UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Adaptive fusion of bright-field microscopy images acquired in different focal planes

Victor Augusto Alves Catanante

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

ICMC USP
SÃO CARLOS

Victor Augusto Alves Catanante

Adaptive fusion of bright-field microscopy images acquired in different focal planes

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. João do Espírito Santo Batista Neto
Co-advisor: Prof. Dr. Odemir Martinez Bruno

USP – São Carlos
September 2020

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

C357a

Catanante, Victor Augusto Alves
Adaptive fusion of bright-field microscopy images acquired indifferent focal planes / Victor Augusto Alves Catanante; orientador João do Espírito Santo Batista Neto; coorientador Odemir Martinez Bruno. -- São Carlos, 2020.
129 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2020.

1. No-reference Image Quality Assessment. 2. Discrete Fourier Transform. 3. Bright-field Microscopy. 4. Multi-focus Image Fusion. 5. Laplacian of Gaussian. I. Batista Neto, João do Espírito Santo, orient. II. Bruno, Odemir Martinez, coorient. III. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Victor Augusto Alves Catanante

Fusão adaptativa em imagens de microscopia de campo claro adquiridas em diferentes planos focais

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. João do Espírito Santo Batista Neto
Coorientador: Prof. Dr. Odemir Martinez Bruno

USP – São Carlos
Setembro de 2020

In memory of Nazir Alves.

ACKNOWLEDGEMENTS

I would like to thank professor João do Espírito Santo Batista Neto for the ideas, support, and patience during the development of this work. I also thank professor Odemir Martinez Bruno for the opportunity to work in the bright-field microscopy lab and for the several important remarks and suggestions about the project. I am deeply indebted to professors Moacir Ponti and Marcelo Vieira, whose assistance and guidance were milestones in the completion of this project. *Indeed*, I thank professor Éder Ritis for helping me to develop the mathematical reasoning part of this work and also for the whole-heartedly friendship. My respectful regards to professor Marina Andretta, for sharing your teaching experience and being an extremely supportive and solicitous person. I am also grateful for all USP professors that shared a bit of their broad knowledge. I would like to recognize the invaluable assistance that Isabella de Cassia Lima Munhoz and Marcelo Barbosa de Andrade provided in teaching me how to deal with microscopes.

I owe my deepest gratitude to Ievgeniia Iermak for being by my side during this, supporting my evolution and even helping with optics and microscopy. I acknowledge my parents, who were always there for me since the beginning and offer unconditional support. I am also grateful to all my friends for the irreplaceable good times we had together. No names here just to keep it fair!

I thank Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (CNPq), process 132795/2018-3 for the financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

“Sempre la pratica dev’essere edificata sopra la buona teorica, della quale la prospettiva è guida e porta, e senza questa nulla si fa bene.”

(Leonardo da Vinci)

RESUMO

CATANANTE, V. A. A. **Fusão adaptativa em imagens de microscopia de campo claro adquiridas em diferentes planos focais**. 2020. 129 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

A microscopia é uma técnica extremamente relevante relacionada a tarefas que lidam com estruturas de ordem micrométrica. Seu uso remonta ao século XVII e tende a avançar paralelamente à evolução conhecimento tecnológico humano. Dentre as diversas aplicações, destacam-se as áreas de ciências biológicas e da saúde, que envolvem estruturas normalmente invisíveis a olho nu. Existem diferenças de profundidade inevitáveis entre os pontos das superfícies e estruturas que produzem desfoque nas imagens; no entanto, é necessário que tais images possuam alta qualidade para análises precisas em aplicações de microscopia. Neste aspecto, a avaliação da qualidade da imagem e a fusão de imagens são exemplos de técnicas que podem ser aplicadas para resolver o problema. Trabalhos recentes em tais áreas mostram que técnicas matemáticas como análise no domínio da frequência, análise multirresolução e redes neurais convolucionais são eficazes para avaliar quantitativamente a qualidade das imagens; paralelamente, os pesquisadores também apresentam muitas técnicas inovadoras para fusão de imagens baseadas em ferramentas clássicas como detecção de bordas ou em estruturas de aprendizado de máquina de última geração. O objetivo deste trabalho é desenvolver um método de duas etapas - uma etapa de avaliação da qualidade da imagem sem referência e uma etapa de fusão da imagem, para realizar a fusão de imagens de microscopia de luz de campo claro adquiridas em diferentes planos focais, além de propor novos conjuntos de dados de imagens de microscopia de campo claro de amostras histológicas de folhas de plantas como referência para testar os algoritmos de avaliação da qualidade e fusão. Análise no domínio da frequência e métodos estatísticos foram utilzados para obter uma métrica de qualidade, e a energia das arestas extraídas com o filtro Laplaciano da Gaussiana foi utilizada como regra de fusão. O coeficiente de correlação de Pearson médio obtido para o método de qualidade de imagem foi de 0.7448, e a frequência espacial média para o método de fusão de imagens foi de 0.0667.

Palavras-chave: Avaliação de qualidade de imagem sem referência, Transformada Discreta de Fourier, Microscopia de Campo Claro, Fusão de Imagens Multifocais, Laplaciano da Gaussiana.

ABSTRACT

CATANANTE, V. A. A. **Adaptive fusion of bright-field microscopy images acquired in different focal planes**. 2020. 129 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Microscopy is an extremely relevant technique related to tasks that deal with micrometric order structures. Its use dates back to the 17th century and tends to evolve in parallel with the evolution of human technological knowledge. Among the various applications, the fields of biological and health sciences stand out, which involve structures normally invisible to the naked eye. There are unavoidable differences in depth between the points of the surfaces and structures which yield out-of-focus blur to images. However, high quality is necessary in order to allow precise analysis in microscopy applications. In this sense, image quality assessment and image fusion are examples of techniques that may be applied to solve the issue. Recent works on such fields show that mathematical techniques such as frequency domain analysis, multiresolution analysis and convolutional neural networks are effective to quantitatively assess the quality of images. At the same time, researchers also present many novel techniques for image fusion, either based on classical tools such as edge detection or based on state-of-the-art machine learning frameworks. The aim of this work is to develop a two-stage method, consisting of a no-reference image quality assessment and an image fusion step, to perform the fusion of bright-field light microscopy images acquired in different focal planes, and propose novel bright-field microscopy image datasets of plant leaf histological samples as a benchmark for testing both quality assessment and fusion algorithms. Frequency domain analysis and statistical methods were used to obtain a quality metric and the energy of edges extracted with the Laplacian of Gaussian filter as the fusion rule. The mean Pearson's correlation coefficient obtained for the image quality method was 0.7448, while the mean spatial frequency for the image fusion method was 0.0667.

Keywords: No-reference Image Quality Assessment, Discrete Fourier Transform, Bright-field Microscopy, Multi-focus Image Fusion, Laplacian of Gaussian.

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AHE	Adaptive Histogram Equalization
CCD	Charge-coupled Devices
CLAHE	Contrast Limited Adaptive Histogram Equalization
CMO	Common Main Objective Stereo Microscope
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
FFT	Fast Fourier Transform
FR-IQA	Full-Reference Image Quality Assessment
FR-IQA	Full-Reference Image Quality Assessment
FT	Fourier Transform
GGD	Generalized Gaussian Distribution
IQA	Image Quality Assessment
IQA	Image Quality Assessment
MLV	Maximum Local Variation
MOS	Mean Opinion Scores
MOS	Mean Opinion Scores
MRI	Magnetic Resonance Imaging
MSSIM	Mean Structural Similarity Measure
MSSIM	Mean Structural Similarity Measure
NR-IQA	No-Reference Image Quality Assessment
NR-IQA	No-Reference Image Quality Assessment
OTF	Optical Transfer Function
p.d.f	probability density function
PCA	Principal Component Analysis
PSF	Point Spread Function
RANSAC	Random Sample Consensus
RGB	Red, Green and Blue
RR-IQA	Reduced-Reference Image Quality Assessment
RR-IQA	Reduced-Reference Image Quality Assessment
SIFT	Scale Invariant Feature Transform
TV	Total Variation

LIST OF SYMBOLS

$f(x,y)$ — Original image as a function of spatial coordinates x and y

$g(x,y)$ — Observed image as a function of spatial coordinates x and y

J_1 — Bessel function of the first kind

$*$ — Convolution

CONTENTS

CHAPTER

1

INTRODUCTION

The microscope is a device that performs extremely important tasks to human knowledge, in theoretical or empirical aspects. It is capable of providing magnified views of small objects and structures. Currently, several variations of microscopes are used to investigate much smaller spaces than those visible to the naked eye (WU; MERCHANT; CASTLEMAN, 2008). Fields such as materials sciences and biology broadly apply microscopy and health professionals use them on large scale for practical procedures, clinical analysis, and research. High throughput microscopy is an important technique for the diagnosis and treatment of genetic diseases. However, to make it acceptable in the clinical environment, it is of great importance to perform high-resolution image acquisition, since low levels of sharpness can directly affect the diagnostic accuracy (QIU *et al.*, 2013).

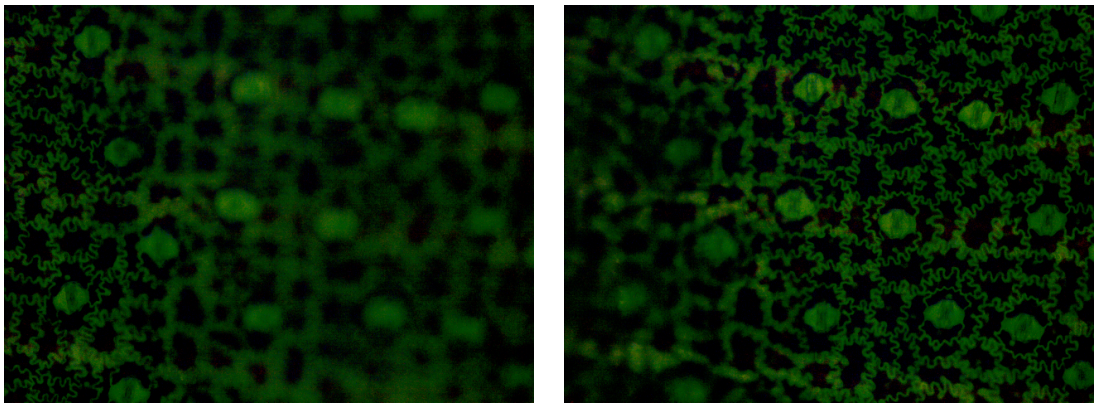
The current advances in microscopy technologies and methods show a natural trend of linking novel microscopy improvements and image processing. This bond dates back to the middle of the 20th century, when some techniques for capturing and manipulating images, primarily developed for televisions, were applied to microscopy images (WU; MERCHANT; CASTLEMAN, 2008). A classic example is noise reduction, which is an important step for cryoelectronic microscopy and also for energy filtering in transmission electron microscopy, before the 3D reconstruction process on Computed Tomography scans. High noise levels hinder the necessary alignment in the reconstruction task (VYAS; YU; PAIK, 2017).

1.1 Motivation

Biological and biomedical analysis procedures using microscopy images also employ image processing algorithms to produce better results. In this scenario, the concept of focus is an element of great relevance. The microscopically analyzed surfaces and structures are *a priori* smooth and homogeneous to the naked eye; when magnified, these images show that those elements are irregular, i.e. they have different depths (when considering an upper view),

textures and topologies. It is, therefore, necessary to constantly adjust the focus to obtain a sharp image. Figure 1 illustrates the problem of differences in depth of focus in histological images of a *Ctenanthe oppenheimiana* specimen with the structure of a slope, acquired in the same scene but with different height settings for the objective lenses. The axial location for Figure 1.(a) was higher than in Figure 1.(b), and this produced clear differences among the blurred and sharp regions of both images.

Figure 1 – Examples of *Cthenante oppenheimiana* partially blurred images, with the leftmost (a) and rightmost parts (b) as sharp due to topological height differences.



(a) (b)

Source: Elaborated by the author.

Sharpness is a concern when it comes to the analysis of microscopy images in order to obtain conclusions. According to Costa *et al.* (2019), blurred regions in sputum smear microscopy images caused by the depth of field limit in the microscope affect the accuracy of bacilli detection. Several proposed methods address the sharpness problem in microscopy images based on image restoration techniques. As stated by Ponti, Nazaré and Thumé (2016), the most frequently used iterative method in microscopic image restoration is the Richardson-Lucy algorithm. Some examples presented by SUN, DUTHALER and NELSON (2005) consist of four classes: *derivative algorithms*, *statistical algorithms*, *histogram-based algorithms* and *intuitive algorithms*. Among these applications, derivative methods deserve more credit. The Fourier Transform proved itself to be effective for low or moderate noise levels; in highly noisy environments, the resulting images were not satisfactory (RICHARDSON, 1972). As a consequence, probabilistic methods based on the Bayes Theorem were developed and provided images with better contrast, higher bandwidth and edge enhancement for confocal fluorescence microscopy samples (PONTI; NAZARÉ; THUMÉ, 2016).

The resulting images from such restoration processes are sharper than the observed ones. However, the algorithms produce degradation as output. An alternative to restoration is to use images from the same object, with different foci, in order to obtain an enhanced depth of field image with low degradation levels. This process is known as image fusion, and even though a significant effort was done by researchers towards the development of techniques in this context,

most of the image fusion methods are applicable but not directly related and built for bright-field microscopy images. The fusion process also require the selection of images that have ate least some sharp region, and according to Koho *et al.* (2016), very few publications that address the problems of measuring the quality of images can be found on microscopy applications. There are many image analysis tools, but not many easily accessible and applicable image quality assessment methods.

1.2 Aims and Hypothesis

This work aims to develop a method to perform the fusion of bright-field microscopy images acquired in different focal planes. We propose three stages to achieve this, described as follows:

- *Dataset acquisition*: Acquisiton of a bright-field microscopy image dataset of leaf samples to evaluate the performance of the methods;
- *No-reference image quality assessment*: Development of a method to quantitatively assess the quality of bright-field microscopy images based on the Fourier transform and descriptive statistics;
- *Multi-focus image fusion*: Fusion of the sharp regions of selected images among the dataset by means of a Laplacian of Gaussian-based method and compose the sharp image.

It is hypothesized that frequency domain information from images may be used to quantify the sharpness of bright-field microscopy images. Simultaneously, we also have the hypothesis that image fusion methods that employ edge detection by means of the Laplacian operator and its derivatives, e.g. the Laplacian of Gaussian, perform well on bright-field microscopy images.

1.3 Contributions

This work yields the bright-field microscopy z-stack datasets as a first contribution. Those may be used to study, develop and test novel no-reference image quality metrics and multi-focus image fusion techniques by image processing and microscopy researchers, as well as industry professionals. The datasets comprise real world high-resolution microscopy images of biological material (particularly, plant leaves) with different illumination techniques and settings, which turns image quality assessment and image fusion into more challenging tasks.

In addition to the datasets, our proposed quality index and fusion algorithms explore well-known mathematical analysis, image processing, signal processing and statistical techniques. This not only it provides significant applications of those tools, but also stimulates studies among

researchers in order to combine new state-of-the-art technologies to strong theoretical concepts in order to develop robust and scalable solutions.

Structure of the document

This monograph is organized as follows:

- Chapter 2 provides the bright-field microscopy basics and the z-stacking technique, i.e. the method to acquire images in different focal planes;
- Chapter 3 comprises relevant information about the bright-field microscopy image formation process and also describes the defocus blur property;
- Chapter 4 provides the theoretical basis of this work, i.e. Fourier transform, image enhancement, registration, fusion and quality assessment, as well as the statistical methods employed in the analysis;
- Chapter 5 presents methods in which this work is based on, by means of a literature review on no-reference image quality assessment and multi-focus image fusion;
- Chapter 6 presents the acquisition and usage protocols of the proposed image datasets and the proposed method;
- Chapter 7 presents our results with the proposed datasets and exposes a discussion concerning the quantitative and qualitative analysis of them;
- Chapter 8 summarizes what has been achieved, suggests some future work on the field and improvements to the methods;
- Appendix A presents fundamentals of optics that may aid the comprehension of bright-field microscopy;
- Appendix B provides details about some of the background concepts;
- Appendix C is the result of mathematical reasoning to prove some properties concerning the method which were hypothesized true.

CHAPTER

2

BRIGHT-FIELD MICROSCOPY FUNDAMENTALS

Microscopes are instruments designed to accomplish several important tasks to human knowledge, capable of magnifying images of small objects and structures which otherwise would not be seen by the human eye. It grants more information about the object of study to the research or the analysis. The first idea of the device was introduced by Romans, which discovered the magnifying property of glass in some sort of biconvex shape. Zacharias Janssen (1588-1632) was responsible for the invention of the first compound microscope with a concave eyepiece and Francisco Fontana (1580-1656) introduced the convex eyepiece version of it (ZÍLIO, 2009). Furthermore, Robert Hooke (1635-1703) and Anton van Leeuwenhoek (1632-1723) were the most prominent science-related men responsible for microscope improvements (WU; MERCHANT; CASTLEMAN, 2008). Due to the development of theories and their empirical proofs of veracity, in addition to the advances in hardware and software power, techniques such as image processing are applied to other fields. This also happens in microscopy, aiming to improve image quality, data reliability, and range of use (BOYDE, 1990).

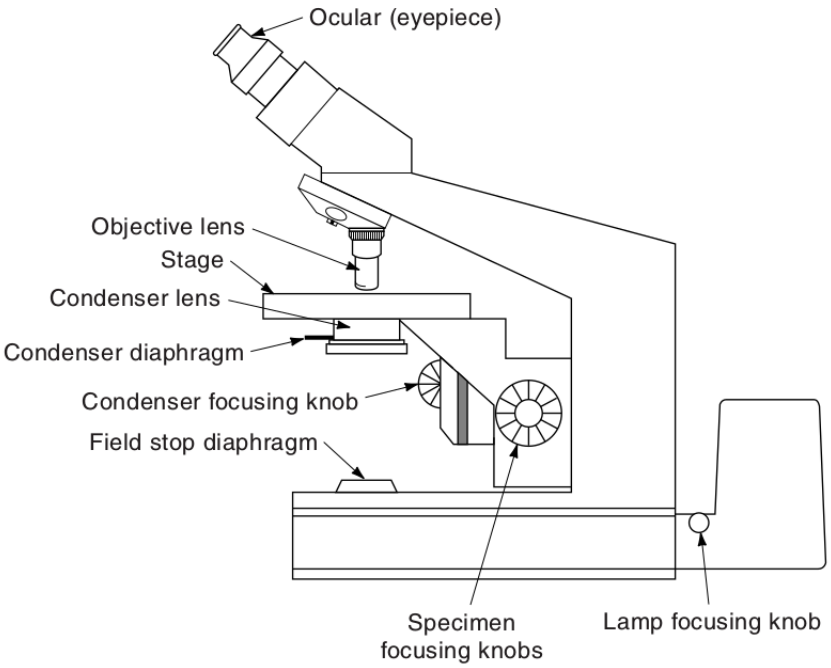
This chapter provides information about bright-field and microscopy concepts in this work, describing the structure of the optical microscope, along with its uses and implications on the acquired images.

2.1 Light microscopy

Historically, the structure of the first microscopes developed by Hooke and Leeuwenhoek had no eyepiece; the compound microscope, developed by Janssen, consists of the magnification lenses and also an eyepiece, which adds more magnification power and delivers the image to the user (LAWLOR, 2019). As stated by Murphy and Davidson (2012), the word *compound* refers to the fact that the objective lens and the eyepiece (or ocular) work together to produce the final

magnification of the image as a product of their magnifications. A graphical representation of the general structure of a compound microscope is shown in Figure 2, and as reported by Bell and Morris (2009), it consists of objective lenses, eyepieces, condensers, the stage and the light source.

Figure 2 – Structure of the basic compound microscope: lenses that capture light rays from the specimen or object (objectives) or where the observer may look through (eyepieces), a collector of light from the light source (condenser), the support for the object (stage) and the light source.



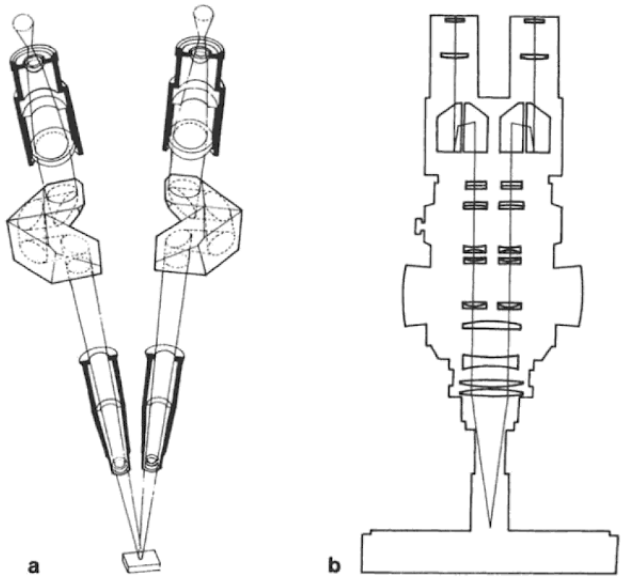
Source: Bell and Morris (2009).

Since light is some sort of radiation, there are several different ways to achieve imaging in microscopes; it can be made by light, polarized light, lasers, X-rays, among others. There are also advanced techniques such as confocal microscopy, which is capable of imaging a very small area of the object, with all the light rays focused on it (ROCHOW; TUCKER, 1994). The choice of the most suitable microscopy depends on the task.

According to Dokland *et al.* (2006), the purpose of light microscopy is to provide magnified images of specimens by means of capturing emitted, reflected or transmitted light in the visible range of the spectrum, or even in the ultraviolet or near-infrared regions. As reported by Lawlor (2019), there are three general styles of light microscope. The *upright microscope* is the easily affordable, easy to use traditional configuration with a light source on the base, a stage and the objective lenses. The *inverted microscope* consists of the inverted configuration of an upright microscope, and offers some advantages when it comes to life sciences applications

such as live cell imaging. Finally, the *stereomicroscope* consists of a fusion of two compound microscopes in a convergent optical system and may have two different objectives and eyepieces or only one objective and two eyepieces (SCHREIER; GARCIA; SUTTON, 2004). The former is named *binobjective-binocular* (Greenough) and the latter *monobjective-binocular*, or Common Main Objective Stereo Microscope (CMO). One of the advantages of CMO microscopes is the higher depth of field, which allows the user to view and investigate biological specimens, relatively small materials and any kind of non-smooth surfaces. Furthermore, it is possible to view and acquire images in three dimensions (ROCHOW; TUCKER, 1994). The structure of both types of stereo compound microscopes is depicted in Figure 3.

Figure 3 – Graphic representation of the basic stereo compound light microscope structure, (a) for the Greenough type and (b) for the CMO.



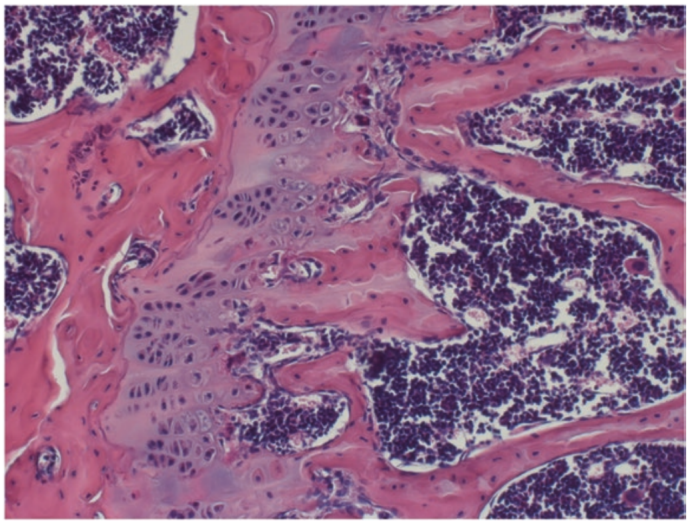
Source: Rochow and Tucker (1994).

2.1.1 Bright-field microscopy

As for the use of light microscopes, there are several techniques and configurations which are achieved by varying the amount of lenses and light sources: bright-field, dark-field, phase-contrast, differential interference and fluorescence (ROANE; PEPPER; MAIER, 2009). As stated by Lawlor (2019), images in bright-field microscopy are characterized by the contrast between the sample and the bright white background, generated by transmitted light. It is commonly used in pathology and histology fields for imaging fixed cells and tissues to reveal their structure, shape, and organization. The amount of light should be controlled, since the sample might suffer substantial changes, e.g. the chlorophyll molecules when illuminated by UV and visible light suffer irreversible breakdown (photodegradation) and generate other photoproducts (PETROVIĆ;

ZVEZDANOVIĆ; MARKOVIĆ, 2017). The Figure 4 provides an example of bright-field microscopy image.

Figure 4 – Example of bright-field microscopy image of bone tissue.

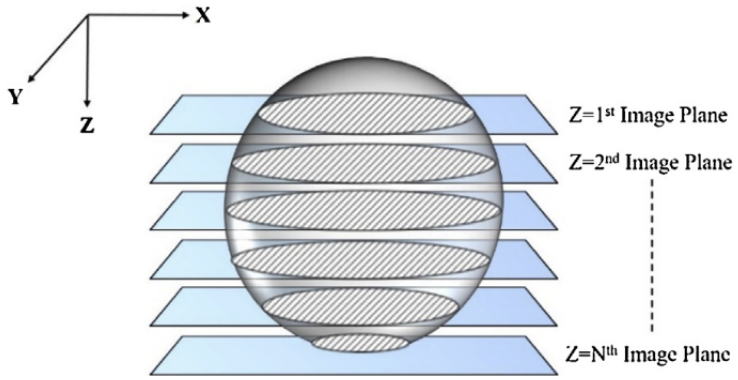


Source: Lawlor (2019).

2.1.2 Z-stacking technique

The *z-stacking* is a procedure to capture images in different positions concerning the *z* axis, named slices, which may create a pseudo 3D image of the sample and consequently retrieve depth information about the specimen (LAWLOR, 2019). Figure 5 presents the scheme of a z-stack acquisition with an arbitrary object:

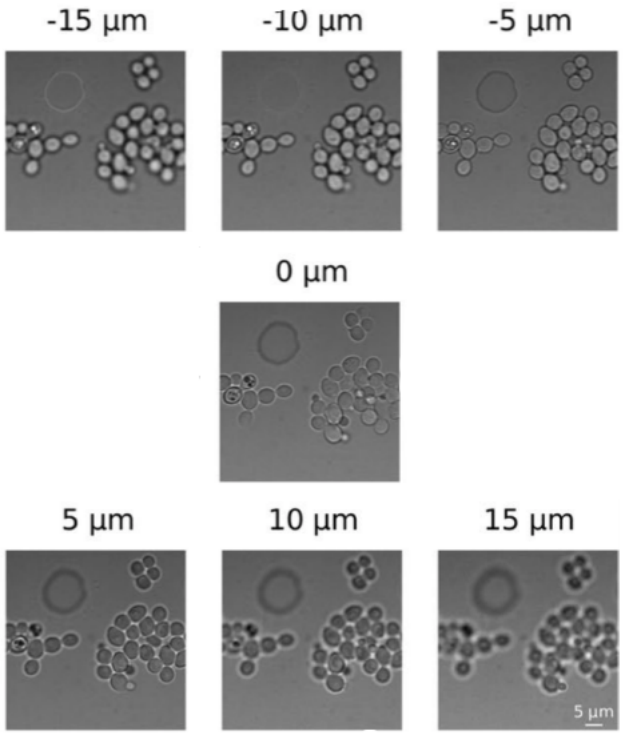
Figure 5 – Scheme of a z-stack image dataset acquisition with an arbritrary object.



Source: Trivedi and Mills (2020).

Objective lenses with large magnifications can only focus on a thin slice of space, which implies that the acquisition of a entirely sharp image of a thick sample will be impossible; the different structures are consequently sharp in different focal planes (GIUSTI *et al.*, 2011). Both the visualization and the acquisition of microscopy images with the z-stacking technique can be done by performing progressive adjustments of the focus knob of the microscope or by motorized stages or objectives, capable of changing the focal plane with more precision. Each focal plane adjustment is done inwards or outwards along the z axis, and the set of acquired slices contain a focused region of the sample. The distance between each slice is dictated manually or automatically. Next, the images must be aligned before any analysis is conducted. Figure 6 represents the differences in focus between the slices of a z-stack.

Figure 6 – Z-stack images of yeast cells, acquired in positions under the focal plane (-15, -10 and -5 μm), exactly on it (0 μm) and above it (5, 10 and 15 μm).



Source: Wei and Roberts (2018).

CHAPTER

3

IMAGE FORMATION AND DEFOCUS BLUR

The human eye constructs images from incident light rays on the *retina*, a very complex set of photoreceptors that converts light into electrical signals which are later interpreted by the brain. The result of this process may be modelled as continuous function of two variables $f(x,y)$ which comprises the *illumination* and the *reflectance* information, i.e. the amount of incident light and the amount of reflected light in the scene, respectively (GONZALEZ; WOODS, 2018).

Digital image processing deals with, as the term suggests, *digital images*, i.e. discrete representations of $f(x,y)$ generated by sensors that transform the illumination and reflectance information into electrical signals. Still according to Gonzalez and Woods (2018), in order to achieve this representation, the signals undergo sampling (signal conversion from continuous to discrete) and quantization (mapping of real-valued intensities to discrete pixel values). For the sake of notation simplicity, $f(x,y)$ denotes the digital image and the term "image" also refer to it throughout this work.

It is clear that the image formation is influenced by several factors: sensor type, scene illumination conditions, and others. In fact, each imaging system such as a camera or a microscope adds its own constraints to the process, e.g. conventional transmitted light microscopy images are only achieved with non-opaque samples (ROTTENFUSSER; WILSON; DAVIDSON, 2020). Although there are many types of microscopy, each with its own imaging procedure, this work limits its scope to bright-field microscopy. Therefore, this chapter summarizes the bright-field microscopy image formation processes and its implications on image quality. Furthermore, it describes blur properties concerning its origins either in image formation or other events.

3.1 Image Formation

As mentioned in Chapter 2, the bright-field microscopy images are formed by either transmitted or reflected light that passes through the sample and reaches the objective lens. The

difference between transmitted light and reflected light microscopes is the illumination system; there is no difference in how both direct light rays leave the specimen (LENG, 2009).

According to Davidson and Abramowitz (2002), the light which reaches the specimen is either undeviated, i.e. does not suffer any disturbances in its direction, or diffracted; the diffracted rays leave the sample with a phase difference in comparison to the undeviated light and cause destructive interference in the eyepiece, which projects a magnified version of this pattern onto a sensor and consequently produces the image.

Furthermore, the diffraction patterns that are captured by objectives have a particular shape. As also stated by Davidson and Abramowitz (2002), the *Airy disks* (also called *Airy patterns*), named after Sir George Biddell Airy (1801 - 1892), are small circular diffraction disks projected by the objectives onto the image plane of the eyepiece diaphragm, which describe the focus profile of the resulting image. The Airy disks, as described by Fowles (1989), follow the Fraunhoffer diffraction pattern, and may be mathematically modelled as an angular distribution of intensity of light diffracted by a circular aperture, given by

$$I(\theta) = I_0 \left[\frac{2J_1(\rho)}{\rho} \right]^2 \qquad \rho = \left(\frac{2\pi \sin\theta}{\lambda} \right) \frac{a}{2}, \tag{3.1}$$

where $I_0 = (C\pi R^2)^2$ is the intensity for $\theta = 0$, C is a constant, R is the radius of the aperture, λ is the wavelength of the light, a is the diameter of the aperture and J_1 is the Bessel function of the first kind and first order (MATHEWS; WALKER, 1970). The Bessel function for the general case of r^{th} order is given by

$$J_r(x) = \sum_{n=0}^{\infty} \frac{(-1)^r}{r!\Gamma(m+r+1)} \left(\frac{x}{2} \right)^{m+2r} \qquad \Gamma(z) = \int_0^{\infty} e^{-u} u^{z-1} du. \tag{3.2}$$

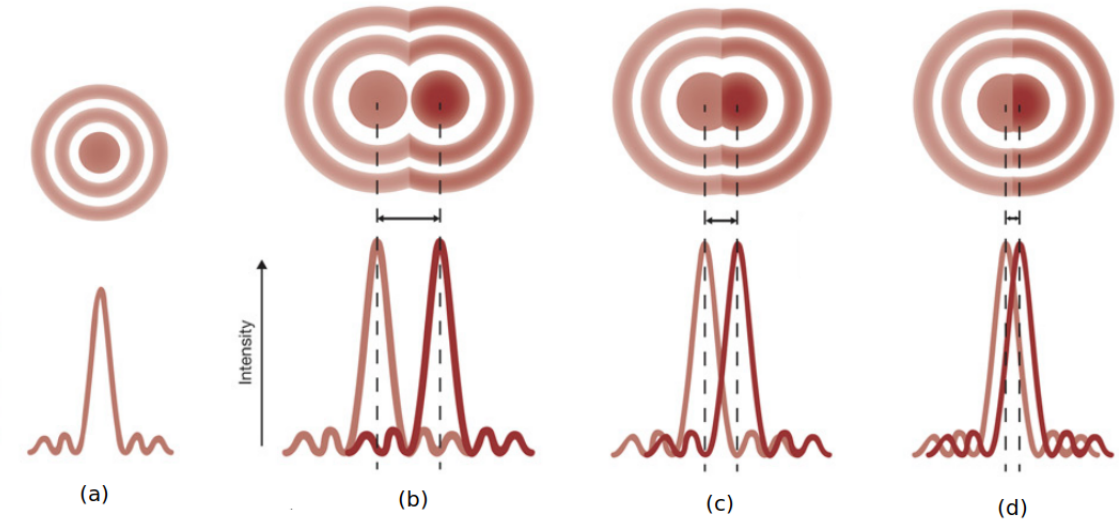
The Airy disks are intrinsically related to the numerical aperture and the definition of *resolution*. The resolution is the minimum distance between two points at which they can be visibly distinguished as two points; optically, it is defined as the minimum distance between two Airy disks that can be distinguished, which is limited by diffraction (LENG, 2009). The resolution of an optical microscope, given by the Rayleigh equation, is described by

$$d = 1.22 \frac{\lambda}{2NA}, \tag{3.3}$$

where d is the space between two adjacent particles that may be distinguished from each other, λ is the wavelength of the illumination and NA is the numerical aperture of the objective (DAVIDSON; ABRAMOWITZ, 2002). It is evident that objectives with higher numerical apertures and shorter wavelengths of visible light will yield better resolution. Figure 7 shows arbitrary examples of Airy patterns, as well as their possible configurations and their consequences to

the image. In Figure 7.(a), the usual shape of Airy patterns is shown, together with its two-dimensional representation as a function of the intensity by an interval. Figure 7.(b) depicts an occurrence of Airy disk overlapping where both points would be properly resolved, i.e. below the Rayleigh limit, and Figure 7.(c) represents the minimum distance in which both points would be distinguished. Finally, Figure 7.(d) represents an unresolved pair of points.

Figure 7 – Arbitrary example of an Airy disk (a), resolved Airy disks (b), Rayleigh limit of resolution (c) and unresolved Airy disks (d).



Source: Adapted from Dunst and Tomancak (2019).

As explained by Goodman (1996), an imaging system, particularly a set of microscope lenses, is said to be *diffraction-limited* if the incident spherical light wave generated from a point-source object is transformed into another spherical wave which converges to an ideal image point, described by the original object point and affected by some sort of isotropic effect, such as magnification in a microscope.

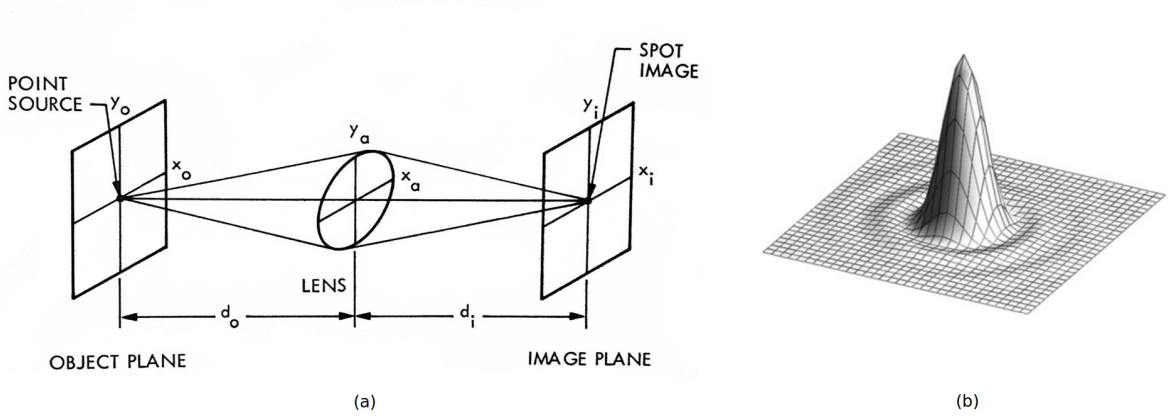
The depth of field was described in Chapter 2 in terms of focal plane distances, but it may also be taken as the *axial resolving power*, a measurement of resolution along the z axis, determined by the numerical aperture and described by the Airy disk profile (DAVIDSON; ABRAMOWITZ, 2002). Similarly to the Rayleigh's equation, the depth of field increases with higher numerical apertures for the objective and shorter wavelengths of the incident light, and it represents a key property concerning the amount of blur in the resulting image.

3.1.1 Point Spread Function and Image Formation Model

When light waves from a point source reach the lenses, they suffer diffraction and refraction, originating a new propagating set of rays that converge to a point in the center of the

image plane in the shape of Airy disks; such shape is called the Point Spread Function (PSF) of the imaging system (also called *impulse response*), and it is intrinsically related to the imaging process (WU; MERCHANT; CASTLEMAN, 2008). Particularly, the bright-field microscopy employs polychromatic nonpolarized incoherent light. Hence, it is possible to relate the Airy disks to the PSF, since those are intensity distributions for each point source of light emanating from the specimen. Figure 8.(a) presents a theoretical scheme of imaging for a point source of light, and Figure 8.(b) depicts the shape of a incoherent PSF.

Figure 8 – Light rays from a point source go through a lens' surface and converge to a point in the image plane (a) and the scheme of a point spread function generated by a focused diffraction-limited system with incoherent light (b).



Source: Adapted from Castleman (1996), WU, MERCHANT and CASTLEMAN (2008).

In Figure 8.(a), the imaging system in in focus, which is given by

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}, \tag{3.4}$$

where f is the focal length of the lens, d_o and d_i are the distances from the point source plane to the lens and the distance from the image plane to the lens, respectively. The intensity of light in the point source is directly proportional to the intensity in the image, what characterizes a *two-dimensional linear system* (CASTLEMAN, 1996). Also according to Castleman (1996), any motion of the point source on its plane moves the image is dictated by the law

$$x_i = -\frac{d_i}{d_o} x_o \qquad\qquad y_i = -\frac{d_i}{d_o} y_o, \tag{3.5}$$

where (x_o, y_o) are the coordinates for the object location on its plane and (x_i, y_i) are coordinates that locate the image on its plane. This implies that the shape of the image will not change according to the object's location, and this property yields *shift invariance* to the system, which

may be called *isoplanatic*. These properties are observed in an ideal imaging system, not in real cases such as an optical microsope. Simple lenses are neither isoplanatic nor linear, however, there are approximations and mathematical tools that allow advanced microscopes to be assumed isoplanatic and linear.

The PSF, as related to the intensity distributions described by the Airy disks, are limited to the area of the aperture. This means that the amount of light that reaches the image plane is truncated by the circular aperture, what is also true for the PSF. The truncation is mathematically represented by the *pupil function*, which is zero outside the boundaries of the aperture and unity otherwise, and might include also information about wave aberrations of the lens (GOODMAN, 1996). As denoted by WU, MERCHANT and CASTLEMAN (2008) with some notation adjustments, the PSF of an incoherent illuminated circular aperture imaging system is the Fourier Transform (explained further in Chapter 4) of the generalized pupil function, given by

$$h_\lambda(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(u,v)e^{j2\pi z\left(\frac{u^2+v^2}{2\lambda L^2}\right)}e^{j2\pi\left(\frac{xu+yv}{\lambda L}\right)}dudv, \tag{3.6}$$

where $h_\lambda(x,y,z)$ is point spread function for a light with wavelength λ, $P(x,y)$ is the pupil function, z is the axial location the focal plane and $L = r/NA$ is the focal length, i.e. ratio between the radius of the circular aperture of the objective and the numerical aperture. The normalized Fourier Transform of the PSF is called Optical Transfer Function (OTF) (CASTLEMAN, 1996).

The image is then formed as a set of impulse responses from each point in the object plane that were magnified by the imaging system. Linear systems possess a general expression, a convolution (which will be explained in Chapter 4) of the input with the system's impulse response, that describes the output (BRIGHAM, 1988). In this sense, the resulting image is a convolution of the PSF with the original image, defined as

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x-u,y-v)f(x,y)dudv, \tag{3.7}$$

where $f(x,y)$ is the original image, $g(x,y)$ is the observed image, $h(x,y)$ is the PSF of the imaging system and u,v are shift parameters.

3.1.2 Discrete Image Formation Model

Digital images follow a discrete model for image formation due to the acquisition process: the spherical waves that leave the objectives reach the surface of Charge-coupled Devices (CCD), sensors which proportionally convert light intensities to electrical signals digitized as pixels (GONZALEZ; WOODS, 2018). The digital images are matrices of pixels that represent light intensities with different channel configurations, where the most common one is the Red, Green and Blue (RGB) image. Therefore, similarly to the image formation model shown in

subsection 3.1.1, the two-dimensional digital image formation is arbitrarily described as a discrete process

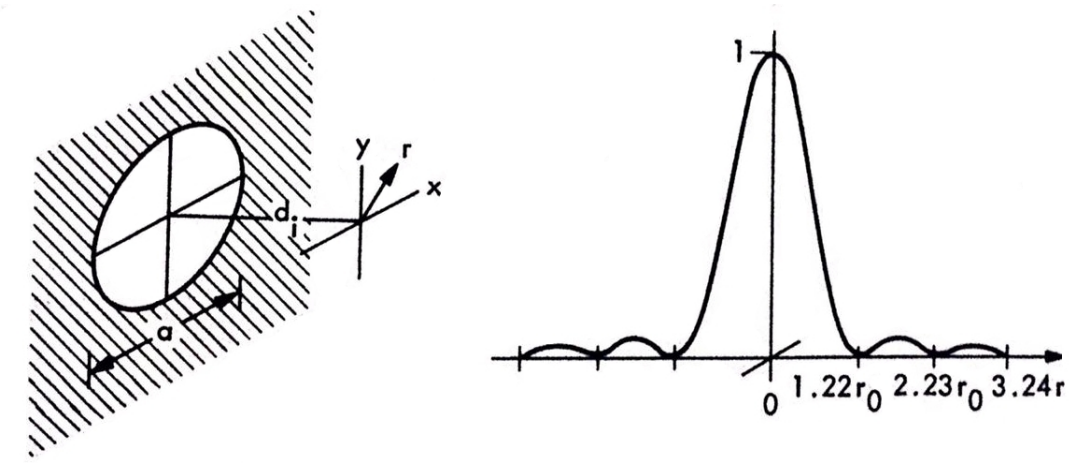
$$g[x,y] = h[x,y] * f[x,y], \tag{3.8}$$

where $*$ denotes the discrete convolution, g, h and f are respectively the observed image, the discrete PSF of the imaging system and the original image, and $x, y \in \mathbb{Z}$. The discrete PSF of an imaging system with incoherent illumination and a circular aperture is given by

$$h(r) = \left[2\frac{J_1[\pi(r/r_0)]}{\pi(r/r_0)}\right]^2, \qquad r = \sqrt{x^2+y^2}, \qquad r_0 = \frac{\lambda d_i}{a}, \tag{3.9}$$

where $h(r)$ is the radially symmetrical PSF, r is the radial distance, r_0 is a scaling factor, J_1 is the Bessel function of first order and first kind, λ is the wavelength of the illumination, a is the diameter of the aperture and d_i is the distance from the lens plane to the image plane. A scheme of the geometric setup of discrete image formation through the PSF is shown in Figure 9.

Figure 9 – Geometric scheme of a lens' circular aperture and arbitrary point spread function profile.



Source: Adapted from Castleman (1996).

The discrete OTF is then the Discrete Fourier Transform (DFT) (explained in Chapter 4) of the PSF in Equation 3.9. The OTF characterizes the intensities of light that emanate from the specimen in terms of frequencies.

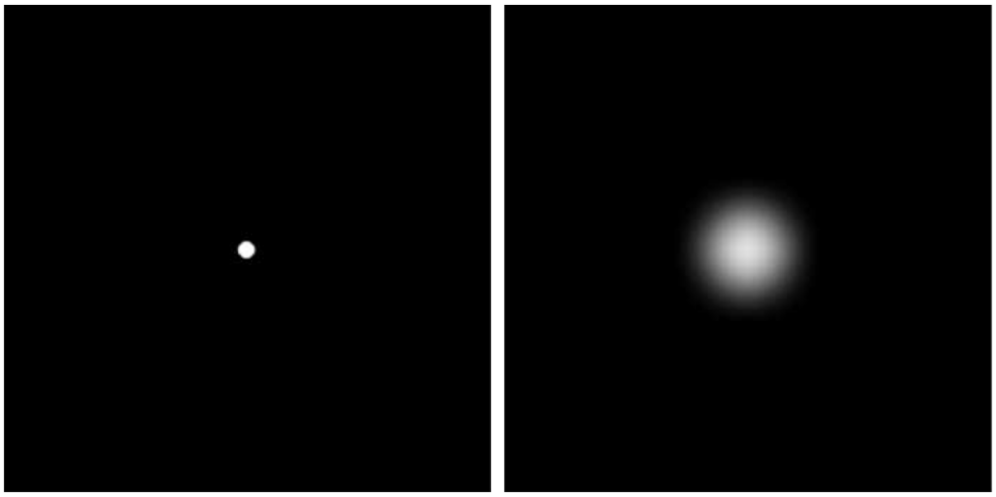
There is another property of image formation and acquisition that influences the image quality: the noise. Pursuant to WU, MERCHANT and CASTLEMAN (2008), imaging is corrupted by intrinsic or extrinsic noise; the former is modelled by a Poisson distribution that influences each photon that reaches the sensor, and the latter is modelled by a Gaussian distribution that sums to the matrix of pixels. Details about noise are out of the scope of this work, since the degradation to be deeply explored is the defocus blur.

3.2 Defocus Blur

The blur effect is one type of degradation that consists of global or local information loss in the image. The defocus blur is caused by the incidence of light within an aperture with significant dimensions, where the source of light is not properly placed in accordance to the focal plane; it is related to the variables of the optical system such as depth of focus, aperture, depth of field, aberrations and so on (JOSHI, 2014). According to Smith (2007), every optical system exhibits blur properties, in higher or lower proportions, due to the depth of focus and its adjustment. Therefore, blurring is unavoidable to a certain extent, hence every imaging device possesses a PSF due to its optics. The PSF is also named *blur kernel*.

Another useful way to mathematically describe the point spread function is through the Dirac Delta. It consists of a generalized function that represents an impulse, i.e. an infinitely high value within an infinitely small period of time (BRACEWELL; BRACEWELL, 2000). Figure 10 shows an arbitrary example of a punctual source of light and its image, which suffers the spreading effect.

Figure 10 – Magnified image of a light impulse (left) and its impulse response function, the PSF (right).



Source: Adapted from Gonzalez and Woods (2009).

Namely, it is a function $\delta(x)$ that is zero-valued for any $x \neq 0$ and is infinity-valued for $x = 0$. This property can be combined with any smooth function $f\colon \mathbb{R}^n \to \mathbb{R}^n$. The continuous Dirac Delta may be written, as stated by Weisstein (2020a), as

$$\delta^2(x,y) = \begin{cases} \infty, & \text{if } x^2+y^2 = 0 \\ 0, & \text{if } x^2+y^2 \neq 0 \end{cases}, \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta^2(x,y)dxdy = 1. \tag{3.10}$$

The discrete version of the Dirac Delta function consists of an infinite sum instead of the integral. This concept of impulse is the point source of light, concerning images. It provides the blur effect on images, as it promotes the diffusion of the acquired information.

CHAPTER

4

THEORETICAL BACKGROUND

This chapter summarizes the relevant theoretical concepts, methods, and tools for the development of our approach to select partially sharp images among a z-stack dataset and merge the best information from each selected image into a high-quality image. The mathematical and image processing concepts needed to develop our method are convolutions, transforms, enhancement, registration, fusion and quality assessment; statistical methods are used as a bridge between yielding a quantitative index of image quality and merging pixels, as well as evaluating the performance.

4.1 Continuous and Discrete Fourier Transform

The Fourier Transform was conceived by Jean Baptiste Joseph Fourier (1768 - 1830) and states that any periodic function can be expressed as the sum of sines and cosines of different frequencies, each multiplied by a different coefficient (GONZALEZ; WOODS, 2018). According to Brigham (1988), the relationship between the different frequency sinusoids and an arbitrary function s to be analyzed is described as

$$S(f) = \int_{-\infty}^{\infty} s(x) e^{-j2\pi fx} dx, \tag{4.1}$$

where $S(f)$ is the Fourier Transform of the $s(t)$ function and $j = \sqrt{-1}$ represents the imaginary unit. Note that the function transformed from the one-dimensional spatial domain to the frequency domain, is represented by f. Similarly, the inverse transform is denoted by

$$s(x) = \int_{-\infty}^{\infty} S(f) e^{j2\pi fx} df. \tag{4.2}$$

As stated by Bracewell and Bracewell (2000), an arbitrary periodic function s with a

period T can be expressed as a Fourier series, given by the expression

$$a_0 + \sum_{1}^{\infty}(a_n \cos 2\pi nft + b_n \sin 2\pi nft), \tag{4.3}$$

where

$$\begin{aligned} a_0 &= \frac{1}{T}\int_{-\frac{1}{2}T}^{\frac{1}{2}T} s(t)dt \\ a_n &= \frac{2}{T}\int_{-\frac{1}{2}T}^{\frac{1}{2}T} s(t)\cos(2\pi nft)dt \\ b_n &= \frac{2}{T}\int_{-\frac{1}{2}T}^{\frac{1}{2}T} s(t)\sin(2\pi nft)dt. \end{aligned}$$

If the function is not periodic, then the Fourier Transform is applied as a continuous function of frequency, i.e. $s(t)$ is represented by the sum of sinusoids of all frequencies (BRIGHAM, 1988). Particularly, this applies to images, which are non-periodic functions. The most common approach is to appraise the image as a section of a periodic function so the use of Fourier Transform makes sense. The Equations 4.1 and 4.2 together are named the *Fourier Transform pair*. Images are represented by two-variable functions, which motivates the use of a two-dimensional Fourier Transform; moreover, as digital images are matrix representations of images, the two-dimensional Discrete Fourier Transform is the most relevant brand of the FT for image processing applications. Consequently, the two-dimensional Fourier Transform pair of a function $s(x,y)$ is given by

$$S(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(x,y)e^{-j2\pi(ux+vy)}dxdy \tag{4.4}$$

$$s(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)e^{j2\pi(ux+vy)}dudv. \tag{4.5}$$

According to Bracewell and Bracewell (2000), this pair of equations represents the analysis of $s(x,y)$ into components of the form $\exp[j2\pi(ux+vy)]$, where the variables u and v represent spatial frequencies. The equation that relates such components to sinusoids is the *Euler's Formula* or *Euler's Identity*, i.e $e^{j\theta} = \cos\theta + j\sin\theta$, where $\theta = 2\pi f$ is a number that represents an angle in radians and $e^{j\theta}$ is the polar form representation of the sinusoids (GONZALEZ; WOODS, 2018). Finally, the two-dimensional DFT form commonly used in image processing is, as reported by Gonzalez and Woods (2018), denoted by

$$S(u,v) = \sum_{x=1}^{M}\sum_{y=1}^{N} s(x,y)e^{-j2\pi(ux/M+vy/)}dxdy \tag{4.6}$$

$$s(x,y) = \frac{1}{MN}\sum_{1}^{M}\sum_{1}^{N} S(u,v)e^{j2\pi(ux/M+vy/N)}dudv, \tag{4.7}$$

where $s(x,y)$ is a discrete function that represents an image of size $M \times N$, x and y, discrete variables that represent spatial coordinates, and u and v are discrete spatial frequencies.

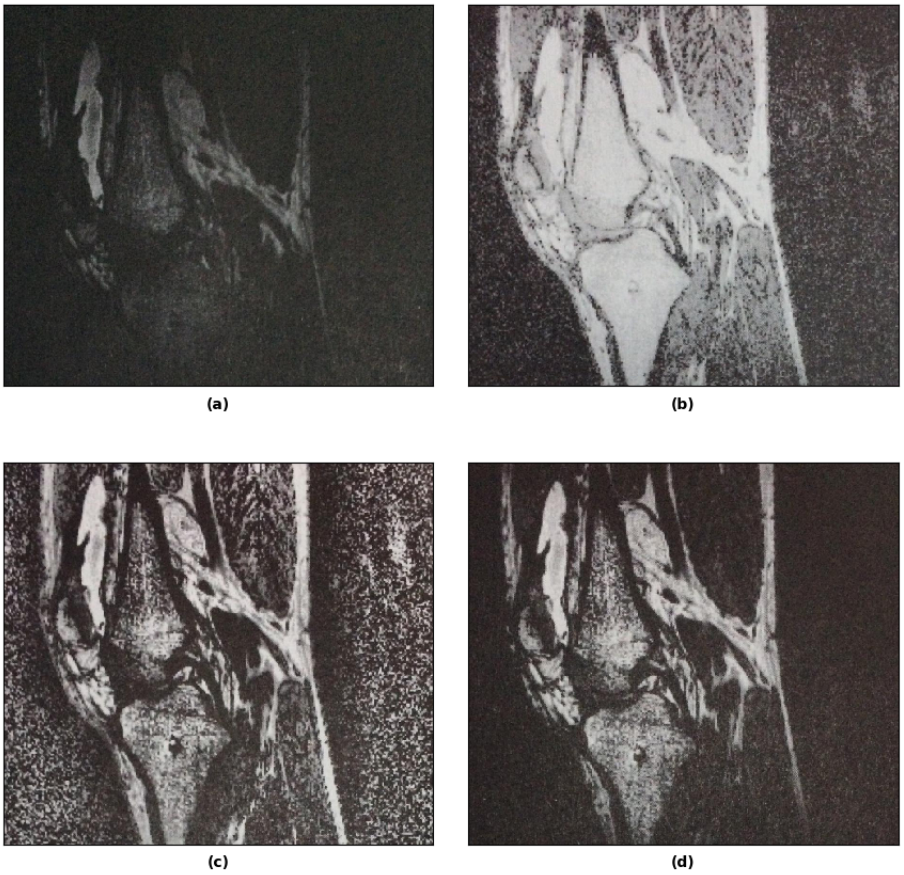
One prominent example of Fourier Transform use is the Convolution Theorem. It states that convolution may be computed by a multiplication in the Fourier domain (BRIGHAM, 1988). This allows much faster computations of convolution in comparison to the spatial domain approach and is frequently used in many applications, such as image filtering and convolutional neural networks. In terms of computational complexity, the conventional two-dimensional Discrete Fourier Transform implementations are of order $\mathcal{O}(n^2)$ for square or zero-padded images and $\mathcal{O}(mn)$ for images of size $m \times n$; for this reason, the Fast Fourier Transform (FFT) is a divide-and-conquer implementation created by Cooley and Tookey in 1965 and reduces the computational complexity to $\mathcal{O}(n \log n)$ (BRACEWELL; BRACEWELL, 2000).

4.2 Contrast Limited Adaptive Histogram Equalization

Contrast enhancement may be described as the process of optimizing the quality of the image in human visual criteria. Histogram equalization is one of the techniques to perform contrast enhancement and works in an image by mapping the distribution of its gray levels to an approximately uniform distribution. The performance of this process is deeply related to the amount of noise in the image since it consists of peaks in the histogram, which unbalances the mapping and enhances noisy structures. One solution to this problem, according to Zuiderveld (1994), is to divide the image into *contextual regions*, i.e. rectangular areas of 8×8 size, compute the optimal contrast for each of the regions and merge the results with bilinear interpolation to avoid boundary effects. This method is known as Adaptive Histogram Equalization (AHE), where the global outlier gray levels do not influence each contextual region contrast enhancement.

The Contrast Limited Adaptive Histogram Equalization (CLAHE) method was proposed to overcome the drawback of noise. As stated by Sonali *et al.* (2019), it is the method that improves the low contrast issue and operates by limiting the contrast enhancement that is usually performed by ordinary histogram equalization or the AHE, which results in the noise enhancement as well. It is accomplished by allowing only a maximum number of pixels in each of the histogram bins and equally distributing the clipped pixels among the whole histogram (ZUIDERVELD, 1994). Figure 11 presents an example of the differences between histogram equalization techniques and their results in a Magnetic Resonance Imaging (MRI) example:

Figure 11 – MRI image of a human knee **(a)**, a simple histogram equalization **(b)**, an adaptive histogram equalization **(c)** and the contrast limited adaptive histogram equalization **(d)**.



Source: Adapted from Zuiderveld (1994).

4.3 Scale-invariant Feature Transform

As originally proposed by Lowe (1999), the SIFT is a feature extraction approach for object and scene recognition. First, the scale-space extrema are detected with a difference-of-Gaussian function, applied in order to identify the invariant scale and orientation keypoints. Each keypoint is selected based on measures of their stability, and one or more orientations are assigned to each keypoint location based on image gradient directions. Finally, a measurement of the local image gradients is performed for the particular scales and neighborhood of each keypoint, followed by a transformation of those into a proper representation.

The custom SIFT implementation described here was proposed by Lowe (2004). The difference-of-Gaussian method for the detection of scale-space extrema is a convolution of an

image $f(x,y)$ with the difference of two nearby scales of distance k, given by

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma))\, f(x,y), \tag{4.8}$$

where $D(x,y,\sigma)$ is the result of the convolution and $G(x,y,\sigma)$ stands for a Gaussian function described as

$$G(x,y,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2+y^2}{\sigma^2}}. \tag{4.9}$$

The difference-of-Gaussians is constructed by convolving the image with several Gaussians separated by the multiplicative factor k, followed by a reduction of the scale-space by doubling σ at each change of scale. Next, the local maxima and minima (extrema) are detected by checking the 8-neighborhood in the current image and 9-neighborhood in the image, in adjacent scales. Having computed the scale-space extrema, the keypoints are localized by fitting a 3D quadratic function of local sample points with the Taylor expansion

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathbf{T}}\frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x}, \tag{4.10}$$

where the derivatives of the difference-of-Gaussians matrix D is computed at the sample point and $\mathbf{x} = (x,y,\sigma)^T$ is the offset for such point. The extremum location is found with the derivative of D with respect to $\mathbf{x}$ by setting it to zero. The value of D at the extremum provides a way to include only stable and good contrast extrema. As well as the low contrast keypoint exclusion, it is necessary to eliminate keypoints which possess a large principal curvature across the edge direction and a small one in its perpendicular direction; this is achieved with a threshold based on the sum and product of the eigenvalues from the trace and determinant of a Hessian matrix computed at the location and scale of each keypoint.

The next step is to provide the orientation of each keypoint concerning local image properties. This yields invariance to image rotation and is done by computing the gradient magnitude and the orientation for each smoothed image sample, denoted by

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{4.11}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right), \tag{4.12}$$

where $m(x,y)$ is the gradient magnitude, $\theta(x,y)$ is the orientation and $L(x,y)$ is the smoothed image, i.e. the observed image convolved with a Gaussian kernel. The obtained information is sampled around each keypoint location at the selected scale and Gaussian blur level, in order to generate the descriptor representation. The samples are then weighted by a Gaussian window

and accumulated into orientation histograms that represent 4×4 subregions. The descriptors are computed from a 16×16 subarray. Finally, the advantages of the custom implementation of SIFT are the invariance to image rotation and scale, robustness across a substantial range of distortions (affine, additive noise and illumination changes) and computational efficiency.

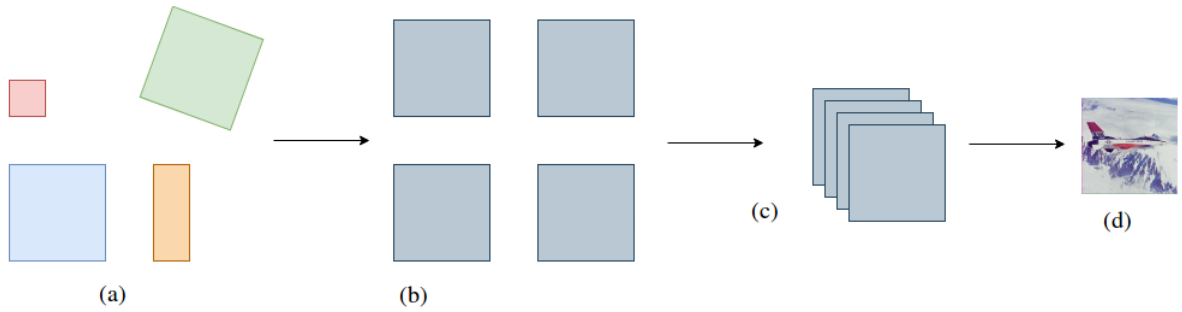
The RANSAC algorithm is a non-deterministic iterative method for fitting a mathematical model to experimental data and also an outlier detector. (FISCHLER; BOLLES, 1981). As for the image registration application in this work, it is used to identify corresponding landmark points in overlapping image tiles (SAALFELD, 2019). Then, the recognized matching points in each image undergo the geometric consistency filtering process with the expected transformation model and the maximal expected error parameters; this last filtering process aims to verify whether all points support the same transformation model and yield the best matches concerning all points for each image of the dataset.

4.4 Image Fusion

Image fusion is a process that merges several images, possibly acquired in diverse conditions or with different cameras, into one image with higher quality, more details and consequently more useful for humans and computer tasks (MITCHELL, 2010). Examples of image fusion applications are noise reduction, edge enhancement, and super-resolution. One traditional use of image fusion occurs in medical imaging fields; the quality of information about illnesses, cells, clinical analysis and several other medical tasks (including the computer-assisted ones) have found profitable results from the image fusion techniques and led themselves to better and faster decisions when it comes to human beings (JAMES; DASARATHY, 2014). There are also relevant applications in remote sensing multispectral images, segmentation of regions in different color spaces, biometry: the pan-sharpening process is the generation of a high-resolution multispectral image from low to high-resolution ones, K-Means segmentation and fusion of pixels in the RGB and the Iris Recognition biometric process with video frames are examples of such tasks, respectively (MITCHELL, 2010). Also according to Mitchell (2010), the general framework for the image fusion procedure consists of four stages: *Multiple Input Images*, *Common Representational Format*, *Fusion* and *Display*. The multiple input images stage is simply the acquisition of the images to be merged. There are several approaches to this: the dataset may be captured from different sensors, under distinct light conditions or angles, with different magnifications, under several focus settings, and with temporal measurements, if the scene changes through time. Figure 35 depicts an arbitrary example of the four stages.

The four arbitrary images in Figure 35.**(a)** represent different images of the same scene, taken at different resolutions, rotation angles, and shapes. In Figure 35.**(b)**, the images are all reshaped, converted to common color space and ready to undergo the processing algorithm which will transform them into feature vectors. Figure 35.**(c)** represents the image fusion by means of

Figure 12 – Image fusion general framework. (a) Multiple Input Images, (b) Common Representational Format, (c) Fusion and (d) Display.



Source: Elaborated by the author.

an arbitrary fusion rule. The resulting image is depicted in Figure 35.**(d)**.

If the acquired dataset images do not share the same features such as dimension, rotation angle, and resolution, then the images should be pre-processed in order to arrive at a common state. This configures the common representational format step, which generates a new and temporary dataset with the same properties, e.g. color space, dimensions, and noise level. The fusion stage employs a decision method to dictate which regions, objects, colors or details will compose the final image; some methods rely on the wavelet transform, for example. Finally, the display stage provides a view of the resulting image, which can be used directly for any further task or even be the input for other image processing operations.

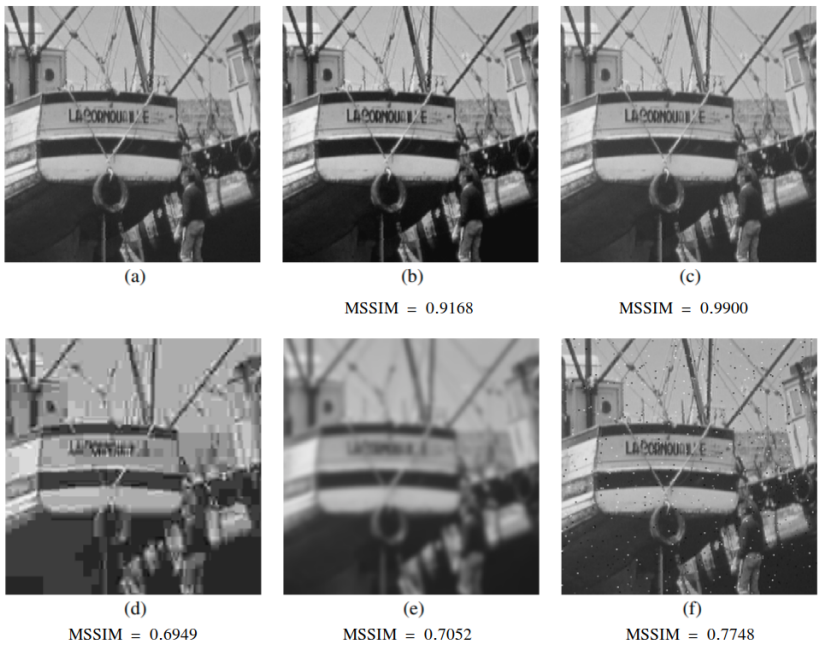
4.5 Image Quality Assessment

Image Quality Assessment (IQA) is the evaluation of image quality as perceived by an average human observer, i.e. how close an image is to a given original or reference image. It is also related to the accuracy of the image acquisition process for an imaging system (BOVIK, 2009). It is known that images are frequently used in health and life sciences, public security systems, remote sensing, and several other fields; hence, there are computational applications that offer some useful service employing image processing. As a result, assessing image quality poses as an important task among those applications for which several techniques are being developed, evolved and deployed.

As stated by Zhou Wang *et al.* (2004), there are three classes of objective image quality metrics that relate to the existence of a no-distortion image (or with a negligible amount of it) for comparison purposes. The Full-Reference Image Quality Assessment (FR-IQA) methods assume

that the reference image is available, while Reduced-Reference Image Quality Assessment (RR-IQA) methods employ a representation of the reference image, such as a set of extracted features. Finally, the No-Reference Image Quality Assessment (NR-IQA) methods, also known as "blind", are those which do not employ a reference image. Figure 36 denotes an example of a full-reference method, the Mean Structural Similarity Measure (MSSIM) method and its output for an image with different types of degradation:

Figure 13 – Example of the MSSIM method output: Original image (a), contrast-stretched image (b), mean-shifted image (c), JPEG compressed image (d), blurred image (e) and salt-pepper impulsive noisy image (f).



Source: Zhou Wang *et al.* (2004).

According to Tang *et al.* (2019), the IQA methods are distributed between the subjective assessment and objective assessment categories. The former is based on a well-defined test environment for random observers to label images and provide the final Mean Opinion Scores (MOS), while the latter is based on the use of strategies such as statistical modeling, machine learning, spatial or spectral image features and so on. It is evident that subjective IQA is demanding; consequently, objective methods are preferred to conduct IQA.

IQA methods are also present within microscopy and its close interaction with image processing. The image acquisition in microscopy techniques may involve lasers, transmitted or reflected light, measurements of atomic force responses, the fluorescence of chemical compounds and several other means. Each technique has an inherent kind of degradation that affects the acquired images or spectra, e.g. the Raman confocal microspectroscopy suffers from the interference of cosmic rays, which yields unexpected peaks in the spectrum.

4.6 Statistics

The measures of the relative standing of an observation describe its location among other values in the distribution, and two examples of these measures are *percentiles* and *z-scores*; also, an observation located outside the range of the distribution is an *outlier* (MENDENHALL; SINCICH, 2016). Percentiles are values that split the data into 100 parts in a sorted dataset, so that the i-th percentile stands for the $i(n+1)/100$ observation, e.g. the 25-th percentile comprises 25% of the data; the z-score, or standard score, is given by

$$z = \frac{x_i - \bar{x}}{\sigma}, \tag{4.13}$$

where x_i is the i-th observation of the variable x, $\bar{x}$ is the mean and σ is the standard deviation of the population or the sample (ZWILLINGER; KOKOSKA, 1999).

The Interquartile Range (IQR) is the length of the interval that contains the middle half of the distribution (DEGROOT; SCHERVISH, 2012). Mathematically, it is the difference between the third (Q_3) and the first (Q_1) quartiles, i.e. the 75th percentile 25th percentile, respectively (DEVORE, 2011). The IQR is described by

$$IQR = Q_3 - Q_1. \tag{4.14}$$

The *kurtosis* is one of the probability distribution shape statistics, which measures the extent of the peak in a distribution, i.e. its "peakedness"; smaller absolute values indicate that the distribution tends to be uniform (ZWILLINGER; KOKOSKA, 1999). First of all, the concepts of *expectation* and *moments* should be described. The expectation of a random variable (and consequently, of a distribution) is a value that summarizes its nature and is given by

$$E(X) = \int_{-\infty}^{\infty} xp(x)dx \qquad E(X) = \sum_{x} xp(x), \tag{4.15}$$

where x is each possible outcome of the random variable X, $p(x)$ is the probability density function for a continuous random variable (left) and the probability function for a discrete random variable (right) (DEGROOT; SCHERVISH, 2012). Still according to DeGroot and Schervish (2012), for a random variable X and every positive $k \in \mathbb{R}$, the expectation $E(X^k)$ is called the k-th moment of X. The r-th moment may be described, according to Zwillinger and Kokoska (1999), as

$$m_r = \frac{1}{n}\sum_{i=1}^{k} p_i(x_i - \bar{x})^r \tag{4.16}$$

for every x_i in the possible outcomes of X. Thus, kurtosis may be defined as the ratio of the fourth moment (Equation B.11 with $r = 4$) by the square of the variance (also Equation B.11 with $r = 2$), denoted by

$$g_2 = \frac{m_4}{(m_2)^2} - 3 \tag{4.17}$$

The -3 constant is inherited from Fischer's approach, where the kurtosis of a normal distribution is zero.

CHAPTER

5

RELATED WORK

This chapter presents a literature review on relevant IQA and multi-focus image fusion techniques, which guided the development of our method. The continuous search for an objective image quality assessment technique covered several strategies, such as subjective testing, statistical modeling, brain science, perceptual modeling, saliency detection, and machine learning (TANG *et al.*, 2019). Every strategy has advantages and trade-offs, according to the nature of the images.

There are challenging features on the multi-focus image fusion part of our method, such as the blurring and illumination patterns, inherent to the bright-field microscopy images. The most trivial pixel-level fusion techniques are simple mathematical operations such as the average or a weighted average of gray levels; those are suitable to the task, but also deliver losses in basic image features, e.g. contrast and saturation (ZHANG; GUO, 2009). Thus, there is a demand for techniques that rely on more sophisticated tools, capable of covering multiple situations: images with different resolutions, light conditions and exposure times, for instance. Multiscale transform approaches, e.g. wavelet domain transforms are good choices and provide a very flexible framework since they depend on the chosen wavelet function (PAJARES; CRUZ, 2004).

5.1 Image Quality Assessment

Defocus blur is an example among all distortions that an image may be acquainted with during its acquisition or processing; this implies that the resulting image will have some sort of degradation in its visual quality, which justifies the need of objective techniques to evaluate and predict the quality of images since the subjective ways are commonly unavailable or inconvenient (Zhou Wang *et al.*, 2004). This section presents a literature review on no-reference IQA methods based on frequency domain analysis, human perception of blur, probabilities, local contrast levels, and gray level variability.

5.1.1 Image sharpness measure for blurred images in frequency domain

De and Masilamani (2013) proposed a DFT-based algorithm to compute an image quality measure of blurred images. This approach is an attempt to quantify sharpness as the amount of high-frequency components in a blurred image is smaller than in a sharper one. The algorithm consists of the following, where $g(x,y)$ denotes the input image of size $M \times N$:

I - Compute the Fourier Transform representation of image $g(x,y)$, denoted by $\hat{g}(m,n)$;

II - Shift the origin of the Fourier coefficients to the center of the matrix;

III - For each Fourier coefficient, compute its magnitude and assign it to a matrix $A(m,n)$

$$A(m,n) = \sqrt{[\mathrm{Re}\,(\hat{g}(m,n))]^2 + [\mathrm{Im}\,(\hat{g}(m,n))]^2}$$

IV - Calculate the maximum value m of $A(m,n)$;

V - Compute the number t of pixels in $\hat{g}$ where $\hat{g} > m/1000$;

VI - Calculate the image quality measure $FM = t/M \times N$.

5.1.2 A no-reference perceptual blur metric

A very simple but efficient perceptual blur metric was proposed by Marziliano *et al.* (2002). It is also based on the fact that high-frequency components are attenuated in a blurred image. Considering that blur affects the edges and texture regions of an image, the technique attempts to measure how scattered they are. First, an edge detector is applied to the luminance component of the image (equivalent to a grayscale-converted image) in order to find the vertical edges. Each row of the image is scanned for pixels corresponding to edges, and the width of each edge is computed by the difference of local extrema closest to the edge. This is done for each edge location, and the global blur measure is the average of all computed widths.

5.1.3 A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB)

Ferzli and Karam (2009) employ a psychometric model, i.e. the assignment of numerical representations to subjective tests to create a perceptual sharpness metric. The subjective tests employ the *Just Noticeable Blur* (JNB) concept, which is the minimum amount of perceived blur around an edge given a contrast higher than the limit of contrast between background and foreground that is possible to notice.

The subjective experiment consists in the evaluation of 27 contrast values (considering the difference between the background and the foreground grayscale values), done by 18 volunteers with normal or corrected vision, with the aim to find a particular standard deviation σ_{JNB} of a 7×7 Gaussian mask. Each contrast value yields a normalized histogram with the volunteers' responses that are treated as the probability of detecting blur as a function of the standard deviation σ, denoted by

$$P = 1 - \exp\left(-\left|\frac{\sigma}{\sigma_{JNB}}\right|^{\beta}\right), \tag{5.1}$$

where σ is the standard deviation of the Gaussian blur filter and corresponds to the blur strength at the considered edge, σ_{JNB} is the standard deviation corresponding to the JNB threshold. The parameters σ_{JNB} and β are chosen by means of a least-square fitting to approximate the probabilities P with the answers from volunteers. The σ_{JNB} value is adjusted to obtain $P = 63\%$.

First and foremost, the image is divided into blocks of dimensions 64×64. Let each block and the image be denoted by R and g, respectively. All the blocks are subject to a Sobel edge detection process and are later classified as "smooth" or "edge" using a threshold. If the number of edge pixels in a block is greater than 0.2% of the total number of pixels in it, then it is classified as an edge block. The smooth blocks are not processed since they are not relevant in terms of blur. Then, for each block, the horizontal edge pixels are retrieved and the width of each edge is computed. With this in hands, the perceptual blur metric for each block R_b is described by the equation

$$D_{R_b} = \left(\sum_{e_i \in R_b} \left|\frac{w(e_i)}{w_{JNB}(e_i)}\right|_{\beta}\right)^{\frac{1}{\beta}}, \tag{5.2}$$

where $w_{JNB}(e_i)$ is the JNB width relative to the contrast of block R_b for all edges e_i. The parameter β is empirically determined to be $3.4 \leq \beta \leq 3.8$. Thus, the general blur measure concerning the whole image corresponds to the probability of all blocks being blurred, given by

$$P_{blur}(I) = 1 - \Pi_{R_b \in g}\left(1 - P_{blur}(R_b)\right), \tag{5.3}$$

where $P_{blur}(R_b)$ may be substituted by $1 - \exp\left(-D_{R_b}^{\beta}\right)$, which results in

$$P_{blur}(I) = 1 - \exp\left(-D^{\beta}\right) \qquad D = \left(\sum_{R_b} |D_{R_b}|^{\beta}\right)^{\frac{1}{\beta}}. \tag{5.4}$$

Finally, D from Equation 5.4 is then normalized by the amount of blocks and represents the final image quality index.

5.1.4 A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD)

A sharpness metric that extends the approach described in 5.1.3 was proposed by Narvekar and Karam (2011). It is based on the cumulative probability of blur detection (CPBD), which applies the same psychometric function framework from JNB. The w_{JNB} value is defined as 5, if the contrast is less or equal to 50 and 3, otherwise. The metric explores the normalized histogram of the probability of blur detection from the whole image, where CPBD is the percentage of edges at which blur is not likely to be detected.

First, the horizontal edges from the image are detected, also divided into blocks of 64×64 and classified as edge blocks or sharp blocks with the same threshold as in Ferzli and Karam (2009). The probability of detection blur for and edge e_i is computed with the Equation 5.5

$$\begin{aligned} P_{BLUR} &= P_{BLUR}(e_i) \\ &= 1 - \exp\left(- \left| \frac{w(e_i}{w_{JNB}(e_i)} \right|^{\beta} \right). \end{aligned} \tag{5.5}$$

Then, the cumulative probability of blur detection is computed from the probability density function of Equation 5.5, and may be described as

$$\begin{aligned} CPBD &= P(P_{BLUR} \leq P_{JNB}) \\ &= \sum_{P_{BLUR}=0}^{P_{BLUR}=P_{JNB}} P(P_{BLUR}). \end{aligned} \tag{5.6}$$

Thus, greater values for the CPBD metric mean sharper images.

5.1.5 S3: A Spectral and Spatial Measure of Local Perceived Sharpness in Natural Images

Vu, Phan and Chandler (2012) proposed one technique to assess image quality that explores both spatial and spectral information from the image. The denomination S_3 represents the combination of S_1, which denotes the spectral measure of sharpness and S_2, which the authors denote as the spatial measure of sharpness. The first step is to convert the RGB image to the grayscale color space with weights 0.2989, 0.5870 and 0.1140 for the red, green and blue channels, respectively. The grayscale image is then divided into blocks denoted by $\mathbf{x}$ of dimensions $m \times m$ and an overlap d around each neighbor block.

In order to compute the $S_1(\mathbf{x})$ measure, they calculate the luminance contrast of each $\mathbf{x}$ block, denoted by CR as follows

$$CR(\mathbf{x}) = l(\mathbf{x}) = (b + k\mathbf{x})^{\gamma} \qquad b = 0.7656;\ k = 0.0364;\ \gamma = 2.2 \tag{5.7}$$

$$CR(\mathbf{x}) = 0 \iff \begin{cases} max(l(\mathbf{x})) - min(l(x)) \leq T_1 \\ \mu_1 \leq T_2 \end{cases} \qquad T_1 = 5;\ T_2 = 2.$$

The S_1 measure for a block is set to zero if the contrast is zero. For the blocks with $CR > 0$, the magnitude spectrum is computed from the two-dimensional DFT, which is defined as the modulus of each complex coefficient and represents how much each frequency is present in the image and is given by $mag(\hat{\mathbf{x}})$. Then, the slope of the magnitude spectrum denoted by $\alpha_{\mathbf{x}}$ is the slope of a line in the standard form $ax + b$ and is obtained by linear regression with

$$\alpha_{\mathbf{x}} = \arg\min_{\alpha} \left\| \beta \hat{\mathbf{x}}^{-\alpha} - mag(\hat{\mathbf{x}}) \right\|_2^2, \tag{5.8}$$

where the L_2 norm is taken over all frequencies. From this framework, $S_1(\mathbf{x})$ is described by

$$S_1(\mathbf{x}) = 1 - \frac{1}{1 + e^{\tau_1(\alpha_{\mathbf{x}} - \tau_2)}}, \tag{5.9}$$

where $\tau_1 = -3$ and $\tau_2 = 2$. Equation 5.9 is a sigmoid function to estimate sharpness considering the slope of the spectral magnitude. This measure does not include the contrast information directly and it may generate inaccurate classifications; this is the reason why the authors propose the use of spatial information. S_2 incorporates the spatial information by analyzing the 8-neighborhoods of pixels in a block $\mathbf{x}$. The Total Variation (TV) is a metric of regularity within the grayscale values and is mathematically described as

$$v(\mathbf{x}) = \frac{1}{255} \sum_{i,j} \left| x_i - x_j \right|, \tag{5.10}$$

where x_i and x_j are 8-neighbors in $\mathbf{x}$. The TV provides a good representation of the absolute differences in each block; this implies that the contrast information is consequently described well. The larger the TV index is, the higher the contrast in the block. The S_2 map is then computed as

$$S_2 = \frac{1}{4} \max_{a \in \mathbf{x}} v(a), \tag{5.11}$$

where a is a 2×2 block of $\mathbf{x}$. Each x is of dimensions 8×8 and the overlap is $d = 4$. The final map for the whole image is the sum of all indices for each of the blocks. Finally, let $\mathbf{X}$ denote the

set of all blocks which form the image. The S_3 image quality index is the weighted geometric mean defined by

$$S_3(\mathbf{X}) = S_1(\mathbf{X})^{\eta} S_2(\mathbf{X})^{1-\eta}, \tag{5.12}$$

where $0 \leq \eta \leq 1$, and $\eta = 0.5$ is the best empirically obtained value for the parameter.

5.1.6 A Fast Approach for No-Reference Image Sharpness Assessment Based on Maximum Local Variation

Bahrami and Kot (2014) proposed the Maximum Local Variation (MLV) metric for image sharpness assessment. The MLV metric improves the idea of the total variation measure. Instead of computing the variations of pixel values among 8-neighborhoods, the authors propose the maximum among all differences in a neighborhood, represented by ψ described by

$$\psi(g(i,j)) = \max(g(i,j) - g(k,l)) \qquad x = \{i-1, i, i+1\};\ y = \{j-1, j, j+1\}, \tag{5.13}$$

where $g(k,l)$ with the bounds for k and l denoted in Equation 5.13 are the 8-neighbors of the pixel $g(i,j)$. In a nutshell, the metric finds the maximum among the 8-neighborhood of a pixel, concerning the ℓ_1 norm as a distance. With this setup, the algorithm begins with the grayscale conversion of an image of dimensions $M \times N$. The neighborhoods are 3×3 blocks along the image and the MLV is computed for all pixels, which results in a map $\Psi(g)$ given by

$$\Psi(g) = \begin{pmatrix} \psi(g(1,1)) & \cdots & \psi(g(1,N)) \\ \vdots & \ddots & \vdots \\ \psi(g(M,1)) & \cdots & \psi(g(M,N)). \end{pmatrix} \tag{5.14}$$

The next step is to apply a statistical technique to analyze how the distribution of Ψ behaves when the image is sharp or blurred. For the sharp regions, i.e. high MLV values or black content, the distribution is closer to hyper-laplacian; for smoother and blurred regions, the distribution is closer to Gaussian. Therefore, the authors chose to parametrize the MLV with a Generalized Gaussian Distribution (GGD), which covers the Gaussian, the laplacian and hyper-laplacian distributions and is denoted as

$$GGD(\Psi(g); \mu, \gamma, \sigma) = \left(\frac{\gamma}{2\sigma\Gamma\left(\frac{1}{\gamma}\right)\sqrt{\frac{\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{3}{\gamma}\right)}}} \right) e^{-\left(\frac{\frac{\Psi(g)-\mu}{\sigma}}{\sqrt{\frac{\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{3}{\gamma}\right)}}} \right)^{\gamma}}, \tag{5.15}$$

where μ is the mean, σ is the standard deviation, γ is the shape parameter and Γ is the gamma function. In a nutshell, the MLV metric is relative to the standard deviation of the GGD distribution: the sharper the image is, the higher the standard deviation. The Ψ matrix is weighted with values obtained with the exponential function $w_{i,j} = e^{\eta_{i,j}}$, where $\eta_{i,j}$ is the rank of $\psi(g(i,j))$ when sorted in ascending order.

5.2 Multi-focus Image Fusion

Either in microscopic or macroscopic scales, imaging devices have a finite depth of field. Blurry regions in images may occur due to several reasons, e.g. the acquisition outside the boundaries of the depth of field (HUANG; ZHONGLIANG, 2007). Therefore, it is nearly impossible to obtain a globally sharp image in microscopic scales, since height differences within the surface of the sample will also be magnified and the depth of field of microscopes is usually small. Image fusion of stack of images, acquired with a variety of focus adjustments, may yield a fully sharp image. In this section, we investigate multi-focus image fusion techniques based on principal component analysis, guided filtering with local linear models, gradients and singular value decomposition.

5.2.1 Pixel-level image fusion using wavelets and principal component analysis

Naidu and Raol (2008) addressed the multi-focus image fusion problem with two fusion rules, based on Principal Component Analysis (PCA) and the Discrete Wavelet Transform (DWT). The Principal Component Analysis is a mathematical procedure based on linear algebra, which performs the extraction of uncorrelated variables that represent data. In other words, the PCA based fusion rule composes the image from the most relevant subset of data from each input image. The proposed PCA-based image fusion algorithm consists of the following steps:

I - Reshape the input images as two-column vectors, which yields a matrix of dimensions $2 \times n$, with n as the original dimension of each image. Let the resulting matrix be denoted as Z;

II - Compute the empirical mean vector of dimensions 1×2, i.e. a vector that contains the mean of observations for each variable;

III - Subtract the empirical mean vector from each column of Z, which results in a matrix of dimensions $2 \times n$. Let the resulting matrix be denoted as Y;

IV - Compute the covariance matrix of Y, i.e. YY^T;

V - Compute the eigenvectors and the eigenvalues of the covariance matrix and sort them by descending order. Let the eigenvectors and the eigenvalues be denoted by V and D, respectively. Both V and D dimensions are 2×2;

VI - Compute the principal components P_1 and P_2 with the first column of V, as

$$P_1 = \frac{V(1)}{\sum V} \qquad P_2 = \frac{V(2)}{\sum V}. \tag{5.16}$$

Finally, the fused image is composed by

$$g_{fused}(x,y) = P_1 g_1(x,y) + P_2 g_2(x,y), \tag{5.17}$$

where g_{fused} represents the fused image and g_1 and g_2 represent the two input images. The process may be extended to several images by applying the method in the first two images of the dataset, then with the fused image and another image and so on. Along with PCA, the authors explored a fusion approach with the DWT. The transform is applied to the input images and decomposes them into subbands of detailed and approximation wavelet coefficients, i.e. high and low-frequency components, respectively. Finally, the fusion rule consists of averaging the approximation coefficients and choosing the maximum among each detailed coefficient in each subband.

5.2.2 Image fusion with guided filtering

The guided filtering, i.e., a windowed linear transform of the input image where coefficients are obtained by linear ridge regression was applied to perform image fusion (LI; KANG; HU, 2013). The authors proposed a three step pixel-level image fusion technique (decomposition, filtering and reconstruction) that also makes use of low-pass filtering, edge extraction, and spatial consistency. The decomposition step transforms the input images into two-scale representations by means of a convolution with an average filter of size 31×31. The first representation is named *base* layer, directly obtained from the average filter convolution, and the second representation is named *detail* layer, obtained by subtracting the base layer from the input image. In order to retrieve the weight map, i.e., a characterization of sharp and blurry regions of the image, edge extraction is performed with a 3×3 Laplacian filtering kernel (which will be further explained in Chapter 6); subsequently, the edges undergo a Gaussian low-pass filter of dimensions 11×11 and $\sigma = 5$. Let S_i^k be the smoothed edge in the pixel k of the ith input image. The weight map is built as

$$P_i^k = \begin{cases} 1, & \text{if } S_i^k = \max\left(S_1^k, S_2^k, \ldots, S_n^k\right) \\ 0, & \text{otherwise} \end{cases}, \tag{5.18}$$

where n is the number of input images. The guided filter is then applied to each image, together with its weight map, which produces final weight matrices W_i^B and W_i^D for the base and detail layers of each image, respectively. Subsequently, all base and detail layers are fused by means of weighted averaging with their respective weight matrices, given by

$$\bar{B} = \sum_{i=1}^{n} W_i^B B_i \qquad \bar{D} = \sum_{i=1}^{n} W_i^D D_i, \tag{5.19}$$

where $\bar{B}$ and $\bar{D}$ are the fused base and detail layers, respectively. The fused image is finally constructed by summing the fused layers.

5.2.3 Multi-scale weighted gradient-based fusion for multi-focus images

Another way to perform edge detection-based image fusion is to extract gradients from each image of a dataset and merge them with the structure tensor, as proposed by Zhou, Li and Wang (2014). The first step consists of obtaining the local gradients of each input image in the form of a structure tensor, i.e. a non-negative second-moment matrix where its eigenvalues represent the intensity changes in a given point, and then a local gradient covariance matrix described as

$$C_\sigma = \begin{pmatrix} \nabla_x^2 * G_\sigma & (\nabla_x \nabla_y * G_\sigma) \\ (\nabla_x \nabla_y * G_\sigma) & \nabla_x^2 * G_\sigma \end{pmatrix}, \tag{5.20}$$

where G_σ is a Gaussian filter, ∇_x and ∇_y are the gradients of the image along x and y directions in a given pixel. The σ parameter represents the standard deviation of the Gaussian function and also the scale of the matrix C_σ. The local image structure is related to the eigenvalues of C_σ, s_1^2 and s_2^2. Let s_1 and s_2 be the square root of such eigenvalues. If one of those is large and the other is small, it means that the pixel consists of an edge; if both are large, then the region is sharp, which indicates a corner. Then, the authors define a structure saliency measure for the image at a selected scale, as

$$Q = \sqrt{(s_1 + s_2)^2 + \alpha(s_1 - s_2)^2}, \tag{5.21}$$

where $\alpha > -1$. This is used as a sharpness measure for the multi-focus images at different scales in the presence of anisotropic blur and also misregistration. The sharpness is first computed at a large scale, i.e. a large σ value, which roughly detects the focused region of each input image.

The next step is to define an unknown region near the boundaries of each focused region and set the gradient weights as 1 for sharp and 0 for blurry. The gradient weights are then calculated by applying the focus measure with a small σ value and then merged with the

gradients. This results in a set of probabilities for each pixel in each image. The composition of the fused image is then guided by the largest probability values.

5.2.4 Image fusion technique using multi-resolution singular value decomposition

Naidu (2011) proposed a fusion technique which is similar to the DWT-based ones, but employs a Multiresolution Singular Value Decomposition (MSVD) instead. The multiresolution SVD performs a factorization of a matrix into eigenvalues and eigenvectors in different scales by the following steps:

I - Let $X = [x_1, x_2, \ldots, x_n]$ be an one-dimensional signal of length n. Reshape X onto a matrix given by

$$X_1 = \begin{bmatrix} x_1 & x_3 & \ldots & x_{n-1} \\ x_2 & x_4 & \ldots & x_n \end{bmatrix}$$

and a scatter matrix $T_1 = X_1 X_1^T$;

II - Factorize the matrix T_1. Let the eigenvector matrix of T_1 be U_1. Then T_1 is transformed into a diagonal matrix S_1^2 as

$$S_1^2 = U_1^T T_1 U_1 = \begin{bmatrix} s_1^2 & 0 \\ 0 & s_2^2 \end{bmatrix},$$

where s_1^2 and s_2^2 are the squares of the singular values. Then, let $\hat{X}_1 = U_1^T X_1$. The top row of $\hat{X}_1$ contains the approximation component, i.e. the largest eigenvalue, and the bottom row contains the detail component, i.e. the smallest eigenvalue. Let φ and ψ represent the top and bottom rows of $\hat{X}_1$, respectively. In order to perform the image decomposition in other scales, the approximation component φ plays the role of X.

The MSVD is a decomposition of 2×2 blocks of the image, rearranged into one-dimensional vectors of dimensions 4×1. The authors proposed a two-level decomposition, which generates 8 images from each input image. At each level, the largest absolute value among the MSVD detailed coefficients will compose the fused image, while at the coarsest level, the average of all approximation is computed and added to the output.

CHAPTER

6

MATERIALS AND METHODS

The literature review presented in Chapter 5 showed that both no-reference image quality analysis and image fusion were addressed by many known techniques such as multiresolution analysis and principal component analysis. However, applications still evolve as new strategies such as machine learning are presented. The idea that motivated the use of the Fourier Transform as the basis for an IQA method is that it is a fast, robust and reliable analysis tool and could produce relevant results if properly explored. Similarly, the image fusion process was developed by exploring many state-of-the-art fusion applications and addressing both computational requirements (both hardware and software) and computational performance, since the method is built to work on normal optical microscope workstations instead of large clusters. As a result, we studied the capabilities of the Laplacian of Gaussian as the basis for our image fusion algorithm.

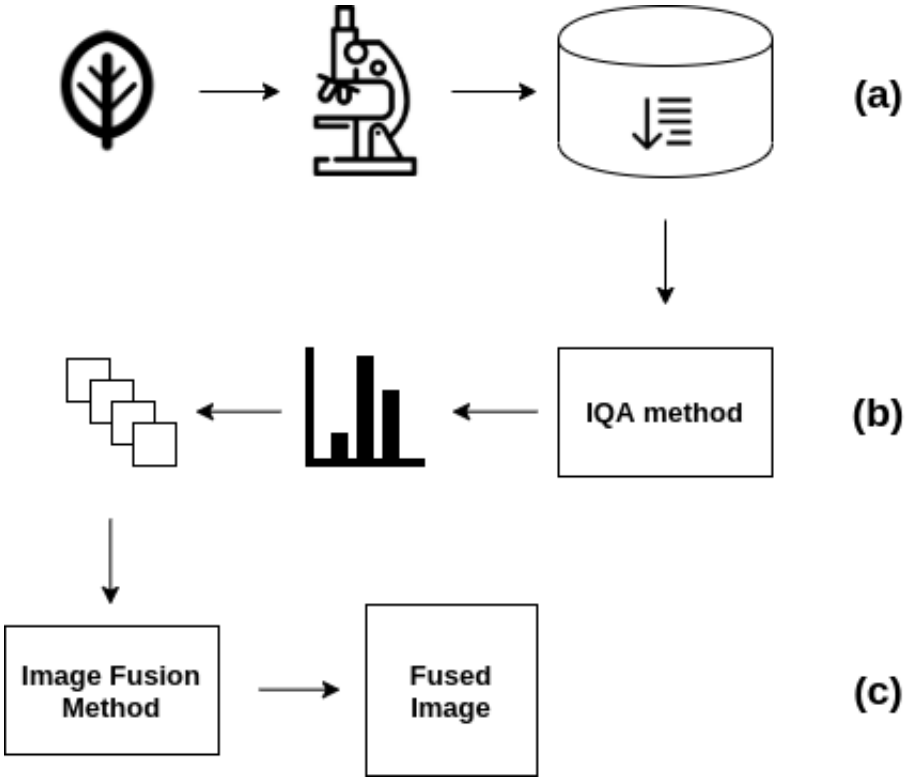
This chapter presents all the details about our no-reference image quality assessment method and our multi-focus image fusion method, as well as relevant information about the proposed datasets, including acquisition details. We also show how the methods relate to each other since it also requires the human knowledge to select the images for registration and posterior fusion, similar to most of the z-stacking systems on microscopy software packages.

6.1 Overview

Our approach, depicted in Figure 14, consists of three parts: a) an image acquisition process based on the z-stacking technique; b) a method for assessing the quality of images that explores frequency domain features and basic statistical analysis tools and outputs a quantitative index that allows the selection of focused images and c) an image fusion procedure, based on the Laplacian of Gaussian filter and energy of edges, which forms the final image. Figure 14.(a) denotes the image acquisition process, which consists of collecting specimens of plants, acquiring images with microscopes and organizing the datasets. Then the images undergo our IQA method in Figure 14.(b), and a subset of the images is selected for registration by means of statistical

methods. Finally, Figure 14.(c) represents the image fusion procedure, which aims to produce a high-quality image.

Figure 14 – Overview of each stage of our approach: **(a)** image acquisition and registration, **(b)** selection of images after IQA and **(c)** image fusion.



Source: Elaborated by the author.

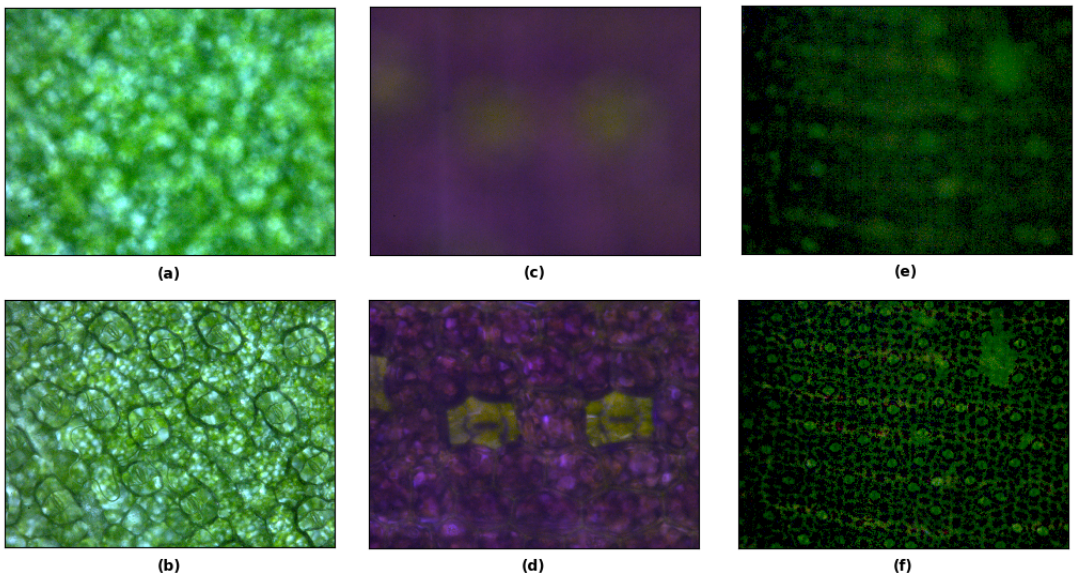
6.2 Proposed dataset

Three bright-field microscopy images datasets were acquired with the ZEISS Axiocam ERc5s camera in the ZEISS SteREO Discovery.v20 and the ZEISS AxioLab A1 stereo microscopes from the Scientific Computing Group (SCG) at São Carlos Institute of Physics (IFSC). The datasets contain images from leaf histological samples of the plants *Callisia repens*, *Tradescantia zebrina* and *Cthenante oppenheimiana*, acquired with different focal planes and with different magnification levels.

A shared feature among the species *Callisia repens*, *Tradescantia zebrina* and *Cthenante oppenheimiana* is the purple abaxial (lower or bottom) leaf surface. This is commonly observed in deeply-shaded understorey plants and can be either transient or permanent, depending on the species and environmental conditions (FILHO; BRUNO, 2018). Several research projects have been conducted by the SCG group on plant leaf images, including biological studies with complex network analysis, where the locations of particular structures of the leaf, i.e. the stomata, were modeled as graphs. The *stoma* (plural *stomata*) is a structure that consists of an aperture

between two cells, named *guard cells*, and controls the exchange of steam, CO_2 and other gases from the inner part of the leaf and the atmosphere (HETHERINGTON; WOODWARD, 2003). Furthermore, the concentration of stomata in leaves of purple plants is high; such stomatal cells are green and create a contrast between the epidermis and the stomata, which yields very good results with optical microscopy imaging (FILHO; BRUNO, 2018). Samples of blurred and sharp images of both datasets are shown in Figure 15.

Figure 15 – Examples of the proposed dataset images: blurred *Callisia* **(a)**, sharp *Callisia* **(b)**, blurred *Tradescantia* **(c)**, sharp *Tradescantia* **(d)** and blurred *Cthenante* **(e)**, sharp *Cthenante* **(f)**.



Source: Elaborated by the author.

6.2.1 Acquisition and Usage Protocols

We will refer to the datasets as *Callisia*, *Tradescantia* and *Cthenante* for notation simplicity. The *Callisia* and *Tradescantia* datasets were acquired with the z-stacking method in the SteREO Discovery.v20 microscope, whereas the *Cthenante* dataset was acquired with the AxioLab A1 microscope. The workstation was connected to the microscopes by means of the ZEISS Axiovision version 4.8 software package. The z-stacks were manually built, i.e. the axial location of the objective was manually changed.

The SteREO Discovery.v20 microscope allows a precise measurement of the objective height, and the slices were acquired with a distance of 10 μm between each other - the maximum manually achievable distance. For both microscopes, the acquisition started with the objective height above the focal plane and, therefore, a completely blurred image. Then, the objective was progressively lowered in the z axis, and at each step, an image was taken; this process was done until the objective height was below the focal plane and the images were blurred again.

The z-stacks were registered for image fusion after the eligible images in each set were chosen. Therefore, after the IQA method was applied, each set of eligible images was aligned with the TrakEM2 package, an ImageJ-based tool for processing and analyzing microscopy images. It includes methods for lens distortion correction, stitching, serial section alignment, correction of section thickness, and contrast adjustment (SAALFELD, 2019). TrakEM2 uses a particular combination of methods. The feature extraction was done with a custom implementation of the Scale Invariant Feature Transform (SIFT), together with a custom extension of the Random Sample Consensus (RANSAC) method for parameter estimation and the geometric consensus filtering process with the expected transformation model and a maximal expected error as parameters (SAALFELD, 2019).

Primarily, a subjective quality index based on the Mean Opinion Score (MOS) was built in order to validate the results. The mean opinion score is the average of values on a predefined scale that an observer assigns to his opinion about the performance of a system (in this case, the imaging system) across a sample of observers (LIU *et al.*, 2019). Practically, it consists of integer numbers in the interval $[1,5]$ where 1 is the worst score and 5 is the best. With the output of three people - two microscopy experts and one from a non-microscopy field - a subjective index was created. Most of the images were classified as 1 and the maximum score was 3. The images were then labeled as 3 - *eligible* and 1 or 2 - *negligible* for the fusion process, respectively. The axial nature of the acquisition allowed for a contiguous set of eligible images in each dataset. Table 1 presents some relevant properties of each dataset.

Table 1 – Information about the proposed datasets.

Dataset	***Callisia***	***Tradescantia***	***Cthenante***
Magnification	50x	200x	100x
Step	10μm	10μm	4μm
Number of images	56	66	55
Eligible	9	2	16
Sharp sequence	41 - 49	50 - 51	30 - 45

Source: Elaborated by the author.

In order to use the proposed datasets for evaluation of new methods, it is necessary to take into account that several images are totally blurred, and an image quality assessment algorithm should yield a low metric number for such images. As the z position approaches the optimal focal plane, the metric may start to produce better values. The images are named with the step of the z-stack starting at 1, so that the eligible images are easily located. The metric should yield the highest quality values are among the range of eligible images, depicted as “Sharp sequence” in Table 1. When the axial position is not exceeds the limits of the sharp sequence, the metric values should decay. The datasets are capable of assessing the monotonicity, accuracy and precision of image quality metrics.

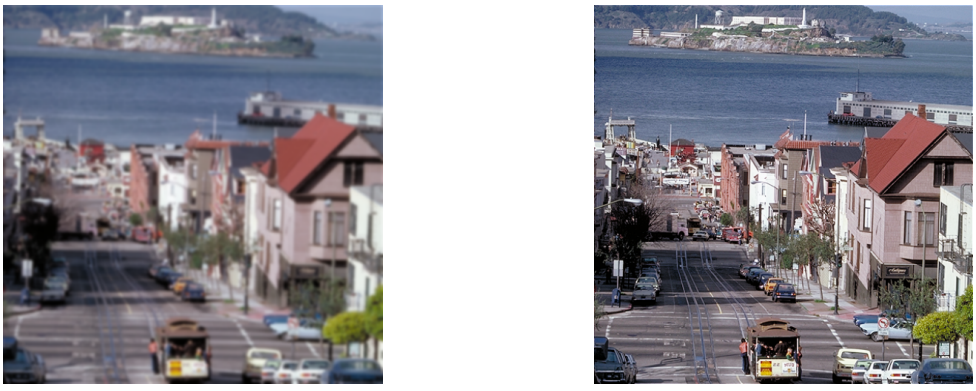
6.3 Benchmark datasets

Benchmark datasets are fundamental in computer vision and image processing research in order to track the performance, accuracy and efficiency of new methods and algorithms. The image quality assessment was also evaluated against literature methods with the well-known Computational and Subjective Image Quality database (CSIQ) image quality assessment database (LARSON; CHANDLER, 2010) and the KonIQ database, the largest image quality assessment database to date (Hosu *et al.*, 2020).

The CSIQ database contains 30 original images of dimensions 512×512. Each image is distorted in four to five different levels separately by JPEG and JPEG-2000 compression, Gaussian blurring, global contrast decrements and additive pink Gaussian noise. The Gaussian blur subset with 150 images was employed in the proposed analysis. The dataset also contains 5000 subjective rates done by 35 different observers, By means of a MOS index.

The KonIQ database was built in order to evaluate the performance of a deep learning method for blind image quality assessment. It consists of 10073 images of dimensions 1024×768 with different labels for brightness, contrast, colorfulness and sharpness. The labels were generated from 1.2 million MOS rating of 1459 observers. One dissimilar feature of KonIQ when compared to CSIQ is that the quality assessment should be done among different scenes with different levels of degradation. Figure 16 shows two samples of the CSIQ database, and Figure 17 presents both blurred and sharp samples from the KonIQ image.

Figure 16 – Samples from the CSIQ database: blurred image (a) and sharp image (b).



(a) (b)

Source: Larson and Chandler (2010).

Figure 17 – Samples from the KonIQ database: blurred image (a) and sharp image (b).

(a) (b)

Source: Hosu *et al.* (2020).

6.4 NR-IQA in Bright-field Microscopy Images Using the Fourier Transform and Kurtosis

The next step is to devise a quantitative representation of sharpness for each image in each dataset. This measure allows for the selection of the eligible images through statistical analysis and the posterior registration of the selected slices. We propose a new method for image quality assessment based on a sampling process of the Fourier spectrum and posterior analysis of the coefficients as a probability distribution using summary and descriptive statistics. This section presents each stage of the method: pre-processing, Fourier spectrum sampling and statistical analysis. Figure 18 shows a diagram of the proposed method.

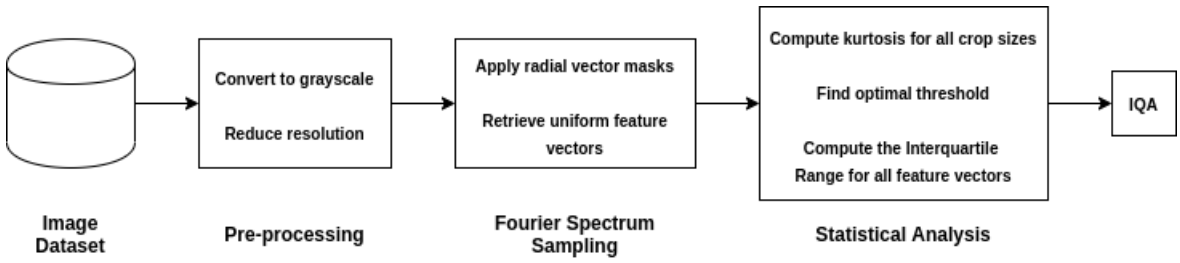


Figure 18 – Pipeline of processing steps of the propose no-reference IQA method.

The DFT was chosen as the basis for the method after tests with the techniques in the literature review, but it was also claimed within the SCG group that it is a simple and robust way to describe image sharpness in terms of a frequency profile. Some tests were done with the discrete wavelet transform, but the response was similar to the DFT. Similarly, after exploring the mathematical properties of the Fourier transform and also the other mathematical functions employed, i.e. the logarithm and the modulus of a complex number, the final feature vector was proven to be a distribution and could be mapped onto a probability space. This discovery brought the idea of statistics to the scope of this work, and by means of visualization techniques such as 2D and 3D plots, it was possible to notice the pattern of blurred and sharp images in the resulting

distribution. The need of a statistical property capable of quantifying such pattern and allowing the method to segregate the blurred and sharp images motivated the research for techniques that deal with the shape of a distribution. The simplest tools were kurtosis, the interquartile range and the z-score, which were empirically proven to be accurate in the proposed setup.

First, each image undergoes a grayscale conversion with the luminance method (PONTI; NAZARÉ; THUMÉ, 2016) - a linear combination of the three channels of a RGB image, as shown in the matrix equation given by

$$I_{luminance} = 0.299R + 0.587G + 0.114B, \tag{6.1}$$

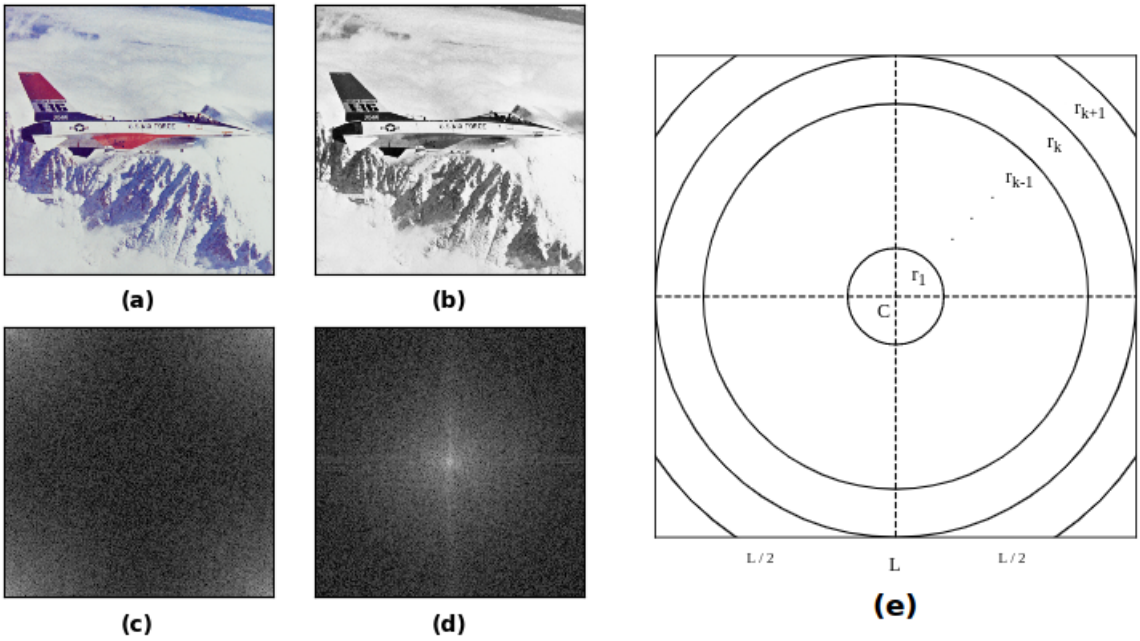
where $I_{luminance}$ is the matrix that represents the grayscale converted image, R, G and B represent the matrices of the red, green and blue channels, respectively.

The CLAHE algorithm is then applied to deliver a more uniform image to the Fourier Transform. This is necessary as microscopy images are influenced by illumination conditions, the focus adjustment itself and also physical properties of the transmitted or reflected light. Next, the DFT is applied to all images and the resulting spectra are analyzed. The Fourier spectrum of a two-dimensional signal such as an image is a matrix with complex coefficients and zeros on each of its four corners and carries information about frequency bands of the image, characterized as a distribution. Our approach requires a shift between the first and third quadrants, and also the second and the fourth quadrants, to the center of the matrix. The unshifted and shifted Fourier spectrum of an grayscale airplane test image are shown in Figures 19.(c) and 19.(d), respectively.

The frequency information may be retrieved from the shifted Fourier spectrum by means of concentric circles drawn over it in the form of masks. Theoretically, the number of masks that may be applied to the spectra is infinite, but the discrete nature of images renders a finite number of masks. Let the matrix of DFT coefficients be graphically represented in Figure 19.(e) and be mathematically denoted as a square of side $L = \max(m,n)$, $k = L/2$ be the maximum radius value for circles within the square and $C = (k,k)$ be the center of the infinite set of concentric circles inscribed in the square. Circles with small radii comprise low-frequency information whereas large radii values cover high-frequency coefficients.

Taking into account the pixel resolution of the input images, it makes sense to sample the information, otherwise, the computational complexity and running times of the algorithm for one image alone would be impractical. Moreover, it is also unrealistic to evaluate every area under each concentric circle to retrieve the frequency content of an image; there is also no standard discrete amount of frequency bands to be evaluated. Therefore, we drove our efforts to comprise as much information about each frequency band as possible and explored the fact that the Fourier spectrum of a real function such as an image is even, i.e. symmetric concerning its origin. Instead of the pure complex coefficients, we represent the Fourier spectrum by the magnitude of its

Figure 19 – Original image **(a)**, luminance grayscale converted image **(b)**, unshifted Fourier spectrum of the grayscale image **(c)**, shifted Fourier spectrum of the grayscale image **(d)** and frequency bands as rings of radius $\{r_i : i \in \mathbb{N}^*\}$ drawn over the 2D spectrum **(e)**.



Source: Elaborated by the author.

coefficients, given by

$$K(m,n) = \log_e \left(1 + \sqrt{[\mathrm{Re}\,(\hat{g}(m,n))]^2 + [\mathrm{Im}\,(\hat{g}(m,n))]^2}\right), \tag{6.2}$$

where $\hat{g}(m,n)$ are the complex Fourier coefficients of an image $g(x,y)$. We propose to sample the spectrum by means of radial lines as masks, i.e. white lines are drawn over a zero-valued matrix with the same dimensions as the Fourier spectrum of the image, which are then element-wise multiplied by the spectrum. The lines are created from the (x_c, y_c) center of the spectrum to points in an approximate radial position, which is calculated by

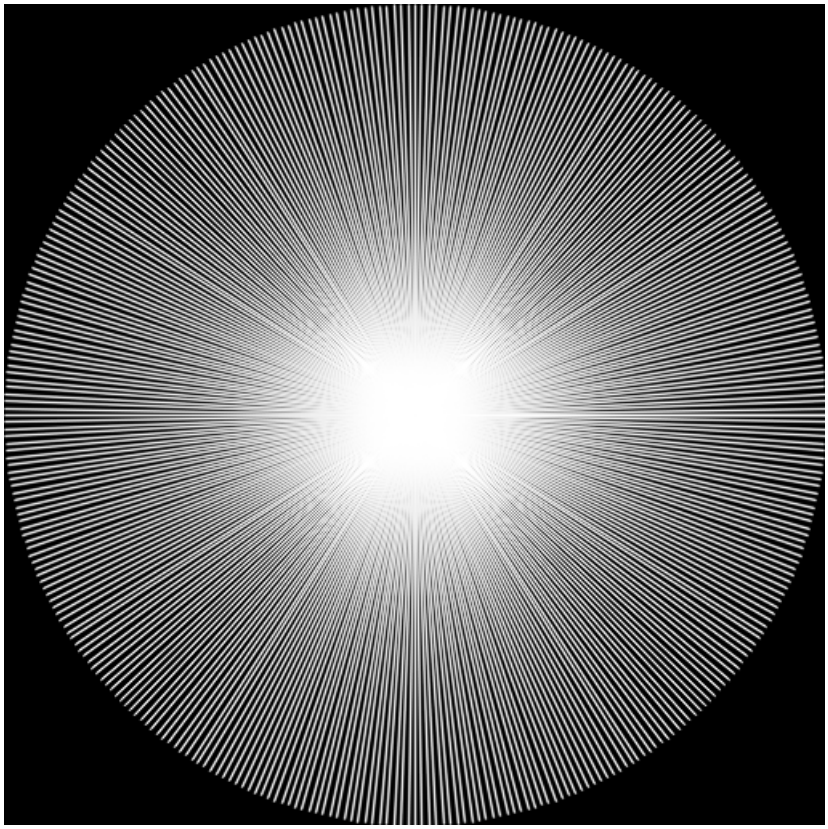
$$P(x,y) = (x_c r_j \cos a_j, y_c r_j \sin a_j) \tag{6.3}$$

with the set of angles $\{a_j\}$ in the radian form, computed as

$$\left\{a_j : a_j = \frac{j\pi}{180}\right\} \qquad j = \{0,1,2,...,360\}. \tag{6.4}$$

Equation 6.3 is a floating-point ordered pair, rounded to the nearest integer value in order to represent a location in the spectrum. The antialiased lines are drawn with a Gaussian filtering process. After all the lines have been generated, the mask is similar to what is shown in Figure 20:

Figure 20 – Final mask of antialised radial lines to sample the Fourier spectrum.



Source: Elaborated by the author.

The extraction of the region of interest from the spectrum with the mask of radial lines results in arrays of complex coefficients that represent samples of the frequency profile of the image. However, despite the antialiasing, the resulting arrays differ in length since the original data is discrete. The vectors go through an element-wise average, which results in a one-dimensional vector as a descriptor of the frequency spectrum, and before such average, we find the smallest vector size among all of them and discard the remaining information in all of them. For comparison purposes, a square image of side L yields a feature vector of size $1/2L$.

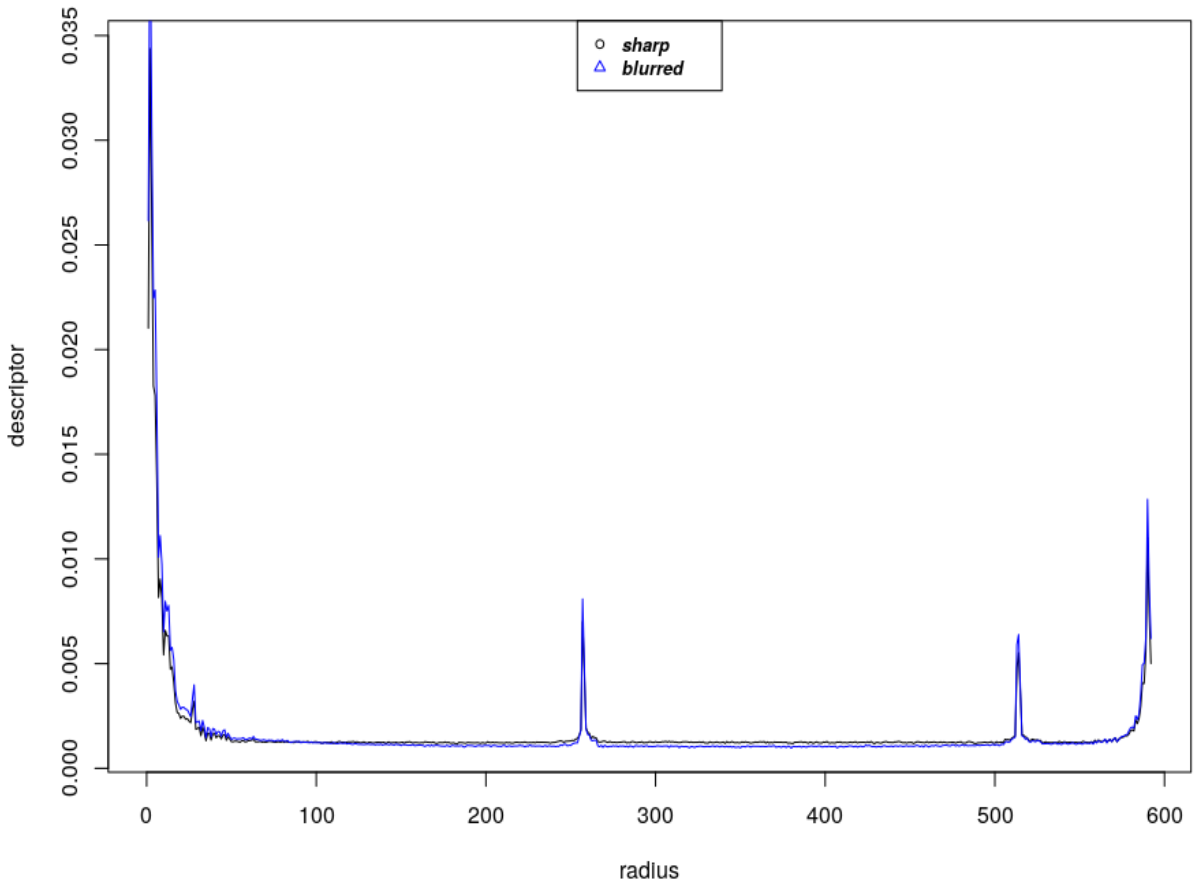
Next, the sharpness information of the image is encoded into a low-dimensional and concise representation that may undergo analysis with statistical tools and the mathematical properties of the Fourier spectrum. The feature vectors from each image need to be transformed into a model that is suitable for techniques such as Descriptive Statistics or Bayesian Inference, and therefore must be mapped onto the probability space $\Omega = [0,1]$. Hence, we apply an operator $T : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$, written as

$$T(x_i) = \frac{x_i}{\sum_{j=0}^{n-1} x_j} \qquad i = \{0,1,...,n-1\}, \tag{6.5}$$

where each x_i is a value of the descriptor which will be mapped onto a probability. Note that each

feature vector after the sampling process is a distribution in the scope of the Distribution Theory, i.e. a continuous linear functional defined on a space of smooth (functions with continuous derivatives up to some desired order over some domain (WEISSTEIN, 2020b)) and compact supported (the functions yields zero outside of a closed and bounded set (ROWLAND; WEISSTEIN, 2020)) functions, with values in the range $[0, \infty)$. Therefore, it makes sense to apply the operator that maps it to a probability. This proposition was proved in Appendix C, which summarizes several definitions from Mathematical Analysis, Measure Theory and Distribution Theory required for the proof. The shape of the frequency distribution represented by each descriptor may vary according to the contents of the image, i.e. lines, circles, edges or complex textures, but it possesses a common representation, as shown in Figure 21 for the "1600" image in the CSIQ database. The black line represents a sharp image (without any degree of Gaussian blur) and the blue line represents the blurriest sample of the "1600" subset.

Figure 21 – Common shape of the descriptors right after computation.



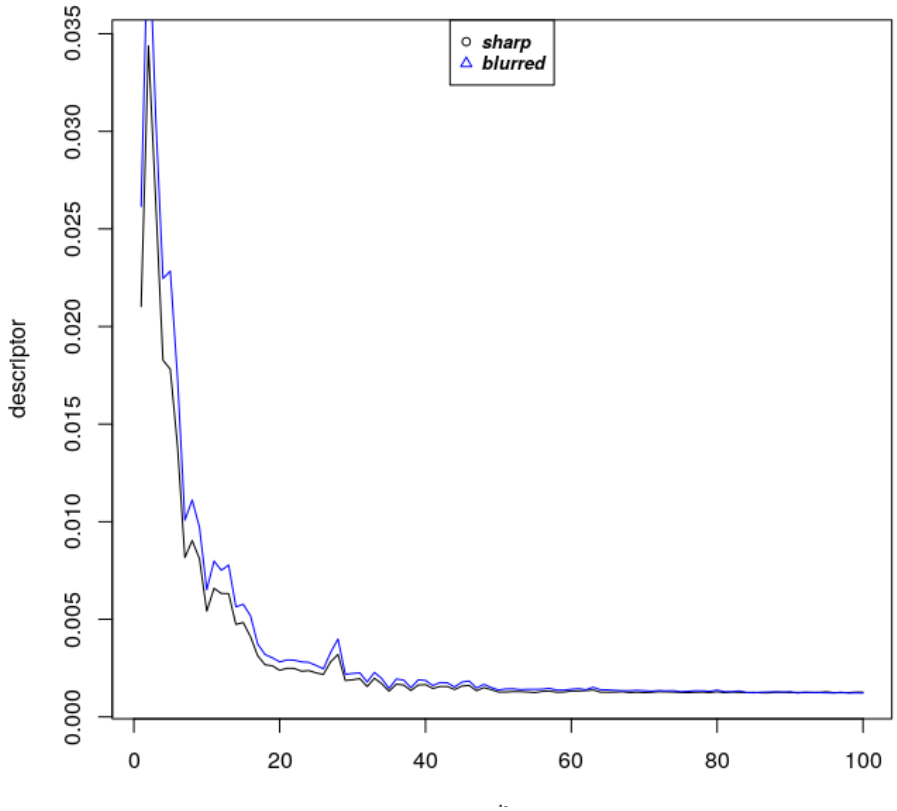
Source: Elaborated by the author.

Information embedded in the low-frequency components of each feature vector corresponds to the Dirac delta distribution within the PSF of the microscope and should be discarded since it is equal to all images and does not resemble blur information; the removal should be done with caution so that the remaining information is enough to represent the blur profile of

the image and allow the further selection by the image quality metric. An optimal threshold should be calculated for "cropping" the data, i.e. only a subset of it will represent the sharpness information. The threshold is chosen to maximize the range of a set of kurtosis values for each feature vector.

Consequently, the data presented by Figure 21 provides an overview of the distribution of frequencies, but it does not depict distinguishable differences between a sharp and a blurred image. However, by maintaining the range of values, selecting a subset of both sharp and blurred feature vectors (in this case, the dimensions of the original image are 512×512 and the feature vectors are of length 592) which starts at the position zero and ends at 100, and finally plotting both subsets, it is possible to notice a difference between the graphs. Figure 22 shows that the blue line (the blurred image) contains larger values within the $[0, 20]$ interval in the x axis in comparison to the black line (the sharp image), providing less "peakedness" to the distribution - which translates itself as lower kurtosis values.

Figure 22 – Low-frequency profiles of a sharp (black line) and a blurry (blue line) image.



Source: Elaborated by the author.

The kurtosis array for every crop size is computed as follows. The crop size initializes as zero and is used to compute the kurtosis of the entire set $\{x_1, x_2, ..., x_n\}$ of all descriptors. The crop size is then incremented by 1, yielding the subset $\{x_2, x_3, ..., x_n\}$. This process is done until the kurtosis for all crop sizes of the feature vectors is computed. Algorithm 1 summarizes the computation of kurtosis for all crop sizes:

Algorithm 1 – Kurtosis computation

```
// X_{c×n}: dataset of n descriptors with size c ∈ C, where
// C = {0, 1, ..., size(descriptor)}

// T(X): operator from equation 6.5 to map the
// descriptors onto probability distributions

X ← T(X)
A ← zeros(c, n)
for each crop size c in C do
    for each descriptor i in {1, 2, ..., n} do
        A[crop][i] ← kurtosis(X[i].subset(0, crop))
    end for
end for
return A
```

Next, the optimal threshold computation is done with algorithm 2. It aims to produce a crop size to maximize the peak to peak value, i.e. the difference between the maximum and minimum values in the dataset. This allows low-frequency information to be discarded from the PSF without loss of the blur.

Algorithm 2 – Find the optimal dataset variability threshold

```
// A_{c×n}: matrix with kurtosis values for all n descriptors that were computed at every crop
// size c ∈ C, where
// C = {0, 1, ..., size(descriptor)}

threshold ← 0
maximum ← ∞
for each crop size c in C do
    row ← {A_{c,1}, A_{c,2}, ..., A_{c,n}}
    a ← max(row)
    b ← min(row)

    if a < 0 or b < 0 then
        continue
    end if

    range ← a − b

    if range > maximum then
        maximum ← range
        threshold ← c
    end if
end for
return threshold
```

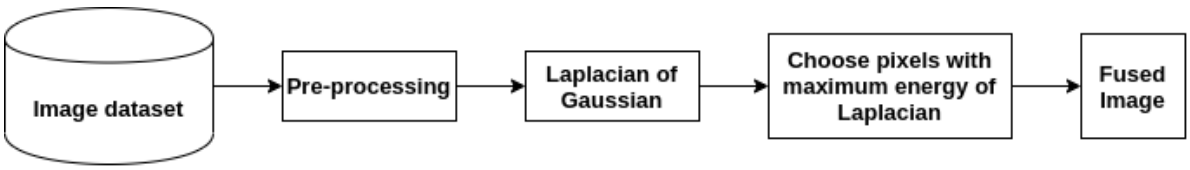
The low-frequency is discarded in each feature vector up to the optimal threshold, and

the logarithm of a final kurtosis value is computed for all of them, in order to enhance the differences between values. In this work, such final kurtosis value of each feature vector of an image corresponds to the amount of details, i.e. high-frequency content on it: the higher the kurtosis value, the sharper the image is. Finally, an estimate of the subset of images that is eligible for the fusion process is then obtained by the z-score transformation by choosing the images that locate at one or more standard deviation units away from the mean.

6.5 Laplacian of Gaussian-based Multi-focus Bright-field Microscopy Image Fusion

The NR-IQA method provided a sharpness measure for each image of the datasets and such numerical representations were analyzed with the z-score; at this point, the estimate of eligible images was prompted to the user and the real eligible ones were chosen. The obtained subset is registered by means of the SIFT-based z-stack alignment tool of the TrakEM2 package. The subsequent stage is to perform the fusion of images in the before-mentioned subset. We propose a multi-focus bright-field microscopy image fusion algorithm with the Laplacian of Gaussian edge detection filter and the energy of edges. Fig. 23 shows a diagram of the proposed method.

Figure 23 – Diagram of the proposed multi-focus bright-field microscopy image fusion algorithm.



Source: Elaborated by the author.

The Laplacian of Gaussian is a standard method for feature extraction or multiscale filtering. In contrast to some of the literature methods presented in Chapter 5, which include several parameters that may affect the accuracy and even computational performance of the method, the Laplacian of Gaussian only requires the standard deviation of the Gaussian function as a parameter. However, some literature methods rely on it during their processing. It was chosen, therefore, due to its simplicity in terms of parameters, usage and efficiency concerning image fusion. The standard deviation of the Gaussian function was empirically set according to the proposed bright-field microscopy image datasets. Furthermore, it might be claimed that the differences in the edges of blurred and sharp images are not a relevant and efficient as a fusion rule, as noisy signals result in false edges. The Gaussian filter in the Laplacian of Gaussian operator is capable not only of smoothing the image but reducing the noise levels, which result in a precise fusion rule.

The grayscale version of the selected images is again used for fusion; since the images were converted with the luminance method, the edges are preserved and therefore the edge detection may be successful. Other methods from the *Luminance* family, e.g. *Luminance*, *Luminance´*, *Luma* and *Decolorize*, are also suitable (KANAN; COTTRELL, 2012).

The images then undergo the Laplacian of Gaussian filter - a spatial filtering algorithm that extracts edges of a smoothed image. The Laplacian is a second-order derivative isotropic linear operator, based on the Laplacian of a function, which defines an edge as the zero-crossing of the second derivative of a function. First and foremost, the Laplacian of a function is denoted by

$$\nabla^2 g(x,y) = \frac{\partial^2 g(x,y)}{\partial x^2} + \frac{\partial^2 g(x,y)}{\partial y^2}, \tag{6.6}$$

where $g(x,y)$ is the input image. For the discrete case (which applies to the digital images), the Laplacian operator is achieved by means of a convolution. The approximation of the second derivatives in each dimension yields the following convolution kernel

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Similarly, instead of extracting edges with the Laplacian filter only, the Laplacian of Gaussian filter performs a Gaussian filtering process to remove noise and smooth the images before retrieving the edges (MARR; HILDRETH, 1980). This approach applies to cases where the quality and reliability of edges obtained by the Laplacian operator are sensitive to noise, and also plays the role of distinguishing the blurry regions from the sharp ones in our approach; pixels belonging to sharp regions will suffer a stronger smoothing effect in comparison to blurry pixels. For our application, an image is convolved with a two-dimensional Gaussian filter, denoted by

$$S(x,y) = g(x,y) * \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2+y^2}{\sigma^2}}, \tag{6.7}$$

where $S(x,y)$ is the smoothed version of the image, $g(x,y)$ is the input image and σ is the standard deviation of the Gaussian function. Then, the Laplacian is applied to the smoothed image S. Both operators are linear and shift-invariant and therefore can be applied in any order, but the Gaussian filter is commonly applied first if the application required separated filters. In our case, the Laplacian of Gaussian operator can be defined as

$$LoG(x,y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{(x^2+y^2)}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}}. \tag{6.8}$$

The σ parameter is responsible for the smoothing magnitude. The convolution kernel is then a combination of both Laplacian and Gaussian filters; higher σ values increase the effective kernel size consequently the smoothing effect. The next step is to retrieve the energy for each region, also based on the fact that blurry regions of an image present less high-frequency components than sharp regions. This stems from the fact that sharp regions are prone to have more edges than smoothed ones, hence a higher energy level. The higher the number the edges in an area of the image, the better the focus on it. Therefore, pixels that correspond to edges are chosen to form the fused image. Their spatial location is used to construct an RGB image by retrieving the content of the three channels from those pixels and assigning them to the final image.

CHAPTER

7

RESULTS

Experiments have been carried out for the two proposed methods: a) the NR-IQA and b) multi-focus image fusion, both with bright-field microscopy. In this chapter, we report the performance of our methods on the created datasets and described in Chapter 6. We also present the quantitative evaluation metrics used. Both IQA and image fusion experiments have been conducted on an Intel Core i7 CPU computer with 8 GB RAM, running Ubuntu Linux 18.04 64-bit.

According to Wang and Li (2011), evaluation metrics such as the Pearson Linear Correlation Coefficient (PLCC), the Spearman's Rank Correlation Coefficient (SRCC) and the Kendall's Rank Correlation Coefficient (KRCC) are quantitative techniques which are more appropriate to our goals. Pursuant to Chen, Smithson and Popovich (2002), the PLCC (also named Pearson's r coefficient) is the most widely used correlation coefficient. It characterizes the degree of the association between linearly related variables and is given by

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 \left(\sum x_i\right)^2 n\sum y_i^2 \left(\sum y_i\right)^2}}, \tag{7.1}$$

where r_{xy} is the r correlation coefficient between the x and y variables, n is the number of observations of both variables, x_i is the value of x for the ith observation and y_i is the value of y for the ith observation. Still, according to Chen, Smithson and Popovich (2002), the KRCC and SRCC are non-parametric tests that also measure the strength of dependence between two variables. They may be denoted as

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n\left(n-1\right)} \tag{7.2}$$

$$\rho = 1 - \frac{6\sum d_i^2}{n\left(n^2-1\right)}, \tag{7.3}$$

where τ and ρ denote the KRCC and SRCC measures, respectively. For both equations, n is the number of observations of both variables. In Equation 7.2, n_c and n_d are the number of concordant and discordant observations, i.e. ranked in the same and opposite ways, respectively. In addition, d_i in Equation 7.3 denotes the difference between the ranks of corresponding variables.

Similarly, we propose the use of Entropy, Spatial Frequency (SF) and the Standard Deviation (STD) for the evaluation of our image fusion method (NAIDU; RAOL, 2008). The Entropy measures the information content of an image and is denoted as

$$He = -\sum_{i=0}^{L} h(i)\log_2 h(i), \tag{7.4}$$

where $h(i)$ is the normalized histogram of the fused image and L is the number of bins of such histogram. The higher the Entropy value, the more details the fused image has, i.e. the sharper it is. The Spatial Frequency presents the overall activity level in the fused image by means of the amount of information in the rows and columns, given by

$$RF = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=2}^{N}[I(x,y)-I(x,y-1)]^2}$$

$$CF = \sqrt{\frac{1}{MN}\sum_{y=1}^{N}\sum_{x=2}^{M}[I(x,y)-I(x-1,y)]^2}$$

$$SF = \sqrt{RF^2+CF^2}, \tag{7.5}$$

where $I(x,y)$ is the fused image and M, N are the image width and height, respectively. Higher values indicate better fusion quality since more activity means less homogeneous regions in our case. Finally, the Standard Deviation measures the contrast in the fused image, and therefore higher values yield higher contrast. It can be computed as

$$STD = \sqrt{\sum_{i=0}^{L}(i-\bar{i})^2 h(i)}, \qquad \bar{i} = \sum_{i=0}^{L} i h(i). \tag{7.6}$$

with $h(i)$ and L as defined in Equation 7.4.

7.1 Image Quality Assessment Evaluation

We compared our results with well-known NR-IQA approaches such as MLV (Bahrami; Kot, 2014), S_3 (Vu; Phan; Chandler, 2012), JNB (Ferzli; Karam, 2009), CPBD (NARVEKAR;

KARAM, 2011), Marziliano *et. al.* (Marziliano *et al.*, 2002) and Kanjar (DE; MASILAMANI, 2013). Table 2 shows the results of the PLCC, SRCC and KRCC metrics for the *Callisia*, *Tradescantia* and *Cthenante* datasets, respectively. For all proposed datasets, the correlation coefficients were computed against the Mean Structural Similarity Index between the resulting image of our fusion method and the images.

Table 2 – Performance comparison of our proposed method and other NR-IQA metrics on the microscopy images datasets, in terms of correlation coefficients between the MSSIM quality score and each method.

Dataset	Index	MLV	S_3	JNB	CPBD	Marz.	Kanjar	**Proposed**
Callisia	PLCC	0.1462	0.0140	0.7336	0.8606	0.8663	0.8167	**0.8155**
	SRCC	0.3099	0.3654	0.9623	0.9658	0.9671	0.9594	**0.9364**
	KRCC	0.2130	0.2558	0.8584	0.8636	0.8701	0.8506	**0.7961**
Tradescantia	PLCC	0.1910	0.0466	0.2867	0.2565	0.2683	0.3400	**0.4708**
	SRCC	0.0249	0.0184	0.6052	0.6023	0.5159	0.6544	**0.5730**
	KRCC	0.0117	0.0107	0.5012	0.4825	0.3930	0.5798	**0.4350**
Cthenante	PLCC	0.0760	0.2346	0.9463	0.8428	0.9490	0.8968	**0.9482**
	SRCC	0.1289	0.2450	0.9617	0.8821	0.9697	0.9650	**0.9466**
	KRCC	0.0976	0.1717	0.8384	0.7212	0.8653	0.8545	**0.8020**
Mean	PLCC	0.1377	0.0984	0.6556	0.6533	0.6945	0.6845	**0.7448**
	SRCC	0.1546	0.2096	0.8430	0.8168	0.8176	0.8596	**0.8187**
	KRCC	0.1074	0.1461	0.7327	0.6891	0.7095	0.7617	**0.6777**

Source: Elaborated by the author.

The difference between our images and benchmark datasets such as LIVE (Sheikh; Sabir; Bovik, 2006) and CSIQ (LARSON; CHANDLER, 2010) is that our bright-field microscopy images are subjected to a non-homogeneous blur kernel and the spherical aberrations are more prominent. Consequently, the labeling of blurred and sharp is subjective, since the notion of quality might be different according to what the images will be used for. In the scope of this work, the reason for assessing image quality is to predict eligible images and select the proper ones to perform the fusion process. The SRCC and KRCC coefficients also evaluate the *monotonicity*, i.e. the property of maintaining the order relation between the sets - it is only nondecreasing or nonincreasing. In this work, monotonicity relates the values obtained by applying our metric on the proposed datasets and the set of labels provided by the subject evaluation of them.

The Kanjar method was also implemented in Python, and the code is also organized in a repository. The methods we used in our comparisons were implemented in MATLAB, C++ and Python programming languages. Details are summarized in Table 3, including the repository with the implementation of our method:

Table 3 – Implementations of the literature IQA methods and ours.

Method	Environment	Implementation
MLV	MATLAB	<https://sites.google.com/site/khosrobahrami2010/publications>
S_3	MATLAB	<http://vision.eng.shizuoka.ac.jp/s3/>
JNB	MATLAB	<https://ivulab.asu.edu/software/jnbm/>
CPBD	Python	<https://pypi.org/project/cpbd/>
Marz.	C++	<https://github.com/PeterWang1986/blur>
Kanjar	Python	<https://github.com/vaugusto92/kanjar-nr-iqa>
Proposed	C++	<https://github.com/vaugusto92/fourier-light-microscopy-nr-ism>

7.2 Image Fusion

The proposed method was implemented in Python programming language with the NumPy, Scipy, scikit-image and Pillow libraries. Following reproducible research ideals, the code is organized and made available in a repository [1].

We compared our proposed method with well-known multi-focus image fusion approaches such as PCA (NAIDU; RAOL, 2008), GF (LI; KANG; HU, 2013), MSWG (ZHOU; LI; WANG, 2014), and MSVD (NAIDU, 2011). The optimal value for the σ parameter was empirically acquired and set to 0.7. Table 4 shows the results of the Entropy, MSSIM and STD metrics for the *Callisia*, *Tradescantia* and *Cthenante* datasets with each fusion approach.

Table 4 – Objective performance evaluation of the proposed method ($\sigma = 0.7$) and other image fusion approaches.

Dataset	Index	PCA	GF	MSWG	MSVD	**Proposed**
Callisia	Entropy	10.9928	11.3332	11.7052	11.4759	**12.1904**
	SF	0.0253	0.0347	0.0441	0.0424	**0.0836**
	STD	0.1966	0.1939	0.1949	0.1968	**0.1987**
Tradescantia	Entropy	8.4619	9.2162	9.2120	8.6751	**9.3011**
	SF	0.0167	0.0249	0.0250	0.0219	**0.0286**
	STD	0.0809	0.0826	0.0825	0.0811	**0.0816**
Cthenante	Entropy	6.7263	6.7577	3.6411	7.7962	**4.3565**
	SF	0.0317	0.0473	0.0645	0.0482	**0.0881**
	STD	0.0810	0.0760	0.0815	0.0827	**0.1117**

Source: Elaborated by the author.

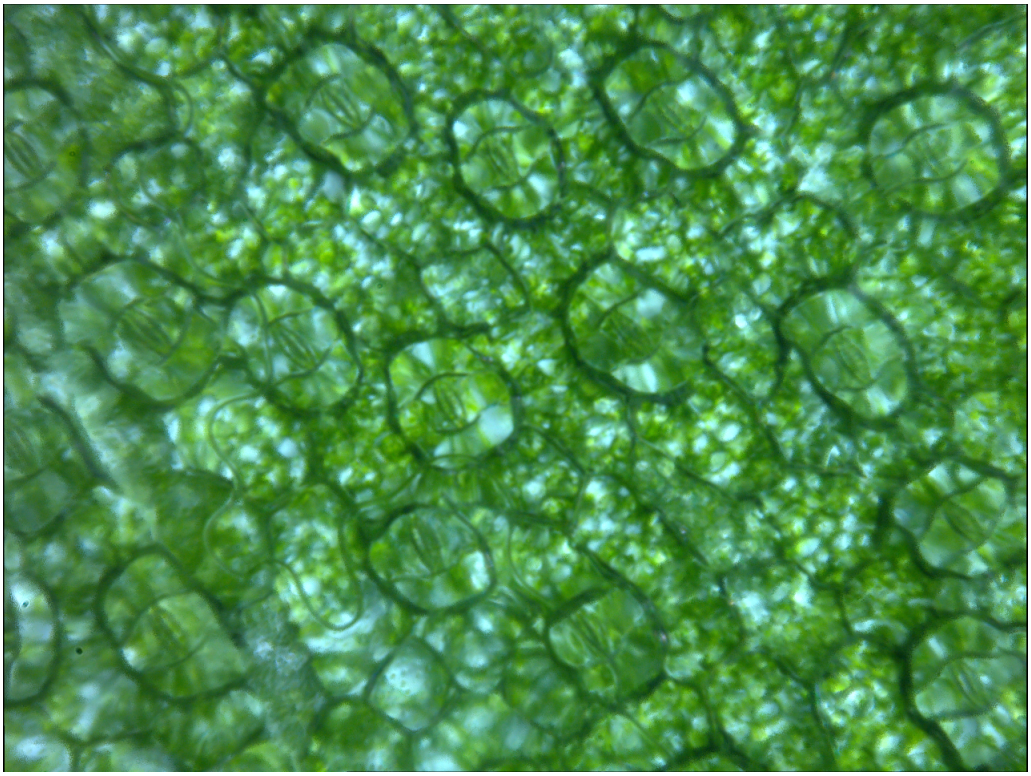
Our image dataset differs from benchmarks such as the Lytro (NEJATI; SAMAVI; SHIRANI, 2015). As seen in Chapter 3, the image formation is subject to anisotropic blur kernels due to the optical properties of the lenses in bright-field microscopy. In the benchmark images, it

[1] <https://github.com/vaugusto92/light-microscopy-image-fusion-prototype>

is possible to point out precisely the blurred and sharp regions, as there are usually two or three images of each scene. Our datasets, on the other hand, do not allow this. Thus, the fusion results for benchmark datasets reach optimal performance, and some of them even contain the reference image for a better comparison. Finally, the fusion rule should be robust to noise and should have a large variation with respect to the degree of blurring (HUANG; ZHONGLIANG, 2007). The Gaussian smoothing procedure provides robustness to noise in our algorithm.

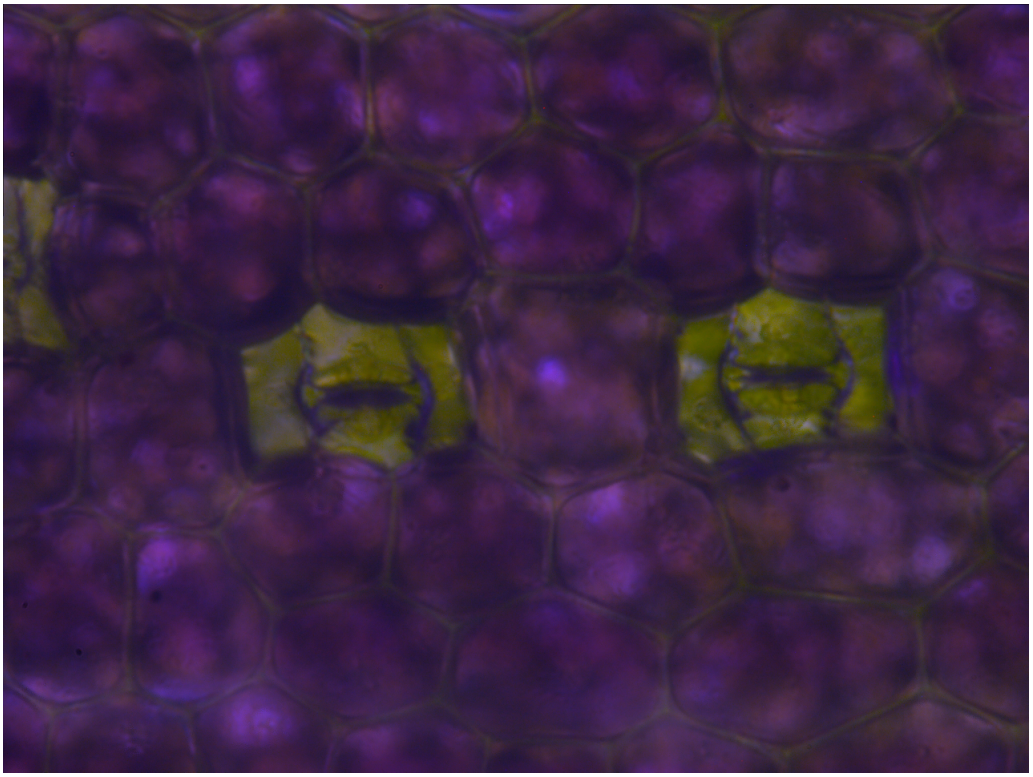
We also present the qualitative results, i.e. the fused images and some samples of each dataset. With the results from our IQA method, we have selected the eligible slices in each dataset and registered with the TrakEM2 alignment tool. The *Callisia* dataset has six slices which present slight differences between the background of the leaf and the stomata. The images in this dataset are the brightest among all images in our datasets. Two eligible slices were obtained in the *Tradescantia* dataset due to its large magnification, where either background sections or stomata are focused. Also, images in this dataset are not as bright as *Callisia* ones, as the specimens present a strong purple color in their abaxial region. Finally, the *Cthenante* dataset presents the strongest difference in focused regions among all datasets and also the darkest images. The results of the fusion process of *Callisia*, *Tradescantia* and *Cthennante* are shown in Figure 24, Figure 25 and Figure 26, respectively, while Figure 27 depicts samples of registered images in each dataset.

Figure 24 – Fused *Callisia* image.



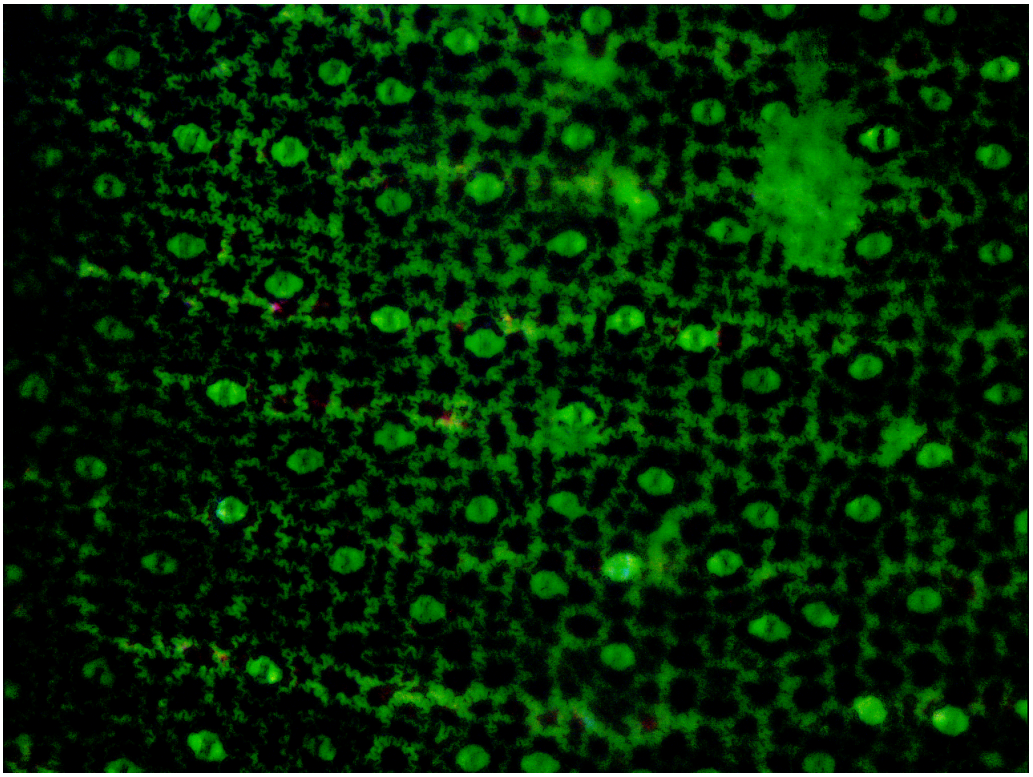
Source: Elaborated by the author.

Figure 25 – Fused *Tradescantia* image.



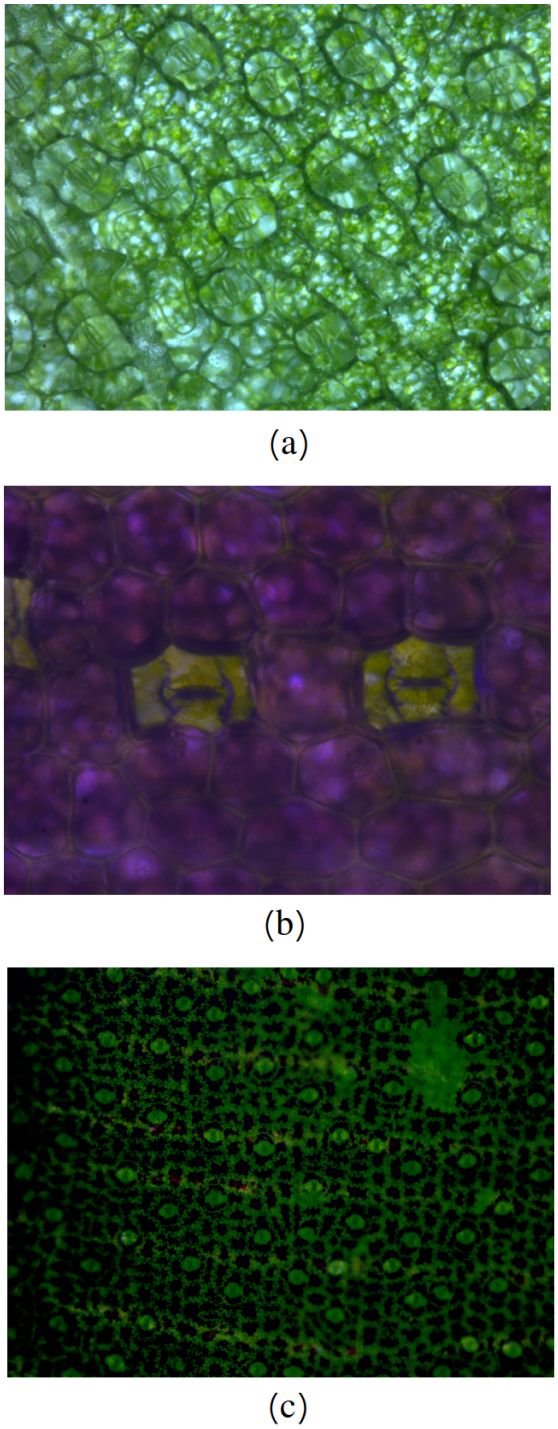
Source: Elaborated by the author.

Figure 26 – Fused *Cthenante* image.



Source: Elaborated by the author.

Figure 27 – Registered slices of our *Callisia* (a), *Tradescantia* (b) and *Cthenante* (c) datasets.



Source: Elaborated by the author.

Each of the images in Figure 27 represent the middle of the z-stack. For example, Figure 27.(c) is blurred on the left and right parts and sharp in the middle section; this shows that the leaf has a slope structure that promoted a clear division between blurred and sharp regions.

CHAPTER

8

CONCLUSIONS

This work proposed a no-reference DFT-based image quality metric that produced reliable results, which highly correlate with the labels obtained by subjective analysis. Additionally, a multi-focus image fusion algorithm for bright-field microscopy z-stacks that explores the energy of edges extracted with a Laplacian of Gaussian operator proved to be an efficient way to perform image fusion. Both of our hypotheses that methods based on frequency domain information and the Laplacian of Gaussian operator were confirmed experimentally in the performance comparison with related methods.

Concerning our IQA method, the implementation is efficient in terms of computational performance and it may be extended to other applications such as auto-focus systems. Future work on this method comprises several possibilities. The improvement of the computational performance of the method by refactoring the implementation and applying parallel programming techniques and the integration of the software with hardware devices may be the basis for new imaging solutions. The study and empirical evaluation of other methods to estimate the set of eligible images rather than z-score, which includes more advanced statistical methods and machine learning methods, may provide even better results.

The most obvious improvement is to implement it in C++ in order to improve computational performance since our Python implementation is a prototype. Next, we should study and evaluate the availability of information in each channel of the RGB images, as it may suppress the need of a grayscale conversion and also may yield better results. Additionally, different smoothing functions and other detection algorithms should be evaluated, and those should also be robust to noise and enhance the difference between blurred and sharp regions. Finally, the research on the availability of different fusion rules rather than the energy of edges and the use of frequency-domain convolution may lead to better results and faster computation, respectively.

The set of mathematical methods and techniques applied in this work are standards in image processing. All of them have been extensively explored during the past few decades, and

this work is another example which underpins the strength of such methods. However, there is still a lot to be explored in terms of tuning parameters and extracting more from the theoretical side of each technique. This approach is important, since it helps to find limitations and also to build novel tools.

An interesting and important insight from this work is the use of a full-reference image quality metric in order to evaluate a no-reference one. This evolves to a brand new area of study, since there are many benchmark methods, e.g. the MSSIM itself, which can be applied in this scenario. The evaluation of a blind metric may be more effective under a combination of several non-blind metrics. The whole IQA area would benefit from a novel validation protocol based on well-known methods.

Publications

- (CATANANTE; BRUNO; NETO, 2020) Victor Augusto Alves Catanante, Odemir Martinez Bruno, and João do Espírito Santo Batista Neto. “Frequency domain kurtosis-based no-reference image quality assessment for bright-field microscopy images”. **To be submitted**.

BIBLIOGRAPHY

Bahrami, K.; Kot, A. C. A fast approach for no-reference image sharpness assessment based on maximum local variation. **IEEE Signal Processing Letters**, v. 21, n. 6, p. 751–755, June 2014. Citations on pages 62 and 84.

BELL, S.; MORRIS, K. **An Introduction to Microscopy**. CRC PressINC, 2009. (Forensics and criminal justice). ISBN 9781420084504. Available: <https://books.google.com.br/books?id=3CFFAQAAIAAJ>. Citations on pages 34 and 101.

BORN, M.; WOLF, E.; BHATIA, A. B.; CLEMMOW, P. C.; GABOR, D.; STOKES, A. R.; TAYLOR, A. M.; WAYMAN, P. A.; WILCOCK, W. L. **Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light**. 7. ed. [S.l.]: Cambridge University Press, 1999. Citations on pages 102 and 103.

BOVIK, A. C. **The essential guide to image processing**. [S.l.]: Academic Press, 2009. Citations on pages 53 and 115.

BOYDE, A. **Modern Microscopies: Techniques and Applications**. Boston, MA: Springer US, 1990. ISBN 978-1-4613-1467-7. Available: <https://doi.org/10.1007/978-1-4613-1467-7_10>. Citation on page 33.

BRACEWELL, R.; BRACEWELL, R. **The Fourier Transform and Its Applications**. [S.l.]: McGraw Hill, 2000. (Electrical engineering series). ISBN 9780073039381. Citations on pages 45, 47, 48, 49, and 109.

BRIGHAM, E. **The Fast Fourier Transform and Its Applications**. Prentice Hall, 1988. (Prentice-Hall Signal Processing Series: Advanced monographs). ISBN 9780133075052. Available: <https://books.google.com.br/books?id=XfJQAAAAMAAJ>. Citations on pages 43, 47, 48, and 49.

BROWN, L. G. A survey of image registration techniques. **ACM computing surveys (CSUR)**, ACM New York, NY, USA, v. 24, n. 4, p. 325–376, 1992. Citation on page 113.

CASTLEMAN, K. R. **Digital Image Processing**. USA: Prentice Hall Press, 1996. ISBN 0132114674. Citations on pages 42, 43, 44, 109, and 111.

CATANANTE, V. A. A.; BRUNO, O. M.; NETO, J. d. E. S. Frequency domain kurtosis-based no-reference image quality assessment for bright-field microscopy images. **arXiv preprint arXiv:2003.11526**, 2020. Citation on page 92.

CHEN, P. Y.; SMITHSON, M.; POPOVICH, P. M. **Correlation: Parametric and nonparametric measures**. [S.l.]: Sage, 2002. Citation on page 83.

ÇINLAR, E. **Probability and stochastics**. [S.l.]: Springer Science & Business Media, 2011. Citation on page 122.

COSTA, M. G. F.; PINTO, K.; FUJIMOTO, L. B.; OGUSKU, M. M.; FILHO, C. F. C. Multi-focus image fusion for bacilli images in conventional sputum smear microscopy for tuberculosis. **Biomedical Signal Processing and Control**, Elsevier, v. 49, p. 289–297, 2019. Citation on page 30.

CROFT, H.; FALCONER, K.; GUY, R. **Unsolved Problems in Geometry: Unsolved Problems in Intuitive Mathematics**. Springer New York, 2012. (Problem Books in Mathematics). ISBN 9781461209638. Available: <https://books.google.com.br/books?id=rdDTBwAAQBAJ>. Citation on page 122.

DAVIDSON, M. W.; ABRAMOWITZ, M. Optical microscopy. In: ____. **Encyclopedia of Imaging Science and Technology**. American Cancer Society, 2002. ISBN 9780471443391. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471443395.img074>. Citations on pages 40 and 41.

DE, K.; MASILAMANI, V. Image sharpness measure for blurred images in frequency domain. **Procedia Engineering**, v. 64, 12 2013. Citations on pages 58 and 85.

DEGROOT, M. H.; SCHERVISH, M. J. **Probability and statistics**. [S.l.]: Pearson Education, 2012. Citations on pages 55, 117, 118, and 119.

DEVORE, J. L. **Probability and Statistics for Engineering and the Sciences**. [S.l.]: Cengage learning, 2011. Citations on pages 55, 116, 117, and 118.

DOKLAND, T.; HUTMACHER, D. W.; NG, M. M.-L.; SCHANTZ, J.-T. **Techniques in Microscopy for Biomedical Applications**. WORLD SCIENTIFIC, 2006. Available: <https://www.worldscientific.com/doi/abs/10.1142/5911>. Citation on page 34.

DUNST, S.; TOMANCAK, P. Imaging flies by fluorescence microscopy: Principles, technologies, and applications. **Genetics**, Genetics, v. 211, n. 1, p. 15–34, 2019. ISSN 0016-6731. Available: <https://www.genetics.org/content/211/1/15>. Citation on page 41.

Ferzli, R.; Karam, L. J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). **IEEE Transactions on Image Processing**, v. 18, n. 4, p. 717–728, April 2009. Citations on pages 58, 60, and 84.

FILHO, H. A.; BRUNO, O. M. Plants with purple abaxial leaves: A repository of metrics from stomata distribution. **bioRxiv**, Cold Spring Harbor Laboratory, 2018. Available: <https://www.biorxiv.org/content/early/2018/04/04/294553>. Citations on pages 68 and 69.

FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, ACM New York, NY, USA, v. 24, n. 6, p. 381–395, 1981. Citation on page 52.

FOLLAND, G. **Real Analysis: Modern Techniques and Their Applications**. Wiley, 2013. (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts). ISBN 9781118626399. Available: <https://books.google.com.br/books?id=wI4fAwAAQBAJ>. Citations on pages 121, 122, 123, 124, 125, and 126.

FOWLES, G. **Introduction to Modern Optics**. Dover Publications, 1989. (Dover Books on Physics Series). ISBN 9780486659572. Available: <https://books.google.com.br/books?id=SL1n9TuJ5YMC>. Citations on pages 40, 101, 102, and 104.

GARLING, D. J. **A Course in Mathematical Analysis: Volume 3, Complex Analysis, Measure and Integration**. [S.l.]: Cambridge University Press, 2014. Citations on pages 126 and 127.

GAUGLITZ, G.; VO-DINH, T. **Handbook of Spectroscopy**. Wiley, 2006. ISBN 9783527605026. Available: <https://books.google.com.br/books?id=RkgGVFlck6QC>. Citation on page 101.

GHEONDEA, A. On locally hilbert spaces. **Opuscula Mathematica**, v. 36, p. 735, 01 2016. Citation on page 123.

GIUSTI, A.; TADDEI, P.; CORANI, G.; GAMBARDELLA, L.; MAGLI, C.; GIANAROLI, L. Artificial defocus for displaying markers in microscopy z-stacks. **IEEE transactions on visualization and computer graphics**, IEEE, v. 17, n. 12, p. 1757–1764, 2011. Citation on page 37.

GONZALEZ, R.; WOODS, R. **Digital Image Processing**. [S.l.]: Pearson/Prentice Hall, 2009. ISBN 9780131687288. Citations on pages 45 and 110.

____. **Digital Image Processing**. Pearson, 2018. ISBN 9780133356724. Available: <https://books.google.com.br/books?id=0F05vgAACAAJ>. Citations on pages 39, 43, 47, 48, 109, 110, 111, 112, and 113.

GOODMAN, J. **Introduction to Fourier Optics**. McGraw-Hill, 1996. (McGraw-Hill Series in Electrical and Computer Engineering: Communications and Signal Processing). ISBN 9780070242548. Available: <https://books.google.com.br/books?id=QllRAAAAMAAJ>. Citations on pages 41 and 43.

GOSHTASBY, A. A. **Image registration: Principles, tools and methods**. [S.l.]: Springer Science & Business Media, 2012. Citation on page 113.

HALLIDAY, D.; RESNICK, R.; WALKER, J. **Fundamentals of Physics Extended, 10th Edition**. Hoboken, New Jersey: Wiley, 2013. ISBN 9781118473818. Citations on pages 102, 103, 104, and 106.

HETHERINGTON, A. M.; WOODWARD, F. I. The role of stomata in sensing and driving environmental change. **Nature**, Nature Publishing Group, v. 424, n. 6951, p. 901–908, 2003. Citation on page 69.

Hosu, V.; Lin, H.; Sziranyi, T.; Saupe, D. Koniq-10k: An ecologically valid database for deep learning of blind image quality assessment. **IEEE Transactions on Image Processing**, v. 29, p. 4041–4056, 2020. Citations on pages 71 and 72.

HUANG, W.; ZHONGLIANG, J. Evaluation of focus measures in multi-focus image fusion. **Pattern Recognition Letters**, v. 28, n. 4, p. 493 – 500, 2007. ISSN 0167-8655. Available: <http://www.sciencedirect.com/science/article/pii/S0167865506002352>. Citations on pages 63 and 87.

JAMES, A. P.; DASARATHY, B. V. Medical image fusion: A survey of the state of the art. **Information Fusion**, Elsevier, v. 19, n. 1, p. 4–19, 2014. ISSN 1566-2535. Citations on pages 52 and 114.

JOSHI, N. Defocus blur. In: ____. **Computer Vision: A Reference Guide**. Boston, MA: Springer US, 2014. p. 171–173. ISBN 978-0-387-31439-6. Available: <https://doi.org/10.1007/978-0-387-31439-6_511>. Citation on page 45.

KANAN, C.; COTTRELL, G. W. Color-to-grayscale: does the method matter in image recognition? **PloS one**, Public Library of Science, v. 7, n. 1, 2012. Citation on page 80.

KOHO, S.; FAZELI, E.; ERIKSSON, J. E.; HÄNNINEN, P. E. Image quality ranking method for microscopy. **Scientific reports**, Nature Publishing Group, v. 6, p. 28962, 2016. Citation on page 31.

LARSON, E.; CHANDLER, D. Most apparent distortion: Full-reference image quality assessment and the role of strategy. **J. Electronic Imaging**, v. 19, p. 011006, 01 2010. Citations on pages 71 and 85.

LAWLOR, D. **Introduction to Light Microscopy: Tips and Tricks for Beginners**. Springer International Publishing, 2019. ISBN 9783030053925. Available: <https://books.google.com.br/books?id=cRLovQEACAAJ>. Citations on pages 33, 34, 35, 36, and 107.

LENG, Y. **Materials characterization: introduction to microscopic and spectroscopic methods**. [S.l.]: John Wiley & Sons, 2009. Citation on page 40.

LI, S.; KANG, X.; HU, J. Image fusion with guided filtering. **IEEE Transactions on Image processing**, IEEE, v. 22, n. 7, p. 2864–2875, 2013. Citations on pages 64 and 86.

LIU, A.; WANG, J.; LIU, J.; SU, Y. Comprehensive image quality assessment via predicting the distribution of opinion score. **Multimedia Tools and Applications**, Springer, v. 78, n. 17, p. 24205–24222, 2019. Citation on page 70.

LOWE, D. G. Object recognition from local scale-invariant features. In: IEEE. **Proceedings of the seventh IEEE international conference on computer vision**. [S.l.], 1999. v. 2, p. 1150–1157. Citation on page 50.

____. Distinctive image features from scale-invariant keypoints. **International journal of computer vision**, Springer, v. 60, n. 2, p. 91–110, 2004. Citation on page 50.

MARR, D.; HILDRETH, E. Theory of edge detection. **Proceedings of the Royal Society of London. Series B. Biological Sciences**, The Royal Society London, v. 207, n. 1167, p. 187–217, 1980. Citation on page 80.

Marziliano, P.; Dufaux, F.; Winkler, S.; Ebrahimi, T. A no-reference perceptual blur metric. In: **Proceedings. International Conference on Image Processing**. [S.l.: s.n.], 2002. v. 3, p. III–III. ISSN 1522-4880. Citations on pages 58 and 85.

MATHEWS, J.; WALKER, R. **Mathematical Methods of Physics**. W. A. Benjamin, 1970. (Addison-Wesley World Student Series). ISBN 9780805370027. Available: <https://books.google.com.br/books?id=1iHvAAAAMAAJ>. Citation on page 40.

MENDENHALL, W. M.; SINCICH, T. L. **Statistics for Engineering and the Sciences**. [S.l.]: CRC Press, 2016. Citations on pages 55, 116, 117, and 118.

MITCHELL, H. B. **Image Fusion: Theories, Techniques and Applications**. 1st. ed. Berlin: Springer Verlag, 2010. ISBN 3642112153, 9783642112157. Citations on pages 52 and 114.

MURPHY, D.; DAVIDSON, M. **Fundamentals of Light Microscopy and Electronic Imaging**. [S.l.]: Wiley, 2012. ISBN 9781118382936. Citations on pages 33, 105, and 107.

NAIDU, V. Image fusion technique using multi-resolution singular value decomposition. **Defence Science Journal**, Defence Scientific Information Documentation Centre, v. 61, n. 5, p. 479–484, 2011. Citations on pages 66 and 86.

NAIDU, V.; RAOL, J. R. Pixel-level image fusion using wavelets and principal component analysis. **Defence Science Journal**, Defence Scientific Information & Documentation Centre, v. 58, n. 3, p. 338, 2008. Citations on pages 63, 84, and 86.

NARVEKAR, N. D.; KARAM, L. J. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). **Trans. Img. Proc.**, IEEE Press, Piscataway, NJ, USA, v. 20, n. 9, p. 2678–2683, Sep. 2011. ISSN 1057-7149. Available: <http://dx.doi.org/10.1109/TIP.2011.2131660>. Citations on pages 60 and 85.

NEJATI, M.; SAMAVI, S.; SHIRANI, S. Multi-focus image fusion using dictionary-based sparse representation. **Information Fusion**, Elsevier, v. 25, p. 72–84, 2015. Citation on page 86.

PAJARES, G.; CRUZ, J. M. de la. A wavelet-based image fusion tutorial. **Pattern Recognition**, v. 37, n. 9, p. 1855 – 1872, 2004. ISSN 0031-3203. Available: <http://www.sciencedirect.com/science/article/pii/S0031320304001037>. Citation on page 57.

PETROVIĆ, S.; ZVEZDANOVIĆ, J.; MARKOVIĆ, D. Chlorophyll degradation in aqueous mediums induced by light and uv-b irradiation: An uhplc-esi-ms study. **Radiation Physics and Chemistry**, v. 141, p. 8 – 16, 2017. ISSN 0969-806X. Available: <http://www.sciencedirect.com/science/article/pii/S0969806X16307927>. Citation on page 36.

PONTI, M.; NAZARÉ, T. S.; THUMÉ, G. S. Image quantization as a dimensionality reduction procedure in color and texture feature extraction. **Neurocomputing**, Elsevier, v. 173, p. 385–396, 2016. Citations on pages 30 and 73.

QIU, Y.; CHEN, X.; LI, H.; CHEN, W. R.; ZHENG, B.; LI, S.; LIU, H. Evaluations of auto-focusing methods under a microscopic imaging modality for metaphase chromosome image analysis. **Analytical Cellular Pathology**, v. 36, p. 37–44, August 2013. Citation on page 29.

RICHARDSON, W. H. Bayesian-based iterative method of image restoration. **J. Opt. Soc. Am.**, OSA, v. 62, n. 1, p. 55–59, Jan 1972. Citation on page 30.

ROANE, T.; PEPPER, I.; MAIER, R. Microscopic techniques. In: ____. **Environmental Microbiology**. [S.l.]: Elsevier Inc., 2009. p. 157–172. ISBN 9780123705198. Citation on page 35.

ROCHOW, T.; TUCKER, P. **Introduction to Microscopy by Means of Light, Electrons, X Rays, or Acoustics**. [S.l.]: Springer US, 1994. (Language of science). ISBN 9780306446849. Citations on pages 34 and 35.

ROTTENFUSSER, R.; WILSON, E. E.; DAVIDSON, M. W. **Contrast Modes in Reflected Light Microscopy**. 2020. Available: <http://zeiss-campus.magnet.fsu.edu/articles/basics/reflectedcontrast.html>. Accessed: 29/03/2011. Citation on page 39.

ROWLAND, T.; WEISSTEIN, E. W. **Compact Support. From MathWorld–A Wolfram Web Resource.** 2020. Citation on page 76.

ROYDEN, H. **Real Analysis**. Macmillan, 1988. (Mathematics and statistics). ISBN 9780024041517. Available: <https://books.google.com.br/books?id=J4k_AQAAIAAJ>. Citations on pages 123 and 124.

SAALFELD, S. Computational methods for stitching, alignment, and artifact correction of serial section data. In: MüLLER-REICHERT, T.; PIGINO, G. (Ed.). **Three-Dimensional Electron Microscopy**. Academic Press, 2019, (Methods in Cell Biology, v. 152). p. 261 – 276. Available: <http://www.sciencedirect.com/science/article/pii/S0091679X19300585>. Citations on pages 52 and 70.

SCHREIER, H. W.; GARCIA, D.; SUTTON, M. A. Advances in light microscope stereo vision. **Experimental Mechanics**, v. 44, n. 3, p. 278–288, Jun 2004. ISSN 1741-2765. Available: <https://doi.org/10.1007/BF02427894>. Citation on page 35.

Sheikh, H. R.; Sabir, M. F.; Bovik, A. C. A statistical evaluation of recent full reference image quality assessment algorithms. **IEEE Transactions on Image Processing**, v. 15, n. 11, p. 3440–3451, Nov 2006. ISSN 1941-0042. Citation on page 85.

SMITH, F.; THOMSON, J. **Optics**. Wiley, 1988. (Manchester Physics Series). ISBN 9780471915355. Available: <https://books.google.com.br/books?id=qNzGQgAACAAJ>. Citation on page 107.

SMITH, W. **Modern Optical Engineering, 4th Ed.** United States of America: Mcgraw-hill, 2007. (McGraw-Hill series on optical and electro-optical engineering). ISBN 9780071476874. Citation on page 45.

SONALI; SAHU, S.; SINGH, A. K.; GHRERA, S.; ELHOSENY, M. An approach for de-noising and contrast enhancement of retinal fundus image using clahe. **Optics & Laser Technology**, v. 110, p. 87 – 98, 2019. ISSN 0030-3992. Special Issue: Optical Imaging for Extreme Environment. Available: <http://www.sciencedirect.com/science/article/pii/S0030399218306637>. Citation on page 49.

SUN, Y.; DUTHALER, S.; NELSON, B. Autofocusing algorithm selection in computer microscopy. **2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005. (IROS 2005).**, IEEE, August 2005. Citation on page 30.

TANG, L.; SUN, K.; BI, J.; WANG, J. Feature comparison and analysis for new challenging research fields of image quality assessment. **Digital Signal Processing**, v. 91, p. 3 – 10, 2019. ISSN 1051-2004. Quality Perception of Advanced Multimedia Systems. Available: <http://www.sciencedirect.com/science/article/pii/S1051200418308832>. Citations on pages 54, 57, and 115.

TIPLER, P.; MOSCA, G. **Physics for Scientists and Engineers: Mechanics, Oscillations and Waves, Thermodynamics**. W. H. Freeman, 2007. ISBN 9781429201247. Available: <https://books.google.com.br/books?id=BMVR37-8Jh0C>. Citations on pages 102, 103, 104, and 105.

____. **Physics for Scientists and Engineers, Volume 2: Electricity and Magnetism, Light**. W. H. Freeman, 2008. ISBN 9781429237925. Available: <https://books.google.com.br/books?id=3_lVDQEACAAJ>. Citations on pages 103 and 107.

TRIOLA, M. **Elementary Statistics, 13th edition**. [S.l.]: Pearson Education, 2017. ISBN 0134462459. Citation on page 116.

TRIVEDI, M. M.; MILLS, J. K. Centroid calculation of the blastomere from 3d z-stack image data of a 2-cell mouse embryo. **Biomedical Signal Processing and Control**, Elsevier, v. 57, p. 101726, 2020. Citation on page 36.

Vu, C. T.; Phan, T. D.; Chandler, D. M. S_3: A spectral and spatial measure of local perceived sharpness in natural images. **IEEE Transactions on Image Processing**, v. 21, n. 3, p. 934–945, March 2012. Citations on pages 60 and 84.

VYAS, A.; YU, S.; PAIK, J. **Multiscale Transforms with Application to Image Processing**. [S.l.]: Springer, 2017. Citation on page 29.

Wang, Z.; Li, Q. Information content weighting for perceptual image quality assessment. **IEEE Transactions on Image Processing**, v. 20, n. 5, p. 1185–1198, May 2011. ISSN 1941-0042. Citation on page 83.

WEI, L.; ROBERTS, E. Neural network control of focal position during time-lapse microscopy of cells. **Scientific Reports**, v. 8, 12 2018. Citation on page 37.

WEISSTEIN, E. W. **Delta Function. From MathWorld—A Wolfram Web Resource**. 2020. Citation on page 45.

____. **Smooth Function. From MathWorld–A Wolfram Web Resource**. 2020. Citation on page 76.

WU, Q.; MERCHANT, F.; CASTLEMAN, K. **Microscope Image Processing**. [S.l.]: Elsevier, 2008. Citations on pages 29, 33, 42, 43, 44, 111, and 112.

ZHANG, Q.; GUO, B. long. Multifocus image fusion using the nonsubsampled contourlet transform. **Signal Processing**, v. 89, n. 7, p. 1334 – 1346, 2009. ISSN 0165-1684. Available: <http://www.sciencedirect.com/science/article/pii/S0165168409000097>. Citation on page 57.

Zhou Wang; Bovik, A. C.; Sheikh, H. R.; Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. **IEEE Transactions on Image Processing**, v. 13, n. 4, p. 600–612, April 2004. Citations on pages 53, 54, 57, 115, and 116.

ZHOU, Z.; LI, S.; WANG, B. Multi-scale weighted gradient-based fusion for multi-focus images. **Information Fusion**, Elsevier, v. 20, p. 60–72, 2014. Citations on pages 65 and 86.

ZÍLIO, S. **Óptica moderna: fundamentos e aplicações**. Centro de Divulgação Científica e Cultural - CDCC, 2009. ISBN 9788588533417. Available: <https://books.google.com.br/books?id=pWrvSAAACAAJ>. Citations on pages 33, 102, 103, 104, and 105.

ZITOVA, B.; FLUSSER, J. Image registration methods: a survey. **Image and vision computing**, Elsevier, v. 21, n. 11, p. 977–1000, 2003. Citations on pages 112 and 113.

ZUIDERVELD, K. Viii.5. - contrast limited adaptive histogram equalization. In: HECKBERT, P. S. (Ed.). **Graphics Gems**. [S.l.]: Academic Press, 1994. p. 474 – 485. ISBN 978-0-12-336156-1. Citations on pages 49 and 50.

ZWILLINGER, D.; KOKOSKA, S. **CRC standard probability and statistics tables and formulae**. [S.l.]: Crc Press, 1999. Citations on pages 55, 117, 118, and 119.

APPENDIX

A

FUNDAMENTALS OF OPTICS

This appendix provides information about optics and microscopy in order to aid the bright-field microscopy and defocus blur comprehension. Plenty of image degradation is due to the system acquisition process; in fact, defocus is a natural occurrence in optics, mainly caused by adjustments of the optical system. The definition of the spectroscopy procedure consists on the interaction between electromagnetic radiation and the matter (GAUGLITZ; VO-DINH, 2006). This concept can be extended to microscopy, which deals with the range of the electromagnetic spectrum of wavelength between 400 and 700 nanometers, i.e. visible light, to create visual representations of the objects (BELL; MORRIS, 2009). Light microscopy is inherently related to optics, and some concepts of the field are directly related to the blurring process; therefore, it is meaningfully important to elucidate them.

A.1 Dual Nature of Light

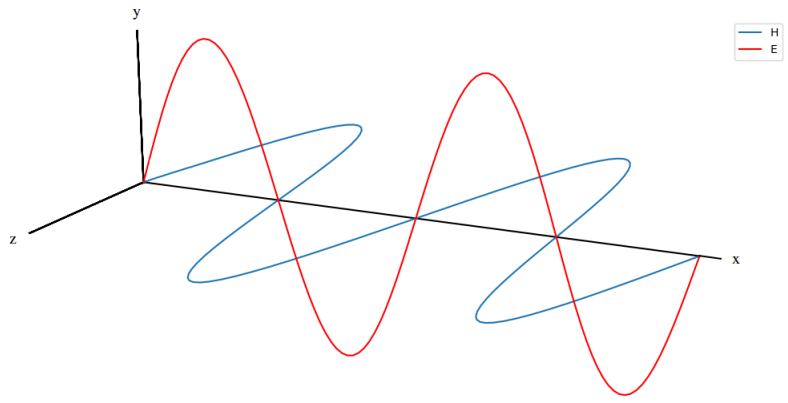
The light was described in different ways according to different geniuses. Isaac Newton (1642-1727) proposed that light had a corpuscular nature, due to the trajectory in which light appeared to travel in a uniform medium in his experiments; Christiaan Huygens (1629-1695) stated in his works that light was traveling in a “wave-like” way and apparently could explain some optical principles such as the interference phenomena (FOWLES, 1989).

The corpuscular approach for explaining the behavior of light was accepted during the 17th and 18th centuries since Newton played a central role in science in that era. The development of electricity and magnetism as solid fields of research and theoretical representations of natural phenomena was happening concomitantly with optics. Michael Faraday (1791-1867) connected magnetism and light for the first time with his studies on light polarization in magnetic field immersed materials. Later, James Clerk Maxwell (1831-1879) established a complete relation between optics and electromagnetism by defining the displacement current density - a relation that involves the polarization of a medium, the intensity of electric fields and the electric permittivity

of vacuum - and writing its differential equations (ZÍLIO, 2009).

Light as an electromagnetic wave is therefore composed of the two vectors and propagates in some particular coordinate direction upon a metric space, e.g. the x coordinate on a three-dimensional Euclidean space. Hence, it is possible to treat light as a wave or particle, according to the application or needs. Light microscopy deals with light as a wave and its related phenomena such as reflection, refraction, interference and diffraction, which will be presented on the following subsections and are useful for a deeper understanding of this work. Posterior studies of Max Planck (1858-1947), Albert Einstein (1879-1955) and Niels Bohr (1885-1962) (FOWLES, 1989) linked the prior discoveries with the quantum theory, and therefore differ from this research's scope. Maxwell enunciated that there were two different vectors which could cause a state of disturbance in the space while dealing with electric charges; those consist of the electric vector **E** and the magnetic induction vector **H**. Together, they construct the electromagnetic field (BORN *et al.*, 1999), as shown in Figure 28.

Figure 28 – Composition of the electromagnetic wave. The red and the blue curves represent the electric (**E**) and induction (**H**) vectorial quantities, respectively.



Source: Elaborated by the author.

A.2 Light Wave Properties and Phenomena

The light waves can be represented by a ray, i.e. a single oriented line which shows the direction of propagation; several waves that propagate in nearly the same direction can be named as a beam (HALLIDAY; RESNICK; WALKER, 2013). Some relevant properties of waves are explained here according to Tipler and Mosca (2007). The frequency (f) is the number of cycles per unit of time, or the reciprocal of the time that a periodic wave takes to execute a complete cycle of oscillatory motion T, given by $f = T^{-1}$. The Amplitude is the maximum displacement

from the equilibrium position, where the wave peak reaches its highest value. Finally, the Phase is a point in time on the cycle of a wave propagation, quantified in degrees.

When rays or beams of light reach surfaces during its propagation, there are four prominent phenomena to consider: reflection, refraction, interference and diffraction. The incident ray of light suffers a split procedure when it reaches a frontier between two homogeneous media. One of the resulting rays reflects within the initial medium and the other one propagates inside the other medium; the first phenomenon is denominated *reflection* and the second, *refraction* (BORN *et al.*, 1999). The speed of light depends on the medium in which it propagates. The *refractive index* is a number that quantifies the speed of light in a particular medium in relation to the speed of light in vacuum, and can be described by

$$n = \frac{c}{v}, \tag{A.1}$$

where c is the speed of light in vacuum and v is the speed of light in the medium. According to Halliday, Resnick and Walker (2013), the reflection law states that the resulting ray propagates within the incidence plane and that the angle of reflection $\theta_1^{'}$ equals the θ_1 angle of incidence; comparably, the refraction law states the same about the incidence plane and relates the incidence θ_1 and refraction θ_2 angles by Snell's law, given by

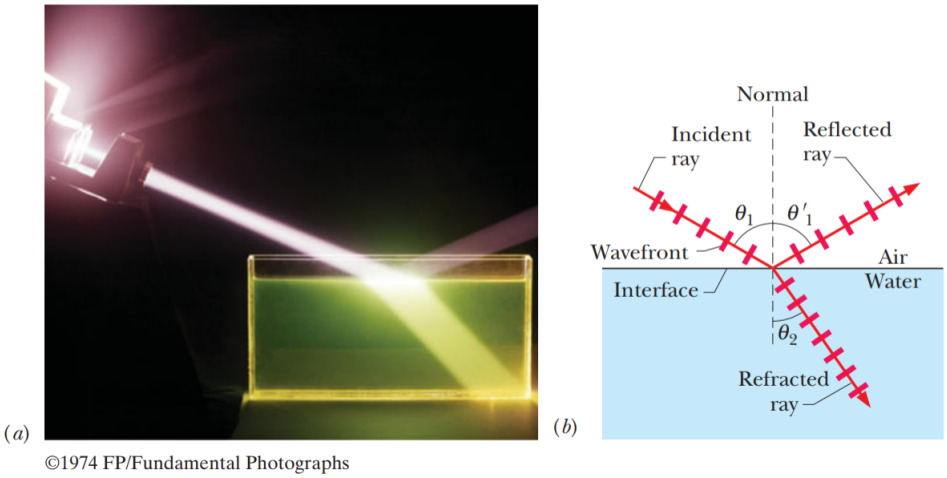
$$n_2 \sin\theta_2 = n_1 \sin\theta_1, \tag{A.2}$$

where n_1 and n_2 are the refractive indices of the media. This framework consists of an approximation and may be considered ideal for didactic purposes. The process that happens in the real situations may involve non-homogeneous media, opaque or translucent media (which blocks the propagation of light or randomly changes the direction of the rays, respectively), and those concepts are relevant to the imaging procedures, e.g. microscopy. Figure 29 depicts the real-world phenomenon and its ideal representation.

When two or more waves of similar or equal frequency superpose in space and form a different intensity pattern, the *interference* of the waves occurred (TIPLER; MOSCA, 2008). Mathematically, when interference is observed with light waves, it consists on a vector addition of the electromagnetic fields (ZÍLIO, 2009). If the interfering waves are in phase, i.e. the difference between the same positions within the wave cycles of the two waves is zero. This process is called *constructive*, which yields a larger amplitude to the new wave. Similarly, if the interfering waves are not in phase, then the process is named *destructive* and yield smaller or null amplitudes to the new wave (TIPLER; MOSCA, 2007). Both types of interference are show in Figure 30.

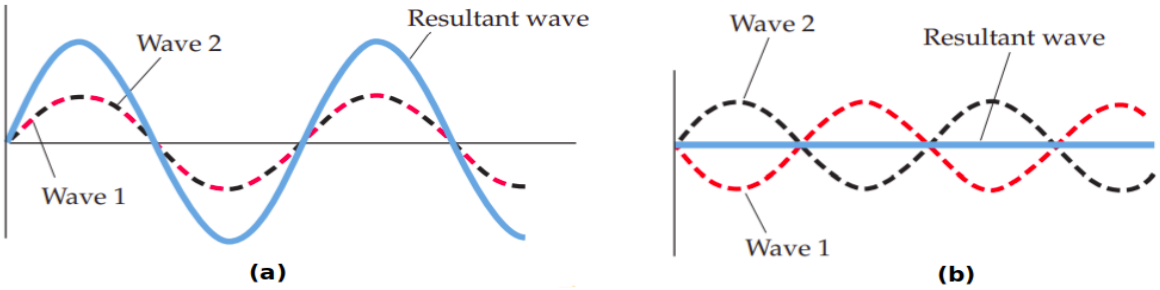
The wave theory of light also contains another important property for the real processes: the *diffraction*, a phenomenon that was discovered by Francesco Maria Grimaldi (1618-1663) and consists of the distortion in a wavefront which focuses on obstacles such as apertures on an object, spheres, disks, or anything with similar dimension to the wavelength of the focusing light

Figure 29 – (a) Example of a beam of light that reflects and refracts when touching the frontier between air and water. (b) Representation of the process with rays.



Source: Halliday, Resnick and Walker (2013).

Figure 30 – (a) Constructive interference and (b) destructive interference.



Source: Adapted from Tipler and Mosca (2007).

(ZÍLIO, 2009). The wavefront deviates and scatters after propagating through the obstacle and transforms itself into circular or spherical waves; this is a relevant property that distinguishes a wave from a particle, since the latter would either propagate without any change in its trajectory or would be blocked by the obstacle (TIPLER; MOSCA, 2007). When a beam of light reaches an opaque object, the waves suffer changes in their direction of propagation, which can be predicted by the fact that all the points in each wavefront (points of identical phase on waves) generate a new wave, as stated by Huygens (FOWLES, 1989). As described by the Huygens principle, each point of the wavefront acts as a source and creates secondary waves which scatter in every direction; each of the new wavefronts is generated by the interference of an infinite number of sources of waves (ZÍLIO, 2009). Diffraction is a phenomenon that always happens as waves pass through apertures, but its magnitude depends on the relative proportions between the aperture

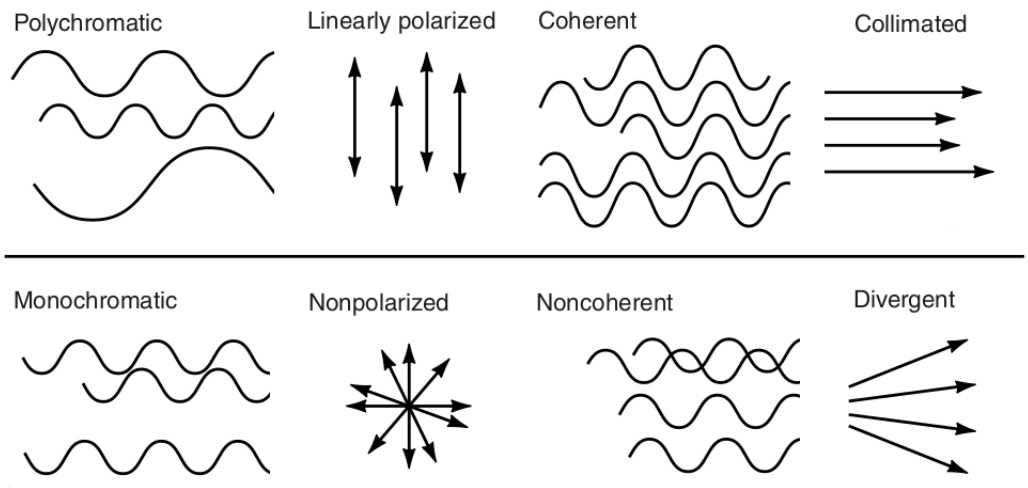
size and the wavelength: when the latter is large in comparison to the former, the diffraction effects are large and, likewise, a relatively small wavelength yields smaller effects (TIPLER; MOSCA, 2007). As stated by Zílio (2009), there are two common types of diffraction:

- *Fresnel Diffraction*: also named near-field diffraction, it occurs when a cylindrical wavefront (the curvature cannot be neglected) passes through an obstacle and diffracts in the near-field, i.e. short distances relative to the path of the diffracted waves' propagation. However, the observation distance is usually finite. It results in different sizes and shapes for the diffraction patterns;
- *Fraunhofer Diffraction*: also named far-field diffraction, it occurs when planar waves (the curvature of the wavefront may be neglected) pass through an obstacle in the far-field, i.e. large distances relative to the path of the diffracted waves' propagation. Practically, the observation distance is infinite. It results in different sizes for observed aperture images.

A.3 Illumination Qualities

The propagating waves, rays or beams of light that illuminate the object in imaging systems carry some attributes which depend on the source and the desired result. Figure 31 depicts some relevant qualities of light.

Figure 31 – Comparison between different qualities of an imaging systems' illumination setup.



Source: Adapted from Murphy and Davidson (2012).

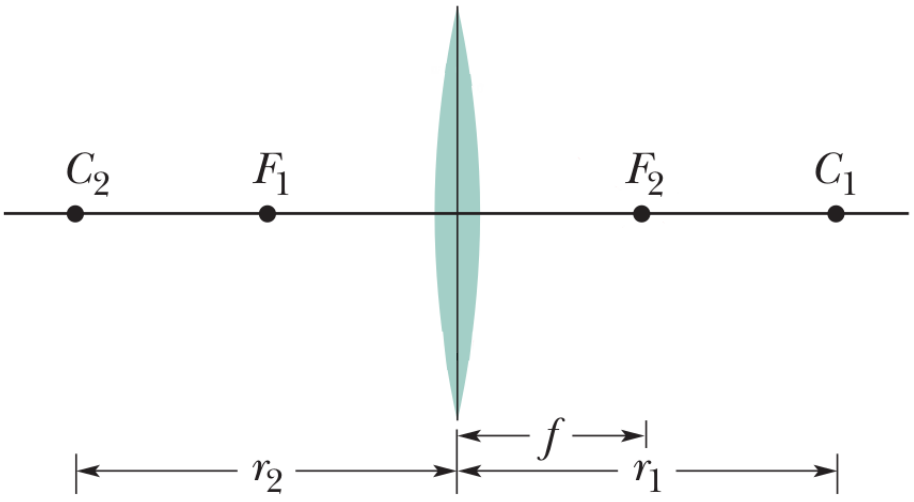
Pursuant to Murphy and Davidson (2012), the attributes in Figure 31 refer to color, polarization, coherence and direction. Every ray in a *monochromatic* beam has the same wavelength, i.e. the same color, and similarly a *polychromatic* beam consists of a mixture of rays with

different wavelengths. The polarization relates to the vibration in the electric vector of light as an electromagnetic wave which may occur in parallel planes (*polarized light*) or not (*nonpolarized light*). The coherence is the relationship between the phase of each wave of a given wavelength: if the phase is the same for all waves, as in a laser beam, the light is named *coherent*, otherwise it is said that the illumination is *noncoherent* or *incoherent*, as in bright-field microscopes. Finally, the waves may propagate in parallel trajectories, i.e. be *collimated*, or may diverge or converge to some point.

A.4 Properties of the Spherical Lenses

As stated by Halliday, Resnick and Walker (2013), *lenses* are objects consisting of a transparent material, with a certain refractive index, that are made of two spherical surfaces on which light propagates and suffers refraction. They are used in optical systems due to their capacity to create images as long as their refractive index is not equal to that of the medium. Yet, in agreement with Halliday, Resnick and Walker (2013), some concepts related to lenses are important in our context and will be shown below. Figure 32 illustrates an arbitrary spherical lens, with principal elements from geometric optics that relate to lenses and its consequent imaging properties.

Figure 32 – Arbitrary scheme of the optical properties of a spherical lens: radii of curvature (r_1, r_2), centers of curvature (C_1, C_2), focal points (F_1, F_2) and focal length (f).



Source: Adapted from Halliday, Resnick and Walker (2013).

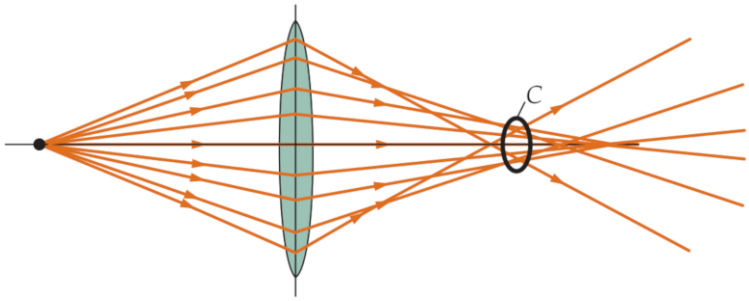
Some properties of the spherical lenses directly influence the image formation process and, consequently, the resulting image quality. The *numerical aperture* (NA), as reported by

Murphy and Davidson (2012), is a measurement in terms of angles that shows how much light the lenses can capture, and it is given by

$$NA = n \sin \theta, \tag{A.3}$$

where θ is the half angle of the cone of specimen light accepted by the objective lens and n is the refractive index between the lens and the specimen. There are optical flaws in lenses that hinder a proper image formation. Those are named *aberrations* (LAWLOR, 2019), and the most relevant types in the scope of this project are the *spherical aberrations*. According to Murphy and Davidson (2012), the spherical aberrations occur when there is a difference in the focal point of incident parallel rays at central and peripheral locations of a spherical lens' surface, which yields a blurred image of either a point source of light or an extended object. It is possible to correct a spherical aberration by changing the shape of the refracting surface, i.e. changing the radius of curvature for the lenses in order to adjust the focal point to one particular distance (SMITH; THOMSON, 1988). The illustration in Figure 33 represents the spherical aberration for a point source of light, where it is possible to see the difference between incident rays in the borders and in the inner regions of the lens. The resulting image is, in this case, a set of concentric circles around a point.

Figure 33 – Arbitrary example of the spherical aberration.



Source: Tipler and Mosca (2008).

APPENDIX

B

THEORETICAL BACKGROUND DETAILS

B.1 Convolution and Image Transforms

As seen in Chapters 2 and 3, the processes of image formation and acquisition concerning linear systems are subject to some operations, i.e. the convolution, the Fourier Transform (FT) and its continuous and discrete versions, that modify the original representation of the scene. The linear system theory provides mathematical tools to explore these operations and others, such as sampling, filtering, and enhancement; it describes the behavior of electrical circuits and optical systems (CASTLEMAN, 1996).

According to Bracewell and Bracewell (2000), the convolution of two arbitrary functions s and t that results in another function r is, with notation adjustments, defined by the integral

$$r(x) = \int_{-\infty}^{\infty} s(u)t(x-u)du, \tag{B.1}$$

where x is the one-dimensional coordinate and u is a shifting parameter. In other words, this procedure is compared to moving a 180 degree-rotated filter mask over the function values and computing the sum of products at each location (GONZALEZ; WOODS, 2018). Hence, the shifting parameter represents the slide of the filter mask over the values. As seen in Chapter 3, convolution is responsible for the image formation and acquisition processes, but it also covers several other applications, e.g. smoothing, sharpening and reducing noise in images. Similarly to the one-dimensional case and pursuant to Castleman (1996), the two-dimensional convolution is denoted by

$$r(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(u,v)t(x-u,y-v)dudv, \tag{B.2}$$

where u and v are shifting parameters, and $*$ denotes the convolution. The *spatial domain* is the subset of the real plane where functions like r, s and t are spanned, and (x,y) as points

within such subset are named spatial variables; consequently, any mathematical operation that employs pixels from this subset is named a spatial domain technique. As for the digital image processing applications, which deal with images as matrices of pixels, the discrete two-dimensional convolution for an image $f(x,y)$ and a function $h(x,y)$ is given by

$$g(x,y) = h(x,y) * f(x,y) = \sum_{m=-a}^{a} \sum_{n=-b}^{b} f(m,n)h(x-m,y-n), \tag{B.3}$$

where $a = (m-1)/2$ and $b = (n-1)/2$, given that the function $h(x,y)$ is considered to be a two-dimensional filter of size $m \times n$ (GONZALEZ; WOODS, 2018).

Convolution, together with several other operations employed in this work, operate directly on the spatial domain by modifying pixel values based on mathematical constraints. Some of those operations may have issues that hinder the operation, e.g. the computation time of a convolution process should be finite, otherwise, its use is impractical; this is one of the reasons why *image transforms* are widely used. They encompass any group of mathematical operations that transfers the input signal or image from its domain to the transform domain (GONZALEZ; WOODS, 2009). Let s be an arbitrary two-variable function, t_f and t_i be the forward and inverse transformation kernels, respectively. The general discrete form of the forward and inverse two-dimensional transforms is denoted by

$$R(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} s(x,y)t_f(x,y,u,v) \qquad s(x,y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} R(u,v)t_i(x,y,u,v), \tag{B.4}$$

where M and N are the dimensions of the image, x and y are coordinates of the image, $u = \{0,1,2,...,M-1\}$ and $v = \{0,1,2,...,N-1\}$ are called transform variables. The t_f function is responsible for the forward domain change and the t_i transfers the image back to the spatial domain. Switching from spatial to frequency domain, for example, allows different operations to be executed that otherwise could not be performed in the spatial domain. The convolution operation, for instance, turns itself into a simple matrix multiplication task on the Fourier Transform domain (which will be detailed in the following sections) and that solves the performance constraint.

B.2 Image Enhancement

Image enhancement is the process of manipulating an image in order to provide a resulting representation that is more suitable for a particular problem, e.g. an enhancement method for medical images may not be efficient for satellite images (GONZALEZ; WOODS, 2018). In microscopy, image enhancement is desirable due to the limited capacity of optical imaging devices and the drawbacks posed by each microscopy technique, e.g. acquisition with different

illumination settings, focal planes, time intervals or spectral channels. Therefore, enhancement algorithms for microscopy should cover all types of information (WU; MERCHANT; CASTLEMAN, 2008). According to WU, MERCHANT and CASTLEMAN (2008), image enhancement techniques are divided into *spatial domain*, *Fourier transform* and *wavelet transform* methods. They will be described in the following paragraphs.

The spatial domain methods are transforms that work directly on pixels of the image (GONZALEZ; WOODS, 2018). Some examples are contrast stretching (redistribution of image gray levels to a wider range), thresholding (image binarization given a preset gray level), histogram equalization, and spatial filtering. Particularly, histogram equalization and spatial filtering play an important role in this work and therefore will be explored further.

Fourier transform domain methods operate with images as a distribution of frequencies since some features are better described by it. Noise, for example, may be suppressed in a sharpening process or reduced by amplifying mid-frequency components and attenuating high-frequency ones. The Wiener Filtering process is a popular example of a frequency domain enhancement method that recovers a noisy signal or image based on estimations of spectral properties from the original image. Other examples of Fourier domain enhancement are band-pass filters and least-squares deconvolution applications.

Finally, there are methods based on the Wavelet Transform, i.e. a mathematical framework that decomposes images or signals into frequency components in different scales. Some operations such as thresholding may be applied to wavelet coefficients. However, the output of the wavelet transform is not always the same; it depends on the chosen wavelet function and consequently should be properly set in order to extract the desired image features.

The image histogram is one of the simplest and most useful tools in image processing and consists of a function that summarizes the gray level content of an image in terms of a frequency distribution (CASTLEMAN, 1996). The histogram equalization consists of a non-linear monotonic mapping to provide an approximation of a uniform distribution to the output image's histogram (GONZALEZ; WOODS, 2018). The output histogram is a normalization of the cumulative histogram of the image, given by

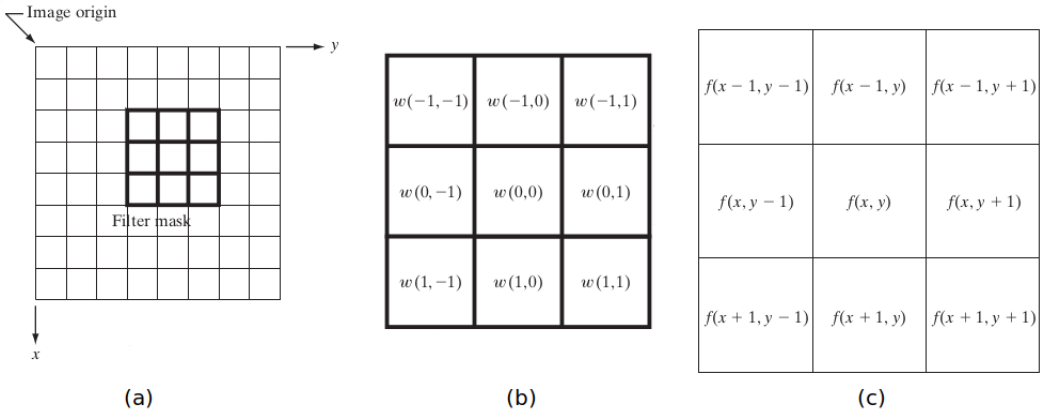
$$hist_{equalized}(r) = \frac{L-1}{MN} hist_{cumulative}(r), \tag{B.5}$$

where $hist_{equalized}(r)$ and $hist_{cumulative}(r)$ are the equalized and cumulative histograms relative to a range L of intensities after the image quantization with r values. $M \times N$ is the image resolution. Since it stems from a sum of probabilities and no new gray intensity levels should be created, the process generates fractional values that are mapped onto integers.

Spatial filtering consists of the convolution of an image with a predefined kernel operator (GONZALEZ; WOODS, 2018). The continuous form may be represented as a convolution over all values of a defined region of the image and the discrete form consists of sliding a weight

mask over the image (WU; MERCHANT; CASTLEMAN, 2008). Figure 34 presents an arbitrary schema of a simple linear spatial filtering procedure:

Figure 34 – Arbitrary example of linear spatial filtering of an image (a) with a 3×3 filter mask (b), which results in filtered sections (c).



Source: Adapted from Gonzalez and Woods (2018).

Examples of discrete spatial filtering in digital image processing are smoothing filters, order-statistic nonlinear filters and sharpening filters (GONZALEZ; WOODS, 2018). Spatial smoothing filters are applied to remove details, edges and lines from an image to reduce noise. The order-statistic nonlinear filters are based on ordering pixels of the image under the filter area and replacing the pixel value in the center of the area with the response from ordering; one example is the *median filter*, which replaces the center pixel with the median of pixels in its neighborhood. Median filters yield significant noise reduction effects if the nature of the noise is random. Finally, the sharpening filters are built to highlight transitions in intensity by spatial differentiation and are used for enhancing edges.

B.3 Image Registration

When images belonging to the same scene are acquired in different conditions such as distinct focus configurations, sensors, or even at different times, they should be geometrically aligned according to a reference image. The process of overlaying two or more images with different acquisition settings is named image registration and plays an important role as a pre-processing step for image fusion, change detection and multichannel image restoration (ZITOVA; FLUSSER, 2003). According to Gonzalez and Woods (2018), magnetic resonance imaging and positron emission tomography systems, for example, are two different medical image modalities that may require images to be registered. Images which were taken in different times such as satellite images also need to be registered.

The image registration methods consist of the following steps, as reported by Zitova and Flusser (2003) and Gonzalez and Woods (2018):

- *Feature detection*: manually or automatically detect distinctive objects, e.g. edges, contours, corners and represent those as *control points*, i.e. points with known locations in the reference and input images;
- *Feature matching*: a relationship between the detected features in each image is established using feature descriptors;
- *Transform model estimation*: parameter estimation for mapping functions that align the input images with the reference image, either by establishing feature correspondence or performing an optimization procedure;
- *Image resampling and transformation*: the image is resampled with interpolation techniques.

Image registration is, in practice, a mapping between two or more images by means of a spatial and an intensity transformation (BROWN, 1992). Some prominent examples of registration methods are the principal axes, multiresolution, optimization-based, boundary, model-based and adaptive methods (GOSHTASBY, 2012). The spatial transformations play an important role in all image registration techniques, and the most common examples are rigid, affine, projective, perspective and global polynomial (BROWN, 1992). Also pursuant to Brown (1992), each of such transformations may be described as:

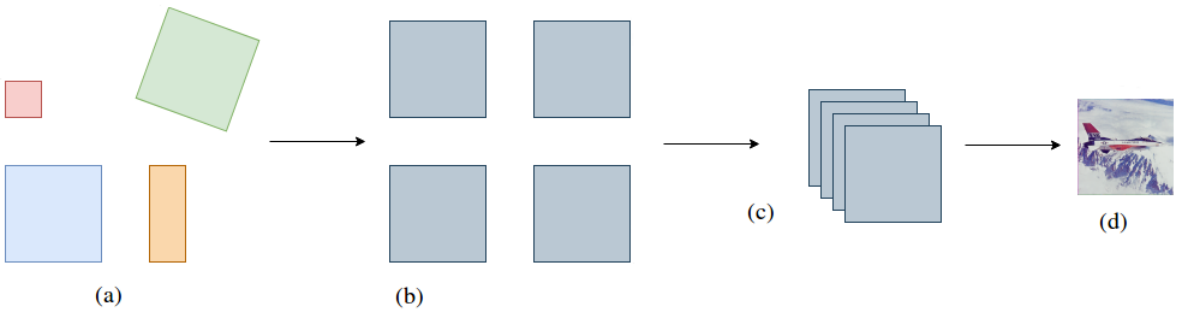
- *Rigid*: accounts for object or sensor movement in which objects in the images retain their relative shape and size. Example: *rigid-body transformation* which combines rotation, translation and scale changes;
- *Affine*: handle more complicated distortions and preserve mathematical properties. Example: *shear transformation*;
- *Projective and Perspective*: the former deals with distortions due to projection of the objects at varying distances to the sensor onto the image plane, and the latter demands prior knowledge about the locations of the objects in the scene relative to the sensor;
- *Polynomial*: cover the broadest range of distortions, as long as they are approximately homogeneous in the image.

B.4 Image Fusion

Image fusion is a process that merges several images, possibly acquired in diverse conditions or with different cameras, into one image with higher quality, more details and

consequently more useful for humans and computer tasks (MITCHELL, 2010). Examples of image fusion applications are noise reduction, edge enhancement, and super-resolution. One traditional use of image fusion occurs in medical imaging fields; the quality of information about illnesses, cells, clinical analysis and several other medical tasks (including the computer-assisted ones) have found profitable results from the image fusion techniques and led themselves to better and faster decisions when it comes to human beings (JAMES; DASARATHY, 2014). There are also relevant applications in remote sensing multispectral images, segmentation of regions in different color spaces, biometry: the pan-sharpening process is the generation of a high-resolution multispectral image from low to high-resolution ones, K-Means segmentation and fusion of pixels in the RGB and the Iris Recognition biometric process with video frames are examples of such tasks, respectively (MITCHELL, 2010). Also according to Mitchell (2010), the general framework for the image fusion procedure consists of four stages: *Multiple Input Images*, *Common Representational Format*, *Fusion* and *Display*. The multiple input images stage is simply the acquisition of the images to be merged. There are several approaches to this: the dataset may be captured from different sensors, under distinct light conditions or angles, with different magnifications, under several focus settings, and with temporal measurements, if the scene changes through time. Figure 35 depicts an arbitrary example of the four stages.

Figure 35 – Image fusion general framework. (a) Multiple Input Images, (b) Common Representational Format, (c) Fusion and (d) Display.



Source: Elaborated by the author.

The four arbitrary images in Figure 35.**(a)** represent different images of the same scene, taken at different resolutions, rotation angles, and shapes. In Figure 35.**(b)**, the images are all reshaped, converted to common color space and ready to undergo the processing algorithm which will transform them into feature vectors. Figure 35.**(c)** represents the image fusion by means of an arbitrary fusion rule. The resulting image is depicted in Figure 35.**(d)**.

If the acquired dataset images do not share the same features such as dimension, rotation angle, and resolution, then the images should be pre-processed in order to arrive at a common

state. This configures the common representational format step, which generates a new and temporary dataset with the same properties, e.g. color space, dimensions, and noise level. The fusion stage employs a decision method to dictate which regions, objects, colors or details will compose the final image; some methods rely on the wavelet transform, for example. Finally, the display stage provides a view of the resulting image, which can be used directly for any further task or even be the input for other image processing operations.

B.5 Image Quality Assessment

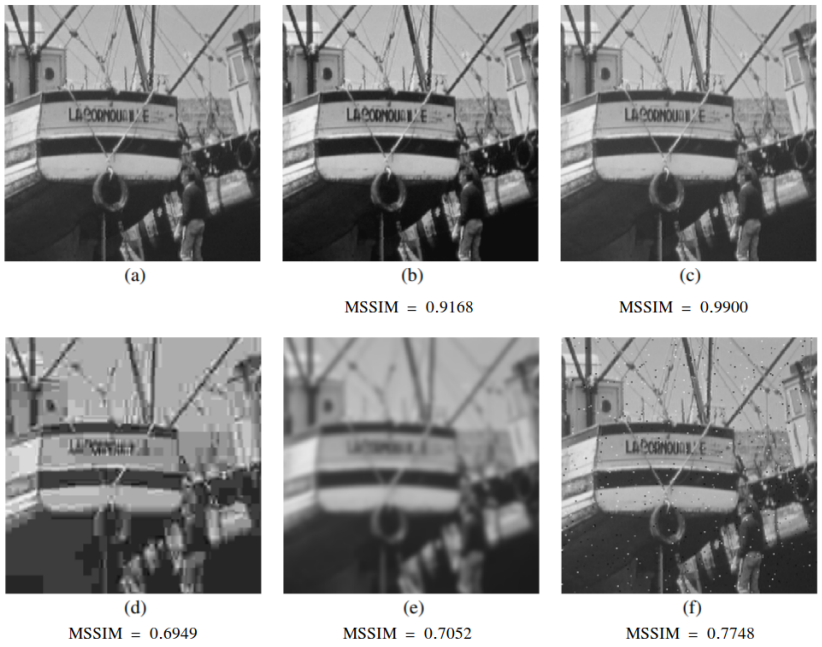
Image Quality Assessment (IQA) is the evaluation of image quality as perceived by an average human observer, i.e. how close an image is to a given original or reference image. It is also related to the accuracy of the image acquisition process for an imaging system (BOVIK, 2009). It is known that images are frequently used in health and life sciences, public security systems, remote sensing, and several other fields; hence, there are computational applications that offer some useful service employing image processing. As a result, assessing image quality poses as an important task among those applications for which several techniques are being developed, evolved and deployed.

As stated by Zhou Wang *et al.* (2004), there are three classes of objective image quality metrics that relate to the existence of a no-distortion image (or with a negligible amount of it) for comparison purposes. The Full-Reference Image Quality Assessment (FR-IQA) methods assume that the reference image is available, while Reduced-Reference Image Quality Assessment (RR-IQA) methods employ a representation of the reference image, such as a set of extracted features. Finally, the No-Reference Image Quality Assessment (NR-IQA) methods, also known as “blind”, are those which do not employ a reference image. Figure 36 denotes an example of a full-reference method, the Mean Structural Similarity Measure (MSSIM) method and its output for an image with different types of degradation:

According to Tang *et al.* (2019), the IQA methods are distributed between the subjective assessment and objective assessment categories. The former is based on a well-defined test environment for random observers to label images and provide the final Mean Opinion Scores (MOS), while the latter is based on the use of strategies such as statistical modeling, machine learning, spatial or spectral image features and so on. It is evident that subjective IQA is demanding; consequently, objective methods are preferred to conduct IQA.

IQA methods are also present within microscopy and its close interaction with image processing. The image acquisition in microscopy techniques may involve lasers, transmitted or reflected light, measurements of atomic force responses, the fluorescence of chemical compounds and several other means. Each technique has an inherent kind of degradation that affects the acquired images or spectra, e.g. the Raman confocal microspectroscopy suffers from the interference of cosmic rays, which yields unexpected peaks in the spectrum.

Figure 36 – Example of the MSSIM method output: Original image (a), contrast-stretched image (b), mean-shifted image (c), JPEG compressed image (d), blurred image (e) and salt-pepper impulsive noisy image (f).



Source: Zhou Wang *et al.* (2004).

B.6 Statistics

Statistics is the science of planning studies and experiments, obtaining, organizing, summarizing, presenting, analyzing, and interpreting data and finally drawing conclusions. *Descriptive statistics* is an important branch that comprises a set of methods which aim to describe relevant characteristics in data (TRIOLA, 2017). The descriptive statistics methods either employ graphical elements such as boxplots, histograms, bar graphs and scatter plots to analyze data or yield numerical summary measures such as means, standard deviations, correlation coefficients and other related indices (DEVORE, 2011). The methods that compose a descriptive statistical approach for data analysis are simple yet powerful tools.

The concepts of *population*, *sample* and *variable* are elementary: a population is a well-defined collection of objects that might be included in the analysis, a sample is a subset of a population and a variable is a feature of the objects which may change from one object to another (DEVORE, 2011). Moreover, a *frequency distribution* is a tool that presents how the data is partitioned among several categories by listing each category and its frequency of data values; a *relative frequency distribution* is a frequency distribution in which each frequency is represented by a proportion, usually as percentage (TRIOLA, 2017). According to Mendenhall and Sincich (2016), the measures of central tendency provides several ways to locate the center

of the relative frequency distribution, and the three most common are the *arithmetic mean*, i.e. the average of the measurements, the *median*, i.e. the middle number when the measurements are ordered in ascending or descending order and the *mode*, i.e. the value that occurs with the greatest frequency. Mathematically, the arithmetic mean is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}, \tag{B.6}$$

where n is the sample size and x_i represents the i-th observation of the variable x (ZWILLINGER; KOKOSKA, 1999).

The measures of variation provide a description of how the values spread along the distribution. Commonly used measures are the range, the variance, and the standard deviation (MENDENHALL; SINCICH, 2016). The range is simply the difference between the largest and the smallest value within the data, which may precisely point out its variability, since it does not comprise the middle values among the distribution (DEVORE, 2011). The variance and the standard deviation are closely related, as the former measures variability based on the squared deviations about the mean and the latter is the positive square root of the variance, as

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad\qquad \sigma = \sqrt{\sigma^2} \tag{B.7}$$

where σ^2 and σ are respectively the variance and the standard deviation, x_i is the i-th observation of the variable x, $\bar{x}$ is the mean, all concerning a sample or the population (ZWILLINGER; KOKOSKA, 1999).

The measures of the relative standing of an observation describe its location among other values in the distribution, and two examples of these measures are *percentiles* and *z-scores*; also, an observation located outside the range of the distribution is an *outlier* (MENDENHALL; SINCICH, 2016). Percentiles are values that split the data into 100 parts in a sorted dataset, so that the i-th percentile stands for the $i(n+1)/100$ observation, e.g. the 25-th percentile comprises 25% of the data; the z-score, or standard score, is given by

$$z = \frac{x_i - \bar{x}}{\sigma}, \tag{B.8}$$

where x_i is the i-th observation of the variable x, $\bar{x}$ is the mean and σ is the standard deviation of the population or the sample (ZWILLINGER; KOKOSKA, 1999).

The Interquartile Range (IQR) is the length of the interval that contains the middle half of the distribution (DEGROOT; SCHERVISH, 2012). Mathematically, it is the difference between

the third (Q_3) and the first (Q_1) quartiles, i.e. the 75th percentile 25th percentile, respectively (DEVORE, 2011). The IQR is described by

$$IQR = Q_3 - Q_1. \tag{B.9}$$

Probability is a common and natural concept among human life, used in expressions such as "It probably will be cold tonight"; however, there is no common formal definition accepted among statisticians and related researchers (DEGROOT; SCHERVISH, 2012). The study of randomness, variability and uncertainty in populations is done by analyzing probabilities, i.e. numerical descriptions of how likely an event is to occur (DEVORE, 2011). Some basic concepts that support the probability theory are described also to the light of Devore (2011), as follows:

- *Experiment*: any activity or process whose outcome is subject to uncertainty;
- *Sample Space*: the sample space of an experiment is the set of all possible outcomes for it;
- *Event*: any collection or subset of outcomes of a sample space;
- *Random Variable*: any rule that associates a number with each outcome in a sample space of some experiment; mathematically, it is a function with the sample space as its domain and the real numbers as its range;
- *Discrete Random Variable*: a random variable with a finite set or a countable infinite sequence of possible values;
- *Continuous Random Variable*: a random variable that yields zero as the probability for every possible outcome or its set of possible values are in a single interval of the real line or all numbers in a disjoint union of intervals.

From these concepts, it is possible to define a distribution in the scope of the probability theory. The *probability distribution* is a collection of all probabilities computed from a discrete or continuous random variable with the set of real numbers; a *discrete* probability distribution is represented by the probability function itself, while a *continuous* probability distribution is represented by a probability density function (p.d.f) (MENDENHALL; SINCICH, 2016).

The *kurtosis* is one of the probability distribution shape statistics, which measures the extent of the peak in a distribution, i.e. its "peakedness"; smaller absolute values indicate that the distribution tends to be uniform (ZWILLINGER; KOKOSKA, 1999). First of all, the concepts of *expectation* and *moments* should be described. The expectation of a random variable (and consequently, of a distribution) is a value that summarizes its nature and is given by

$$E(X) = \int_{-\infty}^{\infty} xp(x)dx \qquad E(X) = \sum_{x} xp(x), \tag{B.10}$$

where x is each possible outcome of the random variable X, $p(x)$ is the probability density function for a continuous random variable (left) and the probability function for a discrete random variable (right) (DEGROOT; SCHERVISH, 2012). Still according to DeGroot and Schervish (2012), for a random variable X and every positive $k \in \mathbb{R}$, the expectation $E(X^k)$ is called the k-th moment of X. The r-th moment may be described, according to Zwillinger and Kokoska (1999), as

$$m_r = \frac{1}{n} \sum_{i=1}^{k} p_i (x_i - \bar{x})^r \tag{B.11}$$

for every x_i in the possible outcomes of X. Thus, kurtosis may be defined as the ratio of the fourth moment (Equation B.11 with $r = 4$) by the square of the variance (also Equation B.11 with $r = 2$), denoted by

$$g_2 = \frac{m_4}{(m_2)^2} - 3 \tag{B.12}$$

The -3 constant is inherited from Fischer's approach, where the kurtosis of a normal distribution is zero.

APPENDIX

C

DEFINITIONS AND PROOFS

This appendix presents the mathematical description of how the feature vectors are built and a proof of the fact that the algorithm yields a probability distribution as the final structure, which allows the use of statistical analysis as the basis of our quality metric. Along with the proof, several definitions and auxiliary theorems that are necessary to the understanding of the proof are described.

Definition 1. (Algebra) (FOLLAND, 2013)

Let X be a nonempty set. A **field** or an **algebra** of sets on X is a nonempty collection $\mathcal{A}$ of subsets of X that is closed under finite unions and complements, as described as follows:

(i) if $\{E_j\}_{j=1}^{n} \in \mathcal{A}$, then $\bigcup_{j=1}^{n} E_j \in \mathcal{A}$;

(ii) if $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.

Definition 2. (σ-algebra) (FOLLAND, 2013)

A **σ-field** or **σ-algebra** $\mathcal{M}$ is an algebra that is closed under countable unions.

Definition 3. (Measure) (FOLLAND, 2013)

Let X be a set equipped with a σ-algebra $\mathcal{M}$. A **measure** on $\mathcal{M}$ is a function $\mu : \mathcal{M} \to [0, +\infty]$, such that

(i) $\mu(\emptyset) = 0$.

(ii) if $\{E_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{M}$, then

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j).$$

From (ii), if $E_j = \emptyset$ for $j > n$, if $\{E_1, E_2, ..., E_n\}$ are disjoint sets in $\mathcal{M}$, then $\mu\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu(E_j)$.

Definition 4. (Dirac Measure) (ÇINLAR, 2011)

Let $(X, \mathcal{M})$ be a measurable space and let x be a fixed point of X. For each subset E in $\mathcal{M}$, put

$$\delta_x(E) = \begin{cases} 1, & \text{if } x \in E \\ 0, & \text{if } x \notin E \end{cases}.$$

Then, δ_x is a measure on $(X, \mathcal{M})$ and is called the **Dirac measure** sitting at x.

Definition 5. (Discrete Measure) (ÇINLAR, 2011)

Let $(X, \mathcal{M}, \mu)$ be a measurable space. Let D be a countable subset of X. For each arbitrarily chosen $x \in D$, let $m(x) > 0$. For $E \in \mathcal{M}$, define

$$\mu_E = \sum_{x \in D} m(x)\delta_x(E).$$

Then, μ is called a **discrete measure** on $(X, \mathcal{M})$.

Definition 6. (Metric and Metric Space) (FOLLAND, 2013)

Let X be a nonempty set. A **metric** on X is a function $\rho : X \times X \to [0, \infty)$ such that

(i) $\rho(x, y) = 0 \Leftrightarrow x = y$;

(ii) $\rho(x, y) = \rho(y, x)$ for all $x, y \in X$;

(iii) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for all $x, y, z \in X$.

A set equipped with a metric is called a **metric space**.

Definition 7. (Open and Closed Ball) (FOLLAND, 2013)

Let (X, ρ) be a metric space. If $x \in X$ and $r > 0$, the **open ball** of radius r about x is

$$B(r; x) = \{y \in X : \rho(x, y) < r\}.$$

In other words, a ball of radius r is the collection of points of distance less than (which makes it **open**) or equal to (which makes it **closed**) r from a fixed point x in a metric space (CROFT; FALCONER; GUY, 2012).

Definition 8. (Topology and Topological Space) (ROYDEN, 1988)

Let X be a nonempty set. A **topology** on X is a family of subsets $\tau \subset \mathcal{P}(X)$, satisfying

(i) $X \in \tau$ and $\emptyset \in \tau$.

(ii) if $O_1, O_2 \in \tau$ imply $O_1 \cap O_2 \in \tau$ (τ is closed under finite intersections).

(iii) if $(O_i)_{i\in I} \in \tau$, then $\cup_{i\in I} O_i \in \tau$ (τ is closed under arbitrary unions).

A **topological space** (X, τ) is a nonempty set together with a **topology** τ on it. The elements of τ are called **open sets** of X.

Definition 9. (Topological Vector Space) (FOLLAND, 2013)

A **topological vector space** is a vector space $\mathbb{V}$ over $\mathbb{K} = (\mathbb{R}$ or $\mathbb{C})$ which is endowed with a topology such that the maps $(x, y) \to x + y$ and $(\lambda, x) \to \lambda x$ are continuous from $\mathbb{V} \times \mathbb{V}$ and $\mathbb{K} \times \mathbb{V}$ to $\mathbb{V}$.

Definition 10. (Locally Convex Topological Vector Space) (FOLLAND, 2013)

A topological vector space is called **locally convex** if there is a base for the topology consisting of convex sets (that is, sets A such that if $x, y \in A$ then $tx + (1-t)y \in A$ for $0 < t < 1$). Most topological vector spaces that arise in practice are locally convex.

Definition 11. (Inductive Limit Topology) (GHEONDEA, 2016)

The inductive limit topology of a topological vector space is the strongest locally convex topology on $\mathbb{V}$ that makes the linear maps on continuous.

Definition 12. (Generated σ-algebra) (FOLLAND, 2013)

Let X be an nonempty set. If $\mathcal{E}$ is any subset of $\mathcal{P}(X)$, there is an unique smallest σ-algebra $\mathcal{M}(\mathcal{E})$ containing $\mathcal{E}$, namely, the intersection of all σ-algebras containing $\mathcal{E}$. Then, $\mathcal{M}(\mathcal{E})$ is called the σ-algebra **generated** by $\mathcal{E}$.

Definition 13. (Borel σ-algebra) (FOLLAND, 2013)

If X is a topological space, the σ-algebra generated by the family of open sets in X (or, equivalently, by the family of closed sets in X) is called the **Borel** σ-algebra on X and is denoted by $\mathcal{B}_X$.

Definition 14. (Closure) (ROYDEN, 1988)

For a set E of real numbers, a real number x is called a **point of closure** of E provided every open interval that contains x also contains a point in E. The collection of points of closure of E is called the **closure** of E and denoted by $\bar{E}$.

Definition 15. (Bounded Set, Totally Bounded Set, Cover and Subcover) (FOLLAND, 2013)

Let the minimal upper bound of a set $A \in \mathbb{R}$ be called the **supremum** and be denoted by $sup(A)$.

Let (X,ρ) be a metric space. We also define the **diameter** of $E \subset X$ as

$$diam(E) = \{sup\rho(x,y) : x,y \in E\}.$$

Then E is called **bounded** if $diam(E) < \infty$.

If $E \subset X$ and $V_{\alpha\,\alpha\in A}$ is a family of sets such that $E \subset \bigcup_{\alpha\in A} V_\alpha$, $\{V_{\alpha\in A}\}$ is called a **cover** of E, and E is said to be **covered** by the V_α's. A **subcover** of E is a subset of $\{V_\alpha\}_{\alpha\in A}$ that still covers E.

Finally, E is called **totally bounded** if, for every $\varepsilon > 0$, E can be covered by finitely many balls of radius ε.

Definition 16. (Sequence) (FOLLAND, 2013)

A **sequence** in a set X is a mapping from $\mathbb{N}$ into X.

Definition 17. (Cauchy Sequence) (FOLLAND, 2013)

A sequence $\{x_n\}$ in a metric space (X,ρ) is called **Cauchy** if $\rho(x_n,x_m) \to 0$ as $n,m \to \infty$.

Definition 18. (Complete Subset) (FOLLAND, 2013)

Let (X,ρ) be a metric space. A subset E of X is called **complete** if every Cauchy sequence in E converges and its limit is in E.

Theorem 1 (FOLLAND, 2013). If E is a subset of the metric space (X,ρ), the following are equivalent:

(i) The subset E is complete and totally bounded;

(ii) **(Bolzano-Weierstrass Property)** Every sequence in E has a subsequence that converges to a point of E;

(iii) **(The Heine-Borel Property)** If $\{V_\alpha\}_{\alpha\in A}$ is a cover of E by open sets, there is a finite set $F \subset A$ such that $\{V_\alpha\}_{\alpha\in F}$ covers E.

Definition 19. (Compact Set) (FOLLAND, 2013)

A **compact set** is a set E which possesses the properties from Theorem 1.

Definition 20. (Support and Compact Support) (FOLLAND, 2013)

Let (X,τ) be a topological space and $C(X)$ be the space of continuous functions. If a function $\varphi \in C(X)$, then the **support** $supp(\varphi)$ of the function is the smallest closed set outside of which φ vanishes, i.e. the closure of $\{x : \varphi(x) \neq 0\}$. If $supp(\varphi)$ is compact, we say that r is **compact supported** and we define

$$C_c(X) = \{\varphi \in C(X) : supp(\varphi) \text{ is compact}\}$$

as the set of all continuous functions with compact support.

Definition 21. (L^p space) (FOLLAND, 2013)

Let $(X,\mathcal{M},\mu)$ be a measure space. If φ is a measurable function on X and $0 < p < \infty$, we define

$$\|\varphi\|_p = \left[\int |\varphi|^p d\mu\right]^{1/p}$$

(allowing $\|\varphi\|_p = \infty$), and we define

$$L^p(X,\mathcal{M},\mu) = \left\{\varphi : X \to \mathbb{C} : \varphi \text{ is measurable and } \|\varphi\|_p < \infty\right\}.$$

We abbreviate $L^p(X,\mathcal{M},\mu)$ by $L^p(\mu)$, $L^p(X)$ or simply L^p. In other words, the L^p space is a function space where a measurable function φ is p-integrable.

Definition 22. (L^2 space) (FOLLAND, 2013)

The $L^2(X)$ is the space of square integrable functions $\varphi : X \to \mathbb{R}$ on a measurable space $(X,\mathcal{M},\mu)$, i.e.

$$\int_{-\infty}^{\infty} |\varphi(x)|^2 d\mu < \infty.$$

Definition 23. (Distribution and $\mathcal{D}'(\mathbb{R}^n)$ space) (FOLLAND, 2013)

Let $U \subset \mathbb{R}^n$ be an open set and $C_c^\infty(U) = \bigcap_1^\infty C_c^k(U)$. A **distribution** on U is a continuous linear functional $\psi : C_c^\infty(U) \to \mathbb{C}$, when $C_c^\infty(U)$ is provided with the inductive limit topology. The space of all distributions on U is denoted by $\mathcal{D}'(U)$ and forms the topological dual space $\mathcal{D}(U) = C_c^\infty(U)$. We set $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^n)$ and we write the functional as $\langle T, v\rangle$ instead of $T(v)$.

Theorem 2 (FOLLAND, 2013). Let $E_k(x) = e^{2\pi ikx}$. Then $\{E_k : k \in \mathbb{Z}^n\}$ is an orthonormal basis of $L^2(\mathbb{T}^n)$.

Proof. Verification of orthonormality is an easy exercise in calculus; by Fubini's theorem it boils down to the fact that $\int_0^1 e^{2\pi ikt} dt$ equals 1 if $k = 0$ and equals 0 otherwise.

Next, since $E_k E_\lambda = E_{k+\lambda}$, the set of finite linear combinations of the E_k's is an algebra. It clearly separates points on $\mathbb{T}^n$; also, $E_0 = 1$ and $\bar{E}_k = E_{-k}$. Since $\mathbb{T}^n$ is compact, the Stone-Weierstrass theorem implies that this algebra is dense in $C(\mathbb{T}^n)$ in the uniform norm and hence in the L^2 norm, and $C(\mathbb{T}^n)$ is itself dense in $L^2(\mathbb{T}^n)$. It follows that $\{E_k\}$ is a basis.

■

Definition 24. (Parseval's Equation) (GARLING, 2014) If $\mathbb{V}$ is a topological vector space, $\varphi, \psi \in \mathbb{V}$, $\hat{\varphi}$ and $\hat{\psi}$ are respectively the Fourier Transforms of φ and ψ, then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(t)\overline{\psi(t)} dt = \sum_{k=-\infty}^{\infty} \hat{\varphi}_k \overline{\hat{\psi}_k}.$$

Particularly, $\frac{1}{2\pi} \int_{-\pi}^{\pi} |\varphi(t)|^2 dt = \sum_{k=-\infty}^{\infty} |\hat{\varphi}_k|^2$.

Corollary 1 (FOLLAND, 2013). Let $\mathbb{T} = [0,1] \times [0,1]$. The Fourier Transform, as defined in Chapter 4, maps $L^2(\mathbb{T})$ one to one onto

$$\ell^2(\mathbb{Z}^2) = \left\{ (\xi_{ij})_{i,j\in\mathbb{Z}} \in \mathbb{C} : \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |\xi_{ij}|^2 < \infty \right\}$$

In other words, the Fourier coefficients of a square integrable function defined within $\mathbb{T}$ are a square integrable sequence of real values.

Proof. (GARLING, 2014)

Parseval's equation implies that the Fourier Transform is an isometric (bijective map between two metric spaces) linear isomorphism (a map which preserves sets and relations) of $L^2(\mathbb{T})$ into $\ell^2(\mathbb{T})$. On the other hand, if $\gamma_n(e^{j\theta}) = e^{jn\theta}$ is an orthonormal sequence, i.e. each vector is orthonormal to all others, in $L^2(\mathbb{T})$ and $\{a_n\}_{n=-\infty}^{\infty} \in \ell_2(\mathbb{Z}^2)$, then $\left(\sum_{i=-n}^{n} a_i\gamma_i\right)_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{T})$. Since $L^2(\mathbb{T})$ is complete, the sequence converges to an element in $L^2(\mathbb{T})$. ■

Next, we provide a mathematical proof that the proposed sampling of the Fourier spectrum and the posterior mapping by means of an operator produces a probability distribution.

Theorem 3. Let $\mathbb{T} = [0,1] \times [0,1]$. The proposed sampling of the Fourier spectrum yields a distribution, which maps $L^2(\mathbb{T})$ into $\ell_2(\mathbb{Z}^2)$.

Proof. Indeed, let the magnitude matrix of Fourier coefficients of a digital image represented by the function $g \in L^2(\mathbb{T})$ be defined as

$$K(m,n) = \log_e\left(1 + \sqrt{[\mathrm{Re}\,(\hat{g}(m,n))]^2 + [\mathrm{Im}\,(\hat{g}(m,n))]^2}\right),$$

where $\hat{g}(m,n)$ are the complex Fourier coefficients just after the transform. Since the modulus of a complex number and the natural logarithm are always positive in this case (since we add 1 to the number, otherwise it could be negative), we have that every element generated by the function K is positive for any outcome of $\hat{g} \in \mathbb{C}$.

The sampling procedure is the element-wise mean of a discrete amount of inradii, taken from a inscribed circle within the matrix. Each inradius is a one-dimensional set S of real numbers from K. Formally, the sampling consists of an operator $T : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ defined as

$$T(x_i) = \frac{\sum_{j=1}^{n} x_{ij}}{n},$$

where x_i is the i-th element of the final sample array and $\sum_{j=1}^{n} x_{ij}$ is the sum of each i-th element for each inradius S. We now need to show that $K \in \ell^2(\mathbb{T})$. Indeed, $L^2(\mathbb{T}) \subset \mathcal{D}'(\mathbb{T})$, due to the following facts:

(i) Every function $\varphi : \mathbb{T} \to \mathbb{C}$, $\varphi \in C_c^{\infty}(\mathbb{T})$ is **locally integrable**, i.e.

$$\int_{\mathbb{T}} |\varphi| dx < \infty.$$

Since $\mathbb{T}$ is compact and φ is continuous, then φ is consequently integrable all over $\mathbb{T}$;

(ii) Every function $\varphi : \mathbb{T} \to \mathbb{C}$, $\varphi \in L^p(\mathbb{T})$ is **locally integrable**, i.e.

$$\int_{\mathbb{T}} |f|^p dx < \infty.$$

It belongs to $L^p(\mathbb{T}^2)$ for all compact subsets from $\mathbb{T}^2$, then φ is called **locally p-integrable**;

(iii) The space $L^2(\mathbb{T})$ is a subset of $L^p(\mathbb{T})$ when $1 \leq p \leq 2$, which is our case, and also $L^p(\mathbb{T})$ is a subset of $\mathcal{D}'(\mathbb{T})$. Thus, $L^2(\mathbb{T}) \subset L^p(\mathbb{T}) \subset \mathcal{D}'(\mathbb{T})$;

(iv) From Corollary 1, it follows that the space $L^2(\mathbb{T})$ is isomorphic and isometric to $\ell^2(\mathbb{Z}^2)$.

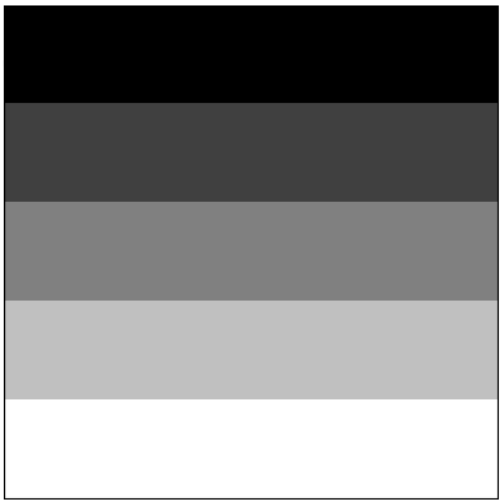
□

We will also show a numerical approach to complete the proof. Let I be an arbitrary matrix which represents a digital image, described by

$$I = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.50 & 0.50 & 0.50 & 0.50 & 0.50 \\ 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \\ 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \end{bmatrix}.$$

Practically, I represents the set of numbers that the g would produce after the acquisition of the following scene by an arbitrary imaging system:

Figure 37 – Arbitrary scene acquired by an arbitrary imaging system.



Source: Elaborated by the author.

The DFT transforms I into a set of Fourier coefficients, as

$$\hat{I} = \begin{bmatrix} 12.5+0i & 0+0i & 0+0i & 0+0i & 0+0i \\ -3.125+4.30119i & 0+0i & 0+0i & 0+0i & 0+0i \\ -3.125+1.01537i & 0+0i & 0+0i & 0+0i & 0+0i \\ -3.125-1.01537i & 0+0i & 0+0i & 0+0i & 0+0i \\ -3.125-4.30119i & 0+0i & 0+0i & 0+0i & 0+0i \end{bmatrix}.$$

Applying $\hat{I}$ in K, we obtain the magnitude matrix of the Fourier coefficients from $\hat{I}$, denoted as M and given by

$$M = \begin{bmatrix} 2.60269 & 0 & 0 & 0 & 0 \\ 1.84318 & 0 & 0 & 0 & 0 \\ 1.45531 & 0 & 0 & 0 & 0 \\ 1.45531 & 0 & 0 & 0 & 0 \\ 1.84318 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

From this result, we may compute the sum of the square of absolute values of M in order to show that the result is finite. Indeed,

$$\sum_{i=1}^{m}\sum_{j=1}^{n} |M|^2 \approx 18 < \infty,$$

what completes the proof. ■


ICMC
USP
SÃO CARLOS