BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering and Communication

MASTER'S THESIS

Brno, 2021   Bc. Sofia Vraňáková

BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING
ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

PROCESSING OF HIGH-RESOLUTION RETINAL IMAGES
ZPRACOVÁNÍ SNÍMKŮ SÍTNICE S VYSOKÝM ROZLIŠENÍM

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR   Bc. Sofia Vraňáková
AUTOR PRÁCE

SUPERVISOR   Ing. Eva Valterová
VEDOUCÍ PRÁCE

BRNO 2021

BRNO FACULTY OF ELECTRICAL
UNIVERSITY ENGINEERING
OF TECHNOLOGY AND COMMUNICATION

Master's Thesis

Master's study program **Biomedical Engineering and Bioinformatics**

Department of Biomedical Engineering

Student: Bc. Sofia Vraňáková   *ID:* 195202

Year of study: 2   *Academic year:* 2020/21

TITLE OF THESIS:

Processing of high-resolution retinal images

INSTRUCTION:

1) Get acquainted with an adaptive optics retina acquisition principle. 2) Elaborate literary research of current approaches for image filtration and registration and blurry image detection with a focus on the adaptive optics retinal images. 3)Design, describe, and implement chosen methods for adaptive optics image processing. 4) Create a testing dataset from provided images and test designed approaches. 5) Optimize the proposed approach to maximize resultant image quality according to reached results. 6) Test the approaches and compare them with the available software. Describe and discuss the results.

RECOMMENDED LITERATURE:

[1] LAZAREVA, Anfisa, Panos LIATSIS a Franziska G. RAUSCHER, 2016. Hessian-LoG filtering for enhancement and detection of photoreceptor cells in adaptive optics retinal images. Journal of the Optical Society of America A. Optical Society of America, 33(1), 84-94. ISSN 1084-7529. Dostupné z: doi:10.1364/JOSAA.33.000084

[2] KOLAR, Radim, Ralf. P. TORNOW, Jan ODSTRCILIK a Ivana LIBERDOVA, 2016. Registration of retinal sequences from new video-ophthalmoscopic camera. BioMedical Engineering OnLine. BioMed Central, 15(1), 1-17. ISSN 1475-925X. Dostupné z: doi:10.1186/s12938-016-0191-0

Date of project specification: 8.2.2021   *Deadline for submission:* 21.5.2021

Supervisor: Ing. Eva Valterová

prof. Ing. Ivo Provazník, Ph.D.
Chair of study program board

ABSTRACT

The master's thesis is focused on the processing of high-resolution retinal images. The aim of this work is to obtain images of higher quality from a sequence of low-quality frames. The frames are first pre-processed by using bilateral filtering and contrast enhancement. The shift between the frames in the imaging sequence is estimated using phase correlation, and these frames are then fused together using the averaging through the frames and the super-resolution technique, more specifically regularization based on bilateral total variance. The resulting median quality scores of obtained images are PIQUE 0.2600, NIQE 0.0701, and BRISQUE 0.3936 for the averaging technique and PIQUE 0.1063, NIQE 0.0507, and BRISQUE 0.1570 for super-resolution technique.

KEYWORDS

adaptive optics, retinal images, phase correlation, super-resolution techniques, bilateral filtering, blurry frame detection

ABSTRAKT

Diplomová práca je zameraná na spracovávanie obrazov sietnice s vysokým rozlíšením. Cieľom práce je zlepšiť výslednú kvalitu výsledných snímkov sietnice získaných zo sekvencie snímkov nižšej kvality. Jednotlivé snímky sú najskôr spracované pomocou bilaterálnej filtrácie a zlepšenia kontrastu. v ďalšom kroku sú odstránené rozmazané snímky a snímky zobrazujúce iné časti sietnice. Posun medzi jednotlivými snímkami v sekvencii sa odhaduje pomocou fázovej korelácie, a tieto obrazy sú potom fúzované do výsledného snímku s vysokým rozlíšením pomocou priemerovania a využitia superrozlišovacej techniky, presnejšie regularizácie pomocou bilaterálneho celkového rozptylu. Výsledné mediánové hodnoty skóre kvality získaných obrazov sú PIQUE 0.2600, NIQE 0.0701, a BRISQUE 0.3936 pre techniku priemerovania, a PIQUE 0.1063, NIQE 0.0507, and BRISQUE 0.1570 pre superrozlišovaciu techniku.

KĽÚČOVÉ SLOVÁ

adaptívna optika, snímky sietnice, fázová korelácia, superrozlišovacie techniky, bilaterálna filtrácia, detekcia rozmazaných snímkov

ROZŠÍRENÝ ABSTRAKT

Zobrazovanie sietnice a vývoj techník na jej analýzu je dôležité z dôvodu, že čo i len malé zmeny na jej povrchu môžu odhaliť poruchy cirkulácie, nervové ochorenia a v neposlednom rade aj očné ochorenia.

Prvý oftalmoskop bol vynájdený českým vedcom Jánom Evangelistom Purkyně, a neskôr bol zlepšený Charlesom Babbagom a Hermannom Helmholtzom. Tento vynález pomohol očným lekárom pri rutinných vyšetreniach. Medzi ďalšie vynálezy patrí fundus kamera, ale všetky zariadenia poskytovali len dvojdimenzionálne obrazy sietnice. Prvým pokusom o zobrazenie sietice ako trojrozmerného obrazu bol vynález stereo fundus kamery. V posledných desaťročiach sa na trojdimenzionálne zobrazovanie sietnice používajú pokročilejšie metódy, napríklad očná koherenčná tomografia, a techniky využívajúce adaptívnu optiku, napríklad skenovacia laserová oftalmoskopia.

Hlavnou myšlienkou adaptívnej optiky je kompenzácia optických aberácií v zobrazovacom systéme oka (rohovka, zrenica a šošovka). Systém na takéto zobrazovanie sa väčšinou skladá z troch častí - zo senzoru vlnoplochy, korektoru vlnoplochy a kontrolného systému. Rôzne aberácie optického systému sa nachádzajú aj v zdravom oku. Niektoré z nich sa dajú odstrániť pomocou očných šošoviek, medzi tieto aberácie patria astigmatizmus a rozostrenie, ale ďalšie aberácie (zväčša nepravidelné) sa dajú korigovať až systémami adaptívnej optiky.

V rámci tejto diplomovej práce je navrhnutý algoritmus na získanie jedného snímku vysokej kvality zo sekvencie snímkov nižšej kvality. Tieto snímky sú najskôr predspracované pomocou bilaterálnej filtrácie a úprave kontrastu, následne sú detegované rozmazané snímky a snímky zachytávajúce rozličné časti sietnice, a tieto sú odstránené. V poslednom kroku sú zvyšné snímky lícované pomocou fázovej korelácie, a výsledkom sú dva obrázky vysokej kvality - jeden je získaný priemerovaním z nalícovanej sekvencie a druhý pomocou superrozlišovacej techniky (regularizácia založená na bilatrálnom totálnom rozptyle).

Popis riešenia

V tejto diplomovej práci boli použité dáta nasnímané kamerou Imagine Eyes v spolupráci s Univerzitou v Lipsku. Sekvencie boli nasnímané od desiatich zdravých subjektov. Pri snímaní subjekty sledovali pohybujúci sa cieľ, preto nasnímané obrázky zachytávajú rôzne časti sietnice. Sada dát obsahuje 119 sekvencií, pričom každá sekvencia obsahuje 40 snímkov s rozlíšením 1279x1279 pixelov. Výsledné snímky získané pomocou softvéru Imagine Eyes majú rozlíšenie 1500x1500 pixelov. Na snímkach je možné pozorovať svetlé body predstavujúce jednotlivé fotoreceptory a tmavé časti s menším počtom fotoreceptorov predstavujú cievne kapiláry.

V prvom kroku je potrebné zaistiť, aby rozsah hodnôt intenzít v snímkach využíval celú stupnicu šedej v rozsahu 0 až 1. Keďže "surové" snímky majú rozsah intenzít od 0 do približne 0,05, je použitá lineárna transformácia pre využitie celej škály šedej.

V ďalšom kroku je odstránené nerovnomerné osvetlenie scény. Snímky sú pri snímaní scény najviac osvetlené v strede a smerom k okrajom množstvo svetla klesá, čo je nežiaduce. Kópia každej snímky je preto silne rozmazaná Gaussovým filtrom s veľkosťou smerodajnej odchýlky 20, a táto kópia je následne odčítaná od pôvodnej snímky. Ďalej je opätovne použitá transformácia histogramu pre využitie celej škály šedej.

Posledným krokom predspracovania snímkov je bilaterálna filtrácia a úprava kontrastu. Bilaterálna filtrácia potlačí šumovú zložku a zachová hrany, a úpravou kontrastu je dosiahnuté zvýraznenie fotoreceptorov a ciev.

Ďalším dôležitým krokom je odstránenie rozmazaných snímkov, ktoré vznikli v dôsledku rýchleho pohybu očí počas snímania sekvencie. Pre detekciu je najskôr na každú snímku aplikovaný Cannyho detektor hrán, a následne binárna maska , ktorá je zložená z núl v strede, len na okrajoch (20 pixelov od každého okraja) sú jednotky, a logická operácia *and*. Ako posledný krok detekcie je vypočítané percentuálne zastúpenie bielych pixelov na okraji obrazu. Predpoklad je taký, že ostré snímky majú hlavný podiel hrán v strednej časti snímky, keďže okraje sú znehodnotené v dôsledku snímania s nerovnomerným osvetlením, a rozmazané snímky obsahujú nadetegované hrany náhodne po celej snímke. Preto ak je percentuálne zastúpenie bielych pixelov reprezentujúcich hrany menej ako 10 %, je snímka považovaná za ostrú a môže sa využívať v ďalších krokoch.

Ďalším problémom sú snímky v rámci jednej sekvencie, ktoré zachytávajú rozdielne časti sietnice. Tieto sú detegované pomocou normalizovanej vzájomnej korelácie. Ak je maximálny koeficient vzájomnej korelácie medzi dvomi nasledujúcimi snímkami menší ako 0,3, sekvencia je rozdelená do dvoch skupín, a tá ktorá obsahuje viac snímkov vstupuje do ďalšej analýzy. V dostupných sekvenciách nastáva len prípad s dvomi skupinami, pri iných dátach by sa musel počítať koeficient vzájomnej korelácie ešte raz v rámci každej skupiny.

Snímky, ktoré prešli predošlou analýzou, sú vo finálnom kroku lícované pomocou fázovej korelácie. Fázová korelácia bola zvolená z dôvodu, že hlavné pohyby medzi snímkami sú spôsobené transláciou, rotáciou a škálovaním, a fázová korelácia dokáže odhadnúť všetky tieto transformácie v potrebnom rozsahu. Pred vstupom do fázy lícovania sú z každého snímku odrezané rozmazané okraje, keďže pri lícovaní by mohlo dôjsť k prekryvu ostrej časti s okrajom a následnej strate informácie o stave

sietnice, a teda snímky v rámci sekvencie majú rozlíšenie 1000x1000 pixelov. Prvý výsledný obraz vysokej kvality je získaný po procese lícovania priemerovaním.

Druhý výsledný obraz je získaný pomocou superrozlišovacej techniky - regularizácie pomocou bilaterálneho celkového rozptylu. V tejto technike sa využívajú transformačné matice získané fázovou koreláciou, a je potrebné vhodne nastaviť paramatere regularizácie - zväčšovací faktor je nastavený na hodnotu 2, šírka rozptylovej funkcie je nastavená na 0,3, regularizačná váha λ je nastavená na hodnotu 0,65, faktor P na hodnotu 2, maximálny počet iterácií je nastavený na hodnotu 7, a maximálna prípustná chyba je nastavená na hodnotu 1.10^{-10}. Do tejto metódy vstupujú snímky bez akejkoľvek filtrácie a úpravy kontrastu, len s korigovaným osvetlením, keďže samotná filtrácia prebieha v rámci regularizácie. Výsledný obraz má rozlíšenie 2000x2000 pixelov. Výsledný obraz je následne kntrastne upravený tak, aby boli zvýraznené štruktúry sietnice.

Výsledky

Oba výsledné obrazy vysokej kvality sú porovnávané s dostupným Imagine Eyes obrazom pomocou metrík na odhad kvality obrazu bez potreby referenčného obrazu - PIQUE, NIQE a BRISQUE. Výsledné skóre týchto metrík pre obraz najvyššej kvality sa blíži k nule, pre nekvalitný obraz sa blíži k jednej. Mediánové skóre pre obraz vzniknutý priemerovaním sú PIQUE 0,2600, NIQE 0,0701 a BRISQUE 0,3936, pre obraz vzniknutý superrozlišovacou technikou sú PIQUE 0,1063, NIQE 0,0507 a BRISQUE 0,1570. V porovnaní s mediánovými skóre pre dostupné Imagine Eyes obrazy - PIQUE 0,5739, NIQE 0,0918 a BRISQUE 0,4538 - sú výsledné obrazy lepšej kvality a teda cieľ práce je splnený. Najkvalitnejšie obrazy vznikli pomocou superrozlišovacej techniky, aj keď v týchto obrazoch sú ťažšie rozlíšiteľné jednotlivé fotoreceptory. Naopak, ak sekvencia snímala časť sietnice, ktorá obsahovala cievy, tak táto technika ich dokázala výrazne zlepšiť aj zo subjektívneho pohľadu. Obrazy vzniknuté priemerovaním dosahujú horšej kvality, čo môže byť spôsobené tým, že po odstránení rozmazaných okrajov došlo k vzniku okrajových artefaktov (čierne pixely na okrajoch). Výhodou tohto obrazu je ale fakt, že fotoreceptory sú viditeľne ohraničené.

Nedostatky uvedených metód by sa dali odstrániť použitím ďalšej techniky lícovania v prípade výsledného obrazu vzniknutého priemerovaním, napríklad lokálnej techniky lícovania, a presnejším odhadom nastavovaných parametrov v prípade superrozlišovacej techniky, napríklad nastavením optimálnych parametrov pre každú sekvenciu zvlášť.

DECLARATION

I declare that I have written the Master's Thesis titled "Processing of high-resolution retinal images" independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the thesis and listed in the comprehensive bibliography at the end of the thesis.

As the author I furthermore declare that, with respect to the creation of this Master's Thesis, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll., Section 2, Head VI, Part 4.

Brno   . . . .
                                                                       author's signature

ACKNOWLEDGEMENT

Contents

List of Figures

Introduction

The retina and techniques for analysing the retina are of great interest due to the ability of the retina to convert incident light into a neural signal that is processed by the visual cortex of the brain. Furthermore, because the retina is metabolically active tissue with a double blood supply it allows for the non-invasive examination of blood flow. As a result the retina can manifest a variety of circulation and brain related diseases as well as eye related diseases. The retina allows the detection and diagnosis of such diseases through high resolution imaging techniques [1].

The optical system of the human eye that transforms light beams into the image, prevents direct examination of the retina. The French physician Jean Mary attempted the first experiment to image the retina by dipping a cat in water. While doing this experiment he discovered that the cat's retinal vessels are visible externally while under water[1]. The first ophthalmoscope was invented by Czech scientist Jan Evangelista Purkyne[2] and it was later reinvented by Charles Babbage [3], to be reinvented one more time by Hermann Helmholtz [4]. With this invention retinal examinations became routine for ophthalmologists. The ophthalmoscope was not very comfortable to use, so the fundus camera was developed [5]. Even though the camera is still used it only provides two-dimensional retinal images. There was a need for three-dimensional images. This lead to the stereo fundus camera which captured the retinal images from different angles and these images were combined by the eye of the observer into a three-dimensional image [6]. Later more advanced techniques to obtain three-dimensional images were introduced, optical coherence tomography and adaptive optics scanning laser ophthalmoscopy these allow the observation of the layers of retinal tissue [1].

The work presented here is centred on the pre-processing and registration of high-resolution retinal images. The chapter 1 summarises the anatomical and physiological description of the human eye. The chapter 2 defines the technique of adaptive optics and how imaging using this technique works, including an overview of its use in retinal imaging. The chapter 3introduces the current approaches to filter and register retinal images. The final chapter 4 discusses how the designed algorithm is implemented and tested.

1 Human eye

Human vision detects a very narrow light spectrum, approximately from 400 nanometers to 750 nanometers. The wavelength of 400 nanometers is perceived as the colour blue and the wavelength of 750 nanometers as the colour red [7]. The full spectrum of the visible light is displayed in figure 1.1. The light travels through the optical system to the retina and photosensitive receptors, which detect and transfer the information about visible light to the brain.
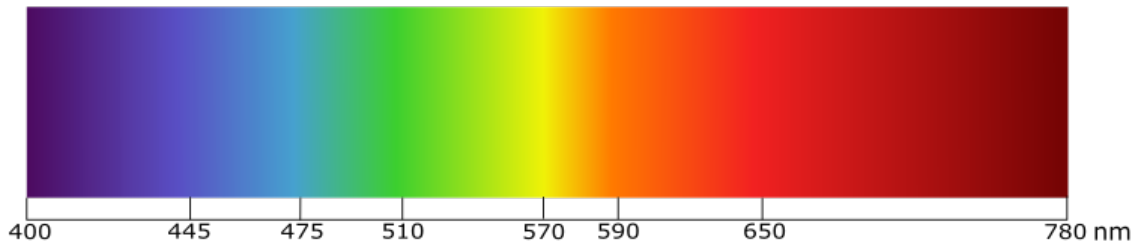
Fig. 1.1: Visible spectrum of the light

1.1 Anatomy of the eye

The eye consists of the eye; extraocular muscles, eyelids, conjunctiva, and lacrimal glands. The eyeball is made of three layers, or coats: the outer layer called *fibrous tunica*, the middle layer named *vascular tunic*, and inner layer called the *retina* [8].

The fibrous tunica consists of the *cornea* and *sclera*. 5/6 of the outer layer is formed by the sclera. It is firm, opaque and white-coloured. The cornea forms the residual 1/6 of the outer layer. It is transparent and in contrast to the sclera innervated. The cornea is also part of the optical system of the eye [8].

The vascular tunic is composed of the *choroid*, *ciliary body*, and *iris*. The choroid is made of a large amount of blood vessels and pigment cells. It is located in the posterior 2/3 of the eyeball. The ciliary body is placed in the continuation of the choroid anteriorly. It includes ciliary muscles which control the shape of *lens* and changes refraction of light ray - accommodation of the eye. The lens is an elastic biconvex structure located behind the iris. The iris is located at the very front of the middle layer. In the middle of the iris is the *pupil*. Light enters the eye through the pupil. The color of the iris depends on the amount of pigmentation [8].

The retina consists of two sections. The first section is in the anterior part and is made entirely of pigment cells. The second section is in the posterior, optical part and is made of bipolar and ganglion cells, and the most importantly *rods* and *cones*.

The place where the optical nerve exits the retina, is called the *optic disc* or *blind spot*. *Fovea* consists only of cones and is the place of the sharpest vision [8].
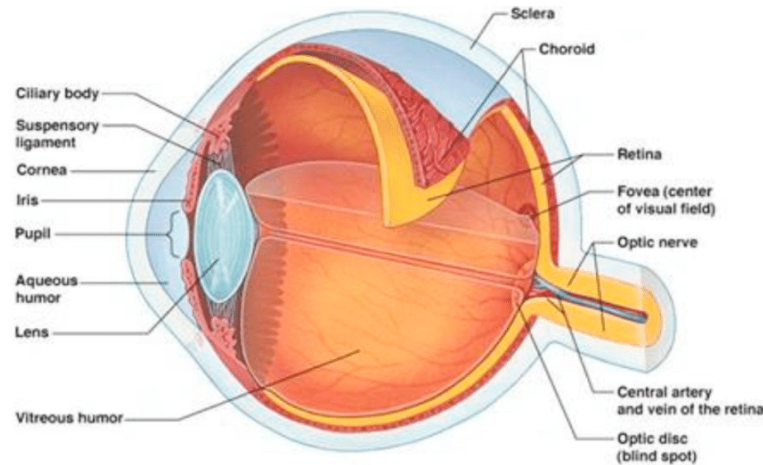
Fig. 1.2: The structure of the eye [9]

Aqueous humor is a clear fluid contained in the eye chambers. The anterior chamber is located between the cornea and the iris, and the posterior chamber is placed between the iris and the lens [8].

Vitreous humor is a gelatinous mass filling the part of the eye behind the lens [8]. The structure of the eye is displayed in figure 1.2.

1.2   Vision

As said in the previous section, the retinal receptors are called rods and cones. The rodes are most sensitive to the wavelength of about 500 nm and are responsible for night (scotopic) vision since they are sensitive to changes of light intensity. There are three types of cones, and each type is sensitive to the different wavelength range - the first type is shortwave (for wavelengths from 415 to 440 nm, which corresponds to blue color), the second type is mediumwave (for wavelengths from 520 to 540 nm, which corresponds to green color) and the last type is longwave (for wavelengths from 550 to 570 nm, which corresponds to reddish color). The cones are responsible for color (photopic) vision since they are sensitive to changes of wavelength [10].

Color vision occurs by mixing of three basic colors - red, green and blue, and thus the other colors and their shades are created. Normal vision is trichromatic. When the sensitivity for a certain color is reduced, it is referred to as anomaly. The loss of sensitivity for a certain color is referred to as anopia. For red color it is called as protanomaly or protanopia, for green color it is called deuteranomaly or

deuteranopia and for blue color it is called tritanomaly or tritanopia. When no type of cones is functional, the vision is greyscaled and it is called monochromacy [10]. In figure 1.3 is displayed, how different types of color blind people see the color spectrum.

Normal vision Protanopia Deuteranopia Tritanopia Monochromacy
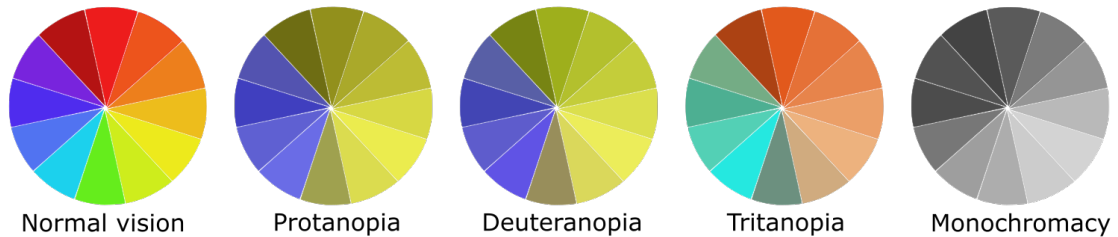
Fig. 1.3: Different types of color blindness

1.3 Eye movements

Eye movements can be generally classified into three main categories - *fixation*, *saccades*, and *smooth pursuit* [11].

Fixation is a movement that consists of stays and jumps from current stay to another. These stays last for about 100 ms. In measured eye-movement data, the positions of gaze are concentrated and the velocity is slow [11].

Saccades are fast eye-movements between two fixation points. A saccade is able to cover angle from 1° to about 40°, but typically between 15° and 20°, and the speed ranges between 30°/s and 500°/s. In measured data, the positions of gaze jump from the last fixation to next with high velocity [11].

Smooth pursuits appear only when eyes are tracking moving targets. This movement is continuous to ensue following targets. Smooth pursuits are usually slower than saccades, speed depends on the speed of targets, usually below 30°/s. In measured data, the positions of gaze change while following the target, and, as said before, the velocity is determined by the speed of moving target [11].

2 Retinal imaging

Imaging of the retina is very important in many medical fields. Besides ophthalmology, which is a logical outcome, retinal imaging is used in other fields as well. Such example can be neurology or cardiology, which may utilize imaging of retinal vascularity and nerve fibers. There are lots of different approaches how to image the retina. In this chapter, the main focus is on the use of adaptive optics [12].

2.1 Adaptive optics

Before the first use of adaptive optics in retinal imaging over 20 years ago [13], it was used in astronomy, more precisely in astronomical telescopes, to improve the loss of resolution from irregularities in atmosphere such as wind and moisture [14].

The main idea is to compensate optical aberrations in the optical system. In ophthalmology, it is the three optical elements of the eye - the cornea, the pupil, and the lens that contribute to the aberrations. The imaging system aided by adaptive optics usually consists of three components - a wavefront sensor, a wavefront corrector, and a control system uniting the previous two components [14].

As displayed in figure 2.1, the distorted wavefront from the source or imaged object is gathered and directed to the wavefront sensor via a deformable mirror (the wavefront corrector). The wavefront sensor measures the distortions in the incoming wavefront, and this information is sent to the control system. The control system analyses this information and computes the commands for the wavefront correctors. After the corrector receives the commands, the surface shape of the deformable mirror is adjusted. This process is iterative, and it is iterating until a certain degree of correction is achieved. As a result, the imaging system obtains images with enhanced quality [15].
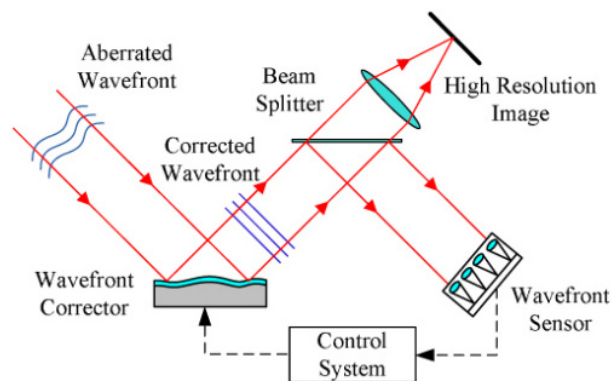
Fig. 2.1: Adaptive optics-aided system [15]

2.1.1 Wavefront sensing

For wavefront sensing, the Hartmann-Shack sensor (HS sensor) is mostly used in ophthalmology. This sensor is composed of an array of the microlenses in front of a detector. When the aberrated wavefront reaches the array of microlens, it produces a distribution of spots on the detector. Each microlens has an intended position on the detector, and the Hartmann-Shack detector evaluates how each detected spot varies from its expected position. The calculated deviation is then used to conclude the wavefront slope and amplitude at each location. This information is used to determine the type and magnitude of distortions across the wavefront [14].

In figure 2.2 is shown, how different type of the wavefront (planar and non-planar) is displayed on the detector. In the case of the planar wavefront, the spots format a regular grid of spots (blue), and in the case of the non-planar distorted wavefront, the spots are located irregularly (red) [16].
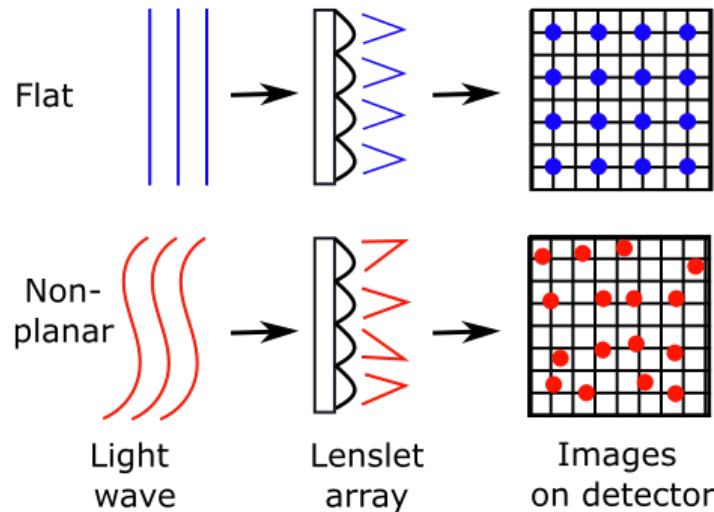
Fig. 2.2: Function of Hartmann-Shack sensor

2.1.2 Wavefront correction

Generally, there are two techniques used for wavefront correction - piston-segmented devices and continuous surface mirrors. The latter are mainly used. The deformable mirror consists of a reflective faceplate deflected by a series of actuators, which are either segmented or continuous. Since the segmented deformable mirror lead to more disruptive effects induced by diffraction, so the continuous deformable mirrors are used in adaptive optics retinal imaging [14].

2.1.3 Zernike polynomials and coeffients

The Zernike polynomials are used for evaluating of the optical quality of an imaging system, because of their ability to describe the wave aberrations [17].

The standard Zernike polynomials are a product of radial polynomials $R_n^m(r)$ and angular functions $\Theta^m(\theta)$

$$Z_n^m(\rho, \theta) = N_n^m R_n^m \Theta^m(\theta), \tag{2.1}$$

where N_n^m is the constant of normalisation, n is the highest power of radial polynomial, and m is the azimutial frequency of the angular function [17].

The radial polynomials are defined in [17] as

$$R_n^m(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s!(\frac{n+m}{2} - s)!(\frac{n-m}{2} - s)!} \rho^{n-2s} \tag{2.2}$$

and angular function is defined in [17] as

$$\Theta^m(\theta) = \begin{cases} cos(m\theta), & m \geq 0, \\ sin(m\theta), & m < 0. \end{cases} \tag{2.3}$$

In the table 2.1, the first nine Zernike coefficients and their corresponding polynomials and optical aberrations are shown.

Tab. 2.1: Zernike coeffients [18]

n	m	No.	Polynomial	Aberration
0	0	0	1	piston
1	1	1	$\rho cos\theta$	x-tilt
1	1	2	$\rho sin\theta$	y-tilt
1	0	3	$2\rho - 1$	focus
2	2	4	$\rho^2 cos2\theta$	astigmatism at 0° and focus
2	2	5	$\rho^2 sin2\theta$	astigmatism at 45° and focus
2	1	6	$(3\rho^2 - 2)\rho cos\theta$	coma and x-tilt
2	1	7	$(3\rho^2 - 2)\rho sin\theta$	coma and y-tilt
2	0	8	$6\rho^4 - 6\rho^2 + 1$	spherical and focus

2.2 Use of adaptive optics in retinal imaging

The normal human eye suffers from many aberrations decreasing image quality attain ability. Despite the ophthalmic lenses are able to correct defocus and astigmatism, still can not correct all aberrations, for example, spherical aberrations, coma, and lots of irregular aberrations. Adaptive optics could possibly correct these aberrations [19].

2.2.1   Adaptive optics flood illumination

The original application of this technique was to understand cone density and arrangement of basic psychophysical responses. In the next years following this application, the use of adaptive optics advance from use in understanding normal vision to exploring pathologies of the retina [14].

Among the advantages of adaptive optics flood illumination is the use of a spinning diffuser to eliminate speckle from the source of the light. However, it has a low axial resolution, which leads to decreasing of longitudinal cone density measurements, and the efficiency was reduced by its long imaging time per frame. The latter was later fixed by the use of the superluminescent diode as a light source [14].

2.2.2   Adaptive optics scanning laser ophthalmoscopy

This technique improves the image contrast of the previous technique. The retina is scanned by a single-point light beam (laser beam), and the scattered light from each point is registered by the detector, and the image is created [14].

In this system there are used different modes of detection. The confocal method is used for imaging the rod receptors and elements of the optic nerve, since the resulting image is of higher contrast and allows retinal sectioning. The non-confocal method is used for visualization of retinal pigment epithelium (known as dark-field imaging because of no need for autofluorescence) and of inner segments of photoreceptors (split-detection imaging). The parallel use of both of these methods is applied in the imaging of the retinal structure, mainly in imaging of the retina with inherited retinal diseases. The main disadvantage of this technique is the tendency for image distortion and the inability of serial image creation [14].

2.2.3   Adaptive optics optical coherence tomography

The latest technological advance in retinal imaging was adding adaptive optics to optical coherence tomography. The native optical coherence tomography offers axial optical resolution at the subcellar level, but the axial and transverse resolution is limited by optical aberrations, so the fusion with adaptive optics system solves this limitation [14].

These systems are used in three-dimensional visualizations of photoreceptors, retinal nerve fibers, and retinal vessels. With the use of this technique complementary with the adaptive optics scanning laser ophthalmology, the resulting image has an ultrahigh axial resolution of tomography and reduced eye movement effects of the scanning laser ophthalmology [14].

2.3 Retinal pathologies detectable by adaptive optics

Retinal imaging via adaptive optics provides high resolution in vivo images, which are used to improve in understanding retinal structure and how is the retina affected by age, sex, or some diseases [20].

2.3.1 Diabetes

It is known that diabetes affects mainly the inner retina, but outer retina is involved in diabetic eye disease as well. It can either be a decrease in cone numbers or an increase in the irregularity of cones [20].

In diabetic eyes is mainly affected retinal circulation. This can be seen as areas where the cones are about 25 % less reflective. The photoreceptors have high metabolic demand and limited support circulation, and diabetes is also neurodegenerative. Both neural and vascular systems of the retina can be imaged by adaptive optics, so it is helpful in monitoring changes in these systems [20].

2.3.2 Myopia and aging

As one is aging, the number of cones in the central 2 degrees of the retina is decreasing. The lower number of cones in this area in older subjects is consistent with other suggestions - the fovea changes shape with aging and loses photopigment [20].

The adaptive optics imaging can show cones that are surviving in myopic eyes over areas that are atrophied [20].

2.3.3 Retinal degenerations

With new researched therapies for retinal degenerations, there is increased demand for imaging, that is able to measure the retinal damage, and adaptive optics imaging can do this [20].

As being said before, adaptive optics imaging is able to make precise measurements of numbers and locations of cones, as well as rodes measurements. On top of this, adaptive optics systems can test visual sensitivity [20].

For example, this type of retinal imaging can be used for discrimination of phenotypic color deficiencies. The cones are either formed, but their outer segment is abnormal and they appear dark in the image, or are missing completely [20].

2.3.4 Measuring of retinal vasculature and blood flow

The retinal blood vessels are used majorly for the detection of many diseases, such as hypertension and diabetes. The measurements used for this are for example vessel caliber, vascular perfusion, blood velocity, oxygen saturation, and autoregulation of vascular parameters [20].

Blood and blood vessels scatter light and filter wavelengths, which leads to decreased visual sensitivity. As a result, the size of the foveal avascular zone is closely related to the amount of inner retinal tissue present at the edge. When the tissue is thicker than normal, the blood vessels are needed to provide the support of the tissue, thus the retina is operating near a critical density of vascularity [20].

Another measurement is the change in vascular wall thickness. The wall thickness is estimated as the width of a bright stripe that goes down the center of retinal arteries. It is clarified by adaptive optics imaging, that the majority of the central bright region arises from the reflections from the moving erythrocytes which are flowing through the vessel [20].

With adaptive optics imaging, it is possible to reveal even mild changes in the retinal vasculature, which can be one of the first indicators of diabetes [20].

2.3.5 Nerve fiber layer

Adaptive optics imaging makes the visualization of the nerve fiber layer in great quality possible. It allows the identification of fibers and shows changes in reflectance with glaucomatous damage, even when the damage is mild [20].

In diseases that cause damage to the nerve fiber layer, for example, a cotton wool spot, adaptive optics imaging can reveal the development of small cell changes [20].

3 Image processing approaches

In retinal imaging there are many different algorithms for the filtering of retinal images, for the registration of the images, and for the detection of blurry images in the sequence.

3.1 Filtering

For removing the noise and blurring from images, the different types of filters are used, depending on the type of noise affecting the image. One of the most common type is an additive Gaussian noise that affects all the pixels in the image independently on their intensity values. The most basic way to reduce the noise in the images affected by this type of noise is the averaging convolutional filter. Another common type is an impulsive noise, also known as salt-and-pepper noise. It affects only isolated pixels or groups of pixels, and its values are very high in comparison with the intensity values of the image, so it turns out like bright pixels or dark pixels in the image. The median filter can be used to remove this type of noise from the image [21].

For the purpose of this thesis, the main focus is on median filtering, bilateral filtering and hybrid filtering (bilateral filter and coherence diffusion).

3.1.1 Median Filtering

The retinal images can be enhanced by the application of a median filter and then by a nonlinear diffusion filter [22]. The median filter is a mask operator, which implies that the neighborhood pixels are considered, as well as the reference pixel itself, to decide which of them is assigned to the output. This operator selects the pixel, which is in the middle position of the sorted sequence of the input pixels. This is in principle a nonlinear operation. The main benefit of the median filtering is that it preserves sharp edges, and removes little objects, which are smaller than the mask [23].

This method can effectively improve the signal-noise-ratio in retinal sequences. Subsequently, the contrast limited adaptive histogram equalization can be used to enhance the contrast of blood vessels [12].

3.1.2 Bilateral filtering

This technique smoothes images while it preserves edges, and therefore meets the requirements for succesful image denoising. It uses two low-pass kernel functions,

the first one is called spatial kernel and its purpose is to assign lower weights to pixels farther from current pixel, and the second one is a range kernel that assigns lower weights to pixels with larger differences in intensity values. The choice of the latter is important, as it allows the edge-preserving effect. The filter is based on a weighted averaging of nearby pixels, and is non-linear, which leads to high-computational cost [24].

3.1.3   Hybrid filtering

The images can be also enhanced and filtered by a hybrid method. First, bilateral filtering is used to reduce a large-scale noise, then, an improved coherence diffusion is used, and lastly, an edge enhancement algorithm is used to improve contrast. Then, it is determined that the noise distribution is of a Gaussian shape and that knowledge is used to choose the bilateral filtering method, rather than the traditional filtering methods, as it has better performance for Gaussian noise. The results of this method depend highly on the noise estimation. If the chosen root mean square of the noise is too high, the bilateral filtering is degraded to Gaussian filter, and edges are blurred, if it is too low, the bilateral filtering is not able to reduce the noise. The coherence diffusion is used to filter any residual noise and enhance coherence structures. In flat image regions, isotropic diffusion is used to supress noise, and near edges, anisotropic diffusion is applied to regularize and ehnance edges. The edge enhancement algorithm is used to improve the contrast of the weak edges. The image edges and brightness are extracted from the image and then are merged into the result of the previous methods (bilateral filtering and coherence diffusion) [25].

3.2   Image registration

Image registration is a method that provides a pair of images (or set of images), that are spatially consistent. This means that in each of these images a particular pixel corresponds to the unique position in the imaged object. One of the images is commonly considered undistorted, so it is taken as a reference image. Sometimes, the reference image is one image from the set, that is distorted the least [23]. To ensure that the corresponding pixels are in the same position throughout the set of images, the image transformation is used [21].

The first step in image registration is estimation of transformation matrix. This matrix contains informations about the movement between the images, such as translation, rotation, scaling and shearing. The transformations that take only translation

and rotation into account are called *rigid* transformations. The transformations that consider also scaling and shearing are named *affine* transformations [21].

Based on the principle of adaptive optics data acquisition the shearing distortion approaches zero, thus can be omitted. In the following subsection, the similarity criteria are introduced, as well as the estimation of the transformation matrix with the use of phase correlation and rigid global registration, and local registration.

3.2.1 Similarity criteria

For computing global similarity, there are many different criteria. The most common are intensity-based and information-based global criteria. The first criteria are based on comparing intensity values in reference and registering images. However, intensity values in registering images are slightly modified in every step as a consequence of interpolation. The examples of intensity-based criteria are correlation coefficient criterion (equation 3.1) and angle (cosine) criterion (equation 3.2). In these equations a stands for reference image, $\bar{a}$ stands for the mean value of reference picture, b stands for registering image and $\bar{b}$ stands for the mean value of registering image. In the case of identical images, the correlation criterion is equal to zero, and the angle criterion is equal to one. The intensity-based criteria are mostly used for the monomodal set of images [23].

$$C_{CC}(\boldsymbol{a}, \boldsymbol{b}) = \frac{\sum_t (a_t - \bar{a})(b_t - \bar{b})}{\sqrt{\sum_t (a_t - \bar{a})^2 \sum_t (b_t - \bar{b})^2}} \tag{3.1}$$

$$C_A(\boldsymbol{a}, \boldsymbol{b}) = \frac{\boldsymbol{ab}}{|\boldsymbol{a}||\boldsymbol{b}|} = \frac{\sum_t a_t b_t}{\sqrt{\sum_t a_t^2}\sqrt{\sum_t b_t^2}} \tag{3.2}$$

The information-based criteria do not work with intensity values, but rather with similar geometrical features, so it is suitable for registration of the multimodal set of images. The most used is mutual information (equation 3.3)), which is calculated based on the joint histogram. In this equation, I_{AB} stands for mutual information, H_A and H_B stand for information in reference and registering images, respectively, and H_{AB} stands for joint information of the union of these images [23].

$$I_{AB} = H_A + H_B - H_{AB} \tag{3.3}$$

Joint histogram

The joint histogram for two equally sized (reference and registering) gray-scale images is a matrix of the size q x r, where q and r is the number of gray shades in reference and registering image, respectively. When the joint histogram of two identical images is created, it looks like one line of nonzero values on the diagonal

of the matrix (figure 3.1, left). If the images are identical, but one of them was transformed by a point operator (for example, contrast enhancement), the nonzero values are located on the curve, which corresponds to the transforming function (figure 3.1, centre). When the images are geometrically transformed, the line is slightly scattered (figure 3.1, right), but the line (and therefore dependence) is still visible [23].
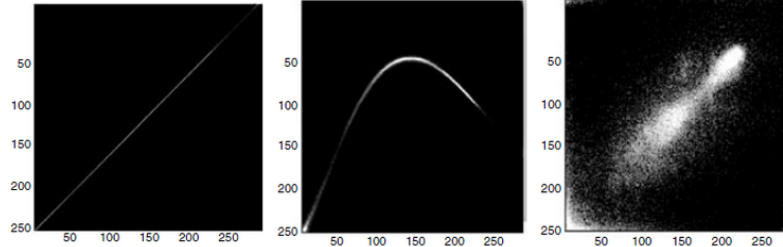

Fig. 3.1: Histogram matrices for different cases - two identical images (left), one of the images has enhanced contrast (centre), geometrically transformed images (right) [23]

3.2.2 Global rigid transformation

In the global registration, the denoised images are registered rigidly. The reference image is one that was taken in the middle of a timed sequence [26].

To determine how much the registered image is shifted from the reference image, the minimizing of the mean square differences (equation 3.4)

$$(x_{shift}, y_{shift}) = \underset{\Delta x, \Delta y}{\operatorname{argmin}}(\sum_{i=1,...N}^{Overlapping area} (I_{ref}(x_i, y_i) - I_{registering}(x_i - \Delta x, y_i - \Delta y))^2/N)$$

(3.4)

is used, where x_{shift} and y_{shift} are mutual shifts in the directions of x and y axes, $I_{reference}$ and $I_{registering}$ are intensity values of reference and registering image, respectively, and N is the total number of pixels in the overlapping area [26].

3.2.3 Phase correlation

This method of global image registration is based on the Fourier transform. For using this approach, the retinal images have to meet the following conditions: they must be of high contrast and the edges must be present in the image. The employment of this method is for the shift estimation based on translation property of the discrete Fourier transform. It can also be used for estimation of rotation and change of scaling, with usage of rotation property and change of scaling property, respectively [12] [21].

In the spectral domain, two displaced images are defined as

$$F_2(u, v) = F_1(u, v) \times e^{u \times x_0 + v \times y_0},$$ (3.5)

where F_1 and F_2 are the spectra of the two images, u and v are the spatial frequences in the directions of x and y axes, and x_0 and y_0 are the displaced positions [12].

The inverse Fourier transform (equation 3.6) of normalised cross spectrum leads to Dirac function at the position (x_0, y_0)

$$\delta(x - x_0, y - y_0) = FT^{-1} \left\{ \frac{F_2(u, v).F_1^*(u, v)}{|F_2(u, v).F_1^*(u, v)|} \right\}$$ (3.6)

where $\delta(x - x_0, y - y_0)$ is the mentioned Dirac function, F_2 is the Fourier spectrum of the second image, F_1^* is the complex conjugate Fourier spectrum of the first image [12].

The accuracy of the shift estimation may be affected by edge effects, which are caused by the periodicity of discrete Fourier transform and different values of pixels on the edges of the image. These effects are eliminated by employing of the window function, which decreases the significance of pixels located on the edge of the image, such as Hann or Hamming function [21].

The input images to the phase correlation have to be identical because otherwise the Dirac impulse will not be formed on the output. When the input images are not identical, for example, they originate from different imaging modalities, in the output matrix are another impulses besides the Dirac impulse, and the shift estimation is impossible [21].

Another limitation of the shift estimation is the inability to distinquish between a large shift in the positive direction and a small shift in the negative direction, so it is important to have apriori knowledge about the direction of the shift [21]. The phase correlation can be also used for estimation of rotation and scaling. For the rotation estimation the rotation property of 2D Fourier transform is used. The rotation angle is estimated from the translation of the amplitude spectra of the pair of images in the polar coordinates. The scaling is estimated similarly, the only diferences are the use of scaling property of the 2D Fourier transform instead of rotation property, and the scaling is estimated using translation of spectra in the log-polar coordinates instead of the simple polar coordinates [21].

3.2.4 Local motion estimation (optical flow method)

This approach consists of several steps - detection of tracking points, Lucas-Kanade tracking, and then the global transformation. The probability map of tracking points is computed from the enhanced image, and it is updated with every registered frame. Then, the Lucas-Kanade algorithm is used for tracking. In this algorithm, one optical flow vector is estimated at each pixel of the image. In the Lucas-Kanade paradigm, the displacement at some pixel is estimated by presuming that it is a rigid motion of a small window centered on the pixel. The size of the window can be really small because large motions are reduced by the phase correlation, so even the smallest motion can be captured. After that, the quality of each tracking point is estimated and finally, the images are registered using the chosen tracking points. This whole process is applied iteratively until the change of parameters between the two iterations is small [12].

3.2.5 Interpolation

After the geometrical transformation (for example, the one mentioned in the previous subsection) is used in discrete space, the transformed pixels have to be assigned back to the discrete grid, as they are often dislocated. If the interpolation is not performed, the resulting images may be distorted or artifacts may appear [23].

The ideal (and hardly used in practice) method of interpolation is interpolation with the use of two-dimensional function *sinc*. However, the *sinc* function is very complex and works in theory with an infinite number of pixels, which increases the time computational complexity. Because of this, other, less accurate methods are used, for example, the nearest-neighbor method of interpolation, the bilinear interpolation, or the bicubic interpolation. The two main requirements for these methods are the good approximation of interpolation properties of function sinc and the use of the smallest number of pixels that are on the input of the interpolation function. These requirements are, logically, the opposites, and each of the methods requires a certain degree of compromise [21].

The nearest-neighbor method of interpolation is the simplest and fastest method out of the methods listed in the previous paragraph. This algorithm assigns the value of the nearest point in the discrete grid to the interpolated value, as seen in figure 3.2 on the left, where the values A', B', C' and D' are the values at interpolated position and A, B, C and D are values of assigned nearest neighbor, respectively. In the resulting interpolated image there are discontinuities due to staircase interpolating function and in the frequency domain, the spectrum of the image is affected by aliasing interference. To overcame these limitations, the input image should be

upsampled, or the use of this function should be the final step before the image is displayed [23] [21].

For better results the bilinear interpolation is used. This method is more complex than the nearest-neighbor method. The new value at the interpolated position is assigned to the value from the curved surface at the corresponding position. The curved surface is computed out of four neighbor pixels, as is displayed in figure 3.2 on the right. The X' is value at interpolated position and A, B, C and D are values of neighboring pixels. In comparison with the nearest-neighbor method, the aliasing interference is suppressed and the interpolating function is continuous. This method is good for displaying, but the use is limited when the following image processing is more complex [23] [21].
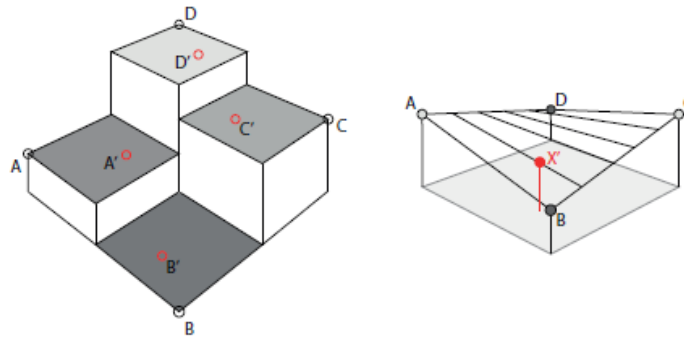

Fig. 3.2: Interpolating value of pixel for nearest-neighbor method (left) and bilinear interpolation (right) [21]

3.3 Detection of blurry images

When registering images, it is assumed that all frames from the sequence are not blurred to ensure the good quality of the final image. However, the eye or head movements result in blurred images, so they must be eliminated before analysis.

The extraction of four features can be used to classify, whether the frame is distorted or not. The features are the mean value of the frame, the entropy of the frame, edge image, and phase correlation. The first feature is dependant on the intensity of the image. The intensity is higher during eye blink because more light is reflected back from the eyelid. The second feature, frame entropy, is computed as a convolution between the original, undistorted frame, and the point-spread function which describes the distortion. The third feature is estimated by convolution of the frame with the Sobel operator to enhance the edges. If the image is blurred, the edges are blurred, too. The last parameter is approached with the use of cross-correlation spectrum

and the inverse Fourier transform between the original scene and the frame. The result of this is a 2D Gaussian function positioned at different locations. The width of this function gives information about the blurriness [27]. For classification were in [27] used three classificators - support vector machine, naive Bayes, and quadratic discriminant analysis classifier.

3.4 Super-resolution reconstruction

The super-resolution reconstruction comes from the fusion of multiple low-resolution images into one single frame with improved and increased spatial resolution and removed noise and blurring. Most super-resolution reconstruction algorithms are based on L1 or L2 norm, which means that they are sensitive to the noise estimation [28].

The popular matrix notation used for formulation of general super-resolution model represented in the pixel domain is defined as

$$\underline{Y}_k = D_k H_k^{cam} F_k \underline{X} + \underline{V}_k \quad k = 1,...N \tag{3.7}$$

where the matrix F_k is the motion between the high-resolution frame (magnified with factor r) $\underline{X}$ and the kth low-resolution frame $\underline{Y}_k$, H_k^{cam} is the blur matrix which models the camera's point spread function, matrix D_k represents the decimation operator, vector $\underline{V}_k$ is the system noise and N is the number of low-resolution frames [29].

There are many methods that do this, such as nonuniform interpolation, frequency domain, maximum likelihood, maximum a posteriori, and projection onto convex sets [28].

3.4.1 Frequency domain methods

These methods are based on the spatial aliasing effect and were among the first to be presented for the super-resolution reconstruction. They are quite simple and not computationally complex, but the model is only for global geometrical transform and linear space-invariant blur, and it is hard to use the spatial domain a priori knowledge for regularization [28].

3.4.2 Projection onto convex sets methods

These methods can include *a priori* information, in contrast to frequency domain methods, but they have high computational cost and slow convergence [28].

3.4.3 Maximum a posteriori regularization methods

These algorithms are based on the stochastic regularization technique of Bayesian maximum a posteriori estimation by minimizing a cost function. The determination of the regularization parameter is crucial for image restoration [28].

There are many approaches for estimating the super-resolution reconstructed image. For the purpose of this thesis, the regularization based on bilateral total variance is explained in detail.

Regularization based on bilateral total variance

Super-resolution reconstruction is an ill-posed problem, meaning that for the under-determined case exist an infinite number of solutions that satisfy the general super-resolution model (equation 3.7). Accordingly, the regularization is necessary, and when considering possible regularization terms, the one which results in a high-resolution image with sharp edges and is relatively simple to implement should be desired [29].

The regularization term balances the lack of measurement information with prior knowledge about the desired high-resolution result and is represented as a penalty factor in the minimization cost function

$$\hat{\underline{X}} = \underset{\underline{X}}{\operatorname{argmin}} \left[\sum_{k=1}^{N} \rho(\underline{Y}_k, D_k H_k F_k \underline{X}) + \lambda \Upsilon(\underline{X}) \right] \tag{3.8}$$

where λ is the scalar regularization parameter, which is used for weighting the first term (similarity cost) against the second term (regularization cost) and Υ is the regularization cost function [29].

One of the most used regularization cost functions is the Tikhonov cost function. This function is based on limiting the total energy of the image or forcing spatial smoothness. The edges and noise are of high-frequency energy, and therefore are removed, and the resulting image does not contain sharp edges [29].

The total variation method, in contrast to the Tikhonov cost function, preserves the edges in the resulting image, because it is based on L_1 norm of the magnitude of the gradient [29].

The bilateral total variation regularization function is easy to implement and preserves edges. It is defined as

$$\Upsilon_{BTV}(\underline{X}) \sum_{l=-P}^{P} \sum_{m=0}^{P} \alpha^{|m|+|l|} \|\underline{X} + S_x^l S_y^m \underline{X}\|_1 \tag{3.9}$$

where S_x^l and S_y^m are operators that shift $\underline{X}$ by l and m pixels in horizontal and vertical directions, respectively, giving several scales of derivatives, α is scalar weight

in range of 0 to 1, and is applied for decaying effect of the regularization terms- If the factor P is equal to one, the regularization is simplified into total variation function [29].

3.5 Image quality assessment

Image quality can be described as the deviation from the ideal or reference model. The decrease of image quality can be affected by the noise and the distortion of the image [30].

Several techniques and metrics can be used for objective image quality assessment. These techniques can be divided into two categories. The categories differ in the availability of the reference image. The first category is a full-reference approach, where the main focus is on the quality assessment of the test image in comparison with the reference image. The second category is a no-reference approach, where the quality is assessed from test image only [30].

3.5.1 Full-reference approach

There are many estimators of image quality in comparison with the reference image, such as Mean Square Error, Structure Similarity Index Method, Feature Similarity Index Method, etc.

Mean Square Error (MSE)

MSE is the most common estimator of image quality. It is a full-reference metric and the value zero means the best quality (the images are identical). MSE is a representation of absolute error [30].

MSE between two images f(x,y) and g(x,y) is defined as

$$MSE = \frac{1}{MN} \sum_{m=0}^{M} \sum_{n=0}^{N} [g(n,m) - f(n,m)]^2 \qquad (3.10)$$

where M and N are the sizes of images vertically and horizontally [30].

Structure Similarity Index Method (SSIM)

SSIM is perception-based model. Image degradation is considered as the change of perception in structural information. It is full-reference metric and the value one means the best quality (images are identical) [30].

Feature Similarity Index Method (FSIM)

FSIM maps the features and measures the similarities between two images - full-reference metric. The value one means the highest similarity (images are identical) [30].

This metric is based on two criteria - phase congruency and gradient magnitude. Phase congruency relies on the features in the spectral domain and is contrast invariant. Gradient magnitude is computed via convolutional masks [30].

FSIM from the two images f_1 and f_2 is defined as

$$S_L(x) = [S_{PC}(x)]^\alpha . [S_G(x)]^\beta \tag{3.11}$$

where α and β are parameters used to adjust the relative importance of S_{PC} and S_G [30].

S_{PC} is similarity based on phase congruency. It is defined as

$$S_{PC} = \frac{2PC_1PC_2 + T_1}{PC_1^2 + PC_2^2 + T_2} \tag{3.12}$$

where PC_1 and PC_2 are phase congruency maps extracted from images f_1 and f_2 respectively, and T_1 is a positive constant [30] .

S_G is similarity calculated from magnitude gradient maps G_1 and G_2 extracted from images f_1 and f_2, and is defined as

$$S_G = \frac{2G_1G_2 + T_2}{G_1^2G_2^2 + T_2} \tag{3.13}$$

where T_2 is a positive constant [30].

3.5.2 No-reference approach

In case the reference picture is not available, some algorithms are trained to evaluate image quality as the human eye does, for example, Perception-based Image Quality Evaluator, Natural Image Quality Evaluator, and Blind Referenceless Image Spatial Quality Evaluator.

Perception based Image Quality Evaluator (PIQUE)

This method originates from an idea of how the human eye perceives the quality of the image. At first, humans see prominent points in the image or regions that are spatially active. After that, the overall image quality increases with local quality at block levels. For computing the PIQUE score, at first, the distortions are estimated only on prominent spatial regions, and then at the local block levels [31].

The PIQUE score lies in the range of 0 to 1. If the score is near zero, it means that the image quality is high. If the score is near one, it represents an image of low quality [31].

Natural Image Quality Evaluator (NIQE)

This method is based on a space domain natural statistic model. Firstly, a multivariate Gaussian model of the features for the image quality assessment is extracted from the database of natural images. Secondly, the distance between this model and a multivariate Gaussian fit of the natural scene features that are extracted from the test image is computed [32].

As in the case of the PIQUE score, the lower NIQE score means better image quality [32].

Blind Referenceless Image Spatial Quality Evaluator (BRISQUE)

This method uses locally normalized luminance coefficients, rather than distortion-specific features as the methods mentioned before. As in the previous method, the model had to be trained on an image database, and then is used to quantify the quality of the test image [33].

As in the previous non-reference methods, the BRISQUE score lies usually in the range of 0 to 1, the lower score meaning better image quality [33].

4   Proposed algorithm

The image sequences used in this work were taken by flood-illumination camera (rtx1, Imagine Eyes, Orsay,France) in cooperation with the University of Leipzig. The retinal images originate from ten healthy subjects. During imaging the subjects focused on a moving light target.Therefore the images are captured in different retina locations. The locations correspond to eye movement involved by eye fixation on movable target.

The raw data are 119 sequences, consisting of 40 frames, eac frame has a resolution 1279 x 1279 pixels. The sequences are acquired with non-uniform lighting. The size of the resulting adaptive images from the software Imagine Eyes is 1500x1500 pixels.The example of an Imagine Eyes image is in figure 4.1 , where the bright spots correspond to individual photoreceptors and the vessels are typically depicted as darker capillary areas with lower photoreceptor occurance. For an accurate estimation of resolution each eye is measured separately, in these images the 1.1 mm^2 corresponds to 4°x4° in the captured scene (computed using the Bennett formula).
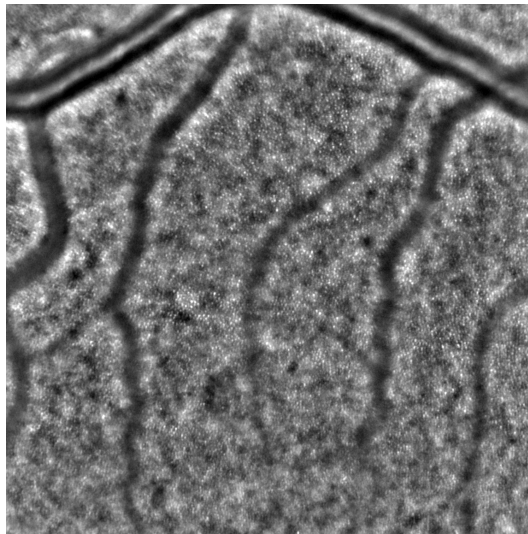

Fig. 4.1: An example of a resulting image obtained by the Imagine Eyes software

The goal of this work is to create one high-quality image from each sequence. The frames in the sequence capture same retinal area, but due to minor and involuntary eye movements are not mutually registered. Furthermore each frame is burdened by uneven illumination and noise and several frames in a sequence are blurred. To obtain the image of the best quality possible, this burdens are removed, and then the frames are registered. The resulting high-quality images are obtained by two methods - averaging of the registered sequence and with the super-resolution technique.

4.1 Image pre-processing

The sequence of raw frames is loaded into the three-dimensional matrix. Each of the images is pre-processed separately.

Firstly, the value range of the processed image is transformed to use the full range from 0 to 1. As is displayed in figure 4.2, the intensity values of the raw image are in the range from 0 to approximately 0.05 and after the transformation, the full range of normalized grayscale values is used.
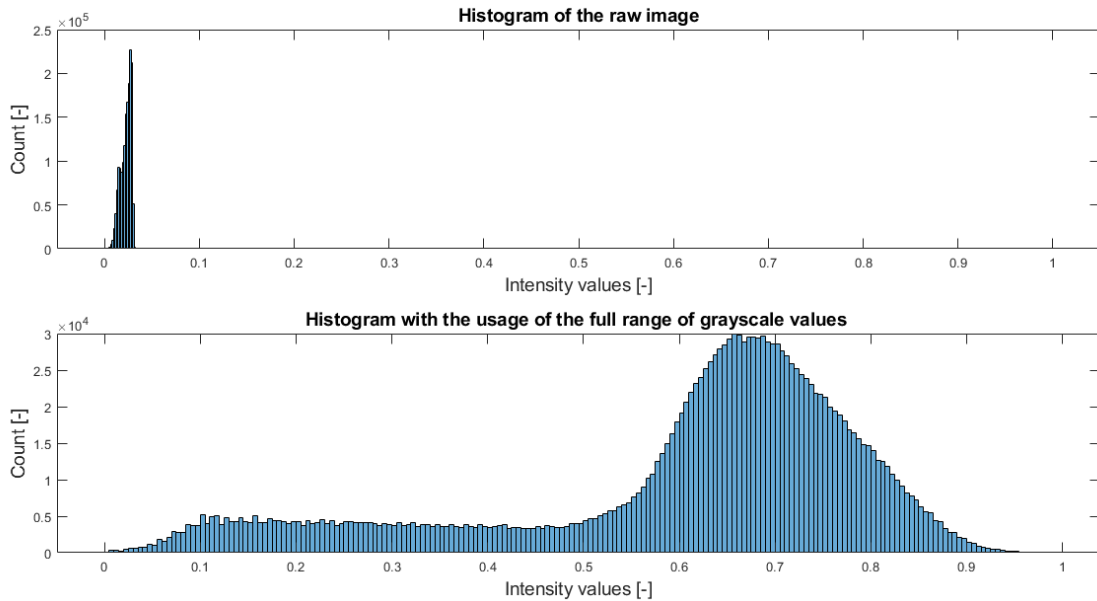

Fig. 4.2: Histogram of raw image (top) and histogram of image with
the usage of full range of grayscale values (bottom)

4.1.1 Correction of non-uniform illumination

The frames are obtained with non-uniform lighting - the central area is very bright and the edges are darker, even after the histogram normalization (figure 4.3 on the left).

To correct this non-uniformity, a copy of the frame is highly blurred by the application of a Gaussian filter where the standard deviation of the Gaussian distribution was empirically determined and set to 20. The Gaussian filter with high standard deviation removes all the small objects in the image like cones or inhomogenities (figure 4.3, centre). However, if the standard deviation is set to a too high value, the usage of the blurred frame could lead to distortion of the frame rather than to frame with uniform illumination.

The last step in this process is the substraction of the blurred frame from the original frame, where the result is normalized again to the scale between 0-1 (figure 4.3,

right). Alternatively, the division could be used instead of substraction, but the results are of higher quality with the use of substraction.

Fig. 4.3: Raw frame (left), blurred frame (centre), frame with uniform illumination (right)

4.1.2 Noise reduction

For noise reduction, the bilateral filter is used. In the developing stage of this step many other denoising options were considered and tested, including median filter, averaging filter and adaptive Wiener filter, but the results after the use of bilateral filter were the most promising, because the others smudged the photoreceptors and that is undesirable for the purpose of this thesis.

After filtering, the contrast is adjusted to saturate the bottom and top 1 % of all pixel values. This step helps to visually distinguish between the white photoreceptors at the top pixel values and the blood vessels and background at the bottom pixel values.

The process of filtering is displayed in figure 4.4. The centre image shows the image after bilateral filtering and the right image shows the image after contrast adjustment. The bilateral filtering suppressed noise in the image and the contrast adjustment highlighted the structures in the image.
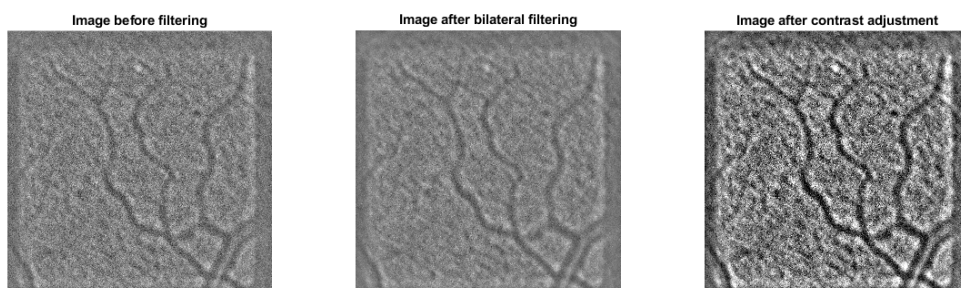

Fig. 4.4: Image before noise reduction (left), image after bilateral filtering (centre), image after contrast adjustment (right)

4.1.3   Detection of blurred frames

One of the key parts of image pre-processing in this thesis is the detection and removal of blurred frames. These frames are mostly the results of fast eye movements which decrease the quality of the resulting image.

To detect these frames, the Canny edge detector is applied first. The edges in the frame are set tovalue of one representing white pixels. The non-edges and background are set to a value of zero representing black pixels. In the non-blurry frames the majority of edges are situated in the centre of the image as a result of uneven illumination of the raw frames. However, in the blurry frames the edges are detected randomly across the whole frame.

In the first step of the detection, a logical mask the same size as the frame is applied to a frame with detected edges. The mask is an area of zeroes in the centre (1239x1239 pixels) and ones on the edges (20 pixels on each side) (figure **??**). When performing a logical operation *and* between the mask and frame, the edges found in the centre of the resulting frame are set to zero. The outer part of the frame has preserved white pixels with detected edges.

After that, the percentage of white pixels in the outer part of the image is computed as:

$$percentage = \frac{number\ of\ white\ pixels\ in\ the\ outer\ part\ of\ the\ image}{number\ of\ all\ pixels\ in\ the\ outer\ part\ of\ the\ image} \qquad (4.1)$$

If the percentage of white pixels is higher than an empirically determined 10 % then the frame is declared blurred and therefore removed.

The following figures show examples of the different types of frame. Figure 4.5 shows a blurry frame and figure 4.6 shows a sharp frame. In each figure the left image is the original, the centre image is after Canny edge detection and the right image is after the mask is applied. The centre images in these figure show that, the sharp frame has edges mostly in the centre of the image while the blurry frame has edges randomly across the whole image.
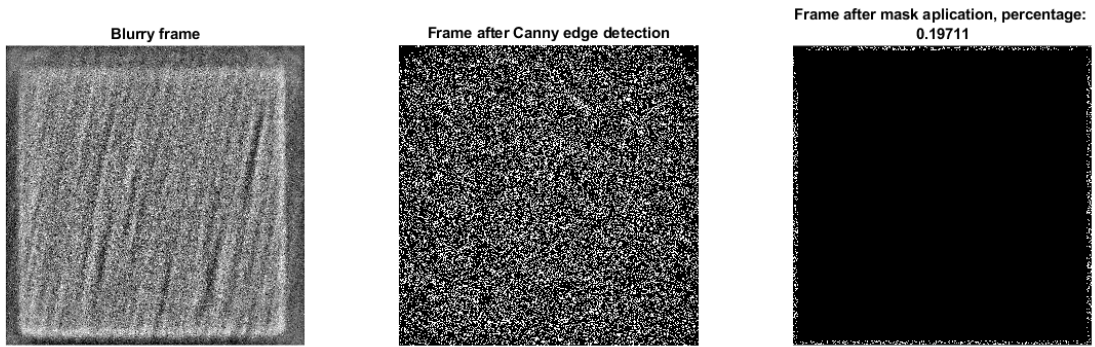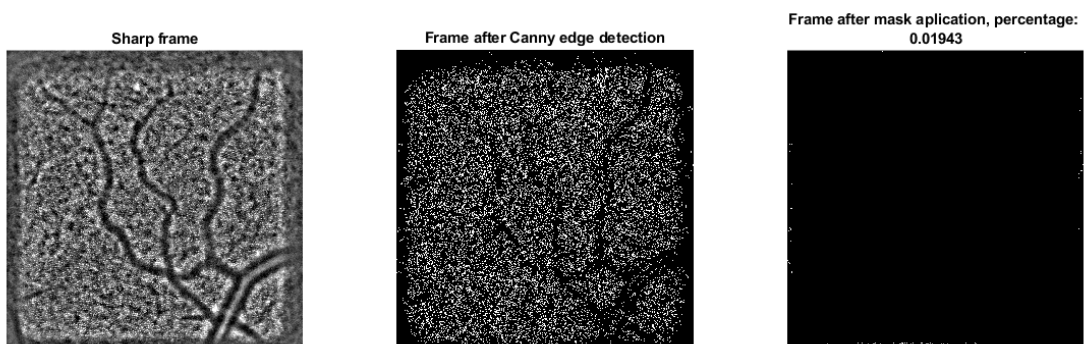
Fig. 4.5: Blurry frame

Fig. 4.6: Sharp frame

During the sequence pre-processing arise another problem, caused by frames that capture different part of retina. This shift within a sequence is initiated by extensive eye movement and disables the registration of the neighboring dissimilar frames. To detect this frames, the normalized cross-correlation between neighboring frames is used (for example, it is computed between the first frame and second frame, then between the second frame and the thirs frame and so on).

As a first step the normalized cross-correlation is computed, then the highest peak with the maximum cross-correlation coefficient is found. When the value is lower than the threshold, the sequence is divided. In the majority of the sequences, the main sequence is divided into two groups, and whichever group contains more frames is used in the next stage of proposed algorithm. To determine whether the frames belong only to two groups, or should be divided into more groups, the normalized cross-correlation coefficients could be computed in a loop until the condition of them not being below the threshold is not met.

The range of the maximum value of the normalized cross-correlation, from one image sequence, is shown in the figure 4.7. Each of the groups is of different color - the

first group is red and the second group is blue. It is obvious that the red group contains more frames, and therefore is used in further analysis.
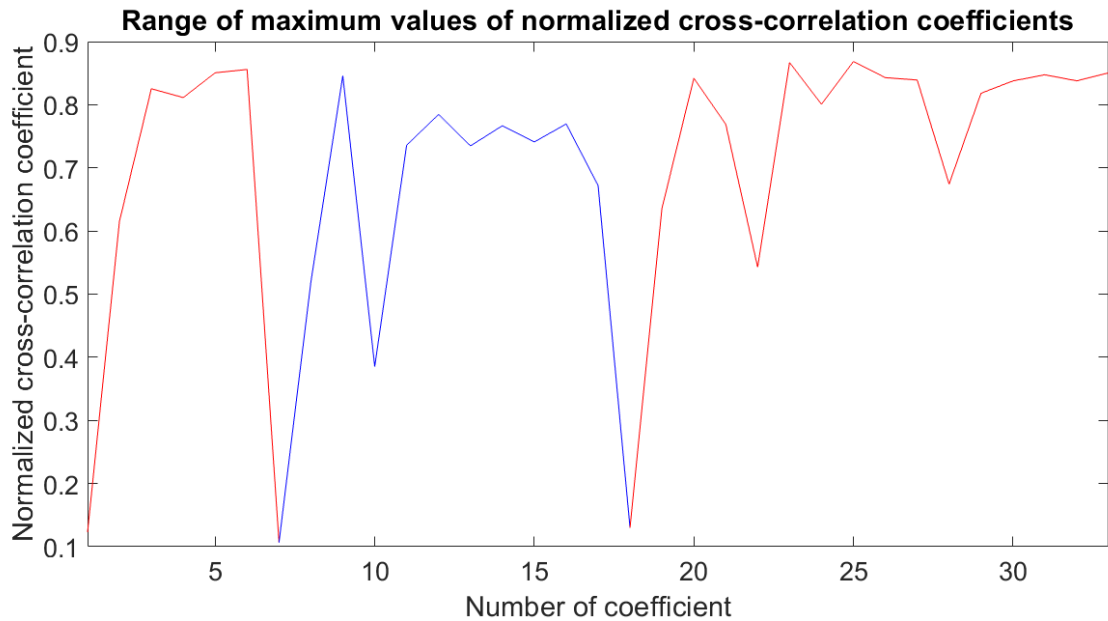

Fig. 4.7: Range of maximum values of normalized cross-correlation coefficients

Figure 4.8 shows two neighboring frames (frame number 7 and frame number 8) and their maximum normalized cross-correlation coefficients. It is obvious that these frames are very different from each other, and the low cross-correlation coefficient only confirms this observation.
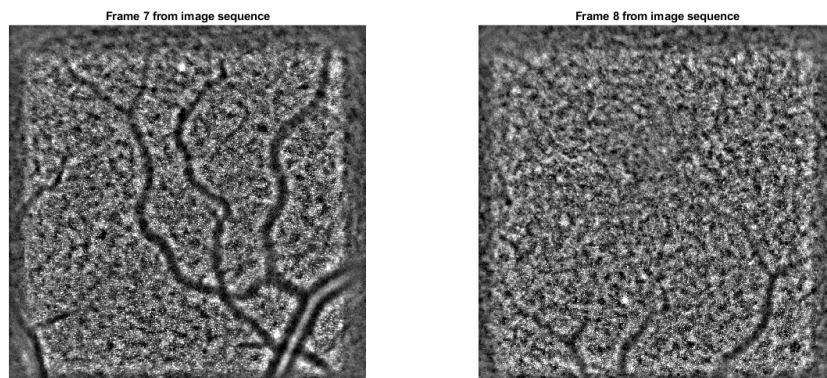

Fig. 4.8: Two neighboring frames with different parts of retina captured. Maximum normalized cross-correlation coefficient: 0.086806

4.2   Image registration

As the frames in a sequence are shifted because of eye movements, an estimation of the shift, rotation and scaling between the frames is needed. During the development of this algorithm, the rigid transformation followed by local registration using the optical flow method was considered. However, this was dropped because of its computational complexity. Also, the properties of the image registration using phase correlation are computed simultaneously, and give more sufficient results. The transformation matrix is computed between the first frame of te sequence and the following frames, so in the sequence of 40 frames, the resulting number of matrices is 39.

In the first step of this stage of algorithm, the blurry edges are cut off all the frames that were selected in the previous stages. The resulting resolution of the frames is 1000x1000 pixels. In the next step, the transformation matrices are computed with phase correlation and applied to the frames. The first resulting image is obtained as the averaged image from the transformed frames (figure 4.9).
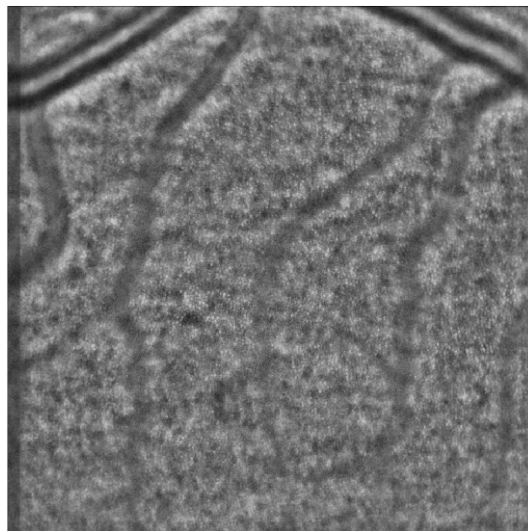

Fig. 4.9: An example of a resulting image obtained by the averaging technique

4.3   Super-resolution reconstruction

Another method of producing a resulting image is with super-resolution reconstruction, specifically bilateral total variation regularization. This was previously explained in section 3.4.3. This method was chosen because it preserves the edges, removes the noise and blurriness while obtaining a single super-resolution image

from multiple low-resolution frames. There are several parameters that are required to set up before reconstruction - magnification factor, in this case it is set to 2, width of the point-spread function that is set to 0.3, regularization weight λ set to 0.00015, factor P set to 2, weight α set to 0.65, maximum iterations set to 7, and measure of absolute tolerance to 1.10^{-10}. All of these parameters were determined empirically. The L1 regularization norm is used. The transformation matrices obtained in the registration via phase correlation are also used in this method.

The sequence of low-resolution frames used in this method are non-registered frames before the application of a bilateral filter and before contrast enhancement, with the blurred edges cut off. For this task a multiframe super-resolution toolbox is used from the Friedrich-Alexander University Erlangen-Nürnberg [34].

In the final step, the contrast of the resulting high-resolution image is adjusted in the way described in section 4.1.2 to highlight the structures in the scene.

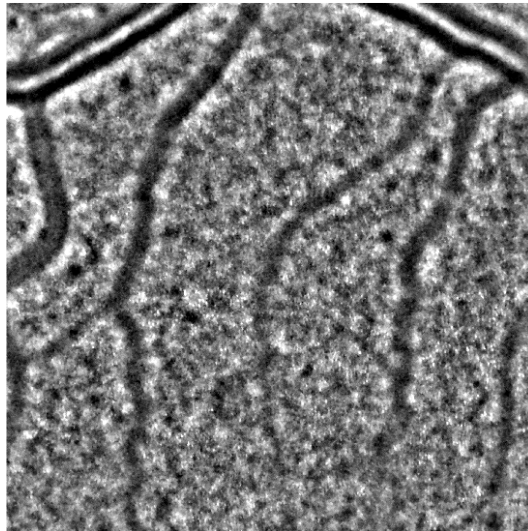The resulting image is displayed in the figure 4.10.

Fig. 4.10: An example of a resulting image obtained by the
super-resolution technique

4.4   Results

As stated in section 3.5, there are many metrics to evaluate the quality of the images. Since there are no reference pictures available to compare with the resulting images, only the no-reference metrics of PIQUE, NIQUE and BRISQUE are used. These are computed from the available Imagine Eyes, the super-resolution and the averaged resulting images. The resolution of the Imagine Eyes image is 1500x1500 pixels and

of the averaged image is 1000x1000 pixels. The resolution of the super-resolution image is 2000x2000 pixels.

The resulting scores are displayed in the boxplots 4.11 for PIQUE scores, 4.12 for NIQE scores and 4.13 for BRISQUE scores. The minimum, median and maximum scores for each set is displayed in table 4.1. From the boxplots and table it is obvious that the highest quality images are obtained from super-resolution technique because the scores are closest to zero.

The boxplots provide visualisations the scores from the whole dataset. The bottom and top of each box are the 25th and 75th percentiles of scores of each method, respectively. The red line in the middle of each box is the median of the scores of the resulting images. The whiskers are lines extending above and below each box. Observations beyond the whisker length are marked as outliers. An outlier appears as a red + sign.
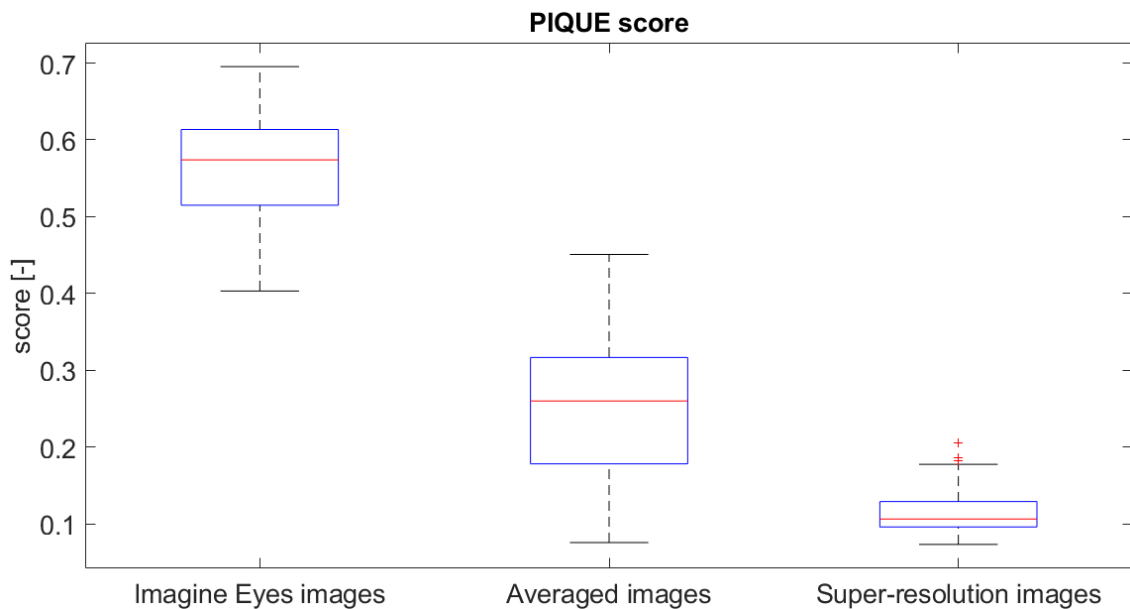

Fig. 4.11: Boxplots of PIQUE scores of the sets of available Imagine
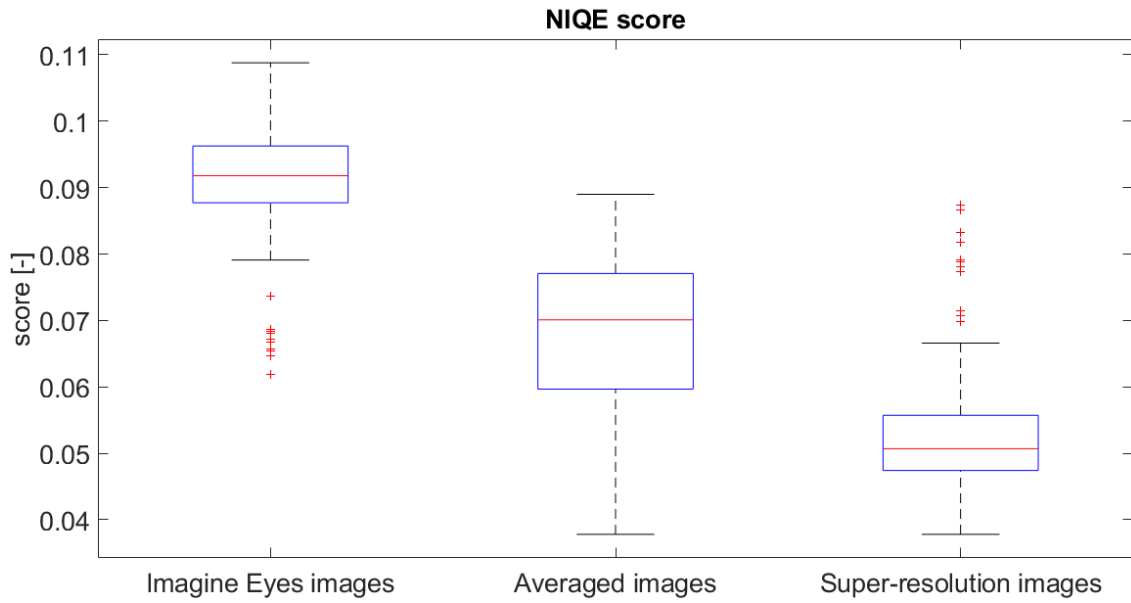Eyes images, averaged images and super-resolution images

42

Fig. 4.12: Boxplots of NIQE scores of the sets of available Imagine Eyes images, averaged images and super-resolution image
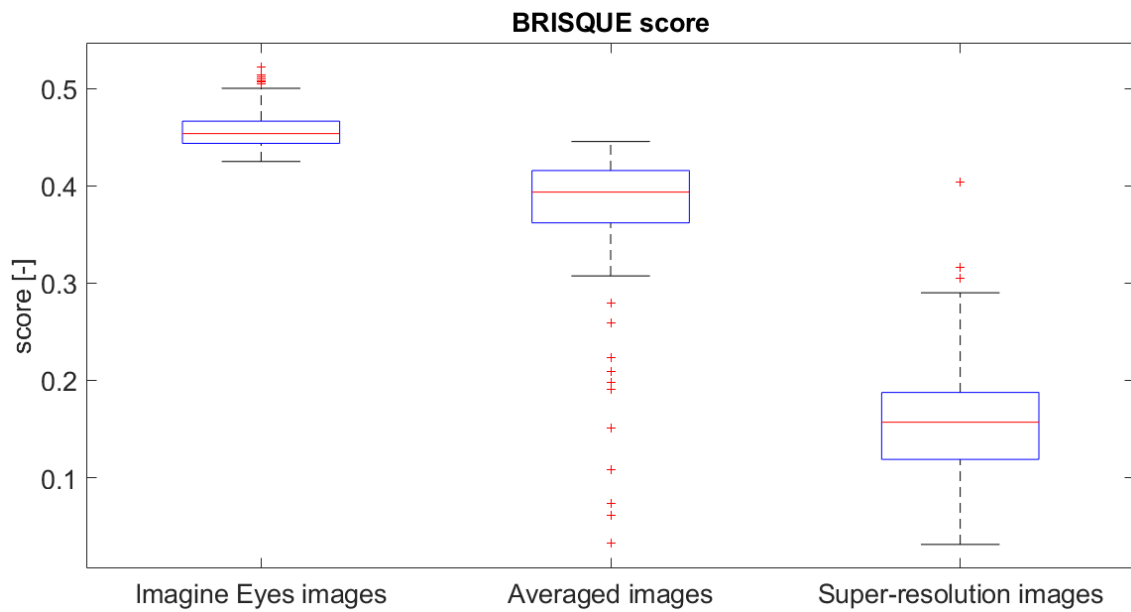

Fig. 4.13: Boxplots of BRISQUE scores of the sets of available Imagine Eyes images, averaged images and super-resolution image

Tab. 4.1: Minimum, median and maximum scores for each set of resulting images

		Imagine Eyes images	Averaged images	Super-resolution images
PIQUE	*maximum*	0.6954	0.4507	0.2056
	median	0.5739	0.2600	0.1063
	minimum	0.4032	0.0756	0.0731
NIQE	*maximum*	0.1088	0.0890	0.0874
	median	0.0918	0.0.0701	0.0507
	minimum	0.0619	0.0378	0.0378
BRISQUE	*maximum*	0.5227	0.4456	0.4043
	median	0.4538	0.3936	0.1570
	minimum	0.4252	0.0323	0.0313

When all three resulting images are displayed in full resolution (figures 4.14 and 4.15), the super-resolution image is also subjectively evaluated as the highest quality image. The averaged image has artefacts of black pixels, as a result of shifts between frames and the cutting of the blurred edges. It is, however of higher quality than the resulting image would be with the retaining of the blurred edges, as the blurred parts would lead to the loss and degradation of the information, and it was decided that for the purpose of this thesis the resulting images with edge artifacts are more desirable than the resulting images with the degraded parts because of blur.
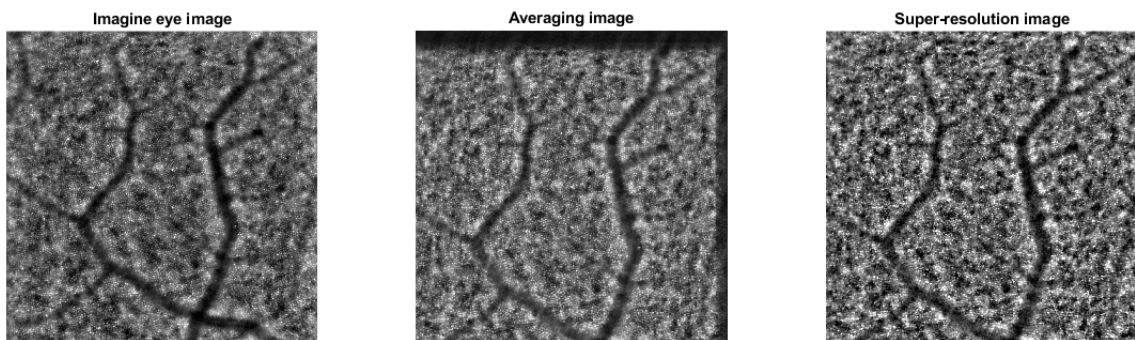

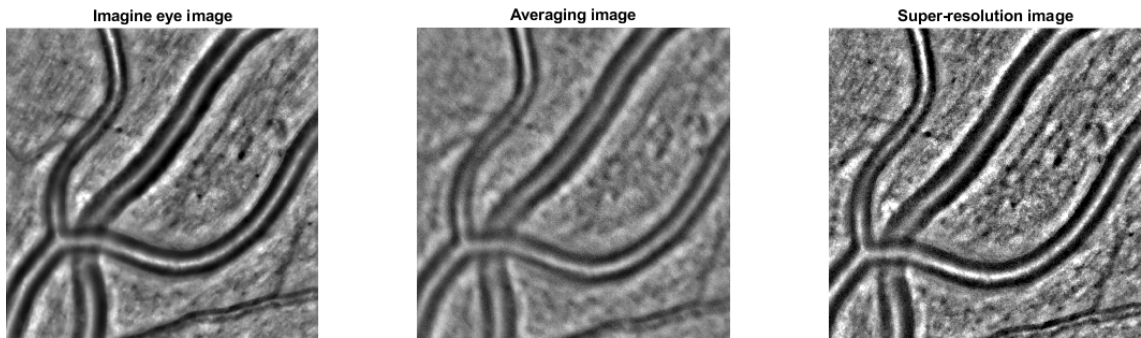Fig. 4.14: Imagine eye image and two resulting images (example one)

Fig. 4.15: Imagine eye image and two resulting images (example two)

When the images are maginified (figures 4.16, 4.17 and 4.18), the imperfections are clearly visible. The Imagine Eyes and averaged images look similar, the photoreceptors are bounded and are clearly visible, but in the super-resolution image, the photoreceptors that are close and have the same illuminance are clustered (as seen in figure 4.16). On the other hand, the areas that are blurred in both the Imagine Eyes and the averaged image are sharpened in the super-resolution image. The super-resolution image is in general less blurry than the other two, since it is resistant to lower accuracy of the image registration. Also, the super-resolution image is able to highlight the blood-vessels (as seen in figures 4.17 and 4.18).

Fig. 4.16: Details of the Imagine Eyes (600x600 px), the averaging (400x400px) and the super-resolution resulting image (800x800px), example one
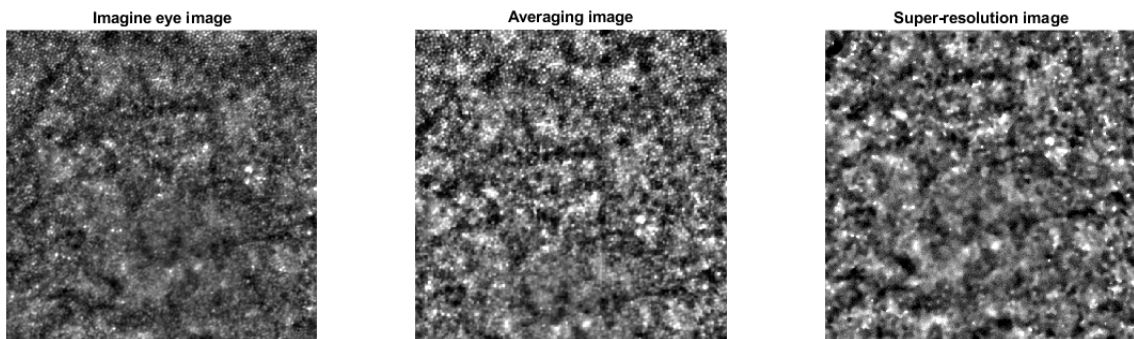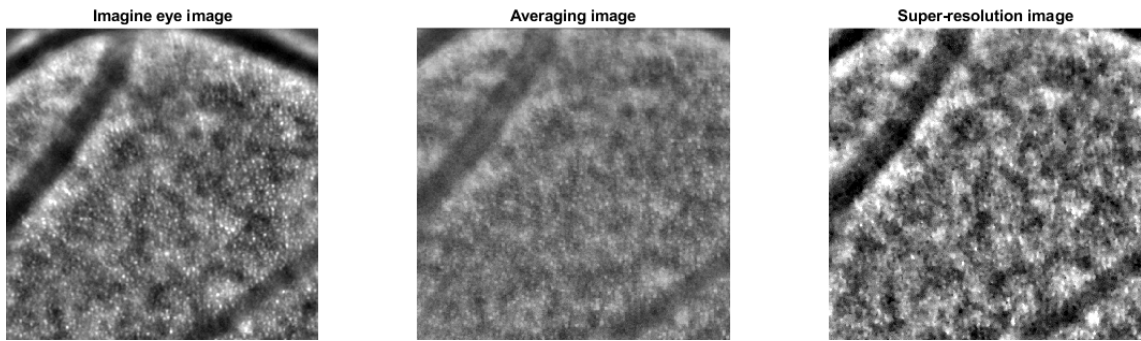
Fig. 4.17: Details of the Imagine Eyes (600x600 px), the averaging (400x400px) and the super-resolution resulting image (800x800px), example two
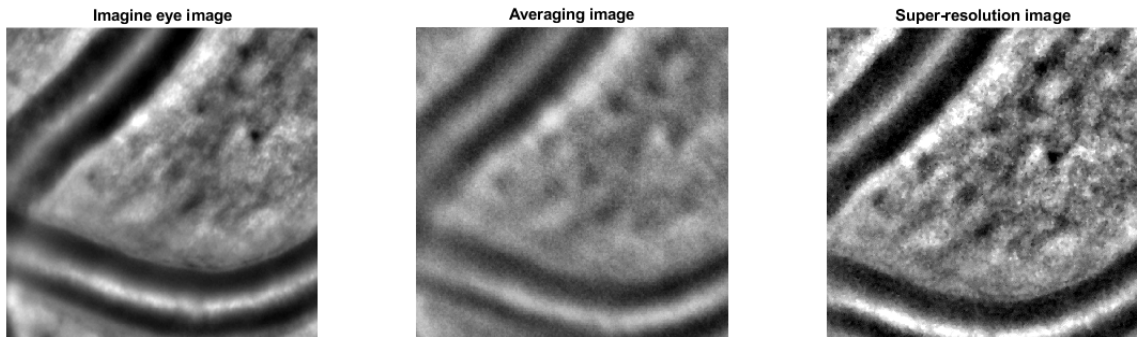

Fig. 4.18: Details of the Imagine Eyes (600x600 px), the averaging (400x400px) and the super-resolution resulting image (800x800px), example three

5   Discussion

In this thesis, the four stages of the final algorithm are proposed. The first stage is for equalization of the non-uniform illumination. The second stage is for detection and removal of the blurred frames. Then the frames captured in different retina area (included within a sequence) are detached. The fourth and final stage is for obtaining high quality resulting images using two separate methods.

The first stage of algorithm is the correction of non-uniform illuminance of the scene. This is done by the substraction of a highly blurred frame (by application of Gaussian filter with high standard deviation) from the original frame. This method is proved quite succesful, but is sensitive to the setting of the standard deviation. Setting of a too high standard deviation causes the formation of the artifacts. Logically, the next step is the transformation of the frames to ensure that the entire range of grey values is used.

The second stage of the algorithm is the detection and removal of blurry frames. This is based on an empirically proven fact that after the application of the Canny detector on frames, less than 10 % of white pixels representing edges is located on the edges of the frame if the frame is sharp. This algorithm succesfully reveals most of the blurred images, but the overall performance could be improved with the use of more sophisticated methods such as measuring the spectrum width.

The third stage of the algorithm is the detection and removal of frames captured after extensive eye movements. These frames captures different retina locations as the rest of the frames. To detect these frames a normalized cross-correlation is used. When the value of the maximum coeffiecient is less than 0.3, that means that these frames are excluded from further analysis. The performance of this algorithm is challenging, mainly because of computational cost of performing normalized cross-correlation for the images of this size.

The fourth and the most important stage of the algorithm is obtaining of resulting single-frame high-quality image. This is done using two methods - averaging and the super-resolution technique.

The resulting image from the first method is obtained by estimating the transformation matrices with phase correlation, then transforming the frames and lastly averaging through third dimension of the transformed frames in sequence. The resulting images are very similar to the available Imagine Eyes images. Although they are sometimes more blurry and usually contains an artifact on the edges. The main advantage of this method is that the photoreceptors are clearly separated. The method is less successful in registration of images focused on blood vessels. This

leads us to possible further improvement of registration technique, which is more reliable andd that is able to estimate even the smallest shifts and rotations. The median of image quality scores obtained by this method are - PIQUE: 0.2600, NIQE: 0.0701, and BRISQUE: 0.3936. The quality scores are not that close to zero as the scores of the images obtained with the second technique, particularly because of the artifacts and overall blurriness, but are still lower than in the Imagine Eyes resulting images obtained by the available software.

The super-resolution technique provides resulting images of the highest quality with median scores - PIQUE: 0.1063, NIQE: 0.0507, and BRISQUE: 0.1570. The bilateral total variation regularization is used to preserve edges, remove the noise and blurriness. The advantage of this method is its resistance to low accurate image registration, in opposite to as the first method, but the photoreceptors in the most of the resulting images are not clearly separated, they are often clustered or smudged, and have no strong boundaries. The performance of this method was excellent in the frames containing blood vessels, as seen in figure 4.15. The disadvantages of this method could be solved by determination of the paramaters for the bilateral total variation regularization for each frame, or at least for each sequence separately.

All these algorithms were tested on the total of 119 image sequences. The first three stages of the final algorithm give good and stable performance for all the sequences. The performance of the last two stages depends on the content of the sequence. The first method gives more accurate results in the sequences with minimum blood-vessels, lots of photoreceptors and slight movements, while the second method is better when used on sequences containing larger movements and blood-vessels. The resulting images obtained from both methods are objectively of better quality than the Imagine Eyes images obtained by available software with median scores - PIQUE: 0.5739, NIQE: 0.0918, BRISQUE: 0.4538. These images are, when evaluated subjectively in some sequences, of better quality, because they have no edge artifacts compared with the averaged images. The photoreceptors are also clearly visible and are not smudged compared with the super-resolution images.

Conclusion

The main focus of this master's thesis is on the use of adaptive optics in retinal imaging and processing of these images. The aim of this thesis is to improve the quality of theses images with the use of phase correlation and super-resolution techniques.

The first chapter contains a description of the human eye from the anatomical and physiological point of view. In the second chapter adaptive optics is defined and the principles of this technique are explained. Also in this chapter is an overview of the use of adaptive optics in retinal imaging. The third chapter contains an explanation of the current approaches in the processing of adaptive optics retinal images. The fourth chapter outlines the design and testing of the designed algorithm for image pre-processing, registration and obtaining the single-frame high-quality result.

The performance of the described algorithm is relatively good. The last part of it is divided into two approaches. The first approach is based on averaging the registered image sequence and gives better results in the image sequences that contain minimum blood-vessels and little movements between the frames. The first approach is also able to display the photoreceptors better. Although the quality scores are slightly worse than scores of the resulting image from the second approach that is based on the super-resolution technique. The second approach is clearly better in the image sequences containing thick blood-vessels and large eye movements. The median quality scores are - PIQUE 0.2600 for averaged images and 0.1063 for super-resolution images, NIQE 0.0701 for averaged images and 0.0507 for super-resolution images, and BRISQUE 0.3936 for averaged images and 0.1570 for superresolution images. Objectively, the images obtained by the second approach are of better quality.

The main benefit of this work is the improvement of the resulting image in comparison with the resulting image obtained from already available software. This could possibly, after some improvements, be used for faster and possibly better diagnosis of retinal pathologies and other disease

In the future this thesis could be extended by better estimating of the used parameters for the super-resolution technique. These parameters were determined mostly empirically, and are the same for all the sequences used. The extension could include estimation for each sequence separately.

Bibliography

[1] ABRAMOFF, Michael D., Mona K. GARVIN and Milan SONKA. Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering* [online]. 2010, 3, 169-208 [accessed 2020-10-28]. ISSN 1937-3333. Available from: doi:10.1109/RBME.2010.2084567

[2] REESE, Paul D. The Neglect of Purkinje's Technique of Ophtha-lmoscopy Prior to Helmholtz's Invention of the Ophthalmoscope. Ophthalmology [online]. 1986, 93(11), 1457-1460 [accessed 2021-5-15]. ISSN 01616420. Available from: doi:10.1016/S0161-6420(86)33556-5

[3] KEELER, C R. Babbage the unfortunate. British Journal of Ophthalmology [online]. 2004, 88(6), 730-732 [accessed 2021-5-15]. ISSN 0007-1161. Available from: doi:10.1136/bjo.2003.018564

[4] KEELER, C. Richard. The Ophthalmoscope in the Lifetime of Hermann von Helmholtz. Archives of Ophthalmology [online]. 2002, 120(2) [accessed 2021-5-18]. ISSN 0003-9950. Available from: doi:10.1001/archopht.120.2.194

[5] A. Gullstrand, "Neue methoden der reflexlosen Ophthalmoskopie", Berichte Deutsche Ophthalmologische Gesellschaft, vol. 36, 1910.

[6] L. Allen, "Ocular fundus photography: Suggestions for achieving consistently good pictures and instructions for stereoscopic photography", Amer. J. Ophthalmol., vol. 57, pp. 13-28, 1964.

[7] WARD, Jeremy P. T. and R. W. A. LINDEN. *Základy fyziologie.* Praha: Galén, c2010. ISBN 9781405177238.

[8] HOLIBKOVÁ, Alžběta and Stanislav LAICHMAN. *Přehled anatomie člověka.* 5. vyd. Olomouc: Univerzita Palackého v Olomouci, 2010. ISBN 978-802-4426-150.

[9] In: *ResearchGate* [online]. [accesed 2020-10-14]. Novel Nanomicellar Formulation Approaches for Anterior and Posterior Segment Ocular Drug Delivery - Scientific Figure, Available from: https://www.researchgate.net/figure/Fig-1-Structure-of-human-eye_fig1_227859518

[10] ROKYTA, Richard. *Fyziologie pro bakalářská studia v medicíně, přírodovědných a tělovýchovných oborech.* Praha: ISV, 2000. Lékařství. ISBN 80-858-6645-5.

[11] YAN, Bei, Tianyi PEI and Xiaojing WANG. Wavelet Method for Automatic Detection of Eye-Movement Behaviors. *IEEE Sensors Journal* [online].

2019, 19(8), 3085-3091 [accessed 2020-10-24]. ISSN 1530-437X. Available from: doi:10.1109/JSEN.2018.2876940

[12] KOLAR, Radim, Ralf. P. TORNOW, Jan ODSTRCILIK and Ivana LIBERDOVA. Registration of retinal sequences from new video-ophthalmoscopic camera. *BioMedical Engineering OnLine* [online]. 2016, 15(1) [accessed 2020-10-25]. ISSN 1475-925X. Available from: doi:10.1186/s12938-016-0191-0

[13] LIANG, Junzhong, David R. WILLIAMS a Donald T. MILLER. Supernormal vision and high-resolution retinal imaging through adaptive optics. Journal of the Optical Society of America A [online]. 1997, 14(11), 1457-1460 [accessed 2021-5-15]. ISSN 1084-7529. Available from: doi:10.1364/JOSAA.14.002884

[14] GILL, Jasdeep S., Mariya MOOSAJEE, Adam M. DUBIS and Ivana LIBERDOVA. Cellular imaging of inherited retinal diseases using adaptive optics. *Eye* [online]. 2019, 33(11), 1683-1698 [accessed 2020-10-25]. ISSN 0950-222X. Available from: doi:10.1038/s41433-019-0474-3

[15] IQBAL, Azhar, Zhizheng WU, Foued BEN AMARA and Ivana LIBERDOVA. Closed-loop control of magnetic fluid deformable mirrors. *Optics Express* [online]. 2009, 17(21), 1683-1698 [accessed 2020-10-25]. ISSN 1094-4087. Available from: doi:10.1364/OE.17.018957

[16] BARANEC, Christoph, Reed RIDDLE, Nicholas M. LAW, et al. Bringing the Visible Universe into Focus with Robo-AO. *Journal of Visualized Experiments* [online]. 2013, (72) [accessed 2020-10-25]. ISSN 1940-087X. Available from: doi:10.3791/50021

[17] DÍAZ, José A., José FERNÁNDEZ-DORADO, Carles PIZARRO and Josep ARASA. Zernike coefficients for concentric, circular scaled pupils: an equivalent expression. *Journal of Modern Optics* [online]. 2009, 56(1), 131-137 [accessed 2020-11-18]. ISSN 0950-0340. Available from: doi:10.1080/09500340802531224

[18] BRYANT, James C. and Catherine WREATH. *Applied Optics and Optical Engineering.* 11. New York: Academic Press, 1992. ISBN 0-12-408611-X.

[19] LIANG, Junzhong, David R. WILLIAMS and Donald T. MILLER. Supernormal vision and high-resolution retinal imaging through adaptive optics. *Journal of the Optical Society of America A* [online]. 1997, 14(11) [accessed 2020-10-28]. ISSN 1084-7529. Available from: doi:10.1364/JOSAA.14.002884

[20] BURNS, Stephen A., Ann E. ELSNER, Kaitlyn A. SAPOZNIK, Raymond L. WARNER and Thomas J. GAST. Adaptive optics imaging of the human retina.

Progress in Retinal and Eye Research [online]. 2019, 68, 1-30 [accessed 2021-5-1]. ISSN 13509462. Available from: doi:10.1016/j.preteyeres.2018.08.002

[21] WALEK, Petr, Martin LAMOŠ and Jiří JAN. *Analýza biomedicínských obrazů: Počítačová cvičení.* Druhé vydání, aktualizované. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačných technologií, Ústav biomedicínskeho inženýrství, 2015. ISBN 978-80-214-4792-9.

[22] KOCH, Edouard, David ROSENBAUM, Aurélie BROLLY, José-Alain SAHEL, Philippe CHAUMET-RIFFAUD, Xavier GIRERD, Florence ROSSANT and Michel PAQUES. Morphometric analysis of small arteries in the human retina using adaptive optics imaging. *Journal of Hypertension* [online]. 2014, 32(4), 890-898 [accessed 2020-11-15]. ISSN 0263-6352. Available from: doi:10.1097/HJH.0000000000000095

[23] JAN, Jiří. *Medical image processing, reconstruction, and restoration: concepts and methods.* Boca Raton, 2006. Signal processing and communications. ISBN 08-247-5849-8.

[24] TU, Wei-Chih, Ying-An LAI a Shao-Yi CHIEN. Constant time bilateral filtering for color images. 2016 IEEE International Conference on Image Processing (ICIP) [online]. IEEE, 2016, 2016, 14(11), 3309-3313 [accessed 2021-5-18]. ISBN 978-1-4673-9961-6. ISSN 1084-7529. Available from: doi:10.1109/ICIP.2016.7532972

[25] BAO, Hua, Changhui RAO, Yudong ZHANG, Yun DAI, Xuejun RAO and Yubo FAN. Hybrid filtering and enhancement of high-resolution adaptive-optics retinal images. *Optics Letters* [online]. 2009, 34(22) [accessed 2020-11-15]. ISSN 0146-9592. Available from: doi:10.1364/OL.34.003484

[26] CHEN, Yiwei, Yi HE, Jing WANG, Wanyue LI, Lina XING, Feng GAO and Guohua SHI. Automated Optical Flow Based Registration for Adaptive Optics Scanning Laser Ophthalmoscope. *IEEE Photonics Journal* [online]. 2020, 12(2), 1-9 [accessed 2020-11-16]. ISSN 1943-0655. Available from: doi:10.1109/JPHOT.2019.2963058

[27] KOLAR, Radim, Ivana LIBERDOVA, Jan ODSTRCILIK, Michal HRACHO and Ralf P. TORNOW. *Detection of distorted frames in retinal video-sequences via machine learning* [online]. 2017-7-28, , 104130A- [accessed 2020-11-18]. Available from: doi:10.1117/12.2284172

[28] PATANAVIJIT, V, S JITAPUNKUL, A. C. BOVIK, Sumohana S. CHANNAPPAYYA and Swarup S. MEDASANI. A Lorentzian Stochastic Estimation for a

Robust Iterative Multiframe Super-Resolution Reconstruction with Lorentzian-Tikhonov Regularization. EURASIP Journal on Advances in Signal Processing [online]. IEEE, 2007, 2015, 2007(1), 4695-4708 [accessed 2021-5-15]. ISBN 978-1-4799-6619-6. ISSN 1687-6180. Available from: doi:10.1155/2007/34821

[29] FARSIU, S., M.D. ROBINSON, M. ELAD, P. MILANFAR and Swarup S. MEDASANI. Fast and Robust Multiframe Super Resolution. IEEE Transactions on Image Processing [online]. IEEE, 2004, 2015, 13(10), 1327-1344 [accessed 2021-5-15]. ISBN 978-1-4799-6619-6. ISSN 1057-7149. Available from: doi:10.1109/TIP.2004.834669

[30] SARA, Umme, Morium AKTER and Mohammad Shorif UDDIN. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. Journal of Computer and Communications [online]. 2019, 07(03), 8-18 [accessed 2021-5-1]. ISSN 2327-5219. Availaible from: doi:10.4236/jcc.2019.73002

[31] VENKATANATH N, PRANEETH D, MARUTHI CHANDRASEKHAR BH, Sumohana S. CHANNAPPAYYA and Swarup S. MEDASANI. Blind image quality evaluation using perception based features. 2015 Twenty First National Conference on Communications (NCC) [online]. IEEE, 2015, 2015, , 1-6 [accessed 2021-5-15]. ISBN 978-1-4799-6619-6. Available from: doi:10.1109/NCC.2015.7084843

[32] MITTAL, A., R. SOUNDARARAJAN, A. C. BOVIK, Sumohana S. CHANNAPPAYYA and Swarup S. MEDASANI. Making a "Completely Blind" Image Quality Analyzer. IEEE Signal Processing Letters [online]. IEEE, 2013, 2015, 20(3), 209-212 [accessed 2021-5-15]. ISBN 978-1-4799-6619-6. ISSN 1070-9908. Available from: doi:10.1109/LSP.2012.2227726

[33] MITTAL, A., A. K. MOORTHY, A. C. BOVIK, Sumohana S. CHANNAPPAYYA and Swarup S. MEDASANI. No-Reference Image Quality Assessment in the Spatial Domain. IEEE Transactions on Image Processing [online]. IEEE, 2012, 2015, 21(12), 4695-4708 [accessed 2021-5-15]. ISBN 978-1-4799-6619-6. ISSN 1057-7149. Available from: doi:10.1109/TIP.2012.2214050

[34] Multi-Frame Super-Resolution Toolbox [online]. [accessed 2021-5-17]. Available from: https://www5.cs.fau.de/research/software/multi-frame-super-resolution-toolbox/