

# TECHNICAL UNIVERSITY OF CRETE

School of Electrical And Computer Engineering

DIPLOMA THESIS

---

## Development and Evaluation of Deconvolution Microscopy Methods

---

*by*  
*Nikoloulas Dimitrios*



*A thesis submitted in partial fulfillment of the requirements  
for the Diploma of Electrical and Computer Engineering*

in the

Electronics Laboratory

Department of Electrical and Computer Engineering

### THESIS COMMITTEE

Professor Balas Costas (Supervisor)

Professor Zervakis Michael

Professor Mania Aikaterini

June 2023





# Acknowledgements

I would like to thank heartfully many people for helping me not only to complete my thesis but most important for their support and cooperation throughout my undergraduate studies.

Firstly, i would like to thank my family and mostly my parents Athanasia and Argyris that always believe, support and encourage me to pursuit my dreams. I would like to thank wholeheartedly my supervisor Professor Costas Balas, who has offered me valuable knowledge and guidance on the area of optoelectronics and imaging diagnostics during my undergraduate studies and most important for his trust, excellent support and cooperation from the begging till the end of my work inside his research group. I would also like to thank Professor Michael Zervakis and Professor Aikaterini Mania for their participation in the examination comittee.

Last but not least, I would like to thank my thesis advisor, Chousos Christos, for giving me the inspiration and the will to learn and to work for new scientific fields and of course my close friends for all the good memories that we have made through the years.



# Abstract

Microscopy is a powerful imaging technique that has enabled researchers to visualize and study the micro- and nano-scale structures of various biological and material systems. However, traditional microscopy techniques are often limited by a narrow depth of field, which means that only a small portion of the specimen can be in focus at a time. To overcome this limitation, multifocus image fusion techniques have been developed to combine multiple images of the same specimen taken at different focal planes to produce a final image that is in focus across the entire depth of the specimen. This thesis explores the application of multifocus image fusion techniques to microscopy, with a focus on the development of novel algorithms that can accurately and efficiently fuse multiple images of a specimen taken at different focal planes.

The diploma thesis focuses on the development of a Quadtree-based multifocus image fusion algorithm that can handle large and complex microscopy images. The proposed algorithm uses a weighted focus measure to determine the focus of different regions of the image and a Quadtree data structure to represent the focus of different regions at different levels of resolution. The Quadtree is then used to guide the fusion process, with regions of the image that are in focus at different focal planes being selected and combined to produce the final fused image. The performance of the algorithm is evaluated using a range of microscopy images and compared with existing state-of-the-art algorithms.



# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
<b>2</b>	<b>Microscopy Principles and Instrumentation</b>	<b>10</b>
2.1	Microscopy . . . . .	10
2.2	Interaction of Light . . . . .	10
2.3	Lenses . . . . .	11
2.3.1	Thin Lenses . . . . .	12
2.3.2	Lens With Important Thickness . . . . .	13
2.3.3	Two Lens System . . . . .	13
2.3.4	Conventional Microscope Lenses . . . . .	14
2.3.5	Using An Image System Instead Of Eyepiece . . . . .	14
2.4	Limits Of Magnification . . . . .	15
2.5	Diffraction . . . . .	15
2.6	Numerical Aperture And Resolution . . . . .	18
2.6.1	Digital Resolution . . . . .	19
2.7	Point Spread Function . . . . .	19
2.7.1	Theoretical Point Spread Function . . . . .	20
2.8	Optical Microscopes . . . . .	24
2.8.1	Compound Microscope . . . . .	24
2.8.2	Dark-field Microscope . . . . .	25
2.8.3	Fluorescence Microscope . . . . .	25
2.8.4	Confocal Microscope . . . . .	29
2.8.5	Why Widefield Confocal Microscopy over Confocal Microscopy	33
<b>3</b>	<b>Multi-Focus Image Fusion</b>	<b>35</b>
3.1	Development of Image Fusion . . . . .	36
3.2	Fundamentals of Image Fusion . . . . .	37
3.2.1	WHY Combine Information of Multiple Images . . . . .	37
3.2.2	Definition of Image Fusion . . . . .	38
3.2.3	Objective of Image Fusion . . . . .	38
3.3	Categories of Fusion Algorithms . . . . .	39
3.4	Steps of Image Fusion . . . . .	40
3.5	Applications . . . . .	41
3.6	Pixel-Level Image Fusion . . . . .	41
3.6.1	Introduction . . . . .	41
3.6.2	Single-Scale Image Fusion . . . . .	42
3.6.3	Multi-Scale Image Fusion . . . . .	43
3.7	Feature-Level Image Fusion . . . . .	48

3.7.1	Introduction . . . . .	48
3.7.2	Feature-Level Fusion Algorithm . . . . .	49
3.7.3	Genetic Algorithm . . . . .	49
3.8	Decision-Level Image Fusion . . . . .	50
3.8.1	Fusion Algorithm Based on Voting Method . . . . .	50
3.8.2	Fusion Algorithm Based on Bayes Inference . . . . .	51
3.8.3	Fusion Algorithm Based on Min-Max Rule . . . . .	52
<b>4</b>	<b>Wide-field Confocal Image Reconstruction</b>	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Proposed Algorithm . . . . .	53
4.2.1	Read-Split Images . . . . .	54
4.2.2	Detecting the focused regions . . . . .	54
4.2.3	Detect focus Blocks in quadtree structure . . . . .	58
4.2.4	Weighted focus measure: Sum of Weighted Laplacian . . . . .	60
4.2.5	Reconstruction of the focused regions . . . . .	62
4.2.6	Final Fused Image . . . . .	63
4.3	Experimental-Results . . . . .	65
4.3.1	Qualitative Comparison . . . . .	65
4.3.2	Quantitative Comparison . . . . .	68
4.4	Lab-Results . . . . .	70
4.4.1	Human Hair Crossed with a Synthetic Fiber . . . . .	70
4.4.2	Convallaria Rhizome . . . . .	72
4.4.3	Sticky Tape Edge and Cotton Fiber . . . . .	73
4.4.4	Quartz Sand . . . . .	75
4.4.5	Sugar Granule . . . . .	76
<b>5</b>	<b>Conclusion</b>	<b>78</b>



# List of Figures

2.1	Interaction of light with a medium . . . . .	11
2.2	General shapes of lenses . . . . .	11
2.3	Image formation of an object . . . . .	12
2.4	Compound of lenses . . . . .	13
2.5	Tube lens eyepiece compound . . . . .	14
2.6	Huygens-Fresnel principle of diffraction of light . . . . .	16
2.7	Fraunhofer diffraction pattern . . . . .	17
2.8	Fraunhofer diffraction pattern . . . . .	17
2.9	Point Spread Function created by Huygens-Fresnel model . . . . .	20
2.10	PSF formed by "infinite" wavelets of the Huygens- Fresnel model. . . . .	20
2.11	Compound Microscope . . . . .	24
2.12	Dark-field Microscope . . . . .	25
2.13	Fluorescence microscope . . . . .	26
2.14	Jablonski energy diagram. . . . .	27
2.15	Widefield (a-c) and laser scanning confocal (d-f) fluorescence microscopy images . . . . .	29
2.16	Diagram of epi-fluorescence laser scanning confocal microscope . . . . .	30
2.17	Confocal Microscopy Scanning Unit . . . . .	31
2.18	Wide-field Confocal Microscopy Diagram . . . . .	33
3.1	Human sensors (a) eyes, (b) ears, (c) nose, (d) tongue, (e) skin . . . . .	35
3.2	Human sensor fusion system for decision-making . . . . .	36
3.3	Single sensor imaging. (a) Single sensor, (b) Single sensor array and its corresponding image represented as a matrix, (c) Digitization process . . . . .	37
3.4	Graphical illustration of image fusion process . . . . .	39
3.5	Steps in image fusion system . . . . .	40
3.6	classification of image fusion methods . . . . .	41
3.7	The multi-scale image decomposition (MSD) . . . . .	44
3.8	Block diagram of the multi-scale image fusion . . . . .	44
3.9	Block diagram of pyramid-based image fusion . . . . .	46
3.10	Block diagram of wavelet-based image fusion . . . . .	47
3.11	Block diagram of filtering-based image fusion . . . . .	48
4.1	Read images and red channel ready for fusion . . . . .	54
4.2	Read images function . . . . .	54
4.3	Situations on a block-pair with focus on shaded regions. Situation 1: one block is fully focused, the other fully defocused; Situation 2: both are partially focused. . . . .	55
4.4	Check the situation . . . . .	55



4.5	Gradients calculations . . . . .	56
4.6	Focus-measures calculation . . . . .	57
4.7	Center the images . . . . .	58
4.8	Algorithm 1 . . . . .	59
4.9	Detecting the focused blocks . . . . .	59
4.10	Modified Laplacian calculation . . . . .	60
4.11	Calculate the weight for ML . . . . .	61
4.12	SWML . . . . .	61
4.13	Morphological Filter . . . . .	62
4.14	Small Region Filter . . . . .	63
4.15	Non part selection math expression . . . . .	64
4.16	Create the first part of the Fused Image from max gradients . . . . .	64
4.17	Select the non parts and create the final Fused Image . . . . .	64
4.18	The multi-focus ‘Disk’ images and its fusion images. (a and b) The source images; (c-i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively. . . . .	65
4.19	The multi-focus ‘Pepsi’ images and its fusion images. (a and b) The source images; (c-i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively. . . . .	66
4.20	The multi-focus ‘Lab’ images and its fusion images. (a and b) The source images; (c-i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively. . . . .	67
4.21	Samples focused in different spots . . . . .	70
4.22	Human Hair Crossed with a Synthetic Fiber Sample RGB Channels .	71
4.23	Human Hair Crossed with a Synthetic Fiber Final RGB Fused Image	71
4.24	Samples focused in different spots . . . . .	72
4.25	Convallaria Rhizome Sample RGB Channels . . . . .	72
4.26	Convallaria Rhizome Final RGB Fused Image . . . . .	73
4.27	Samples focused in different spots . . . . .	73
4.28	Sticky Tape Edge and Cotton Fiber Sample RGB Channels . . . . .	74
4.29	Sticky Tape Edge and Cotton Fiber Final RGB Fused Image . . . . .	74
4.30	Samples focused in different spots . . . . .	75
4.31	Quartz Sand Sample RGB Channels . . . . .	75
4.32	Quartz Sand Final RGB Fused Image . . . . .	76
4.33	Sugar Granule focused in different spots . . . . .	76
4.34	Sugar Granule Sample RGB Channels . . . . .	77
4.35	Sugar Granule Final RGB Fused Image . . . . .	77



# List Of Abbreviations

<b>RoC</b>	Radii Of The Curvature
<b>NA</b>	Numerical Aperture
<b>PSF</b>	Point Spread Function
<b>PCA</b>	Principal Component Analysis
<b>ICA</b>	Independent Component Analysis
<b>IHS</b>	Intensity-Hue-Saturation
<b>EOL</b>	Energy of Laplacian
<b>SML</b>	Sum-Modified Laplacian
<b>SF</b>	Spatial Frequency
<b>MSD</b>	Multi-scale Decomposition
<b>HVS</b>	Human Visual System
<b>DWT</b>	Discrete Wavelet Transform
<b>EPD</b>	Edge-preserving Decomposition Filters
<b>MDFM</b>	Maximum Difference in Focus-Measures
<b>SMDG</b>	Sum of the Maximum Difference in Gradients
<b>ML</b>	Modified Laplacian
<b>WML</b>	Weighted Modified Laplacian
<b>SWML</b>	Sum of the Weighted Modified Laplacian



# Chapter 1

## Introduction

Over the last four centuries, the rise and evolution of optical light microscopes was inevitable, as microscopy constitutes an integral diagnostic and imaging tool for many fields of science, such as medicine, biology, chemistry as well as for many applications in industry. Starting back in 17th century, when Dutch scientist Antonie van Leeuwenhoek developed the bi-convex lenses of small focal length, in order to achieve high magnifications of small objects. Since then, the design of optical microscopes evolved in their today's form of compound setups.

Nowadays, concurrently with optical light microscopy, the need for advanced observations of specimens has led to the development of new setups that are able to reveal specific characteristics and structures of samples under examination. Indicatively, fluorescence microscopes are capable to image specific cellular structures (e.g., intracellular skeleton structures, nucleus organelles, elements of genetic material, etc.), using a combination of appropriate fluorophores and monochromatic illuminations. It is a fact that microscopes are not used as mere observation devices, but employed as major diagnostic instruments in a wide range of biological and medical procedures, such as cytology, pathological anatomy, biopsy etc.

However, despite the technological progress of microscopy, acquiring images with a large depth of field (DOF) remains a challenge. The DOF is defined as the distance range over which the image is in focus. In traditional microscopy, images with a large DOF are obtained by capturing multiple images at different focal planes and merging them into a single in-focus image using image fusion techniques.

One of the most promising image fusion techniques for microscopy is Quadtree-based multi-focus image fusion using a weighted focus-measure. This technique is based on the quadtree decomposition of the input images, where each quadtree node corresponds to a sub-image with a different focus. The weighted focus-measure is then used to assign weights to each sub-image, and the final in-focus image is obtained by fusing the weighted sub-images. The weighted focus-measure can effectively handle the problem of uneven illumination and background noise present in microscopy images.

In this diploma thesis, we aim to investigate the effectiveness of Quadtree-based multi-focus image fusion using a weighted focus-measure in improving the DOF

of microscopy images. This thesis is organized in four main chapters. The first chapter includes an extensive analysis of the theoretical background used for the development of microscopy. The second chapter presents the theory about multi focus image fusion as well as the state of the art, while the third and last main chapter introduces our implementation and the acquired results. Finally the fourth chapter contains the conclusion and the future work.

# Chapter 2

## Microscopy Principles and Instrumentation

### 2.1 Microscopy

Microscopy refers to the use of microscopes to observe and study small objects and structures that are not visible to the naked eye. Microscopes can magnify objects by several orders of magnitude, allowing scientists to explore the intricate details of cells, tissues, and microorganisms. There are several types of microscopy, including light microscopy, electron microscopy, and scanning probe microscopy, each with their own unique capabilities and applications. Microscopy plays a crucial role in many scientific fields, including biology, medicine, materials science, and nanotechnology, and has revolutionized our understanding of the world around us.

### 2.2 Interaction of Light

Light can interact with a medium in many ways, depending on the material of the medium, the angle of the incidence and the wavelength. As for diffraction, this phenomenon is based on the principals of constructive and destructive wave interference. Light behaves in the same way as a mechanical wave encountering an obstacle or passing through a slit. It is defined as the bending of light around the corners of an obstacle or a slit (aperture) into the region of geometrical shadow of the obstacle. In order to exhibit diffraction, this obstacle or this slit must be comparable in size with the wavelength of the encountering wave. From this, also it follows that an obstacle or a slit can have sharp edges.

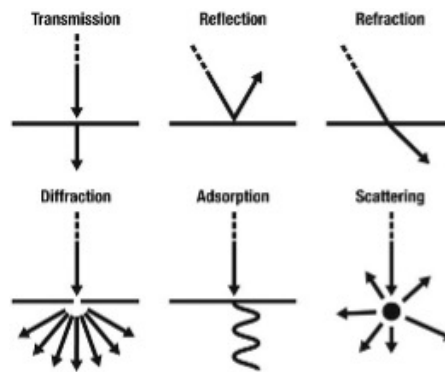


Figure 2.1: Interaction of light with a medium

In microscopy, all these kinds of interaction are significant. Transmission, reflection, refraction are used to build a path for the light rays. Diffraction, absorption and scattering, though, eventually act as degrading factors for the microscoped image. For the purposes of light microscopy, the basic components relative to the optical train are lenses and filters.

## 2.3 Lenses

Lenses come in different shapes. They can be grouped and form a compound system with special properties (i.e better focusing, better magnification or eradication of aberrations). Two are the main characteristics of a lens. The distance over which initially parallel light rays are brought to a focus (focal length  $f$ ) and the magnification. Image formation happens when light rays converge either in real or seem to do so (virtually).[19]

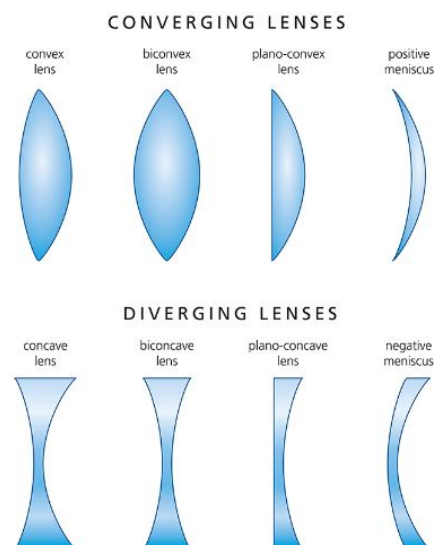


Figure 2.2: General shapes of lenses



- A real image is formed in the plane where light rays converge in real (using converging lenses) and the object to be pictured is placed farther than the focal length of the lens. When the object is placed in between  $f$  and  $2f$  the image is bigger than the object, whereas when it is placed in a distance bigger than  $2f$  the image is smaller. A real image is always inverted and can be formed in a screen. For a converging lens, the focal length is said to be positive, which is the reason why they are also called positive.
- A virtual image is formed in the plane where light rays seem to converge (continuing ,virtually backwards, the rays coming out of the lens). This happens in diverging lenses or when the object to be pictured is placed inside the focal distance of a converging lens. A virtual image is always erect and cannot be projected onto a screen as it on the same side of the lens as the object.

### 2.3.1 Thin Lenses

In the simple case of just one thin convex lens (thickness is negligible), the equation of the thin lens and the provided magnification can be computed, using the metric relations of the formed triangles:[10]

$$\frac{1}{f} = \frac{1}{s_1} + \frac{1}{s_2} \quad (2.1)$$

, where  $s_1$ : object to lens distance,  $s_2$ : real image to lens distance,  $f$ : focal length

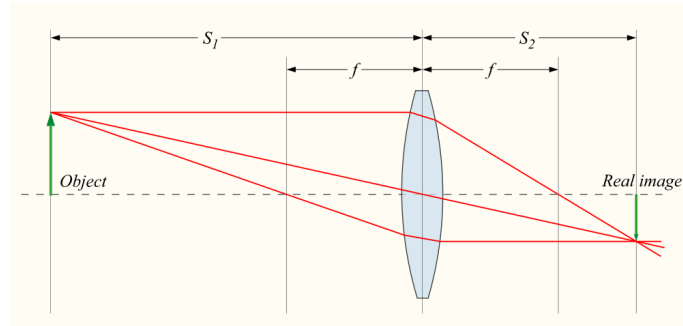


Figure 2.3: Image formation of an object

$$M = \frac{y_i}{y_0} = \frac{s_2}{s_1} \quad (2.2)$$

,where  $y_i$ : image height,  $y_0$ : object height.

Usually, the object is put on the  $f_0$ , so the above equation becomes:

$$M = \frac{y_i}{y_0} = \frac{s_2}{f_0} \quad (2.3)$$

### 2.3.2 Lens With Important Thickness

When the thickness of the lens is not negligible, in order to compute the effective focal length, radii of the curvature (RoC) of each side of the lens, along with the refractive index (r.i) of the lens medium must be known. So the corresponding equation is:

$$\frac{1}{f} = (n - 1) \left[ \frac{1}{R_1} - \frac{1}{R_2} + \frac{(n - 1)d}{nR_1R_2} \right] \quad (2.4)$$

,where  $R_1$ : RoC of closer to the light source lens side,  $R_2$ : RoC of farther to the light source lens side,  $n$ : r.i of the lens medium,  $d$ : lens thickness.

### 2.3.3 Two Lens System

A compound microscope system of two lenses provides better magnification than that of one lens. It consists of at least two converging lenses; the objective and the eyepiece. The former has a focal length  $f_o < 1$  cm and the latter has a  $f_e$  of a few cm. The total magnification  $M$  of the two lenses is given by the product of the magnifications of the individual lenses. Thus, we have:

$$M = M_{objective} \cdot M_{eyepiece} \quad (2.5)$$

The individual magnifications are given by Eq.2.2, which in order lead to:

$$M = \frac{S_2}{S_1} \cdot \frac{S'_2}{S'_1} \quad (2.6)$$

The distances between object-objective lens  $S_1$ , objective lens-real (first) image  $S_2$  and real image-eyepiece lens  $S'_1$  may vary. But the distance between the eyepiece lens

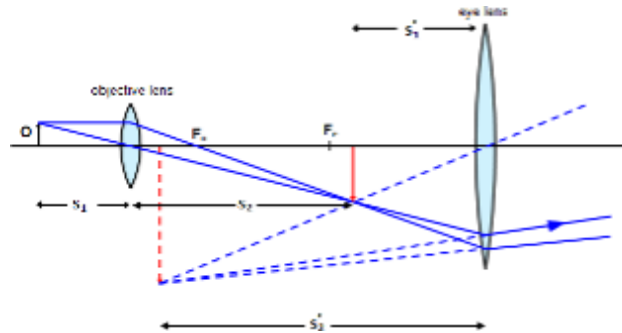


Figure 2.4: Compound of lenses

and the virtual (final) image  $S'_2$  is set to 25 cm. This is because 25 cm is the closest point at which an object can be brought into focus by a "normal" human eye. Closest distance is important, because apparently the image of the object will

be the most detailed it can be. Eventually, the eye perceives the final magnified image as if it were in that distance of 25 cm.

When the object is put on the  $f_0$ , the intermediate image is formed on the  $f_e$ . Hence the above equation becomes:

$$M = \frac{T}{f_0} \cdot \frac{25\text{cm}}{f_e} \quad (2.7)$$

, where T : distance between  $f_0$  and  $f_e$ , usually 16 cm.

### 2.3.4 Conventional Microscope Lenses

In a typical microscope a third lens is added, namely the tube lens, for the reason that the objective lens projects incoming light into infinity. The tube lens forms the real image on its focal point. Light rays originating from the two ends of the object are adequate for the purpose of explaining conventional microscope lenses. To take it from the start: The objective is designed to project the incoming light rays of the object into infinity. Then, the tube lens produces a magnified intermediate image which in turn is captured by the eyepiece and finally shown to the eye. The resulting viewing angle  $\delta_1$  of the case A is much larger than  $\delta_2$  of case B, where the object is seen directly from a distance of approximately 25 cm. Regarding the overall magnification of such a system, it can be computed by the product of the individual magnifications of the objective and the eyepiece lenses.

### 2.3.5 Using An Image System Instead Of Eyepiece

In cases that recording of the specimen in an imaging system of a sensor and a computer monitor is preferred, the eyepiece can be omitted. The magnification is produced by the optical system (i.e objective, tube lens) and the electronic imaging system which provides a magnification factor due to the optical to electronic system adaptor (computer monitor adaptor). The total magnification of the system can be computed as:

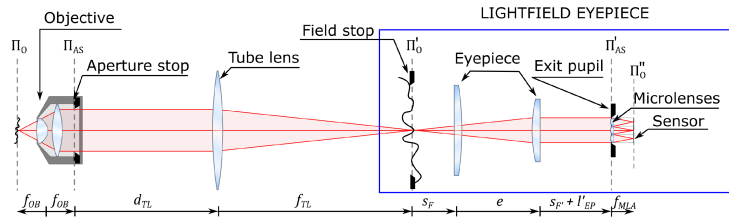


Figure 2.5: Tube lens eyepiece compound

$$M_{overall} = M_{optical} \cdot M_{electronic} \quad (2.8)$$

The optical magnification in case of just one objective and an adapter is given by:

$$M_{optical} = M_{objective} \cdot M_{adapter} \quad (2.9)$$

The electronic magnification is the ratio of the active sensor diagonal and the useful monitor diagonal:

$$M_{electronic} = \frac{diagonal_{monitor}}{diagonal_{sensor}} \quad (2.10)$$

## 2.4 Limits Of Magnification

As it was mentioned above, greater magnification can be achieved using two or more lenses. So, the question is if there are any boundaries to magnifying an object infinitely. The short answer is NO. Despite the fact that with an infinite number of lenses an infinite magnification is achieved, after a point that huge magnification is useless because "deeper" details of the magnified object cannot be resolved further. So the maximum useful magnification for a conventional optical microscope it is known to be 2000x. In terms of the tiniest possible distinguishable area of an object, the conventional optical microscope can produce a virtual image of a 200 nm diameter spot of the object. This is approximately the width of an average-sized bacterium. The smaller this distinguishable area is, the bigger the resolution of the microscope is. However, in practice this resolution of 200 nm cannot be approached due to lenses' imperfections. Further improvements have been accomplished by the Nobel prize-awarded Super Resolution Microscopy which circumvents the aforementioned limit and brings it to the nano-dimension. The explanation to the existence of a limit to the resolving power of the microscope, lies in the concept of diffraction which is eventually the reason why more and more magnification power of lenses lead merely to a magnification of the smallest distinguishable area, without unveiling further details.[15]

## 2.5 Diffraction

Lens-systems have an opening which allows a cone of light to pass through. Here lies the resolution limitation of the microscope (and of every optical device incorporating apertures). As it was mentioned in the early introduction of this section, light bends around the corners of an obstacle or a slit (aperture) into the region of geometrical shadow of the obstacle when the size of the latter is comparable with the wavelength of the incoming light. From this, also it follows that light bends around the sharp edges of an obstacle or an aperture.[21]

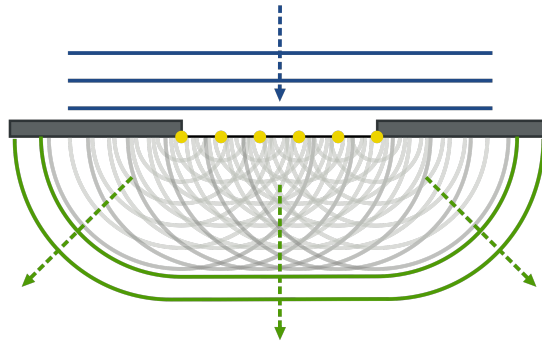


Figure 2.6: Huygens-Fresnel principle of diffraction of light

The significance of the phenomenon of diffraction in the case of image capturing lies in the fact that light does not travel linearly from the aperture towards the sensor (camera or human eye), but instead spreads to a cone angle. As it was mentioned earlier, light diffraction behaves in the same way a mechanical wave behaves entering a slit. The spreading of the post-aperture light can be explained by the Huygens-Fresnel principle. It argues that every single point of a wave front (wavelet) acts like a source of new spherical waves. These emanating waves interfere constructively or destructively with themselves creating a characteristic pattern in the post-aperture area. In Fig.2.6, this principle is visualized with the use of 6 wavelets at the level of the aperture. While these waves propagate the total aggregating wave front takes the form of the green lines, giving the effect of the spreading of light.

When the spreading light eventually "hits" the sensor, a two-dimensional diffraction pattern appears on it. This is the resolved image of the point light source. So, as long as the light bends after it passes through the slit, the resolved image is larger. This means that the final captured image is a degraded form of the true point source. Fig.2.7 exhibits a case where light arrives at the aperture in plane wave fronts. That is accomplished when the light source, the aperture and the sensor are far apart in order for the spherical source waves to expand in such a degree that their wave fronts become plane or close to plane. Another occasion is when curved wave fronts enter the objective lens of the microscope and come out as plane ones. This is the Fraunhofer diffraction. In case where the incoming wave fronts are curved and sensor is relatively close to the aperture, Fresnel diffraction applies. Fraunhofer diffraction helps the understanding of the parameters that affect the level of light bending and the diffraction patterns on the sensor, because of its simplicity compared to Fresnel

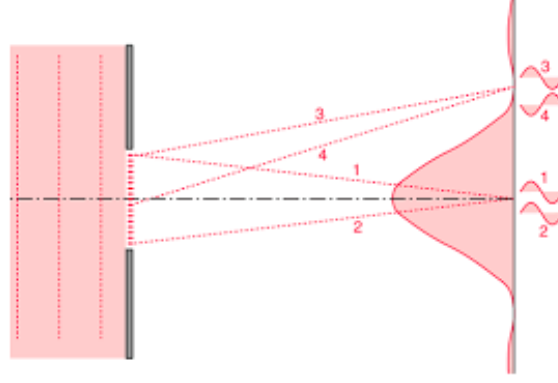


Figure 2.7: Fraunhofer diffraction pattern

diffraction. The figure shows the sensor plane and the intensity of light in each point on it.

A wave of light interferes constructively with another of the same wavelength  $\lambda$  when their path lengths differ by  $m$  multiples of  $\lambda$ . They interfere destructively when their path lengths differ by  $(m + \frac{1}{2})\lambda$ . In Fig.2.7 rays 3 and 4 arrive 180 degrees out of phase on the sensor and lead to a minimum light intensity in that point. Rays 1 and 2 arrive in phase and lead to maximum light intensity. In order to calculate how much the light diffracts, after it passes the aperture, it is sufficient to calculate the angle of the first intensity minimum over the optical axis (aperture middle). From Fig.2.8, using trigonometric small-angle approximations (because Fraunhofer approach considers that aperture to slit length  $L \gg w$  and thus  $\theta' \approx \theta$  and applying the destructive interference condition  $\frac{w}{2} \sin \theta = \frac{\lambda}{2}$ , it can be shown that the angle of first intensity minimum (first trough or dark fringe) is:

$$\sin \theta = 1.22 \cdot \frac{\lambda}{w} \quad (2.11)$$

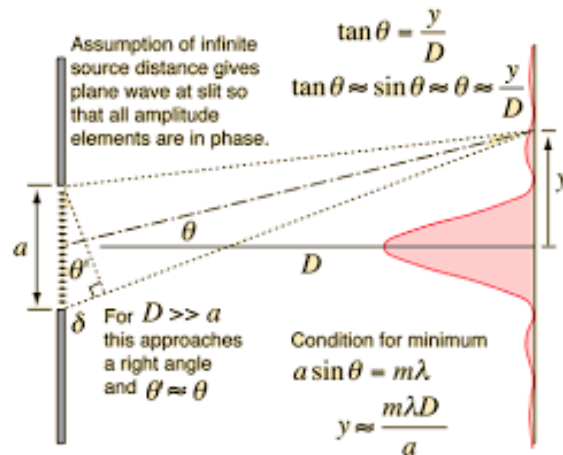


Figure 2.8: Fraunhofer diffraction pattern

The above equation makes clear that the longer the wavelength of the incoming light, the larger the level of light bending after the aperture. Additionally, the bigger the aperture, the smaller the degree to which the light diffracts. Hence, in order to have a better resolved image of the true point source regardless its wavelength, the bigger the aperture is, the better this can be achieved. Bigger aperture means longer lens so as the latter gathers all the light that passes through the aperture. To express how large cone of light a lens can gather, a quantity is introduced.

## 2.6 Numerical Aperture And Resolution

Numerical Aperture of a lens (NA) : characterizes the range of angles over which the lens can accept (or emit) light. It is a function of the refractive index  $n$  of the medium in which the lens is working (i.e air, water, oil) and half the angle ( $\theta$ ) of the largest pencil of light that the lens can accept (or emit).[7]

$$NA = n \cdot \sin \theta \quad (2.12)$$

From this definition along with the aforementioned simple aperture characteristics, it can be concluded that the bigger the numerical aperture, the smaller the diffraction spreading and thus the better the resolution of a point source. But in order to quantify the resolution of the lens-aperture system, the smallest distance between two diffraction patterns, caused by two point sources, must be found so that they both can be resolved (seen separately). This distance is called the Rayleigh resolution limit and is defined as the distance between the crests of the two diffraction patterns, when the crest of the main lobe of the first one meets the first trough of the second one. It is equivalent to the distance between the crest of the main lobe and the first dark fringe. This lateral resolution is a function of wavelength and numerical aperture.

Because of refraction and reflection phenomena that appear when light passes from the coverslip of the specimen to the objective through a gap of air, it is preferable that an immersion oil (high refractive index medium) is used. This allows light to travel straight towards the objective, leading thus to a larger effective NA. So, the refractive index of the immersion oil, also, plays a role in the final formation of the diffraction pattern.

$$r_{lateral} = \frac{1.22\lambda}{2n \sin \theta} = \frac{0.61\lambda}{NA} \quad (2.13)$$

, where  $r$ : lateral resolution of the lens-aperture system,  $\lambda$ : wavelength of light,  $n$ : refractive index of the medium surrounding the point sources,  $\theta$ : half-angle of the cone of light that enters the lens,  $NA$ : effective numerical aperture of the lens-aperture system.

A condenser lens is often used so as to focus the illuminating cone of light onto the sample. This has to be computed too. The lateral resolution is now given by:

$$r_{lateral} = \frac{1.22\lambda}{NA_{obj} + NA_{cond}} \quad (2.14)$$

In a properly configured microscope, the condenser must have equal  $NA$  with the objective. So, this leads to the first resolution equation (Eq.2.13)

### 2.6.1 Digital Resolution

When it comes to capturing an image and translate it to a digital one, the question of how the resolution is related to the image pixels arises. The image 2D signal in order to be aliasing-free and properly represent the analog true signal needs to be sampled accordingly to the Nyquist theorem. The latter states that a signal in general needs to be sample in a rate  $f_s > 2f_{max}$ , with  $f_{max}$  the highest frequency that is wished to be recorded. In an image, frequency has the meaning of periodic changes of luminance of the depicted objects. Thus the corresponding "period" has to do with distance between this changes. The Nyquist sampling rate can be modified to:  $D_s < D_{min}/2$ , where  $D$  corresponds to distance (or size). This gives the imaging sample rate (pixel size) which should be 1/2 the size of the smallest object (resolution) that is wished to be recorded. In practice, an image pixel usually represents distance that is 2.3-3 times smaller than the optical resolution.[20] Let it be that the sampling process is completed. In order to find the physical distance that is represented by a pixel without knowing the sampling rate, the sensor pixel size and the magnification of the optical system must be known: sensor pixel size

$$physical\ distance = \frac{sensor\ pixel\ size}{overall\ magnification} \quad (2.15)$$

## 2.7 Point Spread Function

Fig.2.8 shows the formation of a diffraction pattern on the sensor plane. Because this pattern is the spreaded light initially emanated from a point source, it is given the name Point Spread Function (PSF). The term "function" is valid because the spreading, as it was said, depends on the  $NA$  of the lens-aperture system, the refractive index of the immersion oil (if it exists) and the  $\lambda$  of the source light.

By simulating the single-slit phenomenon in a computer, the PSF can be visualized. Using 2,3 and 9 in-phase wavelets of the Huygens-Fresnel model (Fig.2.12), it can be shown how a PSF is formed across the aperture-sensor plane space. As long as there is no finite number of wavelets, also a simulation of "infinite" ones is shown in Fig.2.13 The resolution on the image plane (lateral resolution) is given by Eq 2.13. The tiniest discernible distance in the optical axis (depth axis) is larger that the lateral one. It can be computed by:[17]



$$r_{axial} = \frac{2\lambda n}{(NA)^2} \quad (2.16)$$

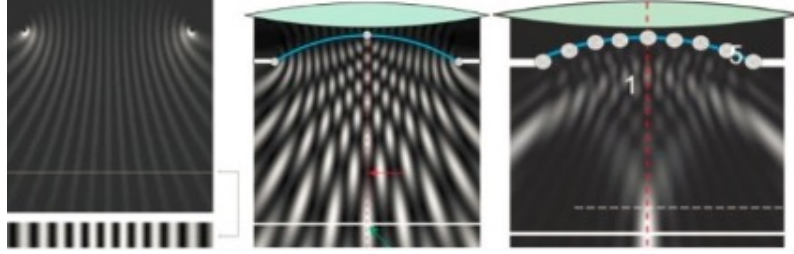


Figure 2.9: Point Spread Function created by Huygens-Fresnel model

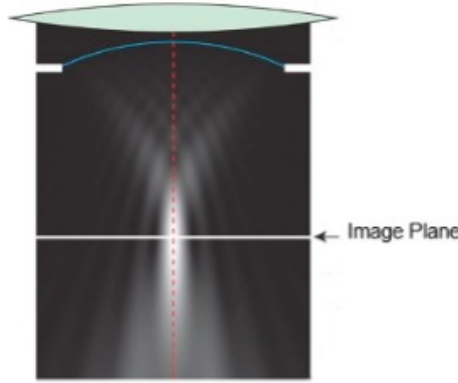


Figure 2.10: PSF formed by "infinite" wavelets of the Huygens- Fresnel model.

This leads to the same conclusion as for the lateral resolution. That is, the shorter the  $\lambda$  or/and the larger the  $NA$ , the better the axial resolution. The PSF can either be calculated theoretically by various models or be measured experimentally.

### 2.7.1 Theoretical Point Spread Function

Taking into account how the wave fronts of light interfere with a propagation medium, a mathematical expression can be built gradually. This expression will be in terms of wave amplitude and phase as a function of the 2D coordinates of the propagating wave of light at a specific, perpendicular to the optical axis, plane  $z$  (depth plane). The emission of waves from a point source can be seen as an input impulse:[18]

$$u_0(x, y) = A_0\delta(x, y) \quad (2.17)$$

, where  $A_0$ : initial wave amplitude of emitting source light.

At first, light travels in free space meaning no lenses or apertures are included. The free-space system has an impulse response, let it be  $h(x, y; z)$ . Thus, the output wave amplitude of this system at the depth plane  $z_0$  just before the lens, will be:

$$u(x, y; z_0) = h(x, y; z_0) * A_0 \delta(x, y) = A_0 h(x, y; z_0) \quad (2.18)$$

, where  $*$  is the sign of convolution.

The affected transmission of light through the lens-aperture system of focal length  $f$  can be described with a pupil function which captures every optical aberration in the amplitude and phase that takes place between the object focus plane and the image plane. These aberrations (deviations from the ideal optical path) happen due to irregularities or misalignments in any component of the imaging system light path, especially the objective lens but also other lenses, mirrors, filters or apertures. Such aberrations are present in various forms. Speaking about rays of the same wavelength: *Astigmatism*, where rays that propagate in two perpendicular planes, have different focal spots. *Coma*, where slanted parallel rays entering a lens, eventually focus on different spots on the image 3D space. *Defocus*, where rays do not focus on the a priori designed focal spot. *Chromatic aberration*, where rays of different wavelengths have different focal spots. *Spherical aberrations*, where rays from different points across the lens, focus on different spots of the optical axis. *Distortion*, where a rectilinear object eventually appears curved. These aberrations can be integrated to a scale-causing factor as for the wave amplitude. Thus an ideal lens-aperture system will have a scaling factor of 1 for every point of the pupil plane. Possible deviations will cause the scaling factor to diminish. Consequently, a pupil function integrates amplitude and phase factors, thus making it a complex function:

$$P(x, y) = p(x, y) e^{j\Phi(x, y; f)} \quad (2.19)$$

So, the wave fronts equation when light enters the free-space just after the exit of the lens becomes:

$$u'(x, y; z_0) = A_0 h(x, y; z_0) p(x, y) e^{j\Phi(x, y; f)} \quad (2.20)$$

Note: no convolution operation here, as the pupil function plays a scaling role.

Finally, when light reaches the image plane, the free-space impulse response function at the image plane  $z_1$  convolves with the so far computed wave front equation. This leads to the wave function at the image plane which is the 2D PSF (on focus):

$$\begin{aligned} PSF(x, y) &= u'(x, y; z_0) * h(x, y; z_1) = \\ &= (A_0 h(x, y; z_0) p(x, y) e^{j\Phi(x, y; f)}) * h(x, y; z_1) \end{aligned} \quad (2.21)$$

The above process can be also done in order to compute the wave function in different planes from the image focus plane. Hence, a 3D PSF is formed. In reality, though, a single 3D model is not sufficient to completely describe image formation throughout the 3D object space. This is because PSF is often a function of the location of the point source in the object space. Thus, in reality every point of the object corresponds to a different PSF. A system, the Point Spread Functions of which behave in this way, is called a shift-variant system.

Different approximations can be made, depending on the nature of diffraction, the possible consideration of shift-variance and the amount and degree of effect of possible aberrations, if the latter are taken into account.

Assuming there are no aberrations and that the Fraunhofer approach is valid, a more specific PSF can be built. A pupil function for a circular aperture of radius  $\alpha$  through which light is transmitted unaffected has the value 1 inside the area encompassed by the aperture in the axes system  $x', y'$ . The polar coordinate translation is:  $r = \sqrt{x'^2 + y'^2}$  and  $\theta = \arctan(\frac{y'}{x'})$ .

At the image plane the axis system  $(x, y)$  is translated in polar coordinates as  $\rho = \sqrt{x^2 + y^2}$  and  $\psi = \arctan(\frac{y}{x})$ .

According to Principles of Optics by Born and Wolf, the diffraction pattern wave function (amplitude) is the Fourier Transform of the pupil function. From this it follows that:

$$D(x, y) = \int_0^\alpha \int_0^{2\pi} e^{-ikr\rho \cos(\theta-\psi)} r d\theta dr \quad (2.22)$$

, where  $k = 2\pi/\lambda$  (wavenumber).

Using the Bessel function of the first kind and order 0, the above equation can be modified as:

$$D(\rho) = 2\pi C \int_0^\alpha J_0(kr\rho) r dr \quad (2.23)$$

, where  $C$  : constant.

PSF is in terms of intensity of light which is proportional to the squared amplitude of the wave function. So, from the above equation arises that:

$$PSF(\rho) = \left| 2\pi C \int_0^\alpha J_0(kr\rho) r dr \right|^2 \quad (2.24)$$

For an aberration-free and shift-invariant in all directions PSF, the 3D Born and Wolf approximation can be useful. The constraint is though, that the observed fluorophore particle is located at the focal plane of the objective lens but right beneath the coverslip. The model is expressed in the Kirchhoff's diffraction integral formula as:

$$PSF(x, y, z) = \left| \int_0^1 J_0 \left[ k \frac{NA}{n_i} \sqrt{x^2 + y^2} \rho \right] e^{-\frac{1}{2}jk\rho^2 z (\frac{NA^2}{n_i})} \rho d\rho \right|^2 \quad (2.25)$$

, where  $PSF$ : scalar,  $J_0$ : Bessel function of the first kind of order 0 (Bessel function appear in problems of wave propagation),  $k$ : wavenumber of the emitted light in vacuum ( $k = 2\pi/\lambda$ ),  $n_i$ : refractive index of the immersion oil.

The Gibson and Lanni PSF model takes into consideration the consequences of the optical path modifications that occur when light rays follow trajectories other than that for which the optics were designed. It integrates the optical path difference (ODP), which is the difference between ideal trajectories and real trajectories, taking into account parameters determining path fluctuation: thickness ( $t_{oil}$ ) and refractive index ( $n_{oil}$ ) of the immersion oil, thickness ( $t_g$ ) and refractive index ( $n_g$ ) of the coverslip and the thickness ( $t_s$ ) and refractive index ( $n_s$ ) of the sample.

Any deviation of any of these parameters from the optimal values for which the optics were designed will cause spherical aberrations. The resulted prolonged focus spot has the effect of an asymmetry in the 3D PSF on the depth axis. In the case of microscopy, this is a common phenomenon which increases as the object focal plane moves deeper into the sample, thus inducing more refractive index mismatches inside the sample as well as between the sample, the coverslip and the immersion medium. Spherical aberration is the reason why PSF is shift-variant on the depth axis.

for the Gibson and Lanni model, it incorporates shift-variance only in the depth axis. It can be seen as a generalization of Born and Wolf in the sense that the fluorophore particle can be located at any depth within the sample. It also considers three optical layers (sample-coverslip-immersion) instead of two (glass-immersion). The model description again in the Kirchhoff's diffraction integral formula is:

$$PSF(x, y, z) = \left| \frac{C}{z_d} \int_0^1 J_0 \left[ k\alpha\rho \frac{\sqrt{X^2 + y^2}}{z} \right] e^{jW(\rho)} \rho d\rho \right|^2 \quad (2.26)$$

, where  $PSF$ : scalar,  $C$ : a normalizing constant,  $z_d$ : tube lens-detector distance,  $W(\rho)$ : phase aberration induced:

$$\begin{aligned} W(\rho) = k \Big\{ & n_s t_s \sqrt{1 - \left( \frac{NA\rho}{n_s} \right)^2} + n_g t_g \sqrt{1 - \left( \frac{NA\rho}{n_g} \right)^2} \\ & + n_{oil} t_{oil} \sqrt{1 - \left( \frac{NA\rho}{n_{oil}} \right)^2} - n_g^* t_g^* \sqrt{1 - \left( \frac{NA\rho}{n_g^*} \right)^2} - n_i^* t_i^* \sqrt{1 - \left( \frac{NA\rho}{n_i^*} \right)^2} \Big\} \end{aligned} \quad (2.27)$$

## 2.8 Optical Microscopes

Optical microscopy, as was mentioned, has played a pivotal role in expanding our understanding of the microscopic world, enabling researchers to explore intricate details of biological specimens, materials, and various scientific phenomena. One of the remarkable aspects of optical microscopy lies in its versatility, with a wide array of microscope types available, each tailored to specific applications and research needs. These different types of optical microscopes employ diverse techniques and technologies to enhance contrast, resolution, and three-dimensional imaging capabilities. From the widely used compound microscope to specialized instruments such as phase-contrast, fluorescence, confocal, polarizing, and darkfield microscopes, each type offers unique advantages and brings valuable insights into the structures and properties of microscopic samples.[13]

### 2.8.1 Compound Microscope

The compound microscope is the most common type of microscope used in laboratories and educational settings. It consists of two or more lenses that work together to magnify the sample. A light source illuminates the specimen from below, and the magnified image is viewed through eyepieces or a camera. Compound microscopes can achieve high magnification and resolution, allowing for detailed examination of various specimens, including cells, tissues, and microorganisms. They are versatile and can be used for a wide range of applications, such as biological and medical research, quality control, and educational purposes.[23]

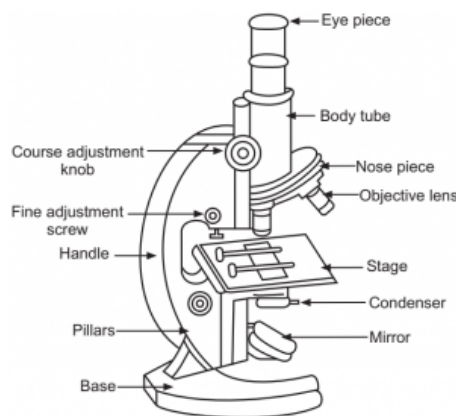


Figure 2.11: Compound Microscope

The compound microscope offers high magnification and resolution, allowing for detailed observation of various specimens. It is versatile and widely used in laboratories and educational settings for a range of applications. The availability of different contrast techniques further enhances its utility. However, compound microscopes have limitations in observing thick or opaque specimens and require careful sample

preparation, which can be time-consuming. Additionally, the limited depth of field may necessitate multiple focusing adjustments.

### 2.8.2 Dark-field Microscope

Dark-field microscopes are specialized optical microscopes that utilize a dark-field illumination technique. In dark-field microscopy, the specimen is illuminated with oblique or off-axis lighting, resulting in a brightly illuminated specimen against a dark background. This technique enhances contrast and allows for the visualization of translucent or unstained specimens that would otherwise be difficult to observe with bright-field microscopy.[6]

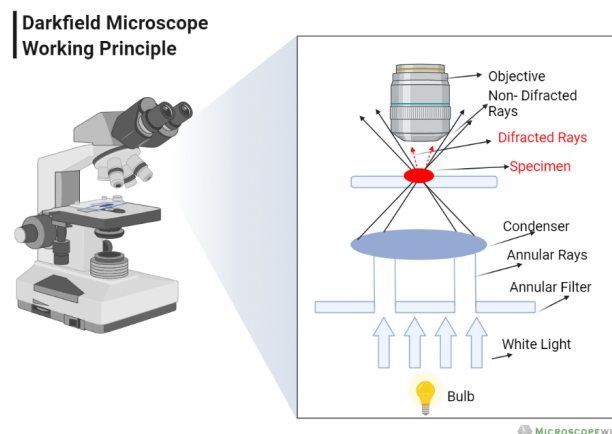


Figure 2.12: Dark-field Microscope

Darkfield microscopy provides excellent contrast, especially for translucent or unstained specimens. The dark background allows the specimen to stand out as bright objects, making it easier to visualize fine details that may be challenging to observe using other illumination techniques. However, it provides a limited field of view due to the use of oblique illumination. This can make it challenging to observe large specimens or obtain an overall view of the sample. It may be needed to acquire multiple images and stitch them together to create a complete picture.

### 2.8.3 Fluorescence Microscope

A fluorescent microscope is a specialized optical microscope. Because specimens are self-illuminated by internal light (fluorescence), they can be seen against a dark background. So, due to the enhanced contrast, it is easier for the human eye to see details not visible in a bright-field microscope. For the latter, there exist ways of being converted to a fluorescent microscope with the use of specific accessories.

When specimens, living or non-living, organic or inorganic, absorb and subsequently re-radiate light, the process is described as photoluminescence. If the emission of light persists for up to a few seconds after the excitation energy (light) is discontinued, the phenomenon is known as phosphorescence. Fluorescence, on the

other hand, describes light emission that continues only during the absorption of the excitation light. The time interval between absorption of excitation light and emission of re-radiated light in fluorescence is of extraordinarily short duration, usually less than a millionth of a second.

### Basic Concept Of A Fluorescence Microscope

For each fluorescent substance, there is a certain range of wavelengths that cause its stimulation. After the absorption of the excitation light, the substance emits light of longer wavelength which is finally captured. In order to illuminate the specimen with the wavelength band which excites it, an excitation filter is placed right after the light source. Similarly, in order to observe the re-emitted fluorescent light, an emission filter is placed right before the eyepiece or the camera sensor. These filters are termed barriers, since they block by absorption the unwanted light. Due to the epi-illumination fluorescence microscope design, there must be a specialized filter that will reflect the excitation light towards the specimen and will allow re-emitted light to pass through it. Such a filter is named a dichroic mirror or else a dichromatic beam-splitter. Reflection occurs via destructive and illuminance-reducing interference with alternating layers of high and low refractive indices, whereas transmission happens via constructive and reinforced interference with the layers. For this reason, such filters belong to the family of interference filters. Specifically, successive layers of dielectric materials, with thickness values ranging between one-quarter and one-half of the target wavelength consist the main body of these filters. In many of the current epi-illumination fluorescence microscopes, the excitation, emission filters along with the dichroic mirror are altogether incorporated into a single fluorescence filter cube. The whole design of the microscope is epigrammatically shown and is termed widefield epi-illumination fluorescence microscope.[9]

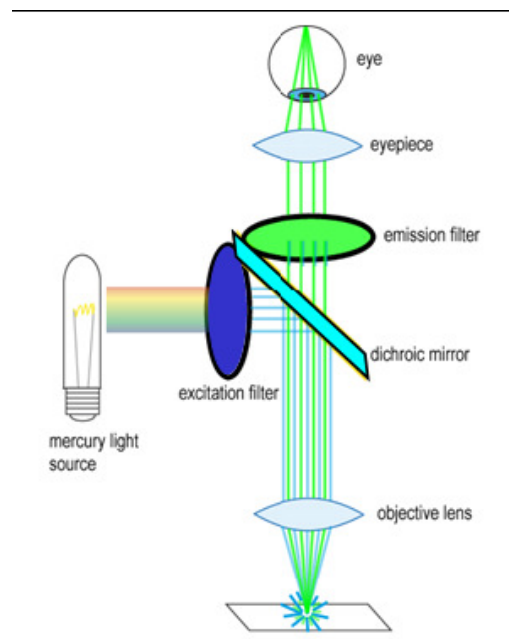


Figure 2.13: Fluorescence microscope

## Fluorescence

Fluorescence may be exhibited naturally by some materials or biological structures (auto-fluorescence/primary fluorescence) as well as exhibited by artificially added fluorescent markers (secondary fluorescence). Fluorescent molecules are also called fluorescent probes, fluorochromes or simply dyes. When they are conjugated to a larger macro-molecule, through absorption or covalent bonds, they are termed fluorophores. Thanks to this attachment, the distribution of the corresponding macro-molecule of a fluorophore can be observed.[5]

Due to heat transfer within fluorophores or interactions between fluorophores and other molecules, the re-radiated light has less energy than the absorbed light. So, the emitted light has a longer wavelength than that of the excitation light. This difference between the maxima of the excitation-emission bands is termed as Stokes shift. Typically, the emission band is a mirror of the excitation one. In general, fluorescence investigations are conducted in the range of ultraviolet to visible spectrum (250 700 nm).

The distribution of electrons and the overall molecular geometry can be determined by electronic states. Several different electronic states exist, depending on the total electron energy and the symmetry of various electron spin states (paired or unpaired spins - opposite or same spins in an orbital). The number of unpaired electron spins in an electronic state defines the state name in the numeration form of singlet, doublet, triplet etc. The ground state of most organic molecules is an electronic singlet, that is no electron spins are unpaired. The excitation process has no effect in changing the spin-pairing, hence the excited states are singlets too. Each electronic state is further subdivided into vibrational and rotational states associated with the atomic nuclei and the bonding orbitals. In fluorescence, various transitions are taking place through different paths among energy levels. Transitions involved in absorption and emission of light by a fluorochrome can be shown in a Jablonski energy diagram.

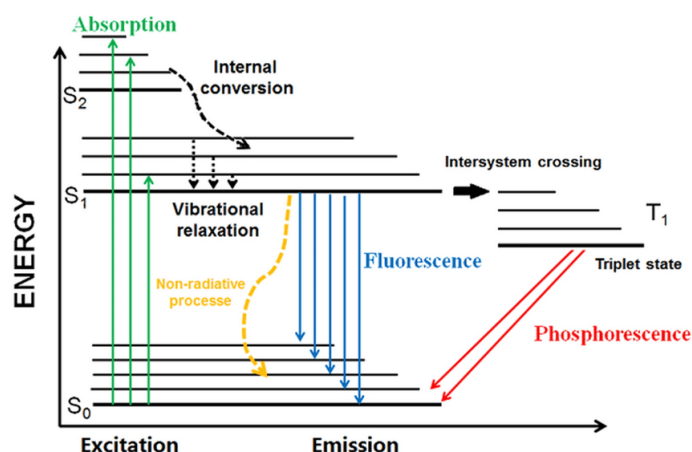


Figure 2.14: Jablonski energy diagram.



Absorption leads the molecule to an excited singlet state depending on the energy of the incident photon which has to be at least equal to the corresponding energy of an excited state. This takes place "instantly", in approximately a femtosecond. Afterwards, several processes will occur with varying probabilities. The most likely will be non-radiative relaxation to the lowest vibrational energy level of the first excited state. This is the vibrational relaxation which can happen directly or through an internal conversion process. From this point forward, many paths can be taken. One is the desired fluorescence emission which is usually accompanied by transitions to higher vibrational energy levels of the ground state. Other roads are energy dissipation non-radiatively as heat or energy transfer due to molecules collision (i.e quenching) or transition to an intermediate triplet excited state at its lowest vibrational state. The latter transition is known as intersystem crossing, again with no radiation. From that point, excitement back to the higher-energy singlet state might occur, resulting in a delayed fluorescence after a radiative relaxation. The other possibility is relaxation to the ground state, even though transition from an excited triplet state to the ground singlet state is forbidden. It might happen again in the form of heat, but also with the emission of a photon, which is termed phosphorescence. Because of this forbidding law, phosphorescence is very rare, with a probability of several orders of magnitude lower than that of fluorescence.

Considering the aforementioned different transition paths, it is clear why fluorescent light is of longer wavelength than that of the absorption light, justifying the Stokes shift. Bearing energy of light is inversely proportional to the wavelength of it. Specifically, radiative relaxations are of high probability when they start from the lowest vibrational state of the first excited state and when they culminate in the highest vibrational state of the ground state. Hence, the emission energy leap is much smaller than the absorption energy leap. Contributing factors to this phenomenon are heat conversions, resonance energy transfer (RET) and quenching processes. In RET, energy is transferred with non-radiative long range dipole-dipole interactions between fluorophores. In quenching, energy might be transferred through collisions between fluorophores and other molecules (usually oxygen, halogens and amines) that are translated as coupling of electronic orbitals between the interacting molecules. Also, formation of non-fluorescent complexes can reduce the population of active excitable molecules, thus limiting the absorption of incident light.

Another unwanted phenomenon, regarding energy and intensity loss, is photobleaching. Fluorochromes can lose the ability of fluorescence permanently due to long exposition to high intensity light which cause chemical destruction to them.

In general, the decay of fluorescence intensity as a function of time in a uniform population of molecules excited with a brief pulse of light is described by an exponential function:

$$I(t) = I(0)e^{-t/\tau} \quad (2.28)$$

, where  $I(0)$ : initial fluorescence intensity,  $\tau$  : time in which fluorescence intensity falls in  $1/e \approx 37\%$  of  $I(0)$ .

The exact decaying degree depends on the particular fluorochrome. In addition to this, other parameters are also used to describe and compare different fluorochromes:

- **Extinction coefficient:** Ability of light to absorb light. Those chromophores that have a high extinction coefficient also have a high probability of fluorescence emission, in the expense though of their intrinsic lifetime (below).
- **Quantum yield:** Gauge of the efficiency of fluorescence emission. It describes the probability that an excited molecule will end up in emitting a photon, either through a direct or a delayed fluorescence process.
- **Fluorescence lifetime:** The characteristic time that a molecule remains in an excited state prior to returning to the ground state.

### 2.8.4 Confocal Microscope

Confocal microscopy offers several advantages over conventional widefield optical microscopy, including the ability to control depth of field, elimination or reduction of background information away from the focal plane (that leads to image degradation), and the capability to collect serial optical sections from thick specimens. The basic key to the confocal approach is the use of spatial filtering techniques to eliminate out-of-focus light or glare in specimens whose thickness exceeds the immediate plane of focus.[3]

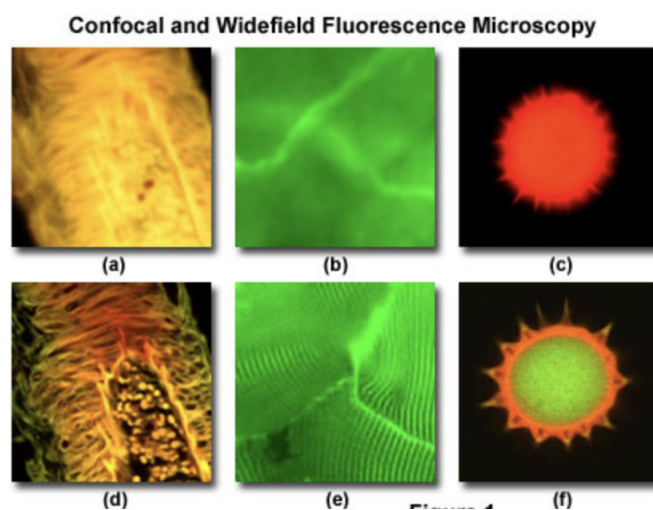


Figure 2.15: Widefield (a-c) and laser scanning confocal (d-f) fluorescence microscopy images

## Principles of Confocal Microscopy

The confocal principle in epi-fluorescence laser scanning microscopy is diagrammatically presented in Fig 2.16. Coherent light emitted by the laser system (excitation source) passes through a pinhole aperture that is situated in a conjugate plane (confocal) with a scanning point on the specimen and a second pinhole aperture positioned in front of the detector (a photomultiplier tube). As the laser is reflected by a dichromatic mirror and scanned across the specimen in a defined focal plane, secondary fluorescence emitted from points on the specimen (in the same focal plane) pass back through the dichromatic mirror and are focused as a confocal point at the detector pinhole aperture.

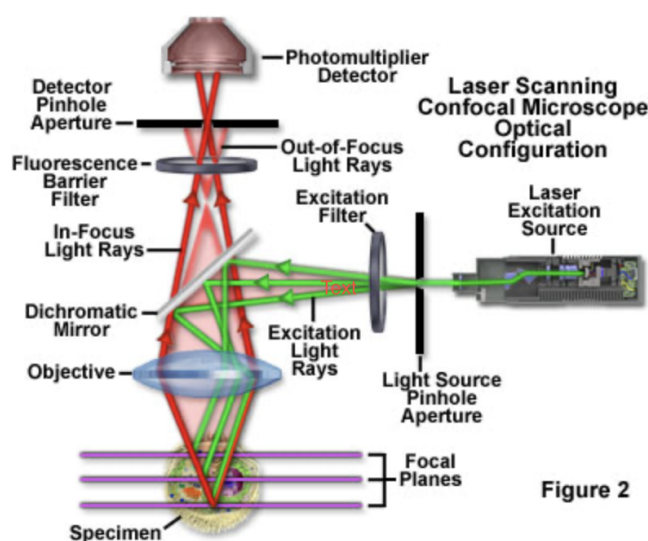


Figure 2

Figure 2.16: Diagram of epi-fluorescence laser scanning confocal microscope

In traditional widefield epi-fluorescence microscopy, the entire specimen is subjected to intense illumination from an incoherent mercury or xenon arc-discharge lamp, and the resulting image of secondary fluorescence emission can be viewed directly in the eyepieces or projected onto the surface of an electronic array detector or traditional film plane. In contrast to this simple concept, the mechanism of image formation in a confocal microscope is fundamentally different. The confocal fluorescence microscope consists of multiple laser excitation sources, a scan head with optical and electronic components, electronic detectors (usually photomultipliers), and a computer for acquisition, processing, analysis, and display of images.

The scan head is at the heart of the confocal system and is responsible for rasterizing the excitation scans, as well as collecting the photon signals from the specimen that are required to assemble the final image. A typical scan head contains inputs from the external laser sources, fluorescence filter sets and dichromatic mirrors, a galvanometer-based raster scanning mirror system, variable pinhole apertures for generating the confocal image, and photomultiplier tube detectors tuned for different fluorescence wavelengths. The general arrangement of scan head components is presented in Fig 2.17.

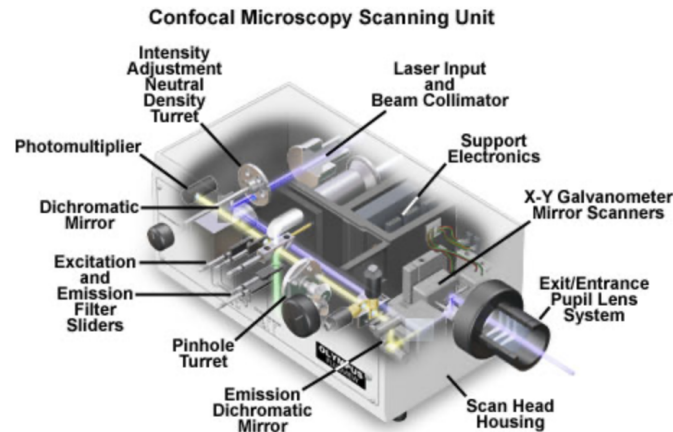


Figure 2.17: Confocal Microscopy Scanning Unit

### Image Generation of a Specimen

In laser scanning confocal microscopy, the image of an extended specimen is generated by scanning the focused beam across a defined area in a raster pattern controlled by two high-speed oscillating mirrors driven by galvanometer motors. One of the mirrors moves the beam from left to right along the x lateral axis, while the other translates the beam in the y direction. After each single scan along the x axis, the beam is rapidly transported back to the starting point and shifted along the y axis to begin a new scan in a process termed flyback. During the flyback operation, image information is not collected. In this manner, the area of interest on the specimen in a single focal plane is excited by laser illumination from the scanning unit.

As each scan line passes along the specimen in the lateral focal plane, fluorescence emission is collected by the objective and passed back through the confocal optical system. The speed of the scanning mirrors is very slow relative to the speed of light, so the secondary emission follows a light path along the optical axis that is identical to the original excitation beam. Return of fluorescence emission through the galvanometer mirror system is referred to as descanning. After leaving the scanning mirrors, the fluorescence emission passes directly through the dichromatic mirror and is focused at the detector pinhole aperture. Unlike the raster scanning pattern of excitation light passing over the specimen, fluorescence emission remains in a steady position at the pinhole aperture, but fluctuates with respect to intensity over time as the illumination spot traverses the specimen producing variations in excitation.

Fluorescence emission that is passed through the pinhole aperture is converted into an analog electrical signal having a continuously varying voltage (corresponding to intensity) by the photomultiplier. The analog signal is periodically sampled and converted into pixels by an analog-to-digital (A/D) converter housed in the scanning unit or the accompanying electronics cabinet. The image information is temporarily stored in an image frame buffer card in the computer and displayed on the monitor. It is important to note that the confocal image of a specimen is reconstructed, point

by point, from emission photon signals by the photomultiplier and accompanying electronics.

### **Advantages and Limitations**

Confocal microscopy is a powerful imaging technique that offers several advantages over traditional microscopy methods. However, it also has certain limitations that should be considered. Here are the advantages and limitations of confocal microscopy:[2]

One of the significant advantages of confocal microscopy is its ability to capture optical sections at different depths within a specimen, enabling the reconstruction of three-dimensional images. This feature is particularly useful for studying biological samples with complex structures, such as tissues and cells. Additionally, confocal microscopy can provide real-time imaging, allowing us to observe dynamic processes in living cells and organisms. The technique also offers excellent optical resolution, surpassing the limits of traditional light microscopy, and can resolve structures down to the subcellular level.

Another advantage of confocal microscopy is its ability to selectively image specific regions within a specimen. By using fluorescent dyes or labeled molecules, we can target specific cellular components or molecular structures and visualize them with high specificity. This capability is crucial in many biological applications, such as studying protein localization, cell signaling pathways, and intracellular dynamics.

Despite its numerous advantages, confocal microscopy does have some limitations. First, it requires specialized equipment, including a laser light source, scanning mechanism, and high-performance detectors, which can be costly. The complexity of the instrument and the need for skilled operators further add to the challenges of implementing confocal microscopy. Additionally, the technique relies on the use of fluorescent labels or dyes, which may introduce potential artifacts or interfere with the natural behavior of the specimen.

Another limitation is the limited penetration depth of confocal microscopy. While it excels at imaging thin sections and surface structures, it may struggle with imaging thick specimens or samples with high light scattering properties. The penetration depth is largely determined by the wavelength of the excitation light and the optical properties of the specimen. In some cases, additional techniques such as optical clearing or multiphoton microscopy may be employed to overcome these limitations.

Lastly, confocal microscopy typically requires a longer acquisition time compared to wide-field microscopy. The scanning mechanism used to acquire optical sections across the specimen can result in slower image acquisition rates. This limitation may pose challenges when studying fast biological processes or capturing rapid cellular events.

### 2.8.5 Why Widefield Confocal Microscopy over Confocal Microscopy

Widefield confocal microscopy, is a variation of confocal microscopy that utilizes a spinning disk with multiple pinholes to achieve optical sectioning and improved imaging speed compared to traditional confocal microscopy.[14]

#### Principle

Widefield confocal microscopy uses a spinning disk that contains an array of pinholes, usually arranged in a spiral pattern as shown in Fig. 2.18. The disk is placed in the illumination path of the microscope, in front of the objective lens. When the disk spins rapidly, each pinhole sequentially allows light to pass through, creating a pattern of illuminating points on the sample. The light emitted from the sample is then collected by the objective lens, and only the light from the illuminated points corresponding to the open pinholes is detected.

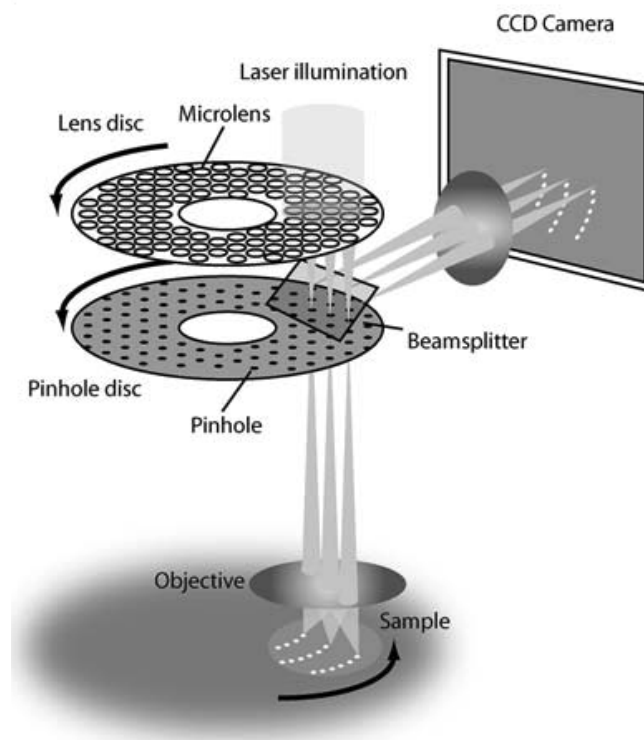


Figure 2.18: Wide-field Confocal Microscopy Diagram

### Advantages over Confocal Microscopy

**Faster Imaging:** Widefield confocal microscopy enables faster imaging compared to traditional confocal microscopy. Since multiple pinholes are used simultaneously, it allows for the capture of multiple image points in parallel. This results in increased imaging speed, making widefield confocal microscopy suitable for dynamic processes and live cell imaging.

**Reduced Phototoxicity:** Widefield confocal microscopy typically uses lower laser power compared to traditional confocal microscopy, reducing the potential for phototoxicity. The shorter exposure times due to faster imaging also contribute to reduced photodamage, making it more suitable for long-term live cell imaging experiments.

**Improved Signal-to-Noise Ratio:** By using multiple pinholes, widefield confocal microscopy rejects out-of-focus light and reduces background noise, resulting in improved signal-to-noise ratio. This leads to enhanced image quality, better contrast, and increased resolution compared to conventional widefield microscopy techniques.

**Cost-Effectiveness:** Widefield confocal microscopy is generally more cost-effective compared to traditional laser scanning confocal microscopy. The spinning disk technology is less complex and expensive than the galvanometer-based scanning systems used in laser scanning confocal microscopes. This makes widefield confocal microscopy a viable option for researchers with budget constraints.

In summary, widefield confocal microscopy offers faster imaging, reduced phototoxicity, improved signal-to-noise ratio, and cost-effectiveness compared to traditional confocal microscopy. In the following chapters will present the algorithm for reconstructing an image based on widefield confocal microscopy

# Chapter 3

## Multi-Focus Image Fusion

Human beings possess wonderful sense to appreciate visuals. Eye plays a key role in supporting various human activities. An image capture of a visual scene always conveys much more information than any other description adhered to it.

Human beings have five sensing capabilities or systems as shown in Fig. 3.1. They are eyes, ears, nose, tongue, and skin. These sensors are able to acquire independent information. Eyes can visualize a scene. Ears can sense the data by listening to sounds. Nose can smell the odor of an object. Tongue can sense the taste of an object. Skin can sense the texture and size of the object. As shown in Fig. 3.2, all these five sensing systems act as sensors. Human brain collects data from these individual sensors and fuses or combines information for compact representation or better description about a scenario. This compact data is useful for decision-making and task execution.[8]

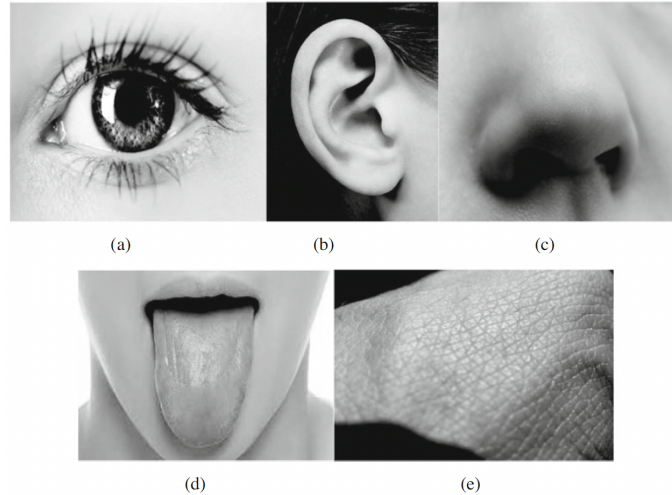


Figure 3.1: Human sensors (a) eyes, (b) ears, (c) nose, (d) tongue, (e) skin

Data fusion is a process of combining information from several sources for optimal or compact representation of a huge data supporting better description and decision-making. Human brain is the best example of a data fusion system.



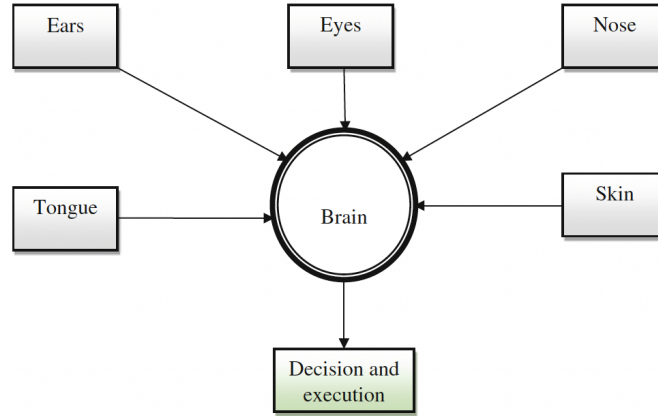


Figure 3.2: Human sensor fusion system for decision-making

Even when we take one sensor, for example, eye, it can derive many useful details of a scene by looking at the scenario more than once. The brain can integrate visuals and give details hidden in a single view. Multiple views will always improve the decisions.

Image fusion has a long history and development phase. Research on image fusion has been carried out by adopting various mathematical tools and techniques. It is very important to introduce concepts of image fusion in a systematic way. To this end, in this chapter, fundamental concepts, categories, types, and applications of image fusion are covered.

### 3.1 Development of Image Fusion

During 1990s, major attention in image processing was drawn by pyramid decomposition based on techniques. In 1980, Burt and Julesz were the first authors who proposed an image fusion algorithm based on pyramids for binocular images. Then, they proposed another fusion algorithm for multi-modal images. Later, Burt and Kolczynski developed an improved approach to fuse images based on pyramid decomposition. Some other well-known image fusion work from Alexander Toet for fusing visible and infrared images depends on various pyramids and wavelet transforms for surveillance applications. Later, neural networks were successfully adopted for visible and infrared image fusion by Ajjimarangsee and Huntsberger. Furthermore numerous transform-based methods such as stationary wavelet, complex wavelet, curvelet, contourlet, and non-sub-sampled contourlet were utilized for fusion. Optimization-based methods were also adopted for image fusion. Recently, many researchers implemented image fusion methods based on filtering-based decompositions. Cross bilateral filter, guided image filter, rolling guided image filter, anisotropic diffusion, and weighted least square filter based techniques are a few notables among them.[12]

## 3.2 Fundamentals of Image Fusion

An image is a two-dimensional quantity. It can be viewed as the combination of illumination and reflectance. Illumination stands for the amount of light from the source falling on the object, and reflectance corresponds to the amount of light that is reflected from the same object.

A sensor is a device which converts incoming energy into an electrical signal as shown in Fig. 3.3a. In case of imaging sensors, reflected energy will be converted into a corresponding electrical signal. As displayed in Fig. 3.3b, the sensor array gives a large number of signals. The sampling in the spatial domain is performed by the sensor array. These signals are quantized to obtain a digital image representation. This entire process is termed as digitization which is shown in Fig. 3.3c. Thus, the visual information present in a scene can be captured as a digital image  $f(x, y)$  using a sensor array as shown in Fig. 3.3b.

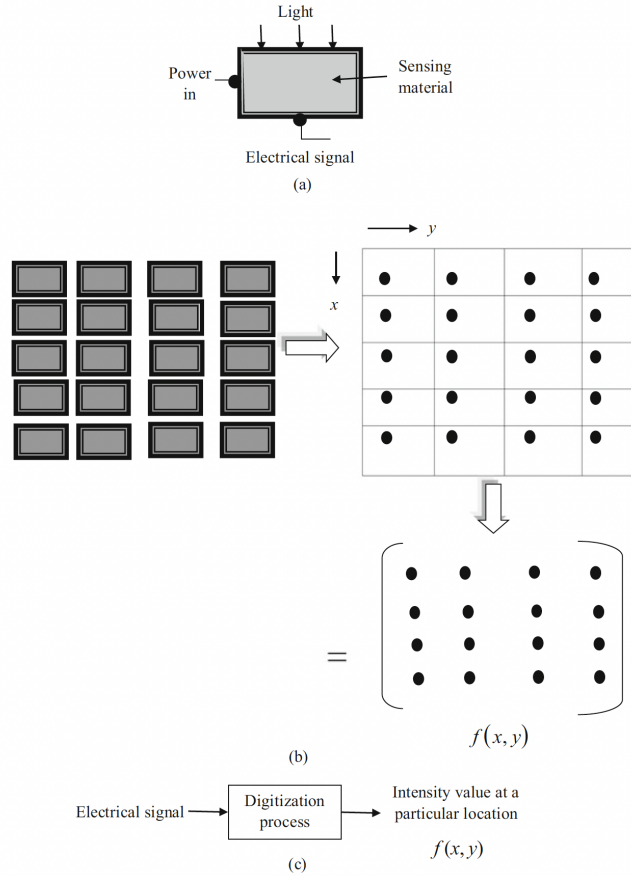


Figure 3.3: Single sensor imaging. (a) Single sensor, (b) Single sensor array and its corresponding image represented as a matrix, (c) Digitization process

### 3.2.1 WHY Combine Information of Multiple Images

Single-sensor image capture may not always provide complete information about a target scene. Sometimes we need two or more images of the same scene for better

visual understanding. These images may be captured by using a single sensor or by using multiple sensors of different modalities, depending on the application. These image captures provide complementary or visually different information. A human observer cannot reliably combine and observe a composite image from these multiple image captures. Useful or complementary information of these images should be integrated into a single image to provide a more accurate description of the scene, than any one of the individual source images.

### 3.2.2 Definition of Image Fusion

Image fusion has numerous definitions. Among them, some well-known definitions are presented here below.

- Produce a single image from a set of input images. The fused image should have complete information which is more useful for human or machine perception.
- Generation of a result which describes the scene better than any images captured in a single shot.
- Image fusion is a process of combining images, obtained by sensors of different wavelengths simultaneously viewing of the same scene, to form a composite image. The composite image is formed to improve image content and to make it easier for the user to detect, recognize, and identify targets and increase his situational awareness.
- Image fusion is the process of merging or combining or integrating useful or complementary information of several source images such that the resultant image provides a more accurate description about the scene than any one of the individual source images.

### 3.2.3 Objective of Image Fusion

Image fusion process can also be explained using set theory representation as transferring of information of two sets. The Venn diagram representation of this process is shown in Fig.3.4. Each set stands for the information contribution corresponding to that particular image. Sets  $A$  and  $B$  represent the information contribution by the two source images  $A$  and  $B$ , respectively. Set  $F$  corresponds to the information contribution by the fused image  $F$ .

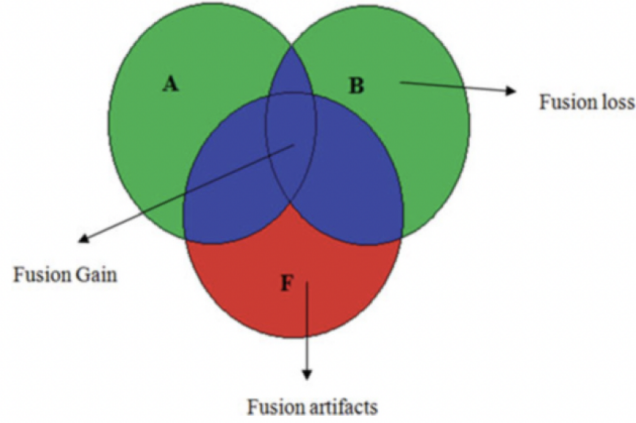


Figure 3.4: Graphical illustration of image fusion process

Ideally, this fused image should contain all the information from source images. However, it is not possible. In practice, not all the source image information is transferred into the fused image. Only required and necessary information will be transferred. In Fig. 3.4, blue portion represents the information transferred from source images in the fused image which is simply referred to as “fusion gain or fusion score.” Green portion indicates the information lost during the fusion process which is termed as “fusion loss.” This information in source images is not present in the fused image, and red portion corresponds to unnecessary information (fusion artifacts) introduced in the fused image.

The main objective of an image fusion algorithm is to generate a visually good fused image with less computational time, by maximizing the fusion gain and minimizing the fusion loss and fusion artifacts.

### 3.3 Categories of Fusion Algorithms

Image fusion algorithms are broadly divided into three categories: pixel, feature, and decision levels.

Pixel-level fusion is performed on each input image pixel by pixel. Pixel-level fusion methods can be implemented in the spatial domain or in a transform domain. In the spatial domain, these methods can be implemented pixel by pixel. However, transform domain methods work by a coefficient. For a small change in the frequency coefficient, the whole resultant image will be effected. To obtain a better fused image without artifacts, best transform technique with a suitable fusion rule should be chosen.

At feature level, fusion is executed on the extracted features of source images. Feature-level fusion schemes usually consider segmented regions based on different properties such as entropy-level, variance-level, and activity-level measurements. These algorithms give a robust performance in the presence of noise.

At decision level, fusion is performed on probabilistic decision information of local decision makers. These decision makers are in turn derived from the extracted features. These fusion techniques integrate information from source images based on decision maps derived from the features.

### 3.4 Steps of Image Fusion

An image fusion system consists of eight fundamental steps as shown and explained below. These are: image acquisition, pre-processing, image registration, image fusion, post-processing, fusion performance evaluation, storage, and display.

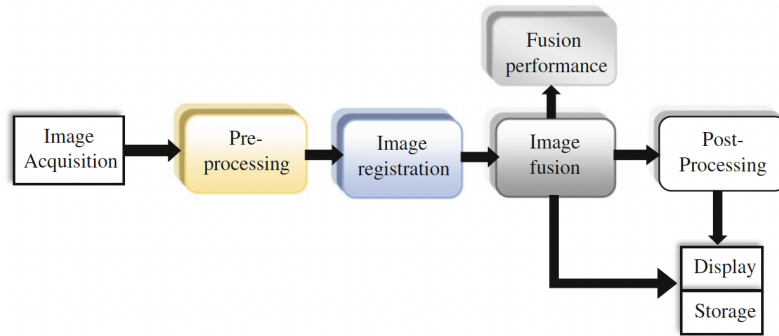


Figure 3.5: Steps in image fusion system

- During image acquisition stage, visually different or complementary images will be captured using a single sensor or multiple sensors of different modalities.
- In the pre-processing step, noise or artifacts introduced in the source images during image acquisition process are removed or reduced.
- Image registration is the process of aligning or arranging more than one images of the same scene according to a coordinate system. In this process, one of the source images will be taken as a reference image. It is also termed as the fixed image. Then geometric transformation will be applied on remaining source images to align them with the reference image.
- Fusion process can be performed at three levels: pixel, feature, and decision level. Pixel-level fusion schemes are preferable for fusion compared to another level of approaches because of their effectiveness and ease of implementation.
- During the fusion process, some required information of source images may be lost, and visually unnecessary information or artifacts may be introduced into the fused image. Hence, fusion algorithms need to be assessed and evaluated for better performance. This performance analysis can be carried out by evaluating them qualitatively by visual inspection and quantitatively using fusion metrics.
- In post-processing, fused images are further processed depending on the application. This processing may involve segmentation, classification, and feature extraction.

- Source images, fused images, post-processing results, and their corresponding data will be stored with help of storage devices such as hard disks and flash memories.
- Finally, fused images and post-processing results such as segmented images, features, and classification results can be displayed using devices like LCD and LED monitors.

## 3.5 Applications

Image fusion finds applications in various fields such as digital photography, medical imaging, remote sensing, concealed weapon detection, military, night vision, autonomous vehicles, visual inspection in industrial plants, ambient intelligence, and person re-identification.

## 3.6 Pixel-Level Image Fusion

### 3.6.1 Introduction

If source images are combined by performing pixel-wise operations, then it is referred to as pixel-level image fusion. The main objective of any pixel-level image fusion algorithm is to generate a visually good fused image with less computational time along with the following properties:[11]

- It has to transfer complementary or useful information of source images into the composite image.
- It should not lose source image information during the fusion process.
- It should not introduce artifacts into the fused image.

In the view of image fusion objective various fusion algorithms have been proposed. As shown in Fig.3.6, image fusion is broadly classified as single-scale and multi-scale fusion methods.

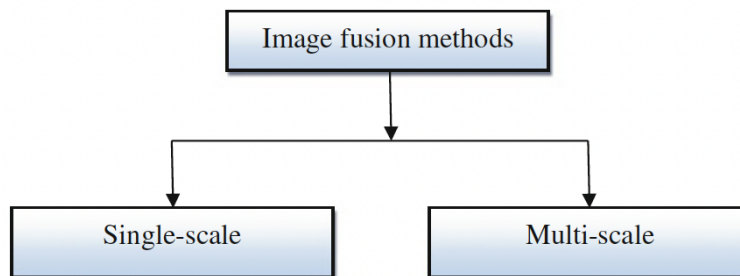


Figure 3.6: classification of image fusion methods

### 3.6.2 Single-Scale Image Fusion

The fusion is performed on present scale source images without further decomposition. These are also referred to spatial domain techniques. Simple operators such as average, weighted average, minimum, maximum, and morphological operators are used for fusion.

In the simple average method, the resultant fused image  $F(i, j)$  is obtained by calculating the pixel-wise average operation on input images  $A(i, j)$  and  $B(i, j)$ , as follows:

$$F(i, j) = (A(i, j) + B(i, j))/2 \quad (3.1)$$

In the weighted average method, the resultant fused image  $F(i, j)$  is obtained by computing the pixel-wise weighted average operation on input images, as follows:

$$F(i, j) = \sum_{i=0}^m \sum_{j=0}^n wA(i, j) + (1 - w)B(i, j) \quad (3.2)$$

,where  $w$  is the weight factor.

In the selective maximum method, the resultant fused image  $F(i, j)$  is obtained by applying the pixel-wise maximum operation on input images.

$$F(i, j) = \max(A(i, j), B(i, j)) \quad (3.3)$$

In the selective minimum method, the resultant fused image is obtained by calculating the pixel-wise minimum operation on input images.

$$F(i, j) = \min(A(i, j), B(i, j)) \quad (3.4)$$

These methods are easy to implement. But, they may produce brightness or color distortions into the fused image.

Principal component analysis (PCA), independent component analysis (ICA), and intensity-hue-saturation (IHS) are some of the well-known methods in the spatial domain category. These methods may suffer from the spectral distortion and give desirable results only for a few fusion datasets.

Focus measure-based approaches are famous in this class. Here, source images are divided into blocks, and various focus measures are employed to select the best among image blocks. Variance, energy of image gradient, energy of Laplacian (EOL), sum-modified Laplacian (SML), and spatial frequency (SF) are various focus measures successfully used for the fusion. It is observed that SML gives superior performance compared to other focus measures. However, it is computationally expensive.

To overcome these problems, multiple iterations to find an optimal solution are made (fused image). These optimization methods may over smooth the fused image because of multiple iterations.

### 3.6.3 Multi-Scale Image Fusion

Multi-scale fusion methods are developed to overcome drawbacks of the single-scale fusion. Multi-scale decomposition (MSD) extracts the salient information (visually significant information) of source images for the fusion purpose. MSD methods perform better than single-scale fusion methods due to following facts:

- Human visual system (HVS) is sensitive to changes in the saliency information such as edges and lines. These features can be well extracted and fused with help of the MSD.
- It offers better spatial and frequency resolution.

In the MSD, source images are decomposed into approximation and detail coefficients/layers at several scales. For example, as displayed in Fig.3.7, a given source image  $I$  is decomposed into approximation coefficient  $C_A^1$  and detail coefficient  $C_D^1$  at level 1 ( $L_1$ ).  $C_A^1$  at  $L_1$  is further decomposed into  $C_A^2$  and  $C_D^2$  at level 2 ( $L_2$ ). In general, approximation and detail coefficients at  $n$ th decomposition level can be represented as  $C_A^n$  and  $C_D^n$ , respectively. This MSD process is called as analysis, and it is represented as  $\psi$ . Its inverse process is termed as the synthesis which is indicated as  $\psi - 1$ .



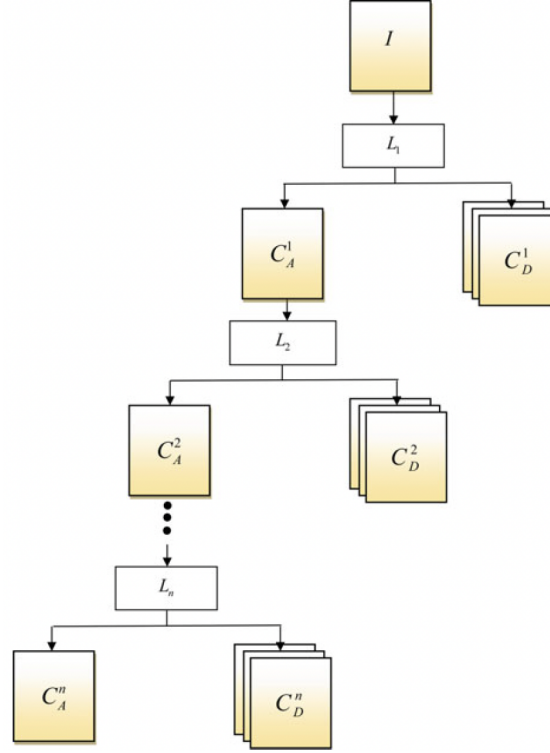


Figure 3.7: The multi-scale image decomposition (MSD)

In the multi-scale image fusion, source images  $I_1$  and  $I_2$  are decomposed into approximation and detail coefficients as shown in Fig. 3.8.

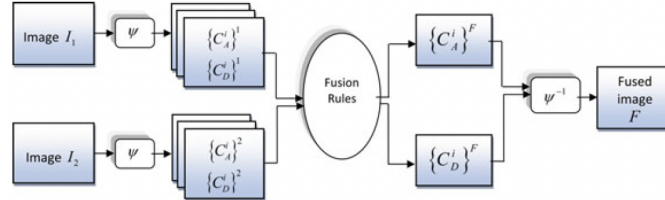


Figure 3.8: Block diagram of the multi-scale image fusion

Approximation and detail coefficients of the source image  $I_1$  at  $i$ th level are represented as  $\{C_A^i\}^1$  and  $\{C_D^i\}^1$ , respectively. Similarly for  $I_2$ , they can be represented as  $\{C_A^i\}^2$  and  $\{C_D^i\}^2$ . This decomposition process is called as analysis. Fusion is performed on these decomposed coefficients to obtain final approximation  $\{C_A^i\}^F$  coefficient and detail coefficient  $\{C_D^i\}^F$  by employing various fusion rules.

All these fused coefficients at different levels will be combined to obtain the fused image  $F$ . This process is termed as synthesis. Multi-scale image fusion methods are further classified as:

- Pyramid-based fusion.
- Wavelet transform-based fusion.
- Filtering-based fusion.

### Pyramid-based Fusion

The basic idea of these methods is explained as follows: First, decompose source images into successive sub-images using some operations such as blurring and down sampling. Next, apply fusion rules on these decomposed sub-images. Finally, reconstruct the fused image from these fused sub-images. The general block diagram of the pyramid based fusion is depicted in Fig. 3.9. As shown in the figure, source images  $I_1$  and  $I_2$  are blurred using linear filter and down sampled by 2 along rows and columns. This process is given as:

$$\begin{aligned} \{C_P^{i+1}\}^1 &= \left[ \{C_P^i(x, y)\}^1 * w(x, y) \right], \{C_P^{i+1}\}^2 \\ &= \left[ \{C_P^i(x, y)\}^2 * w(x, y) \right], i = 0, 1, 2, \dots N \end{aligned} \quad (3.5)$$

,where  $\{C_P^{i+1}\}^1$  represents sub-images obtained from pyramid decomposition of the source image  $I_1$  at  $(i+1)$  - th level which depends on its previous level sub-image  $\{C_P^i(x, y)\}^1$ .  $\{C_P^0\}^1$  represents the input image  $I_1$ . The convolution operation is represented by  $*$   $w$  is a linear filter, and  $N$  represents the number of levels. The same is true for the source image  $I_2$  as well. Various fusion rules can be employed on these decomposed sub-images to obtain fused sub-images  $\{C_P^i\}^F$  at various levels  $L_i$ . Then pyramid is reconstructed back from these fused sub-images to get the fused image  $F$ .

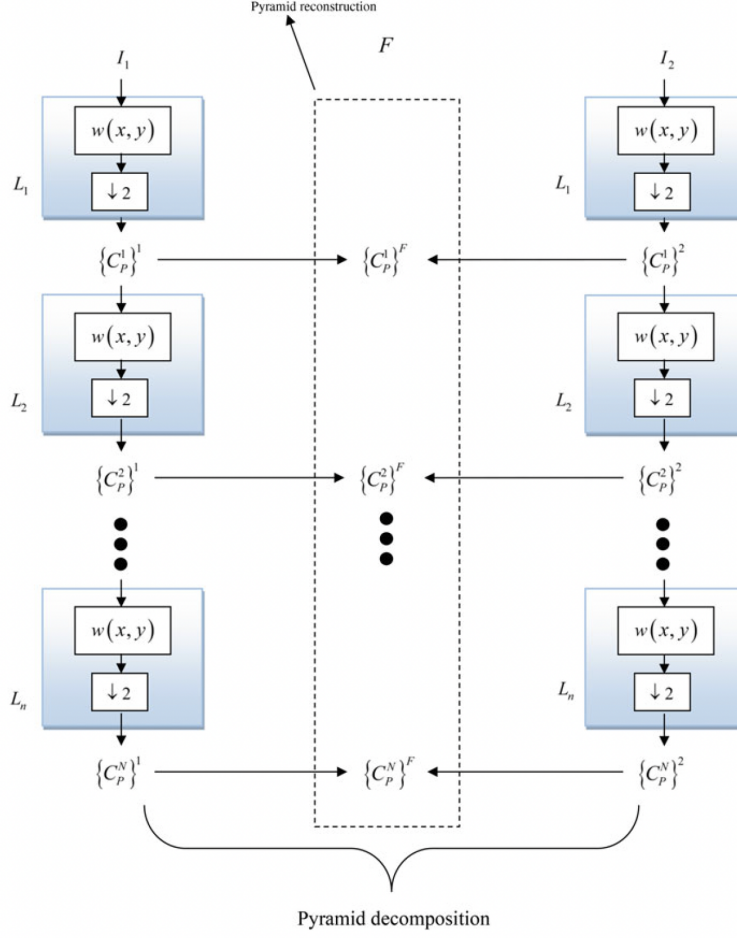


Figure 3.9: Block diagram of pyramid-based image fusion

### Wavelet Transform Fusion

One of the most known in this category is discrete wavelet transform (DWT) decompositions. In DWT, each source image is decomposed into wavelet coefficients at various levels. By using different fusion rules, these wavelet coefficients are fused. Finally, inverse wavelet transform is applied on these final fused wavelet coefficients to obtain the desired fused image. A general block diagram of wavelet-based fusion is shown in Fig. 3.10.

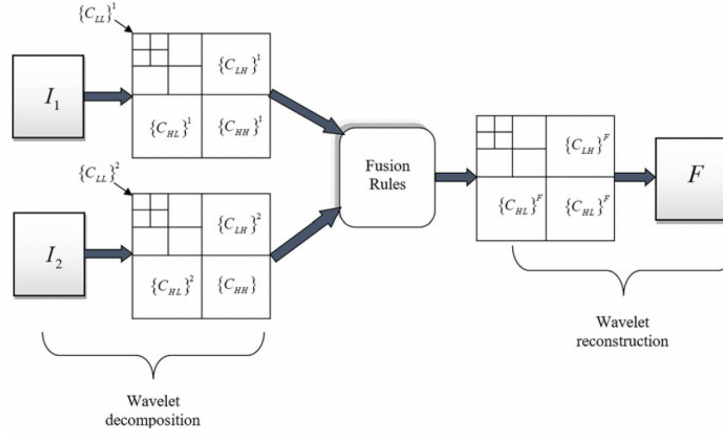


Figure 3.10: Block diagram of wavelet-based image fusion

As shown in the picture above source image  $I_1$  is decomposed into a set of four sub-images in various directions.  $\{C_{LL}\}^1$  is an approximation coefficient of  $I_1$  which represents the low frequency content.  $\{C_{LH}\}^1$ ,  $\{C_{HL}\}^1$ , and  $\{C_{HH}\}^1$  indicate detail coefficients in horizontal, vertical, and diagonal directions, respectively. Approximation coefficient  $\{C_{LL}\}^1$  at level 1 will be further decomposed into approximation coefficient and detail coefficients in horizontal, vertical and diagonal directions at level 2, and so on. The same thing is true for source image  $I_2$  as well. By employing various fusion rules, wavelet coefficients of source images at various levels will be combined. From these fused wavelet coefficients, final fused image  $F$  is generated by employing inverse wavelet transform on them.

### Filtering-Based Fusion

These techniques use image filtering techniques such as edge-preserving decomposition filters (EPD) and non-EPD to perform decomposition process. In this category of fusion, each source image is decomposed into approximation/base layers containing large-scale variations in intensity and detail layers containing smallscale variations in intensity. Separate fusion rules are employed to combine these decomposed base and detail layers. Finally, these final base and detail layers are combined to generate the fused image. The block diagram of these methods is shown in Fig. 3.11.

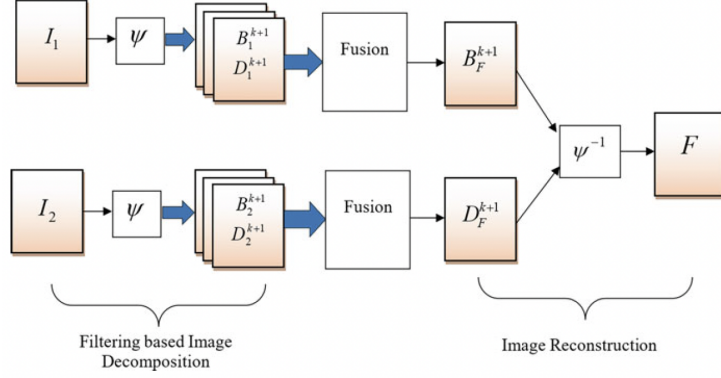


Figure 3.11: Block diagram of filtering-based image fusion

Here,  $\psi$  and  $\psi^{-1}$  represent the filtering-based image decomposition and reconstruction process, respectively. First, two source images  $\{I_n\}_{n=1}^2$  are decomposed into base layers containing large-scale variations in intensity and detail layers containing small scale variations in intensity as

$$\{B_n^{k+1}\}_{n=1}^2 = \{B_n^k\}_{n=1}^2 * w, k = 0, 1, \dots, k \quad (3.6)$$

,where  $B_n^{k+1}$  is the base layer of  $n$ -th source image at  $k + 1$  level which depends on its previous level base layer  $B_n^k$ .  $B_n^0$  represents the  $n$ -th input image  $I_n$ . The convolution operation is represented by  $*$  and  $w$  is an image filter.

## 3.7 Feature-Level Image Fusion

### 3.7.1 Introduction

Feature-level fusion is an intermediate-level fusion process that uses the feature information extracted from the original information of each source for comprehensive analysis and processing. The idea of feature-level fusion is to first extract useful features from the original multi-sensor imaging and then merge these features into new feature vectors for further processing. Typical image features include edges, corners, lines, and the like. Compared with pixel-level fusion, feature-level fusion has more information loss and less computation. Feature-level fusion is a fusion of information at the intermediate level. It not only retains a sufficient number of important information but also compresses the information, which is beneficial to real-time processing.[16]

### 3.7.2 Feature-Level Fusion Algorithm

#### Serial Strategy

It is assumed that  $A$  and  $B$  are two feature spaces defined in the pattern sample space  $\Omega$ . For any sample  $\Gamma \in \Omega$ , the corresponding feature representation vector is  $\alpha \in A$  and  $\beta \in B$ . The serial fusion strategy chains the two feature representation vectors into a large vector  $\gamma$ .

$$\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (3.7)$$

Obviously, if  $\alpha$  is the  $n$  dimension and  $\beta$  is the  $m$  dimension, the synthesized vector  $\gamma$  is the  $(n + m)$  dimension. In this way, all serially synthesized vectors form a feature space of a  $(n + m)$  dimension.

#### Parallel Strategy

It is assumed that  $A$  and  $B$  are two feature spaces defined in the pattern sample space  $\Omega$ . For any sample  $\Gamma \in \Omega$ , the corresponding feature representation vector is  $\alpha \in A$  and  $\beta \in B$ . The parallel fusion strategy represents these two features into a complex vector  $\gamma$ .

$$\gamma = \alpha + i\beta \quad (3.8)$$

Among them,  $i$  is an imaginary unit. It should be noted that if the dimensions of  $\alpha$  and  $\beta$  are inconsistent, the dimension is consistent by zeroing the low dimensional vectors. For example,  $\alpha = (a_1, a_2, a_3)^T$  and  $\beta = (\beta_1, \beta_2)^T$  first convert  $\beta$  to  $(\beta_1, \beta_2, 0)^T$  and then synthesize vector  $\gamma = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2, \alpha_3 + i0)^T$ .

Compared with the serial fusion strategy, the parallel fusion strategy reduces the dimension of the fused vector. More importantly, it introduces the concept of unitary space, which transforms the fusion problem of two real vector spaces into a mathematical problem of a complex vector space.

### 3.7.3 Genetic Algorithm

The basic idea of genetic algorithms is based on this. The algorithm interprets the possible solutions of the problem into 0 and 1 code strings called chromosomes. Given a set of initial chromosomes, the genetic algorithm manipulates them using genetic operators to generate a new generation. The new generation of chromosomes may contain better solutions than previous generations. Every chromosome needs to be evaluated by its fitness function. The goal of the genetic algorithm is to find the most suitable chromosome.

The main calculation process of the genetic algorithm is as follows:

- Begin.
- Determine fitness function and coding rules.
- Identify genetic algebra, population size, crossover rate, and mutation rate.
- Randomly generate the initial population and calculate its fitness.
- Repeat
- Selecting two (or more) chromosomes in the population to perform crossover at the crossover rate.
- Choosing chromosomes in the population to mutate at the mutation rate.
- Calculate the fitness of each chromosome.
- Select offspring groups (properly breed; those who are disqualified).
- Until the specified algebra is reached or a satisfactory result is obtained.
- End.

## 3.8 Decision-Level Image Fusion

Decision-level fusion is a high-level information fusion. It can be performed by following the four steps. They are: First is the multi-sensor imaging processing. Second is the decision generation. Third is the convergence in the fusion center. Final step is the concluding fusion process. In the information processing architecture, the peak of the fusion level is the decision-level fusion, also known as high-level fusion. In general, decision-level fusion is more perfect than the others and can better overcome the shortcomings of each sensor. For other fusion levels, the failure of one sensor means the failure of the entire system. Compared with pixel-level fusion and feature-level fusion, the decision-level fusion has the best real-time performance, but the main drawback of this method is the information loss.[4]

### 3.8.1 Fusion Algorithm Based on Voting Method

Voting methods are the simplest techniques for multisensory comprehensive labeling. Each sensor provides an input statement that observes the identity of the entity, which is then searched by voting to find a statement that “more than half the sensors agree” and announce the result of the vote: a union description. Of course, sometimes it may be necessary to introduce weighting methods, thresholding techniques, and other methods of determination, thus increasing the complexity of the voting method to a certain extent. Voting is very useful when accurate prior probabilities are not available, especially for real-time blending.

### 3.8.2 Fusion Algorithm Based on Bayes Inference

The Bayesian theory was published by Thomas Bayes in 1763. Its basic principle is that given a priori likelihood estimation of a hypothesis, the Bayesian method can update the hypothetical likelihood function with the advent of new evidence

Bayesian inference is an important method for dealing with stochastic patterns. Assume that  $X$  is a group of information sources,  $X = \{x_1, x_2, \dots, x_R\}$ , and the target is judged to  $\omega_j$  according to the maximum posteriori probability (MAP).

$$Z \rightarrow \omega_j$$

If

$$P(\omega_j|x_1, \dots, x_R) = \max_k P(\omega_k|x_1, \dots, x_R) \quad (3.9)$$

According to Bayesian theory, the maximum posterior probability (MAP) can be expressed as:

$$P(\omega_k|x_1, \dots, x_R) = \frac{p(x_1, \dots, x_R|\omega_k)P(\omega_k)}{p(x_1, \dots, x_R)} \quad (3.10)$$

$p(x_1, \dots, x_R)$  is the joint probability density, which can be expressed as:

$$p(x_1, \dots, x_R) = \sum_{j=1}^m p(x_1, \dots, x_R|\omega_j)P(\omega_j) \quad (3.11)$$

Assuming that information sources are statistically independent, we get:

$$p(x_1, \dots, x_R|\omega_k) = \prod_{i=1}^R p(x_i|\omega_k) \quad (3.12)$$

By the equations above we get:

$$P(\omega_k|x_1, \dots, x_R) = \frac{P(\omega_k) \prod_{i=1}^R p(x_i|\omega_k)}{\sum_{j=1}^m P(\omega_j) \prod_{i=1}^R p(x_i|\omega_j)} \quad (3.13)$$

Generally, the Bayesian approach is actually a special case of D-S evidence theory, so all data fusion problems using the Bayesian approach can be replaced by D-S evidence theory.



### 3.8.3 Fusion Algorithm Based on Min-Max Rule

Bayesian inference and summation rules construct the basic framework of sensor fusion. According to the following formula, we can deduce other fusion strategies:

$$\prod_{i=1}^R P(\omega_k|x_i) \leq \min_{i=1}^R P(\omega_k|x_i) \leq \frac{1}{R} \sum_{i=1}^R P(\omega_k|x_i) \leq \max_{i=1}^R P(\omega_k|x_i) \quad (3.14)$$

#### Maximum Rules

According to the summing rule and Eq 3.14, we can get the maximum rule:

$$Z \rightarrow \omega_j$$

If

$$(1 - R)P(\omega_j) + R \max_{i=1}^R P(\omega_j|x_i) = \max_k [(1 - R)P(\omega_k) + R \max_{i=1}^R P(\omega_k|x_i)] \quad (3.15)$$

Assuming the same probability of occurrence in any one class, we get

$$Z \rightarrow \omega_j$$

If

$$\max_{i=1}^R P(\omega_j|x_i) = \max_k \max_{i=1}^R P(\omega_k|x_i) \quad (3.16)$$

#### Minimum Rules

According to Bayesian inference and Eq 3.14, we can get the maximum rule:

$$Z \rightarrow \omega_j$$

If

$$P^{-(R-1)}(\omega_j) \min_{i=1}^R P(\omega_j|x_i) = \max_k P^{-(R-1)}(\omega_k) \min_{i=1}^R P(\omega_k|x_i) \quad (3.17)$$

Assuming the same probability of occurrence in any one class, we get

$$Z \rightarrow \omega_j$$

If

$$\min_{i=1}^R P(\omega_j|x_i) = \max_k \min_{i=1}^R P(\omega_k|x_i) \quad (3.18)$$

# Chapter 4

## Wide-field Confocal Image Reconstruction

### 4.1 Introduction

In the past decades, different genres of multi-focus image fusion algorithms have been proposed. Most of these algorithms are done either in the spatial domain or transform domain.

In this diploma thesis a quadtree-based multi-focus image fusion algorithm is going to be proposed. The procedure in order to conclude to the fused image is explained below.

We make use of an effective quadtree decomposition strategy to decompose the source images into blocks with optimal sizes in a quadtree structure. In this structure the focused blocks are detected by using a new weighted focus-measure, named as the sum of the weighted modified Laplacian. Naturally, the focused blocks are merged to form the focused regions. Then, the focused regions are reconstructed to verify the consistency by using two sequential filters.

This algorithm is simple yet effective, because of the quadtree decomposition strategy and the new weighted focus-measure. To do the comparison, the algorithm is compared with several existing fusion algorithms, in both the qualitative and quantitative ways. The experimental results show that the proposed algorithm yields good results.[22]

### 4.2 Proposed Algorithm

The main goal of this algorithm is to generate one fully focused image from the given rgb multi-focus images. The process of the proposed algorithm is: (1) read the rgb multi-focus images; (2) split each image in channels, red-green-blue, in order to process them; (3) decompose the source images into blocks with optimal sizes, according to an effective quadtree decomposition strategy; (4) find out the focused blocks by measuring the focus of each block-pair in the tree structure; (5) reconstruct the focused regions by using two sequential filters.[1]

### 4.2.1 Read-Split Images

In the first part of the algorithm we read the given multi-focus images, save them into an array and then split them in red, green and blue channel. Then the fusion algorithm we be applied in each channel.

```
% Read images and get the dimensions
[rgbImage] = read_photos(myFolder);
rgbImage=double(rgbImage);
num = size(rgbImage,4);

% Set Black pic to re-create the rgb fused image
allblack = zeros(size(rgbImage,1),size(rgbImage,2),'uint8');

%-----Red Channel-----%
for i = 1 : num
    redChannel(:, :, i) = rgbImage(:, :, 1, i); % Red channel
end

% Quadtree based image fusion
[redfimg, red_decision_map] = Quadtree_Fusion(redChannel, num, level);

% Show the redChannel fusion image
just_red = cat(3, redfimg, allblack, allblack);
figure();
imshow(just_red);
title('Red Channel Of Fused Image')
```

Figure 4.1: Read images and red channel ready for fusion

```
function [imageArray] = read_photos(myFolder)
if ~isdir(myFolder)
    errorMessage = sprintf('Error: The following folder does not exist:\n%s', myFolder);
    uiwait(warndlg(errorMessage));
    return;
end
filePattern = fullfile(myFolder, '*.png');
imageFiles = dir(filePattern);
for k = 1:length(imageFiles)
    baseFileName = imageFiles(k).name;
    fullFileName = fullfile(myFolder, baseFileName);
    fprintf(1, 'Now reading %s\n', fullFileName);
    imageArray(:, :, :, k) = imread(fullFileName);
end
end
```

Figure 4.2: Read images function

### 4.2.2 Detecting the focused regions

Detecting the focused regions from the source images is the most crucial part of the algorithm. To achieve the optimal subdivision of the source images, and get all the focused regions, we propose an effective quadtree decomposition strategy. For convenience, the proposed strategy is described for two source images and the strategy can be easily extended to three or more source images.

Two source images represent the root block-pair at the first level of the quadtree structure. If the decomposition criterion is fulfilled in the root block-pair, then each block in the pair is subdivided into four quadrants. And this process is repeated on the smaller block-pairs, until the decomposition criterion is not fulfilled in all the block-pairs.

Two situations may occur in a block-pair: (1) one block is fully focused and the other is fully defocused; (2) both of the blocks are partially focused.

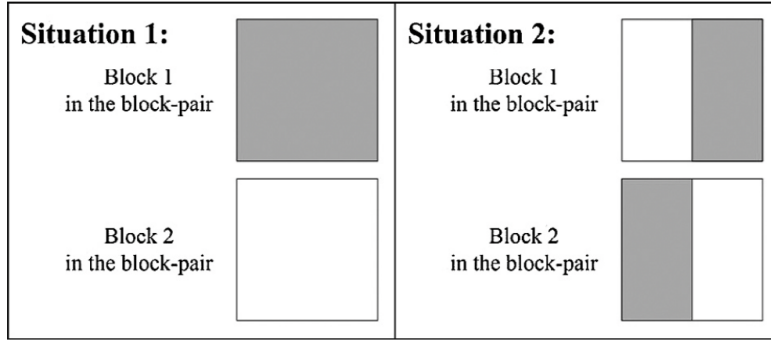


Figure 4.3: Situations on a block-pair with focus on shaded regions. Situation 1: one block is fully focused, the other fully defocused; Situation 2: both are partially focused.

To find out the situation, the maximum difference in focus-measures (MDFM) and the sum of the maximum difference in gradients (SMDG) are firstly calculated on the block-pair. If  $MDFM \geq 0.98 * SMDG$ , then the block-pair fits Situation 1, otherwise, this block-pair fits Situation 2. If the block-pair fits Situation 1, then the fully focused block can be found out in this pair; otherwise, the block-pair fits Situation 2, each block in the pair should be further subdivided into four quadrants.

```
% Compute MDFM: maximum difference in the focus measures
MDFM = abs(minBlock - maxBlock);

% Extract the maximum and minimum gradients of each pair of blocks
[minblockV, Sind] = qtgetblk(minGrad, S, dim);
[maxblockV, Sind] = qtgetblk(maxGrad, S, dim);

maxSum = sum(sum(maxblockV));
minSum = sum(sum(minblockV));
% Compute SMDG: the sum of maximum difference of the gradients
SMDG = abs(maxSum - minSum);

% Compute the threshold T for discriminate whether the block-pair should be split
T = SMDG * 0.98;

% The block-pairs should be split
result = (MDFM > 0) & (MDFM < T);
```

Figure 4.4: Check the situation

In order to measure the focus of the block-pairs, the gradients of the source images should be firstly calculated. The gradient used here is the weighted modified gradient. We find out the maximum gradient map and the minimum gradient map of the source images, respectively.

$$GM_{max}(x, y) = \max(GM_1(x, y), GM_2(x, y)) \quad (4.1)$$

$$GM_{min}(x, y) = \min(GM_1(x, y), GM_2(x, y)) \quad (4.2)$$

, where  $GM_{max}$  and  $GM_{min}$  are the maximum and minimum gradient map of all source images, respectively.

```
%----- Computing the minGrad, maxGrad and sumGrad at each pixel -----
minGrad = NormGrads(:,:,1);
maxGrad = minGrad;
for ii = 2 : num
    minGrad = min(minGrad, NormGrads(:,:,ii));
    maxGrad = max(maxGrad, NormGrads(:,:,ii));
end
```

Figure 4.5: Gradients calculations

In Eq 4.1 and 4.2,  $GM_1$  and  $GM_2$  are the gradients of the source images, respectively. Because  $GM_{max}$  at each location is chosen as the greater gradient of  $GM_1$  and  $GM_2$  at the corresponding location,  $GM_{max}$  approximates the gradient map of a fully focused image. Similarly,  $GM_{min}$  at each location is selected as the minimum gradient of  $GM_1$  and  $GM_2$  at the corresponding location, then it approximates the gradient map of a fully defocused image.

Then we need to select out the gradient map, maximum gradient map and minimum gradient map for the block-pair at the block area. But we have already computed the max and min Gradients so the maps are just copied from the corresponding part of  $GM_i$ ,  $GM_{max}$  and  $GM_{min}$ , respectively. By this way, the gradients of the sub-regions need not to be recomputed.

Then we need to find out the maximum and minimum focus-measures of the blocks in the pair, respectively.

$$FM_{max} = \max(FM_1, FM_2) \quad (4.3)$$

$$FM_{min} = \min(FM_1, FM_2) \quad (4.4)$$

, where  $FM_{max}$  and  $FM_{min}$  are the maximum and minimum focus measures of all the block in the pair, respectively.

```

% Compute the maxFM and minFM
sumBlock = sum(sum(blockV));
minBlock = sumBlock;
maxBlock = sumBlock;
for i = 2 : num
    [blockV, Sind] = qtgetblk(NormGrads(:,:,i), S, dim);
    tempBlock = sum(sum(blockV));
    minBlock = min(minBlock, tempBlock);
    maxBlock = max(maxBlock, tempBlock);
end
    
```

Figure 4.6: Focus-measures calculation

For a partially focused block, the focus-measure of this block will become greater, with the enlarging of the focused area in the block. And, the focus-measure of this block will become smaller, with the shrinking of the focused area in the block. Thus, if a block-pair fits in Situation 1, in which a block is a fully focused and another is fully defocused,  $FM_{max}$  of this block-pair would equal to the focus-measure of the fully focused block and  $FM_{max}$  should be the extreme maximum focus-measure of this block. Meanwhile,  $FM_{min}$  of this block-pair would equal to the focus-measure of the fully defocused block and  $FM_{min}$  should be the extreme minimum focus-measure of this block.

Then, the maximum difference in focus-measures ( $MDFM$ ) for the block-pair can be calculated as the difference between maximum focus-measure  $FM_{max}$  and minimum focus-measure  $FM_{min}$  as expressed below.

$$MDFM = FM_{max} - FM_{min} \quad (4.5)$$

Generally, if  $MDFM$  of a block-pair equals to the difference in the focus-measures of a fully focused block and the corresponding fully defocused block, then this block-pair should fit Situation 1, in which there is a fully focused block and a fully defocused block.

Finally, the sum of the maximum difference in gradients ( $SMDG$ ) can be calculated as show in Fig 4.4 and below:

$$SMDG = \sum \sum GM_{max}(x, y) - \sum \sum GM_{min}(x, y) \quad (4.6)$$

However in the real images, there may exist some noise pixels in any one of the corresponding blocks. As a result, for a block-pair in Situation 1,  $MDFM$  of the block-pair usually appears a little less than the corresponding  $SMDG$ . To suppress this defect, we choose  $0.98 * SMDG$  as the threshold value to discriminate the situation. For a block-pair, if  $MDFM \geq 0.98 * SMDG$ , this block-pair fits Situation 1. Otherwise, this block-pair fits Situation 2.

### 4.2.3 Detect focus Blocks in quadtree structure

As mentioned above, the  $m$  source images are input as the root block-set at the first level of the quadtree structure. For the root block-set, if  $MDFM \geq 0.98 * SMDG$ , then this block-set fits Situation 1 and the fully focused block can be found out by measuring the focus of the blocks in this set. Otherwise, this block-set fits Situation 2 and each block in the set should be further subdivided into four smaller blocks at the second level. This process is repeated on the smaller block-sets, until all focused blocks are found out or the maximum level of the quadtree structure is reached. Finally, the detected focused blocks are merged naturally to form the focused regions.

The source images are decomposed to the minimum  $2 \times 2$  sized blocks, and the corresponding level is the maximum level, as show below.

$$maxlevel = \log\{max(M - 1, N - 1)\} + 1 \quad (4.7)$$

, where  $M \times N$  is the size of the source images;  $\log$  is the logarithm operator.

Before the quadtree decomposition, the dimensions of the source images should be extended to a squared size and a power of 2.

```
% Center the images and the gradient maps on the extended images or maps
dx = ceil((NormDim - p1) / 2);
dy = ceil((NormDim - p2) / 2);

NormGrads = zeros(NormDim, NormDim, num);
NormGrads(dx + 1 : dx + p1, dy + 1 : dy + p2, : ) = Grads;
|
```

Figure 4.7: Center the images

The process of detecting the focused blocks in a quadtree structure is algorithmically summarized in Algorithm 1.

---

```

(1) Read the  $m$  source images as the root block-set;
(2) Compute the gradient map of each source image;
(3) Calculate the maximum possible level (max level) of the
quadtree structure;
(4) Decompose the source images into a quadtree structure:
Set current level  $l$  equal to 1;
while ( $l < \text{max level}$ )
  if the number of the block-sets at level  $l$  equals to zero
  then
    break;
  end
  for each block-set at level  $l$ 
    Compute the corresponding SMDG and MDFM;
    if  $\text{MDFM} \geq 0.98 \times \text{SMDG}$  then
      Compute the focus-measure of each block in the set;
      if only one block has the maximum focus-measure
      then
        Extract this block from the corresponding source
        image to the fusion image;
      end if
    else
      Subdivide each block in this set into four quadrants;
    end if
  next block-set
   $l = l + 1$ ;
loop

```

---

Figure 4.8: Algorithm 1

```

% Define a tag for most FM images
FM_tag = zeros(NormDim, NormDim);

% Begin loop
while (level <= maxlevel)

    len = length(find(S == dim));
    if len ~= 0
        values = zeros(num, len);
        for i = 1 : num
            % Extrat the corresponding blocks at the current level
            [BlockGrad, S_idx] = qtgetblk(NormGrads(:, :, i), S, dim);

            % Compute focus-measure of the blocks
            values(i, :) = sum(sum(BlockGrad));
        end

        % Calculate the maximum focus-measure of the corresponding blocks
        maxVals = max(values);

        % Find the image indices of the blocks with maximum focus-measure
        is_equal = zeros(1, len);
        eq_count = is_equal;
        index = is_equal;

        for i = 1 : num
            is_equal = (maxVals == values(i, :));
            eq_count = eq_count + is_equal;
            index = index + is_equal * i;
        end

        index = index .* (eq_count == 1) + (-1) * (eq_count > 1);

        % Detecting the focused blocks
        blocktag = zeros(size(BlockGrad));
        for i = 1 : len
            blocktag( : , : , i) = index(1, i);
        end

        % Set the fusion tag as the image indices of the focused blocks
        FM_tag = qtsetblk(FM_tag, S, dim, blocktag);
    end

    % Next scale
    dim = dim/2;
    level = level + 1;
end

```

Figure 4.9: Detecting the focused blocks



#### 4.2.4 Weighted focus measure: Sum of Weighted Laplacian

In order to detect the focused regions precisely, a weighted focus-measure called the sum of the weighted modified Laplacian is proposed.

For the multi-focus images, a focused block should preserve more sharp edges than the corresponding defocused blocks. In images, the gradients represent the edge information. However in the homogeneous regions, the corresponding gradients are small. The focus-measures of both the focused block and the corresponding defocused blocks would also be small. Therefore, the focus-measure of a focused block could not be clearly distinguished from the corresponding defocused blocks by using some commonly used focus-measures. And, the focus-measure of a block in the homogenous regions could be easily affected by the noise pixels.

To overcome this problem, we weight the gradient at each location in a local window. By this way, the strength of the final gradient at each location would be enhanced by the nearby gradients. It is proved experimentally that the sum of the modified Laplacian can outperform the other focus-measures. Thus a weighted version should be powerful and it is created the sum of the weighted modified Laplacian.

So we start with calculating the Laplacian. The second derivatives in the x and y directions may have the opposite signs, so that they tend to cancel each other. Therefore, a small value may be obtained at a high gradient location. To suppress this defect, the modified Laplacian is proposed (ML), which takes the absolute values of the second derivatives in the Laplacian. The discrete approximation to ML can be expressed as:

$$ML(x, y) = |2f(x, y) - f(x - step, y) - f(x + step, y)| + |2f(x, y) - f(x, y - step) - f(x + step, y + step)| \quad (4.8)$$

, where *step* is a variable spacing, which is used to adapt to the possible variations.

```
% Get the dimensions of the image
[M, N] = size(g);

% A fast way to compute the SML by using the matrix operations
mat1 = gpad(step + 1 : M + step, step + 1 : N + step);

mat2 = gpad(1 : M, step + 1 : N + step);
mat3 = gpad(2 * step + 1 : M + 2 * step, step + 1 : N + step);

mat4 = gpad(step + 1 : M + step, 1 : N);
mat5 = gpad(step + 1 : M + step, 2 * step + 1 : N + 2 * step);

% Two components
part1 = abs(2 * mat1 - mat2 - mat3);
part2 = abs(2 * mat1 - mat4 - mat5);

result = part1 + part2;

% Threshold
result(result < T) = 0;
```

Figure 4.10: Modified Laplacian calculation

Then, the sum of the modified Laplacian (SML) at the location  $(x, y)$  can be calculated in a small window around  $(x, y)$ , as shown below:

$$SML(x, y) = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} ML(i, j), ML(i, j) > T \quad (4.9)$$

, where  $N$  determines the size of the window and  $T$  is a threshold value.

To improve the detection accuracy of the focused blocks in the homogenous regions, we present a weighted version of  $ML$ , named as the weighted modified Laplacian (WML).

First, the  $ML$  gradients of the source images are calculated according to Eq 4.8. Then, the  $WML$  gradient at the pixel location  $(x, y)$  is calculated as the weighted sum of the  $ML$  gradients in a small local window. The  $WML$  can be mathematically expressed as:

$$WML(x, y) = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} \frac{ML(i, j)}{\sqrt{(i-x)^2 + (j-y)^2}}, ML(i, j) > T \quad (4.10)$$

```
% Weight operator: H
H = ones(sz, sz);
center = c + 1;
for ii = 1 : sz
    for jj = 1 : sz
        dis = sqrt( (ii - center) ^ 2 + (jj - center) ^ 2 );
        H(ii, jj) = 1 / (dis + 1);
    end
end
```

Figure 4.11: Calculate the weight for ML

Finally, the sum of the weighted modified Laplacian (SWML) at the location  $(x, y)$  can be calculated as:

$$SWML(x, y) = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} WML(i, j) \quad (4.11)$$

, where  $N$  determines the size of the window.

```
% First, compute the modified Laplacian gradients
mlg = ML(img, step, T);
% Then, sum the ML gradients with local weights
result = Local_Weight(mlg, bsz);
```

Figure 4.12: SWML

### 4.2.5 Reconstruction of the focused regions

Although the quadtree detection method and the weighted focus-measure  $WML$  are very effective, there are still some small spurious regions inside the detected focused regions. Thus we use two sequential filters to reconstruct the focused regions, one morphological filter and one small region filter.

#### Morphological Filter

In the detected focused regions, there may be some small lines or burrs in the connection portions, and some nearby regions may be disconnected by the inappropriate decisions.

The morphological filter is used to do a minor adjustment of the focused regions. Thus, a small size of the disk structuring element would be enough for the deletion of the small burrs and the connection of the nearby regions.

```
for ii = 1 : N
    Imap( : , : , ii) = (map == ii);
end
for ii = 1 : Iter
    for jj = 1 : N
        temp = Imap( : , : , jj);
        se = strel('disk', ii + 2);

        temp = imopen(temp, se);
        temp = imclose(temp, se);
        Imap( : , : , jj) = temp;
    end
end
% Equal map
sum_map = sum(Imap, 3);
% non_map = sum_map > 1 | sum_map < 1;
def_map = (sum_map == 1);
Fusion_map = zeros(size(map));
for ii = 1 : N
    Fusion_map = Fusion_map + Imap( : , : , ii) * ii;
end
result = Fusion_map .* def_map;
```

Figure 4.13: Morphological Filter

### Small Region Filter

there may exist some small isolated regions inside a focused region. To handle this problem, we utilize a small region filter to delete the small isolated regions or fill the holes in the focused regions detected from each source image.

```
% Define the size of the small region
P = small_size;
for ii = 1 : num
    ptmap = (decision_map == ii);

    % Process the Positive image, delete the small patches
    pL = bwlabel(ptmap, 8);

    % Compute the area of each component.
    pS = regionprops(pL, 'Area');

    % Remove small objects.
    ppL = ismember(pL, find([pS.Area] >= P));

    % Process the Negative image, delete the small patches
    ntmap = ~(ppL > 0);
    nL = bwlabel(ntmap, 8);

    % Compute the area of each component.
    nS = regionprops(nL, 'Area');

    % Remove small objects
    npL = ismember(nL, find([nS.Area] >= P));
    tmap = ~(npL > 0);
    filtered_map(:, :, ii) = tmap;
end

% Find if there are confused pixels
sCount = sum(filtered_map, 3);
Tag = (sCount == 1);
result = zeros(size(decision_map));
for ii = 1 : num
    result = result + ii * filtered_map(:, :, ii) .* Tag;
end
```

Figure 4.14: Small Region Filter

### 4.2.6 Final Fused Image

The final fusion image is produced as the following steps.

- The finally detected focused regions are copied directly from the source images to the fusion image.
- There may exist some regions belonging to no image or more than one image, here the value of each pixel in these regions is copied as the pixel from the source image  $i$  which has the maximum gradient at this location. If the gradients of different source images at a pixel location have the same value, then the value of the pixel at this location is copied as the average value of the pixels from all the source images. The mathematical expression is shown below.

$$I_F(x,y) = \begin{cases} \frac{1}{m} \sum_{i=1}^m I_i(x,y), & \text{if } GM_p(x,y) = GM_q(x,y), \\ p \neq q, & p, q \in [1, m] \\ I_i(x,y), & i = \arg \max_j (GM_j(x,y)), \\ & j \in [1, m] \end{cases}$$

Figure 4.15: Non part selection math expression

```
% Initialize the fusion image
fimg = zeros(p1,p2);

% Firstly, fusing the defined part, copied directly according to the decision map
for ii = 1 : num
    fimg = fimg + channel(:,ii) .* (Fusion_decision_map == ii);
end

% Secondly, fusing the non-defined part, copied by maximum selection method
max_tag = zeros(p1, p2, num);
img_tag = zeros(p1, p2);

% Find the pixels with the maximum gradients from each gradient map
for ii = 1 : num
    tag = (Grads(:,ii) == maxGrad);
    max_tag(:,ii) = tag;
    img_tag = img_tag + tag .* ii;
end
```

Figure 4.16: Create the first part of the Fused Image from max gradients

```
% The nonpart images and maximum selection
non_part = (Fusion_decision_map < 1);
nonImgs = channel;
part1 = zeros(p1,p2);
for ii = 1 : num
    nonImgs(:,ii) = nonImgs(:,ii) .* non_part;
    part1 = part1 + nonImgs(:,ii) .* max_tag(:,ii);
end

% Finding the positions where more than one Grad(i) have the maxGrad
max_num = sum(max_tag, 3);

% The single and multiple positions
single_num = (max_num == 1);
multi_num = 1 - single_num;

% If there are more than one image grad(i) equal to maxGrad
part2 = sum(nonImgs, 3) ./ num;

% As to the whole nonpart
nonPart = part1 .* single_num + part2 .* multi_num;

% Final Fused image, which is combined by the fusion images of the defined
% part and the non-part
fimg = uint8(fimg + nonPart);
```

Figure 4.17: Select the non parts and create the final Fused Image

### 4.3 Experimental-Results

To show the effectiveness of the algorithm, we compare our algorithm with some existing algorithms, contrast gradient based algorithm, Gradient pyramid based algorithm, DWT based algorithm, PCA based algorithm, and shift invariant DWT based algorithm (SIDWT).

The experiments have been performed on several commonly used multi-focus image sets such as the clock, disk, lab etc. The performance of the algorithms is evaluated by both the qualitative and quantitative methods. Some comparison examples are shown below.

Finally some "lab" results of the algorithm are going to be presented. All the sample images, that are going to be fused, were taken by a microscope and provided by the electronics lab in order to full-fill this thesis and see the actual effectiveness of the algorithm.

#### 4.3.1 Qualitative Comparison

For the multi-focus image fusion, good visual effect means effective fusion. To do the qualitative comparison, the performance of the algorithms should be evaluated by measuring the vision quality of the fusion images. Here, three comparison examples are going to be presented.

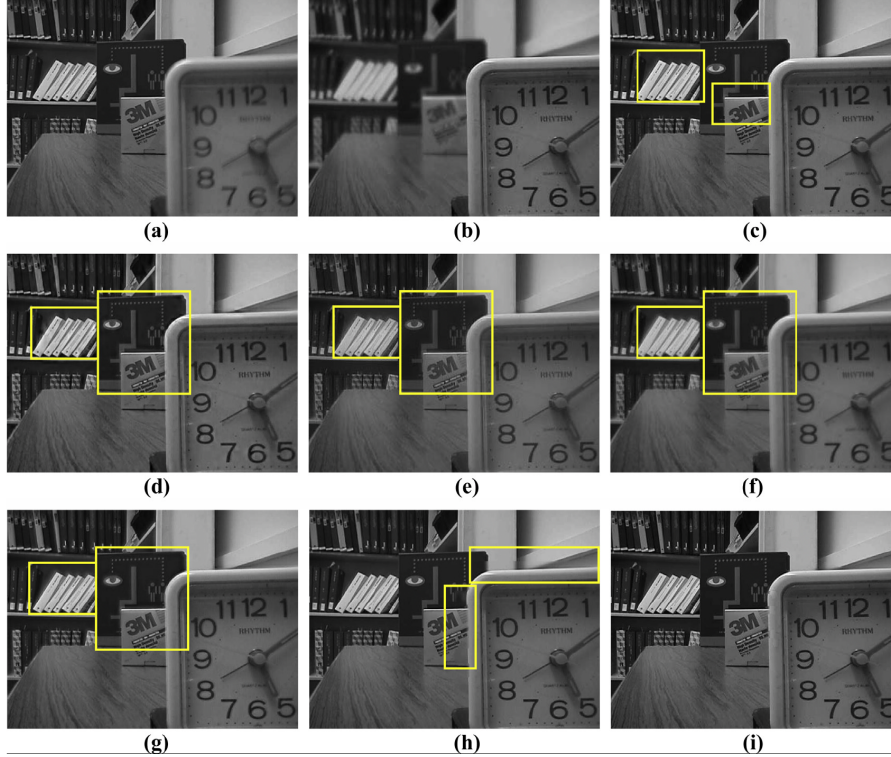


Figure 4.18: The multi-focus 'Disk' images and its fusion images. (a and b) The source images; (c-i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively.

Fig 4.18 shows the fusion results of the ‘Disk’ image set. It is obvious that the fusion images by the Gradient pyramid based algorithm (Fig. 4.18(c)) and PCA based algorithm (Fig. 4.18(d)) appear low in contrast. The fusion images by the Gradient pyramid based (Fig. 4.18(c)), PCA based (Fig. 4.18(d)), contrast pyramid based (Fig. 4.18(e)), DWT based (Fig. 8(f)) and SIDWT based (Fig. 4.18(g)) algorithms suffer from the blurring effects, labeled by the yellow and orange rectangles. De’s algorithm, its fusion image (Fig. 4.18(h)) yields blocking artifacts and blurring edges, labeled by the yellow rectangles. The fusion image (Fig. 4.18(i)) produced by the proposed algorithm shows the best visual effect.

As known, the contrast pyramid based, DWT based, gradient contrast based, PCA based and SIDWT based algorithms are the transform domain techniques, and they would have a smooth effect on the details of the image. Thereby, their fusion images would suffer from the blurring effects around the edge area. Our proposed algorithm could well extract the focused regions from the source images, and maintain them in good consistency, therefore the fusion images would be produced with good visual quality.

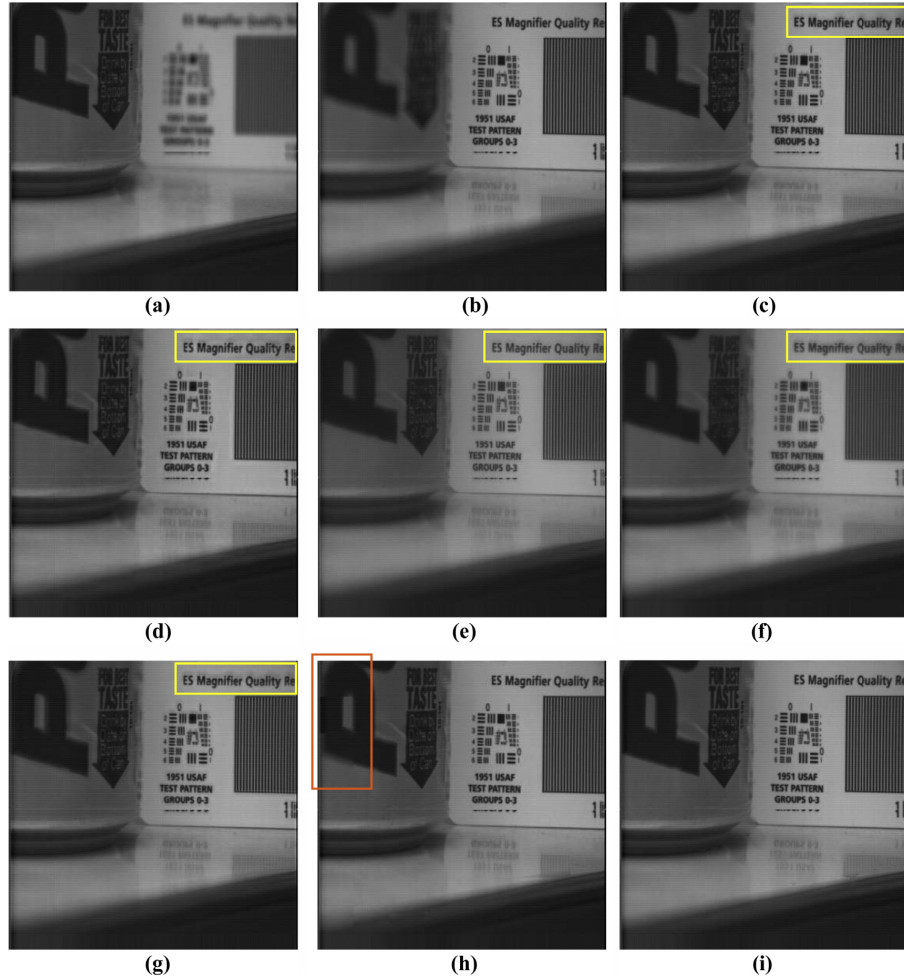


Figure 4.19: The multi-focus ‘Pepsi’ images and its fusion images. (a and b) The source images; (c–i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively.



Fig 4.19 shows the fusion results of the ‘Pepsi’ image set. The fusion images of the Gradient pyramid based algorithm (Fig. 4.19(c)) and PCA based algorithm (Fig. 4.19(d)) appear to be reduced in contrast. The fusion images of the contrast pyramid based (Fig. 4.19(e)), DWT based (Fig. 4.19(f)) and SIDWT based (Fig. 4.19(g)) algorithms have the blurring effects around the text regions, labeled by the yellow rectangles. Also, the fusion image of De’s algorithm (Fig. 4.19(h)) has the blocking artifacts in the character ‘P’ area, labeled by the yellow rectangle. In the end, the fusion image (Fig. 4.19(i)) of the algorithm is fully focused and good for the human visual perception.

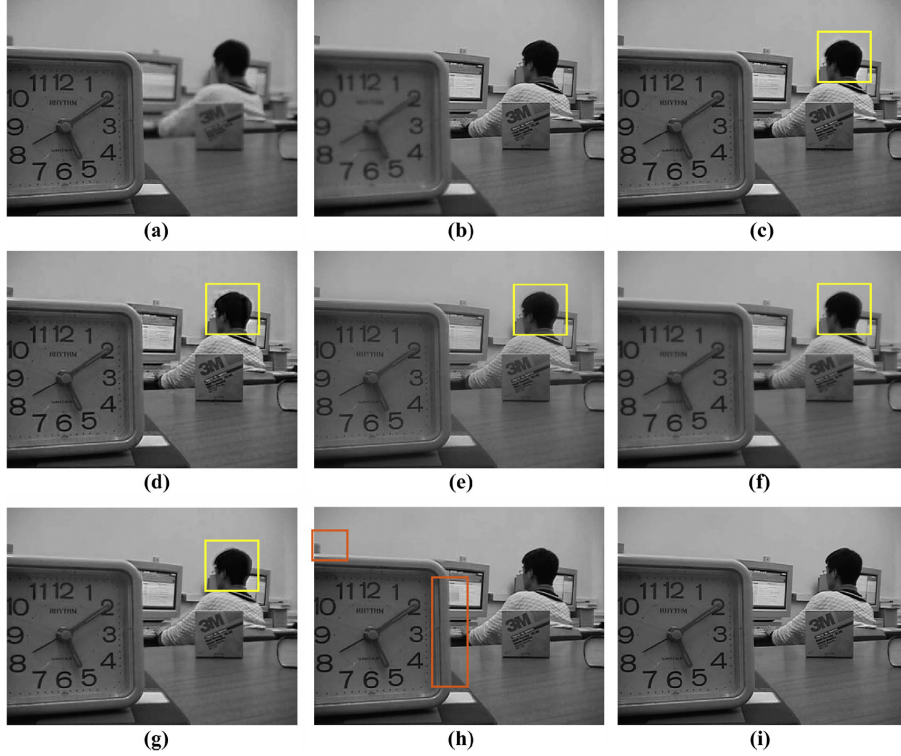


Figure 4.20: The multi-focus ‘Lab’ images and its fusion images. (a and b) The source images; (c–i) the fusion images by using contrast pyramid based algorithm, DWT based algorithm, gradient pyramid based algorithm, PCA based algorithm, SIDWT based algorithm and ours algorithm respectively.

Fig 4.20 shows the fusion results of the ‘Lab’ image set. Here we surely conclude the same thing that our algorithm produces fusion image with better visual quality.

These examples show that the algorithm performs better than the other algorithms. Also, it indicates that the algorithm is competent for the fusion of the multi-focus images.



### 4.3.2 Quantitative Comparison

To assess the fusion images objectively, the fusion algorithms are also evaluated by three quantitative indices, which are gradient similarity metric ( $Q_{GM}$ ), edge based similarity metric ( $Q^{AB/F}$ ) and normalized mutual information ( $Q_{MI}$ ).

#### Gradient similarity metric: $Q_{GM}$

The first metric is the gradient similarity between the source images and the fusion image. It is based on the observation that a pixel has a high gradient value when it is focused

#### Edge based similarity metric: $Q^{AB/F}$

An edge based similarity metric  $Q^{AB/F}$  evaluates the amount of the edge information transferred from the source images to the fusion image.

#### Normalized mutual information: ( $Q_{MI}$ )

Mutual information ( $MI$ ) is a quantitative measure of the mutual dependence of two variables. It shows the measurement of the information shared by two images.

### Quantitative comparison results

The three quantitative indices are calculated according to their definitions, and the results are shown in the following tables. The bold values are the maximum values in the corresponding rows. If the values of the three indices are larger, the performance of the corresponding algorithm is better.

Normalized mutual information.

Image sets	Contrast pyramid	DWT	Gradient pyramid	PCA	SIDWT	De's algorithm	Proposed algorithm
Clock	0.9724	0.8827	0.8670	0.9703	0.9436	1.2075	<b>1.2244</b>
Disk	0.8275	0.7404	0.7466	0.8373	0.8033	1.1373	<b>1.1459</b>
Flower	0.7468	0.6659	0.6661	0.7566	0.7127	1.0171	<b>1.1255</b>
Lab	1.0015	0.9324	0.8953	1.0271	0.9935	1.2280	<b>1.2456</b>
OpenGL	0.9810	0.8914	0.7856	1.0279	0.9575	1.2714	<b>1.2906</b>
Pepsi	0.9643	0.8925	0.8656	0.9759	0.9289	1.2086	<b>1.2483</b>

Comparison of the computational costs (s).

Image sets	Contrast pyramid	DWT	Gradient pyramid	PCA	SIDWT	De's algorithm	Proposed algorithm
Clock(500 × 503)	0.0523	0.0750	0.0800	0.0075	0.3072	1.5036	<b>0.9946</b>
Disk(640 × 480)	0.0613	0.0906	0.0973	0.0092	0.3922	2.3228	<b>1.7222</b>
Flower(764 × 576)	0.0846	0.1322	0.1412	0.0141	0.5952	3.1220	<b>1.8370</b>
Lab(640 × 480)	0.0630	0.0895	0.1018	0.0097	0.3961	2.2095	<b>1.8393</b>
OpenGL(758 × 569)	0.0895	0.1270	0.1424	0.0130	0.5713	2.7748	<b>1.8204</b>
Pepsi(512 × 512)	0.0520	0.0790	0.0836	0.0080	0.3208	1.5156	<b>1.0290</b>

Edge based similarity metric.

Image sets	Contrast pyramid	DWT	Gradient pyramid	PCA	SIDWT	De's algorithm	Proposed algorithm
Clock	0.6880	0.6619	0.6597	0.6137	0.7003	<b>0.7384</b>	0.7373
Disk	0.6849	0.6503	0.6440	0.5365	0.6822	0.7340	<b>0.7383</b>
Flower	0.6444	0.6222	0.6222	0.6197	0.6583	0.6821	<b>0.6962</b>
Lab	0.6910	0.6616	0.6585	0.5651	0.6839	0.7442	<b>0.7470</b>
OpenGL	0.6946	0.6784	0.6707	0.5893	0.7058	0.7289	<b>0.7308</b>
Pepsi	0.7548	0.7293	0.7342	0.6362	0.7445	<b>0.7847</b>	0.7847

It can be seen from Table 1 that, the proposed algorithm performs the best in the most of the cases, which implies that the proposed algorithm can well preserve the sharp gradients from the source images.

The comparison results in Table 2 show that the proposed algorithm performs better than other algorithms in most of the cases. It indicates that the proposed algorithm could well preserve the prominent texture information from the source images.

It is clear from Table 3 that, the proposed algorithm yields the best results in all the cases. This is because the proposed algorithm can well combine the independent information from each source image to form one single image. Here, the independent information is referred as the focused regions detected from each source image.

### Comparison of the computational costs

Finally, we have compared the computational costs of the fusion algorithms. The following Table 4 shows that, the transform domain algorithms (the contrast pyramid based, DWT based, gradient contrast based, PCA based and SIDWT based algorithms) are all running fast. The spatial domain methods (the proposed algorithm and De's algorithm) are running slightly slower than the transform domain algorithms.

Gradient similarity metric.

Image sets	Contrast pyramid	DWT	Gradient pyramid	PCA	SIDWT	De's algorithm	Proposed algorithm
Clock	0.9735	0.9751	0.9621	0.9543	<b>0.9819</b>	0.9778	0.9762
Disk	0.9725	0.9755	0.9535	0.9239	0.9757	0.9725	<b>0.9774</b>
Flower	0.9648	0.9669	0.9445	0.9322	0.9645	0.9607	<b>0.9671</b>
Lab	0.9807	0.9818	0.9646	0.9463	0.9831	0.9813	<b>0.9834</b>
OpenGL	0.9556	0.9642	0.9376	0.8853	<b>0.9662</b>	0.9592	0.9608
Pepsi	0.9852	0.9854	0.9670	0.9449	0.9822	0.9860	<b>0.9862</b>

Although the proposed algorithm has a little higher computational cost than the transform domain algorithms, the proposed algorithm outperforms all the transform domain methods and the De's algorithm from both the qualitative and quantitative evaluations. Therefore, the proposed algorithm is competent for the multi-focus image fusion.

## 4.4 Lab-Results

It is known that the lens of a camera can only focus at one distance or on a plane at a time. Thereby, only the objects falling in the depth of field (DOF) of the lens appear to be sharp, while the objects out of the DOF are blurred. As a result, the multiple objects of interest may not be all focused in one single image.

The samples that are going to be processed have been taken by a microscope of the electronics lab. We will definitely see that many pictures are focused on different spots. Then the algorithm that has been implemented will be applied on these images and the final results that will be produced, are very promising. Pictures of a human hair crossed with a synthetic fiber, a convallaria rhizome, quartz sand and sugar granule will be some of them.

Firstly we will show some of the on-focus and out-of-focus images. Then the fused image in each channel, red green and blue respectively and the final RGB fused image will be represented.

### 4.4.1 Human Hair Crossed with a Synthetic Fiber

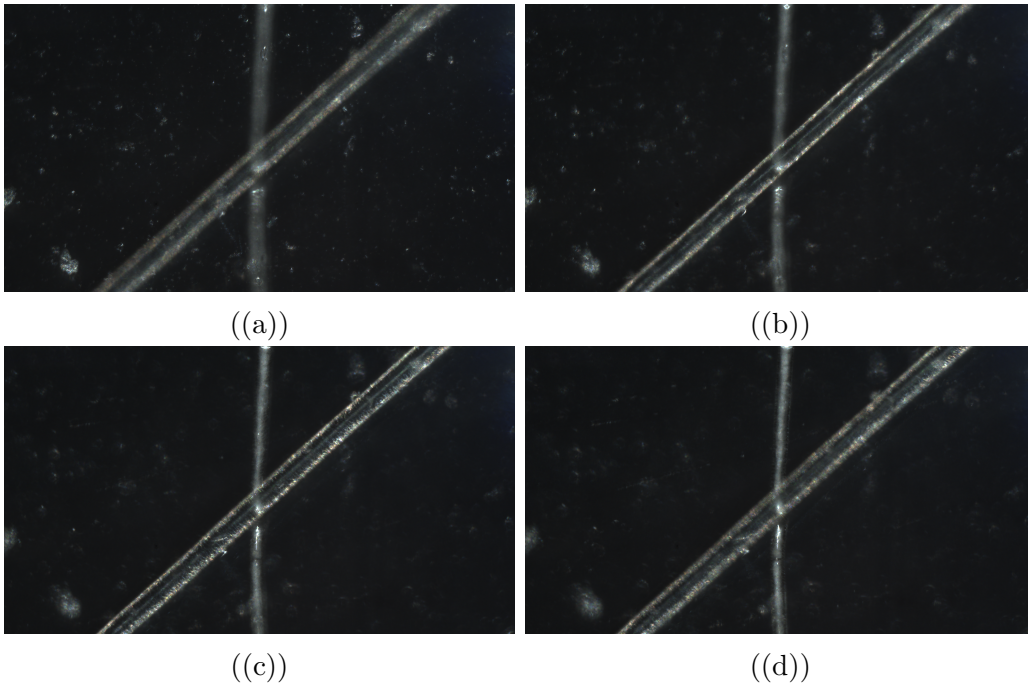


Figure 4.21: Samples focused in different spots

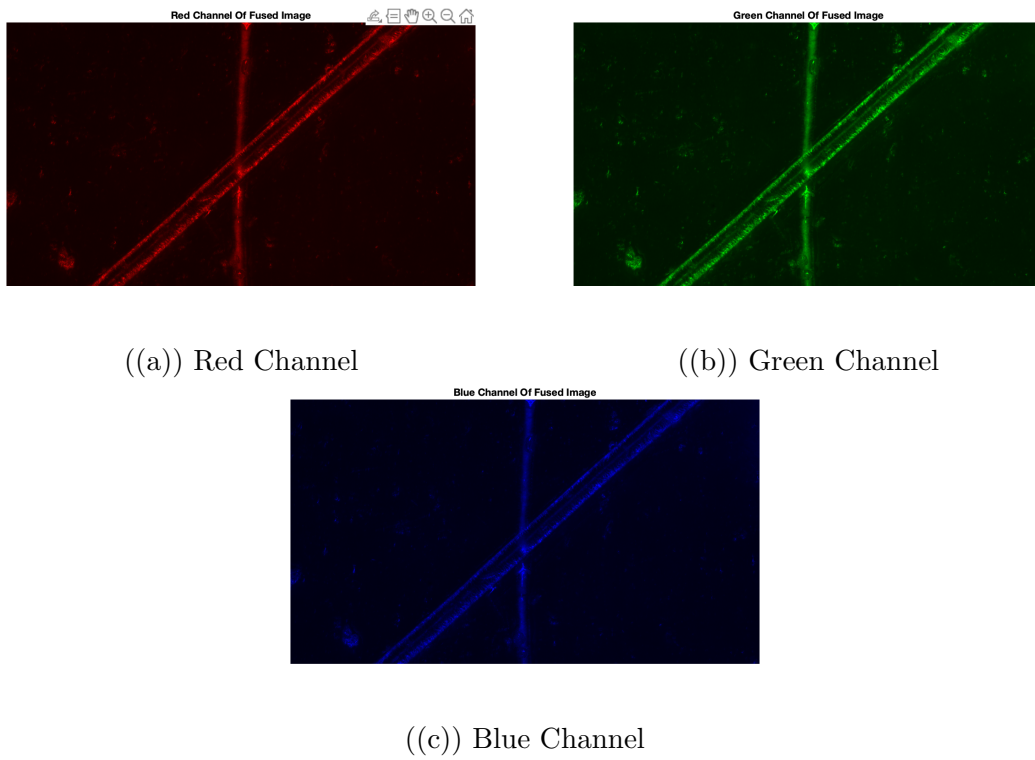


Figure 4.22: Human Hair Crossed with a Synthetic Fiber Sample RGB Channels

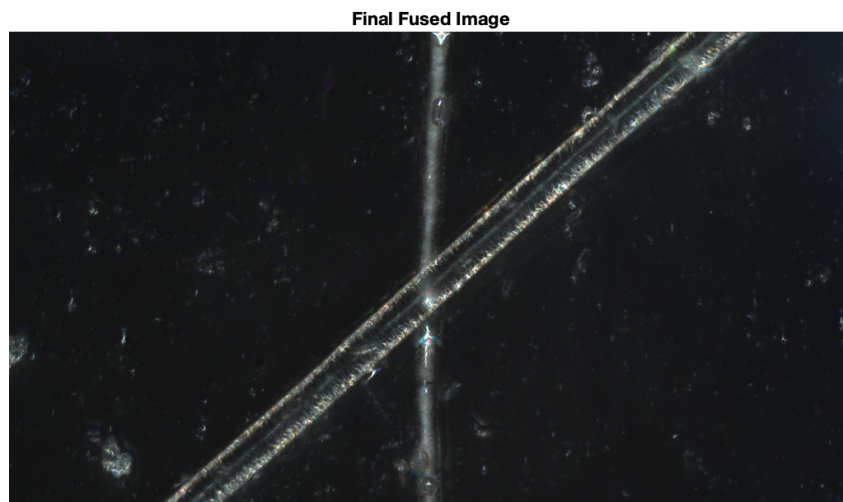


Figure 4.23: Human Hair Crossed with a Synthetic Fiber Final RGB Fused Image

### 4.4.2 Convallaria Rhizome

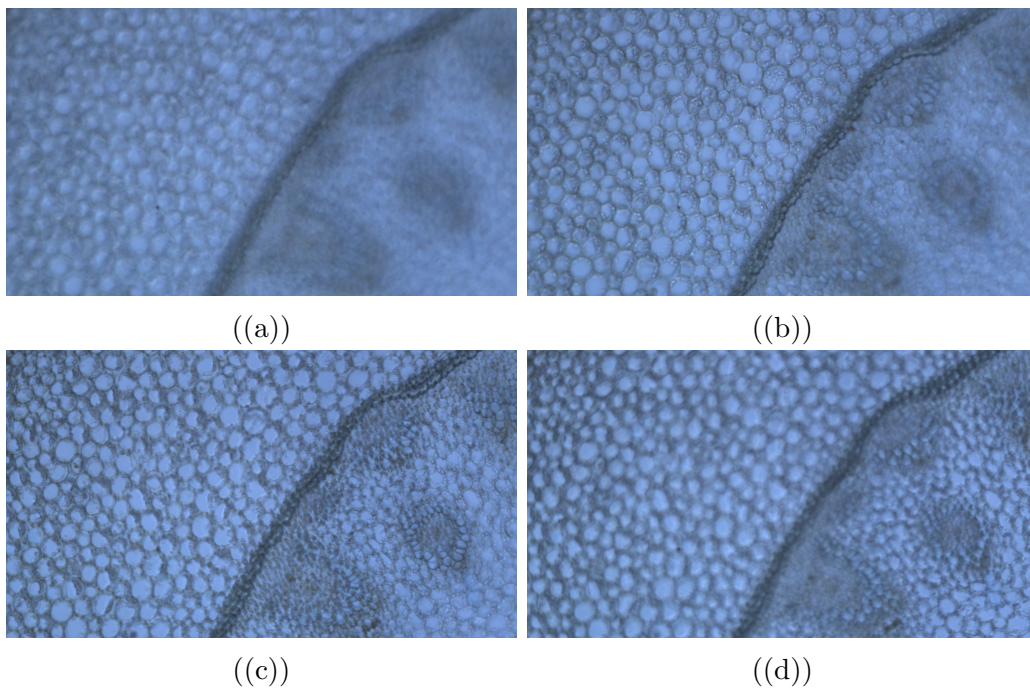


Figure 4.24: Samples focused in different spots

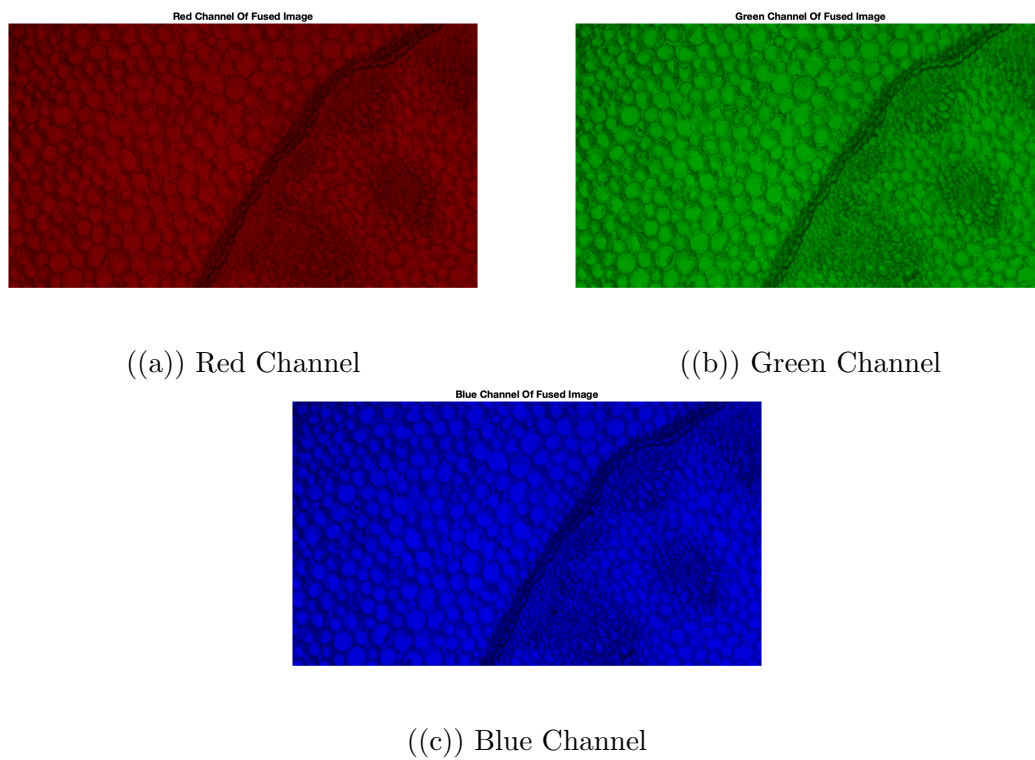


Figure 4.25: Convallaria Rhizome Sample RGB Channels



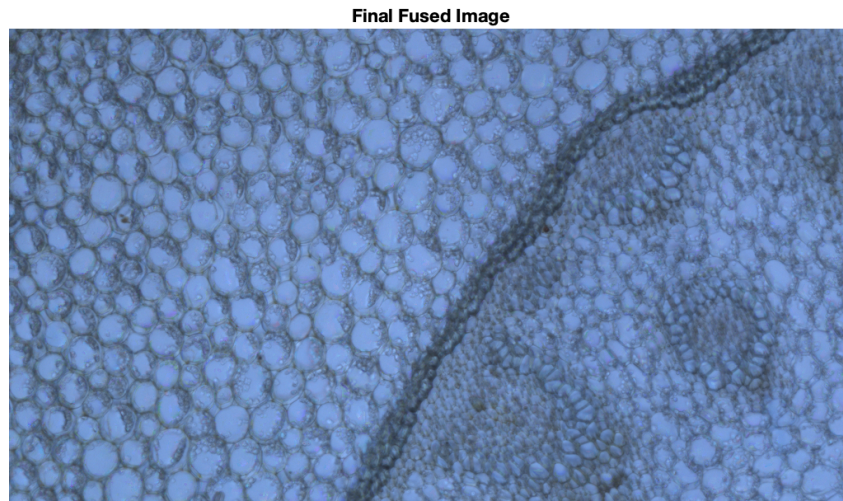


Figure 4.26: Convallaria Rhizome Final RGB Fused Image

#### 4.4.3 Sticky Tape Edge and Cotton Fiber

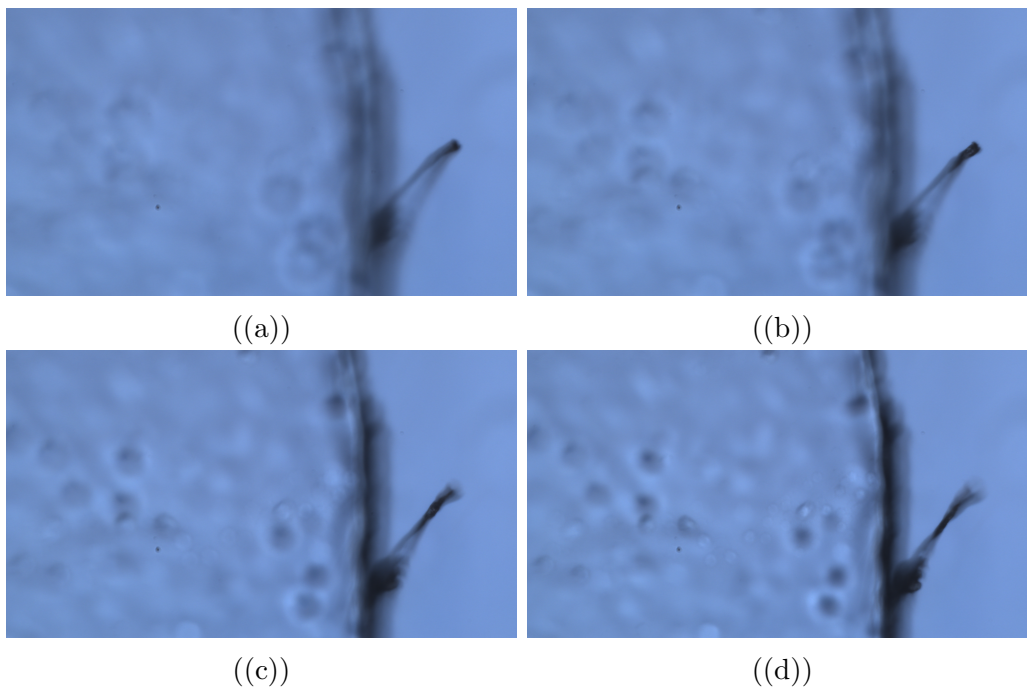


Figure 4.27: Samples focused in different spots

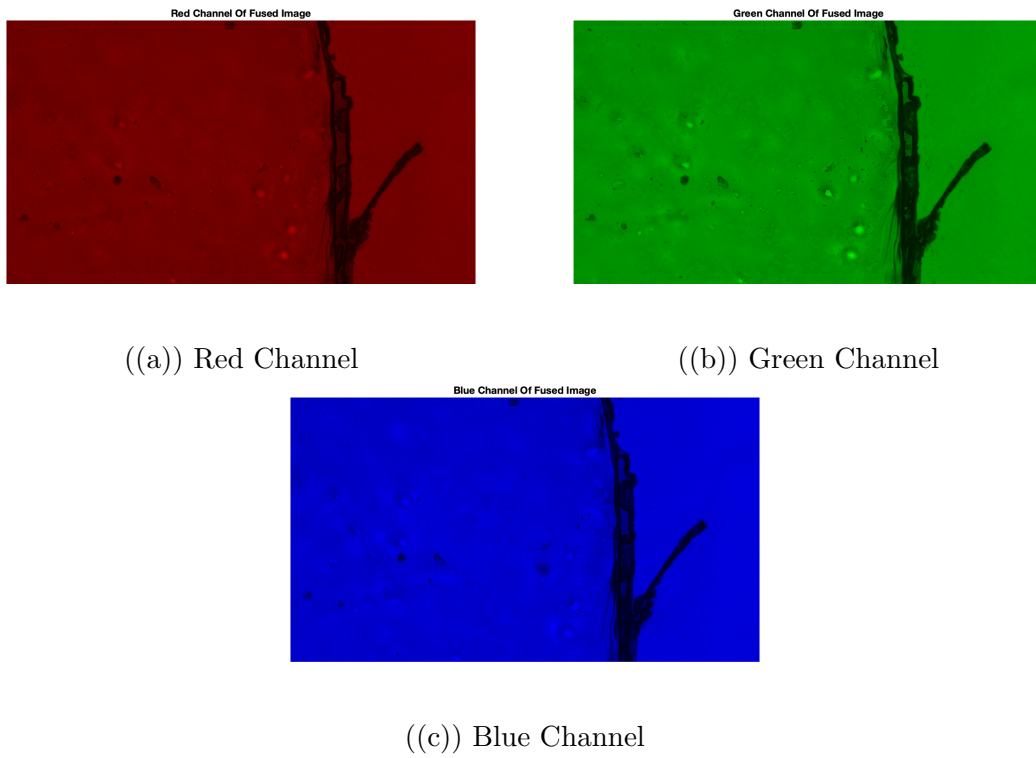


Figure 4.28: Sticky Tape Edge and Cotton Fiber Sample RGB Channels

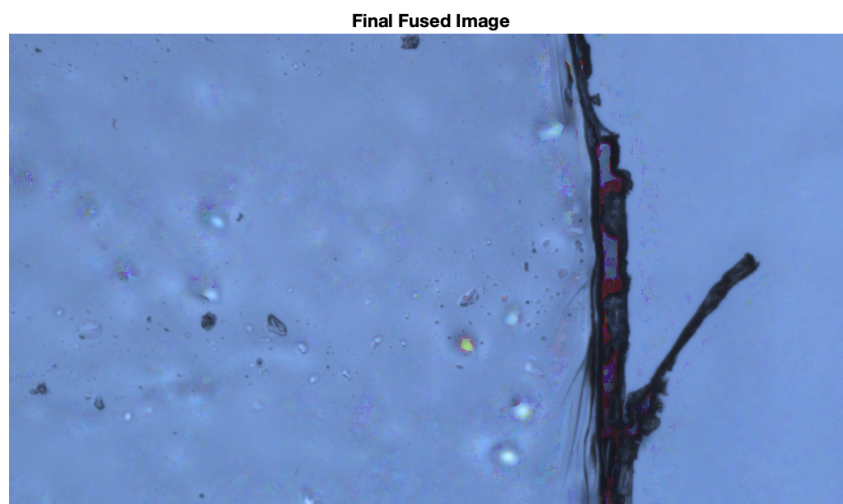


Figure 4.29: Sticky Tape Edge and Cotton Fiber Final RGB Fused Image

#### 4.4.4 Quartz Sand

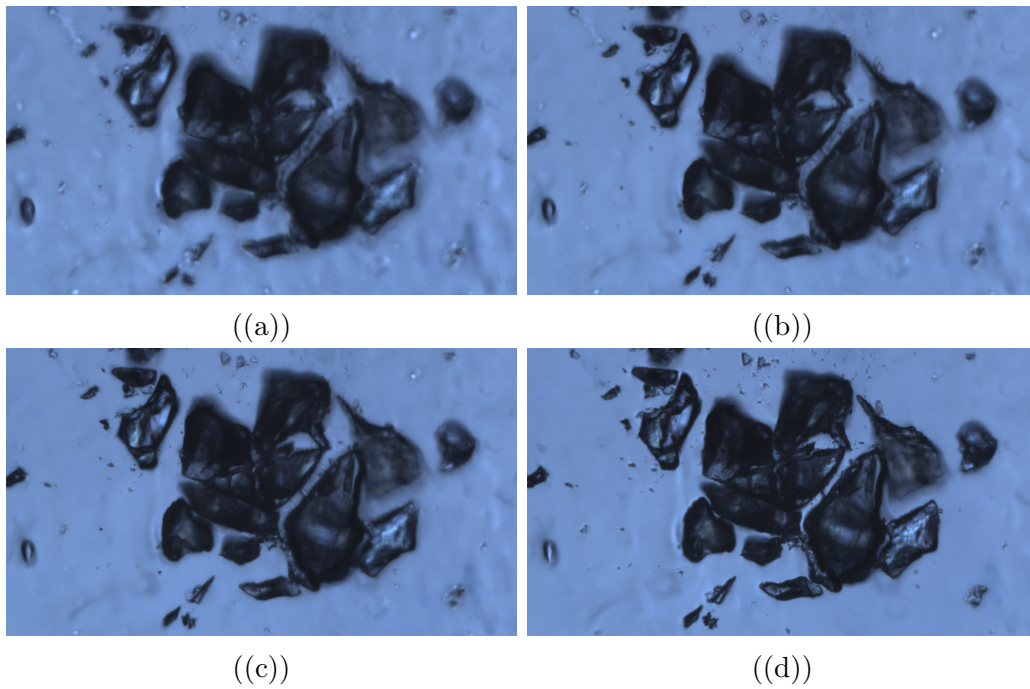


Figure 4.30: Samples focused in different spots

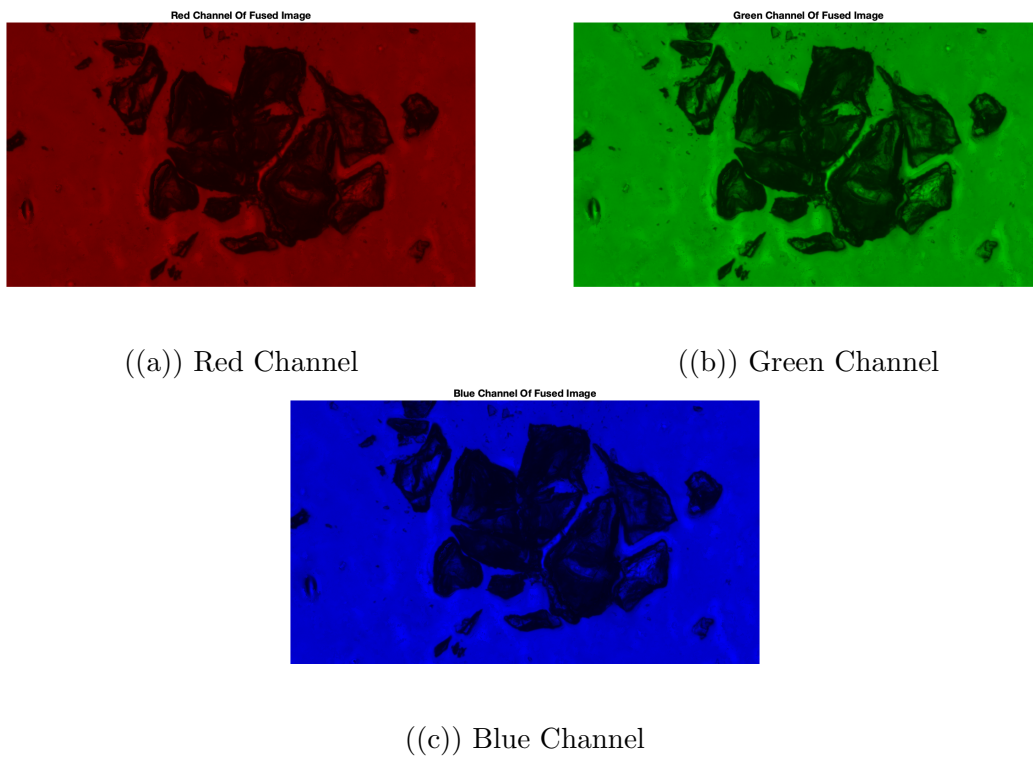


Figure 4.31: Quartz Sand Sample RGB Channels



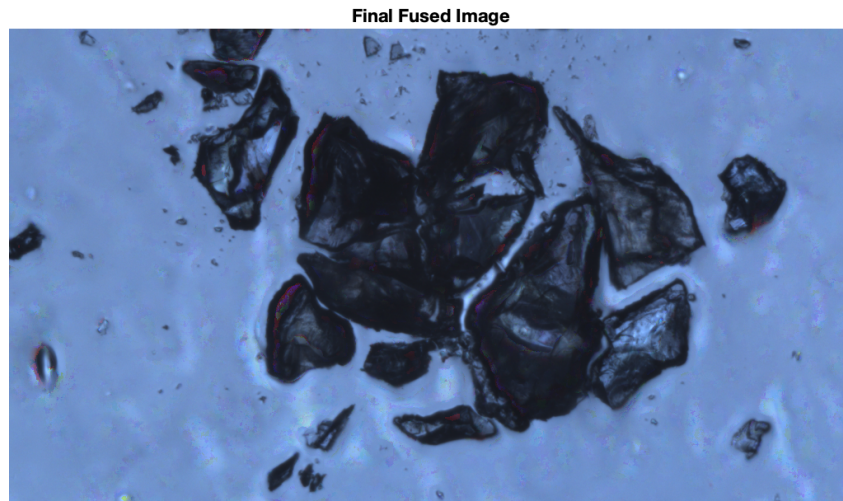


Figure 4.32: Quartz Sand Final RGB Fused Image

#### 4.4.5 Sugar Granule

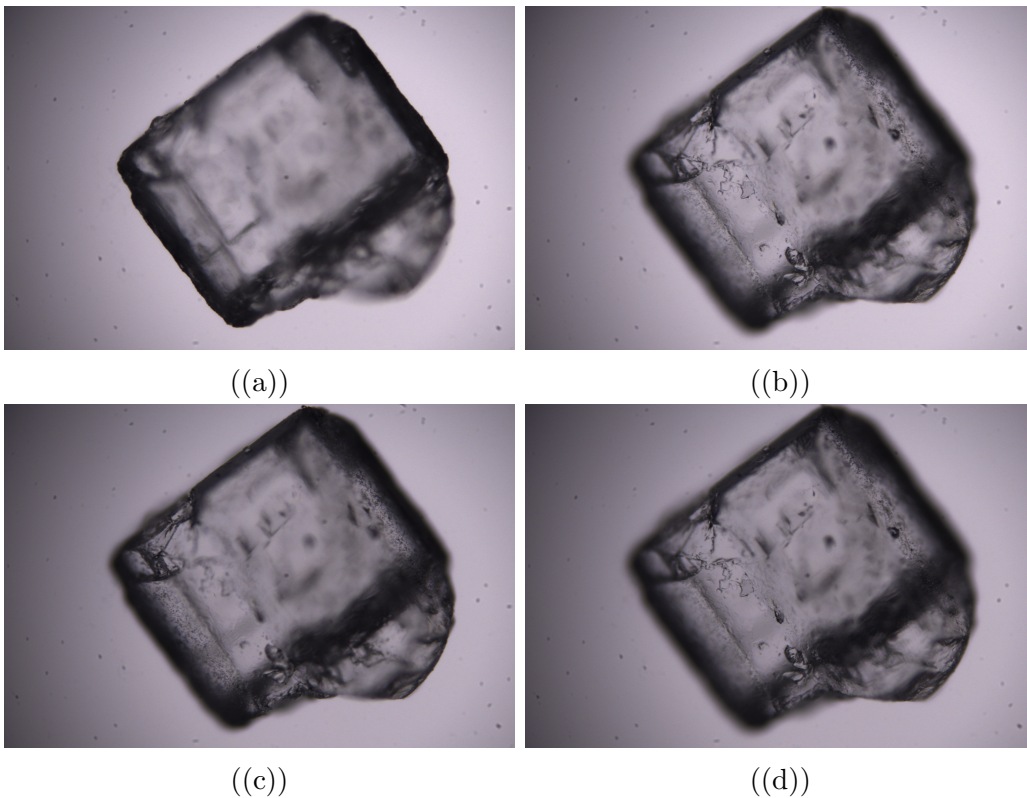


Figure 4.33: Sugar Granule focused in different spots

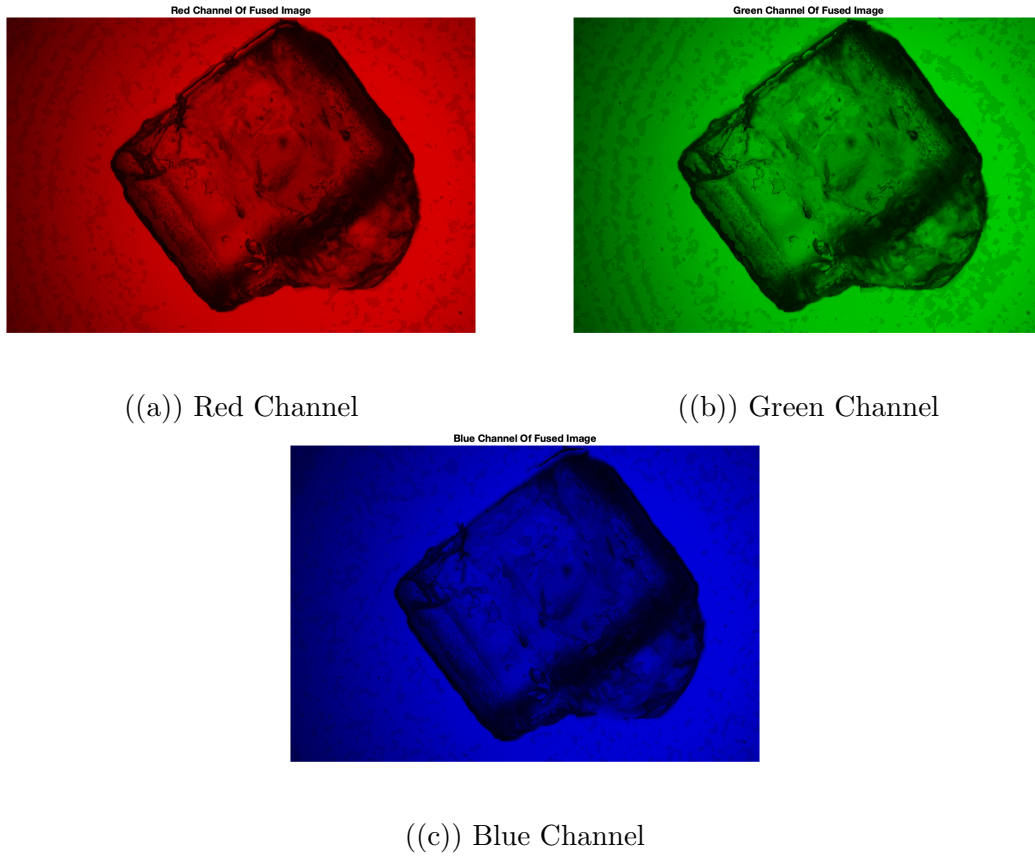


Figure 4.34: Sugar Granule Sample RGB Channels

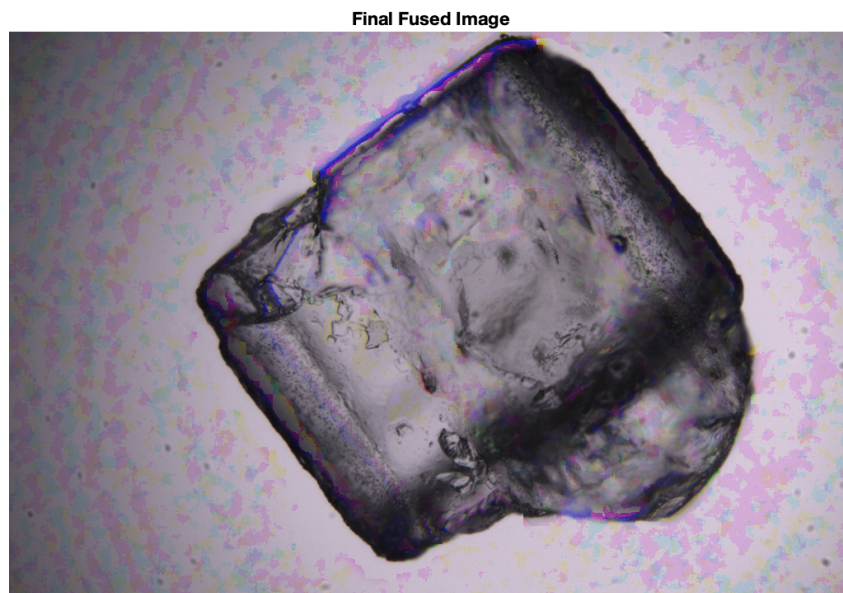


Figure 4.35: Sugar Granule Final RGB Fused Image

# Chapter 5

## Conclusion

To conclude this thesis presents a simple and effective quadtree-based algorithm for multi-focus image fusion. In this work, an effective quadtree decomposition strategy and a weighted focus-measure were presented, thus the focused regions could be detected from the source images in a quadtree structure, effectively and precisely. Also the detected focused regions could be well extracted from the source images and reconstructed to produce the fusion image.

Because *SWML* is sensitive to the homogeneous regions, the focused regions could be detected more accurately by using *SWML* than the commonly used focus-measures. In the algorithm, all source images are decomposed to the maximum level of the quadtree structure, therefore the focused regions could be detected precisely in the tree structure. Moreover, the algorithm is simple, effective and reliable, because of the quadtree decomposition strategy and the weighted focus-measure. The experimental results show that the algorithm performs very well for multi-focus image fusion even though there are some failures using the Matlab function `cov` to reconstruct the final RGB image.

# Bibliography

- [1] Xiangzhi Bai **and others**. “Quadtree-based multi-focus image fusion using a weighted focus-measure”. in *Information Fusion*: 22 (2015), **pages** 105–118. ISSN: 1566-2535. DOI: <https://doi.org/10.1016/j.inffus.2014.05.003>. URL: <https://www.sciencedirect.com/science/article/pii/S1566253514000669>.
- [2] *Confocal Microscope Principle, Uses, Parts, Advantages, And Disadvantages*. — *microbiologynote.com*. <https://microbiologynote.com/confocal-microscope-principle-uses-parts-advantages-and-disadvantages/>.
- [3] *Confocal Microscopy - Introduction — Olympus LS — olympus-lifescience.com*. <https://www.olympus-lifescience.com/en/microscope-resource/primer/techniques/confocal/confocalintro/>. [Accessed 19-Jul-2023].
- [4] *Decision-Level Image Fusion — link.springer.com*. [https://link.springer.com/chapter/10.1007/978-981-15-4867-3\\_4](https://link.springer.com/chapter/10.1007/978-981-15-4867-3_4).
- [5] *Fluorescence — chem.libretexts.org*. [https://chem.libretexts.org/Bookshelves/Physical\\_and\\_Theoretical\\_Chemistry\\_Textbook\\_Maps/Supplemental\\_Modules\\_\(Physical\\_and\\_Theoretical\\_Chemistry\)/Spectroscopy/Electronic\\_Spectroscopy/Radiative\\_Decay/Fluorescence](https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Electronic_Spectroscopy/Radiative_Decay/Fluorescence).
- [6] Alec De Grand. *What Is Darkfield Microscopy?* — *olympus-lifescience.com*. <https://www.olympus-lifescience.com/en/discovery/what-is-darkfield-microscopy/>.
- [7] A.S. Holik. “Optical Microscopy”. in *Encyclopedia of Materials: Science and Technology*: by editor K.H. Jürgen Buschow **and others**. Oxford: Elsevier, 2001, **pages** 6458–6463. ISBN: 978-0-08-043152-9. DOI: <https://doi.org/10.1016/B0-08-043152-6/01142-6>. URL: <https://www.sciencedirect.com/science/article/pii/B0080431526011426>.
- [8] *Image Fusion — link.springer.com*. <https://link.springer.com/book/10.1007/978-981-15-4867-3>.
- [9] *Introduction to Fluorescence Microscopy — microscopyu.com*. <https://www.microscopyu.com/techniques/fluorescence/introduction-to-fluorescence-microscopy>.
- [10] Michael J Kidger. “Fundamental optical design”. in *SPIE-International Society for Optical Engineering*: 2001.
- [11] Shutao Li **and others**. “Pixel-level image fusion: A survey of the state of the art”. in *Information Fusion*: 33 (2017), **pages** 100–112. ISSN: 1566-2535. DOI: <https://doi.org/10.1016/j.inffus.2016.05.004>. URL: <https://www.sciencedirect.com/science/article/pii/S1566253516300458>.

- [12] *Multi-focus image fusion techniques: a survey - Artificial Intelligence Review* — *link.springer.com*. <https://link.springer.com/article/10.1007/s10462-021-09961-7#citeas>.
- [13] *Optical microscope - Wikipedia* — *en.wikipedia.org*. [https://en.wikipedia.org/wiki/Optical\\_microscope](https://en.wikipedia.org/wiki/Optical_microscope). [Accessed 19-Jul-2023].
- [14] John Oreopoulos, Richard Berman **and** Mark Browne. “Chapter 9 - Spinning-disk confocal microscopy: present technology and future trends”. **in** *Quantitative Imaging in Cell Biology*: **by** editor Jennifer C. Waters **and** Torsten Wittman. **volume** 123. *Methods in Cell Biology*. Academic Press, 2014, **pages** 153–175. DOI: <https://doi.org/10.1016/B978-0-12-420138-5.00009-4>. URL: <https://www.sciencedirect.com/science/article/pii/B9780124201385000094>.
- [15] David W Piston. “Choosing objective lenses: the importance of numerical aperture and magnification in digital optical microscopy”. **in** *The Biological Bulletin*: 195.1 (1998), **pages** 1–4.
- [16] *Pixel- and Feature-Level Image Fusion Concepts and Algorithms — 19 — M* — *taylorfrancis.com*. <https://www.taylorfrancis.com/chapters/mono/10.1201/9781439800058-19/pixel-feature-level-image-fusion-concepts-algorithms-jitendra-raol>.
- [17] *Point Spread Function (PSF) — Scientific Volume Imaging — svi.nl*. [https://svi.nl/Point-Spread-Function-\(PSF\)](https://svi.nl/Point-Spread-Function-(PSF)).
- [18] Kurt Rossmann. “Point spread-function, line spread-function, and modulation transfer function: tools for the study of imaging systems”. **in** *Radiology*: 93.2 (1969), **pages** 257–272.
- [19] José Sasián. *Introduction to lens design*. Cambridge University Press, 2019.
- [20] *Spatial Resolution in Digital Imaging — microscopyu.com*. <https://www.microscopyu.com/tutorials/spatial-resolution-in-digital-imaging>.
- [21] *The Diffraction Barrier in Optical Microscopy — microscopyu.com*. <https://www.microscopyu.com/techniques/super-resolution/the-diffraction-barrier-in-optical-microscopy>.
- [22] Jiwei Wang **and** others. “Multi-focus image fusion based on quad-tree decomposition and edge-weighted focus measure”. **in** *Signal Processing*: 198 (2022), **page** 108590. ISSN: 0165-1684. DOI: <https://doi.org/10.1016/j.sigpro.2022.108590>. URL: <https://www.sciencedirect.com/science/article/pii/S0165168422001311>.
- [23] *What is a Compound Microscope? — Microscope World Blog — microscope-world.com*. <https://www.microscopeworld.com/p-3470-what-is-a-compound-microscope.aspx>.