers*, Vol. 33, pp. 990–993, Wiley Online Library, 2002.

39. Tang, J., J. Kim and E. Peli, "Image enhancement in the JPEG domain for people with vision impairment", *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 11, pp. 2013–2023, 2004.

40. Kim, J., A. Vora and E. Peli, "MPEG-based image enhancement for the visually impaired", *Optical Engineering*, Vol. 43, No. 6, pp. 1318–1329, 2004.

41. Fullerton, M. and E. Peli, "Post-transmission digital video enhancement for people with visual impairments", *Journal of the Society for Information Display*, Vol. 14, No. 1, pp. 15–24, 2006.

42. Atabany, W. and P. Degenaar, "A robust edge enhancement approach for low vision patients using scene simplification", *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International*, pp. 1–4, IEEE, 2008.

43. Perona, P. and J. Malik, "Scale-space and edge detection using anisotropic diffusion", *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 12, No. 7, pp. 629–639, 1990.

44. Canny, J., "A computational approach to edge detection", *IEEE Transactions on pattern analysis and machine intelligence*, Vol. PAMI-8, No. 6, pp. 679–698, 1986.

45. Wolffsohn, J. S., D. Mukhopadhyay and M. Rubinstein, "Image enhancement of real-time television to benefit the visually impaired", *American Journal of Ophthalmology*, Vol. 144, No. 3, pp. 436–440, 2007.

46. Al-Atabany, W. I., M. A. Memon, S. M. Downes and P. A. Degenaar, "Designing and testing scene enhancement algorithms for patients with retina degenerative disorders", *Biomedical Engineering Online*, Vol. 9, No. 1, p. 27, 2010.

47. Fleck, M. M., "Some defects in finite-difference edge finders", *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 14, No. 3, pp. 337–345, 1992.

48. Fairman, H. S., M. H. Brill and H. Hemmendinger, "How the CIE 1931 color matching functions were derived from wright–guild data", *Color Research & Application*, Vol. 22, pp. 11–23, 1997.

49. Schanda, J., *Colorimetry: understanding the CIE system*, Wiley-Interscience, 1 edn., 2007.

50. Liang, J., B. Grimm, S. Goelz and J. F. Bille, "Objective measurement of wave aberrations of the human eye with the use of a hartmann–shack wave-front sensor", *JOSA A*, Vol. 11, No. 7, pp. 1949–1957, 1994.

51. Hartmann, J., "Bermerkungen über den bau und die justierung von spektrographen", *Zeitschrift für Instrumentenkunde*, Vol. 20, p. 47, 1900.

52. Shack, R. V., "Production and use of a lecticular hartmann screen", *J. Opt. Soc. Am.*, Vol. 61, pp. 656–661, 1971.

53. Alonso, M., Jr, A. Barreto, J. G. Cremades, J. A. Jacko and M. Adjouadi, "Image pre-compensation to facilitate computer access for users with refractive errors", *Behaviour & Information Technology*, Vol. 24, No. 3, pp. 161–173, 2005.

54. Mohammadpour, S., A. Mehridehnavi, H. Rabbani and V. Lakshminarayanan, "A pre-compensation algorithm for different optical aberrations using an enhanced wiener filter and edge tapering", *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pp. 935–939, IEEE, 2012.

55. Alonso, M., A. Barreto, J. A. Jacko, M. Adjouadi and M. Choudhury, "Improving computer interaction for users with visual acuity deficiencies through inverse point spread function processing", *Proceedings. IEEE SoutheastCon, 2005.*, pp. 421–427, IEEE, 2005.

56. NobelPrize.Org, *The Nobel Prize in Physics 1953*, `https://www.nobelprize.org/prizes/physics/1953/summary/`, accessed at September 2018.

57. Zernike, F., "Diffraction theory of the knife-edge test and its improved form, the phase-contrast method", *Monthly Notices of the Royal Astronomical Society*, Vol. 94, pp. 377–384, 1934.

58. Thibos, L. N., R. A. Applegate, J. T. Schwiegerling and R. Webb, "Standards for reporting the optical aberrations of eyes", *Journal of Refractive Surgery*, Vol. 18, No. 5, pp. S652–S660, 2002.

59. Wyant, J. C. and K. Creath, "Basic wavefront aberration theory for optical metrol-

ogy", *Applied optics and optical engineering*, Vol. 11, No. s 29, p. 2, 1992.

60. Lakshminarayanan, V. and A. Fleck, "Zernike polynomials: a guide", *Journal of Modern Optics*, Vol. 58, No. 7, pp. 545–561, 2011.

61. Malacara, D., *Optical shop testing*, Wiley Series in Pure and Applied Optics, Wiley-Interscience, 2 edn., 1992.

62. Kintner, E., "A recurrence relation for calculating the zernike polynomials", *Optica Acta*, Vol. 23, pp. 499–500, 1976.

63. Chong, C.-W., P. Raveendran and R. Mukundan, "A comparative analysis of algorithms for fast computation of zernike moments", *Pattern Recognition*, Vol. 36, No. 3, pp. 731–742, 2003.

64. Prata, A. and W. Rusch, "Algorithm for computation of zernike polynomials expansion coefficients", *Applied Optics*, Vol. 28, No. 4, pp. 749–754, 1989.

65. Belkasim, S., M. Ahmadi and M. Shridhar, "Efficient algorithm for fast computation of zernike moments", *Circuits and Systems, 1996., IEEE 39th Midwest symposium on*, Vol. 3, pp. 1401–1404, IEEE, 1996.

66. Chong, C.-W., P. Raveendran and R. Mukundan, "An efficient algorithm for fast computation of pseudo-zernike moments", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 17, No. 06, pp. 1011–1023, 2003.

67. Papakostas, G. A., Y. S. Boutalis, C. Papaodysseus and D. K. Fragoulis, "Numerical stability of fast computation algorithms of zernike moments", *Applied Mathematics and Computation*, Vol. 195, No. 1, pp. 326–345, 2008.

68. Dai, G.-m., "Modal compensation of atmospheric turbulence with the use of zernike polynomials and karhunen–loève functions", *JOSA A*, Vol. 12, No. 10, pp. 2182–2193, 1995.

69. MacRae, S. and M. Fujieda, "Slit skiascopic-guided ablation using the nidek laser", *Journal of Refractive Surgery*, Vol. 16, No. 5, pp. S576–S580, 2000.

70. Mrochen, M., M. Kaemmerer, P. Mierdel, H. E. Krinke and T. Seiler, "Principles of tscherning aberrometry", *Journal of Refractive Surgery*, Vol. 16, No. 5, pp. S570–S571, 2000.

71. Molebny, V. V., I. G. Pallikaris, L. P. Naoumidis, I. H. Chyzh, S. V. Molebny and V. M. Sokurenko, "Retina ray-tracing technique for eye-refraction mapping", *Ophthalmic Technologies VII*, Vol. 2971, pp. 175–184, International Society for Optics and Photonics, 1997.

72. Molebny, V. V., S. I. Panagopoulou, S. V. Molebny, Y. S. Wakil and I. G. Pallikaris, "Principles of ray tracing aberrometry", *Journal of Refractive Surgery*, Vol. 16, No. 5, pp. S572–S575, 2000.

73. Rozema, J. J., D. E. Van Dyck and M.-J. Tassignon, "Clinical comparison of 6 aberrometers. Part 1: technical specifications", *Journal of Cataract & Refractive Surgery*, Vol. 31, No. 6, pp. 1114–1127, 2005.

74. Gomez, A., A. Del Rey, C. Bautista, A. Ferrandiz, D. Gonzales and S. Burgos, "Principles and clinical applications of ray-tracing aberrometry (part 1)", *J Emmetropia*, Vol. 3, pp. 96–110, 2012.

75. Smith, G., B. K. Pierscionek and D. A. Atchison, "The optical modelling of the human lens", *Ophthalmic and Physiological Optics*, Vol. 11, No. 4, pp. 359–369, 1991.

76. Kasthurirangan, S., E. L. Markwell, D. A. Atchison and J. M. Pope, "In vivo study of changes in refractive index distribution in the human crystalline lens with age and accommodation", *Investigative Ophthalmology & Visual Science*, Vol. 49, No. 6, pp. 2531–2540, 2008.

77. Goldberg, K. A. and K. Geary, "Wave-front measurement errors from restricted concentric subdomains", *JOSA A*, Vol. 18, No. 9, pp. 2146–2152, 2001.

78. Schwiegerling, J., "Scaling zernike expansion coefficients to different pupil sizes", *JOSA A*, Vol. 19, No. 10, pp. 1937–1945, 2002.

79. Dai, G.-m., "Scaling zernike expansion coefficients to smaller pupil sizes: a simpler formula", *JOSA A*, Vol. 23, No. 3, pp. 539–543, 2006.

80. Campbell, C. E., "Matrix method to find a new set of zernike coefficients from an original set when the aperture radius is changed", *JOSA A*, Vol. 20, No. 2, pp. 209–217, 2003.

81. Marcos, S., S. A. Burns, E. Moreno-Barriusop and R. Navarro, "A new approach to the study of ocular chromatic aberrations", *Vision Research*, Vol. 39, No. 26, pp. 4309–4323, 1999.

82. Watson, A. B., "Computing human optical point spread functions", *Journal of Vision*, Vol. 15, No. 2, pp. 26–26, 2015.

83. Thibos, L. N., M. Ye, X. Zhang and A. Bradley, "The chromatic eye: a new reduced-eye model of ocular chromatic aberration in humans", *Applied Optics*, Vol. 31, No. 19, pp. 3594–3600, 1992.

84. Duane, A., "Studies in monocular and binocular accommodation, with their clinical application", *Transactions of the American Ophthalmological Society*, Vol. 20, p. 132, 1922.

85. Khan, A., J. M. Pope, P. K. Verkicharla, M. Suheimat and D. A. Atchison, "Change in human lens dimensions, lens refractive index distribution and ciliary body ring diameter with accommodation", *Biomedical Optics Express*, Vol. 9, No. 3, pp. 1272–1282, 2018.

86. Ostrin, L., S. Kasthurirangan, D. Win-Hall and A. Glasser, "Simultaneous measurements of refraction and a-scan biometry during accommodation in humans", *Optometry and Vision Science*, Vol. 83, No. 9, pp. 657–665, 2006.

87. Kim, Y.-T., "Contrast enhancement using brightness preserving bi-histogram equalization", *IEEE transactions on Consumer Electronics*, Vol. 43, No. 1, pp. 1–8, 1997.

88. Moniruzzaman, M., M. Shafuzzaman and M. F. Hossain, "Brightness preserving bi-histogram equalization using edge pixels information", *2013 International Conference on Electrical Information and Communication Technology (EICT)*, pp. 1–5, IEEE, 2014.

89. Fattal, R., D. Lischinski and M. Werman, "Gradient domain high dynamic range compression", *ACM Transactions On Graphics (TOG)*, Vol. 21, pp. 249–256, ACM, 2002.

90. Fairchild, M. D. and G. M. Johnson, "Meet iCAM: a next-generation color appearance model", *Color and Imaging Conference*, Vol. 2002, pp. 33–38, Society for Imaging Science and Technology, 2002.

91. Johnson, G. M. and M. D. Fairchild, "Rendering HDR images", *Color and Imaging Conference*, Vol. 2003, pp. 36–41, Society for Imaging Science and Technology, 2003.

92. Kuang, J., G. M. Johnson and M. D. Fairchild, "iCAM06: A Refined image appearance model for hdr image rendering", *Journal of Visual Communication and Image Representation*, Vol. 18, No. 5, pp. 406–414, 2007.

93. Rizzi, A., T. Algeri, G. Medeghini and D. Marini, "A proposal for contrast measure in digital images", *Conference On Colour In Graphics, Imaging, And Vision*, Vol. 2004, pp. 187–192, Society for Imaging Science and Technology, 2004.

94. Rizzi, A., G. Simone and R. Cordone, "A modified algorithm for perceived contrast measure in digital images", *Conference on Colour in Graphics, Imaging, and Vision*, Vol. 2008, pp. 249–252, Society for Imaging Science and Technology, 2008.