Restoration Methods For Biomedical Images In Confocal Microscopy

von
Master of Engineering
Nabil Aly Mohamed Aly Lashin

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
– Dr.-Ing. –

genehmigte Dissertation

Berlin 2005
D 83

Restoration Methods For Biomedical Images In Confocal Microscopy

von
Master of Engineering
Nabil Aly Mohamed Aly Lashin

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
– Dr.-Ing. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr.-Ing. Stefan Jänichen
Gutachter:	Prof. Dr. rar. nat. Klaus Obermayer
Gutachter:	Prof. Dr.-Ing. Olaf Hellwich

Tag der wissenschaftlichen Aussprache: 12 July 2005

Berlin 2005
D 83

To my parents

Acknowledgements

This work was carried out in the Neuronal Information Processing Group of Prof. Klauss Obermayer, at the Technical University of Berlin. I am grateful to all the group members for their help during different stages of this work.

Zusammenfassung

Diese Doktorarbeit stellt neue Lösungen zum Problem Bildrestauration im biomedizinischen Bereich vor. Das Konfokal-Mikroskop ist eine verhältnismäßig neue Bildungstechnik, die als Standardwerkzeug in biomedizinischen Studien eingesetzt wird. Diese Technik dient zum Sammeln einer Reihe von 2D Bildern der einzelnen Abschnitte innerhalb eines Probestücks, um eine 3D Darstellung des Gegenstandes zu erzeugen. Trotz seiner verbesserten Belichtungseigenschaften unterliegen die beobachteten Bilder Störungen augrund der begrenzten Grösse der Punktantwort (PSF) und das Poisson-Rauschens. Bildrestaurationstechniken versuchen diese Störungen herauszurechnen und das Originalbild zu rekonstruieren. Diese Doktorarbeit beginnt mit der Beschreibung des Konfokal-Mikroskops und den Quellen von Artefakten. Dann werden die vorhandenen Bildwiederherstellungsmethoden vorgestellt und verglichen. Die Arbeit ist in drei Teile gegliedert: Im ersten Teil wird eine neue begrenzte blinde Dekonvolutionsmethode eingeführt. Durch eine passende Re-Parametrisierung wird dabei a priori Wissen eingebaut. Für die PSF wird ein parametrisches Modell, mit einem begrenzten Satz von Basisunktionen benutzt, um Nicht-Negativität, zirkulaäre Symmetrie und Limitierung der Frequenzbandbreite sicher zu stellen. Für das Bild stellt die quadratische Re-Parametrisierung die Nicht-Negativität sicher. Die Entfaltungsmethode wird anhand von simulierten und realen Konfokal-Mikroskopie Daten ausgewertet. Der Vergleich mit einem nicht-parametrisierten Algorithmus zeigt, dass die vorgeschlagene Methode verbesserte Leistung und schnellere Konvergenz erreicht. Im zweiten Teil der Arbeit wird eine neue Methode eingeführt, die versucht die anisotrope tiefabhängige Unschärfe zu beheben. Wenn röhrenförmige Gegenstände -wie Neuronen- abgebildet werden, sind die aufgenommenen Bilder degradiert und die Extraktion der genauen Morphologie der Neuronen wird erschwert. Es wird eine neue Methode vorgeschlagen, mit der sich die PSF ohne irgendein Vorwissen über das Belichtungssystem aus dem augenommenen Bild schätzen läßt. Diese Methode, die auf der Schätzung des ursprünglichen Gegenstandes basiert ist für Fälle verwendbar, in denen der abgebildete Gegenstand eine bekannte Geometrie hat. Mit der vorgeschlagenen Dekonvolutionsmethode werden geometrische Verzerrungen beseitigt und die wiederhergestellten Bilder sind für weitere Analysen besser verwendbar. Im dritten Teil wird eine neue Methode zur adaptiven

Regularisierung vorgeschlagen. Diese vorgeschlagene Technik passt ihr Verhalten abhängig von den lokalen Intensitätsgradient im Bild an. Die neue Technik wird getestet und mit der “total variation” und der Tikhonov Regularisierungtechnik verglichen. Die Experimente zeigen, dass mit dem adaptiven Verfahren, die Qualität der rekonstruierten Bilder verbessert wird.

Abstract

This thesis introduces new solutions to the problem of image restoration in biomedical fields. The confocal microscope is a relatively new imaging technique that is emerging as a standard tool in biomedical studies. This technique is capable of collecting a series of 2D images of single sections inside a specimen to form a 3D image of the object. Moreover, the use of laser light increases the resolving capabilities of the microscope. Despite of its improved imaging properties, the observed images are blurred due to the finite size of the the point spread function and corrupted by Poisson noise due to the counting nature of image detection. Image restoration aims at reversing the degradation and recovering an estimate of the true image.

This thesis starts with the description of the confocal microscope and the sources of degradation. Then, the existing image restoration methods are studied and compared. The work done in this thesis is divided into three parts:

In the first part, a new constrained blind deconvolution method is introduced. In this method, re-parameterization is used to strictly enforce apriori knowledge. For the PSF, a parametric model based on a set of constrained basis functions is used. This re-parameterization ensures circular symmetry, and band-limitedness. For the image, quadratic re-parameterization ensures non-negativity. The deconvolution method is evaluated on both simulated and real confocal microscopy data sets. The comparison with a non-parameterized algorithm shows that the proposed method exhibits improved performance and faster convergence.

In the second part, a new method to correct the effect of anisotropic, depth-variant blur is introduced. When objects of tubular-like structure, like neurons, are imaged, the acquired images are degraded and the extraction of accurate morphology of neurons is hampered due to these anisotropic deformations. A new method to estimate the PSF from the acquired image without any prior knowledge about the imaging system is proposed. This method which is based on the estimation of the original object and is suitable for cases in which, the object being imaged has a known geometry. Using the proposed deconvolution method, geometric distortions are eliminated and the restored images are more suitable for further analysis.

In the third part, a new method for adaptive regularization is proposed. The proposed technique adapts its behavior depending on the local activities in the

image, as reflected in the magnitude of the intensity gradient. The new technique is tested and compared to both the total variation and the Tikhonov regularization techniques. Experiments show that, using the adaptive technique, the quality of the restored images is improved.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Confocal Microscopy

Fluorescence imaging techniques are widely used in Biomedical sciences to observe microscopic specimens. The structure of interest within the specimen is injected by a fluorescence dye which emits photons when excited by a specific wavelength of light. The conventional wide field microscopy produces images that are blurry due to the contribution of light from above and below the focal plane. Moreover, using this technique it is not possible to observe single sections inside a thick specimen without cutting the specimen into thin slices. Confocal microscopy is a relatively new imaging technique that is becoming more and more a standard tool, widely used in biomedical sciences. In confocal microscopy, a highly focused beam of light is used to illuminate only the point at focus inside the specimen. By scanning the focal point in the lateral direction (XY) a 2D image of only one section located at focus is collected. By moving the focus step by step along the optical axis (Z) and repeating the lateral scanning, a series of 2D images are collected. These 2D images are combined to construct a 3D image of the whole specimen. The ability to acquire images of only one slice (optical section) of a whole specimen is known as optical sectioning. This non-invasive 3D imaging capability is fundamental to many studies in biological and medical fields. Moreover, the images recorded by the confocal microscope exhibit higher resolution and are less blurred in comparison to images acquired by the conventional wide field microscope. This is mainly due the significant rejection of out of focus light, using a pinhole in front of the (PMT)

detector. Despite of these advantages of the confocal microscope over conventional imaging techniques, still some imperfections contribute to degrade the acquired images. The imaging properties of the microscope are characterized by its point spread function (PSF). The PSF is the image produced by the microscope, of an infinitesimally small point source of light. Ideally, this image should be a point as well. However, due to diffraction and aberrations, the point is smeared to a spot of finite size. Due to the finite size of the PSF, confocal images are blurred. Another source of degradation is the noise that arises during the image detection process. For the correct interpretation and analysis of confocal microscopy images, it essential to remove this degradation. In order to do so, the recorded images need to be further processed using image restoration techniques.

1.2 Image Restoration

The acquired image is a degraded, blurred and noisy, version of the true image. Image restoration techniques aim at reversing the degradation undergone by the image to recover the true image. A model of the degradation is essential to these algorithms, specifically, knowledge of the PSF and noise characteristics. The problem of image restoration has been extensively studied and several algorithms are developed to solve this problem. Classical linear techniques, restore the true image by filtering the observed image using a properly designed filter. Examples are inverse filtering, Wiener filtering and Tikhonov-Miller algorithm. The naive direct inverse filtering leads to noise amplification in the restored image. Wiener and Tikhonov-Miller filtering techniques avoid the problems of direct inversion using a stabilized filter to produce relatively better restorations. The main drawback of these techniques is that negative intensities occur in the restored image. Since negative intensities have no physical interpretation, the restored image must be constrained to be non-negative. The problem of image restoration is ill-posed in the sense that it admits infinitely many solutions. More prior knowledge has to be considered in order to limit the domain of admissible solutions. The prior knowledge is usually in the form of constraints such as: non-negativity, limited support and band-limited-ness. The iterative methods for image restoration, provide a way to impose the constraints on the restored image after each iteration to

avoid convergence to an infeasible solution. An example of maximum likelihood (ML) algorithms is the Richardson-Lucy (RL) algorithm which is widely used for the restoration of astronomical image. The RL algorithm is constrained but unregularized and has to be stopped before complete convergence. Regularization can be incorporated into the formulation of the ML methods, using some penalty functional that imposes prior assumptions about the true image. This technique is known as penalized ML. The classical Tikhonov form of regularization has been used in several restoration algorithms [51]. Good's roughness penalty has been used in [45]. Recently, total variation has been successfully used for image restoration [4]. It should be noted that the specific form of the regularization functional imposes assumptions on the restored image, and hence should be carefully selected. In the Bayesian framework, the prior probability density is formulated to reflect knowledge about the true image. This prior probability density is then modified, using Bayes rule into a posterior probability density. The maximum *a posteriori* methods (MAP) estimate the true image as that image corresponding to the mode of the posterior probability density.

1.3 Blind deconvolution

Most image restoration algorithms assume that the PSF is known. In many situations this is not the case. In confocal microscope, a PSF can either be calculated or measured. However, during the image acquisition the light passes through the specimen and the PSF is affected by the optical characteristics of the specimen. Especially in the case of imaging inside deep structure, or when the optical properties of the specimen (refractive index) are significantly inhomogeneous, the variation of the PSF is so great that each specimen is imaged with a totally different PSF. When only partial information about the PSF is known, blind deconvolution algorithms are used to estimate both the true image and the PSF from the recorded image.

1.4 Scope of The Thesis

This thesis is concerned with the problem of image restoration. New restoration methods are developed to improve over the existing techniques. The main topics tackled in this thesis can be outlined as

- Constrained blind deconvolution algorithm.
  In this algorithm re-parameterization is used to achieve two goals, a parametric model of the PSF is used to reduce the number of parameters to be estimated and hence improve the rate of convergence. In the design of the parametric model, the general imaging constraints like: non-negativity, band-limitedness, are explicitly taken into account in order to avoid infeasible estimates. The performance of the proposed algorithm is evaluated using both synthetic and confocal images.

- Restoration of neurons 3D images.
  A new method to extract the PSF form the observed image is proposed. Based on the assumption that, the cross-sectional shape of neurons is, on the average, circular, the PSF can be estimated from the inspection of the geometric distortion of the cross-sectional shape in the recorded image. The estimated PSF shows dependence on the depth inside the specimen. A fast implementation of a depth-variant restoration algorithm is presented.

- Adaptive regularization.
  Based on an analysis of the the widely used regularization functionals (Tikhonov and total variation), a new adaptive regularization technique is proposed. Using the proposed technique, the quality of the restored images is improved. The proposed technique is tested using both artificial and confocal images.

1.5 Organization of The Thesis

- Chapter 2
  This chapter is meant to be an introduction to the confocal microscopy. In this chapter the design of the microscope and its components are discussed.

Limitations and sources of degradations are outlined. Finally, the chapter ends by introducing example confocal data sets that are used in this study.

- Chapter 3

  This chapter is devoted to a review of restoration techniques. In this chapter, the advantages and shortcomings of linear restoration methods and Maximum likelihood algorithms are studied. Finally the Bayesian approach to the image restoration problem, with different prior models is introduced. Simulations are performed to compare different algorithms.

- Chapter 4

  In this chapter a novel approach to strictly impose constraints on the restoration problem is introduced. The constraints on the PSF are implemented using a parametric model. A new constrained parametric blind deconvolution algorithm is introduced. Simulations are performed to evaluate the performance of the proposed algorithm and compare it to the existing algorithms.

- Chapter 5

  In this chapter a new method for the restoration of confocal microscopy images of neurons is presented. We start by studying the effect of the refractive index mismatch on the PSF. Following this, a method to estiamte the PSF from the images of objects of known geometry, like neurons, is introduced. Finally, a computationally efficient depth-variant restoration algorithm is developed and tested on confocal images. The chapter ends with a discussion of the introduced methods.

- Chapter 6

  This chapter investigates the effect of the regularization functional on the restored image. Two regularization functionals (TV and Tikhonov) that are widely used in image restoration algorithms are analyzed. Based on this analysis, a new adaptive regularization functional is suggested. The adaptive regularization avoids the undesirable features of existing approaches while retaining their desirable features. Simulations are performed to evaluate the proposed technique and compare it to the existing approaches.

- Chapter 7

In this chapter the overall conclusions of this work are outlined together with a discussion of possible future research directions.

Chapter 2

Confocal Microscope

Introduction

In conventional optical microscopy the whole specimen is uniformly illuminated, and the light emitted is collected by the objective lens to form the image. This setup hampers the acquisition of 3D images of reasonable quality, as the light emitted from the whole volume of the specimen contributes to the recorded image. When focusing at a certain plane inside the specimen, the light emitted from both the in-focus and out-of-focus planes are collected leading to a loss of resolution. To avoid this problem thick specimens need to be physically cut into several thin specimens that are imaged separately. Afterwards, the acquired images are combined into a single image stack.

The confocal microscope is designed to improve on the performance of the conventional optical microscope. In confocal microscopy, a highly focused cone of light illuminates the specimen. The out-of-focus points within the specimen are much less illuminated. The light emitted is then collected and refocused again by the objective lens on an aperture, the pinhole, in front of the detector (see Fig.2.1). The light emanating from out-of-focus points is mostly rejected by the pinhole, thus not reaching the detector. This ability to discriminate between in-focus and out-of-focus emitted light enables the use of confocal microscope to acquire images of thin slices within the specimen volume. Moreover the resolution of the recorded images is much improved due to the rejection of out-of-focus light. Two dimensional images are produces by scanning the in-focus plane laterally.

A 3D image stack is then collected by focusing into different planes within the specimen. This optical slicing capability saves several man-hours of work and is free of the artifacts that occur during the physical slicing of specimens. However, the thickness of specimen is limited by several factors like, aberrations, intensity attenuation, and bleaching of the dye.

2.1 Principle of Operation

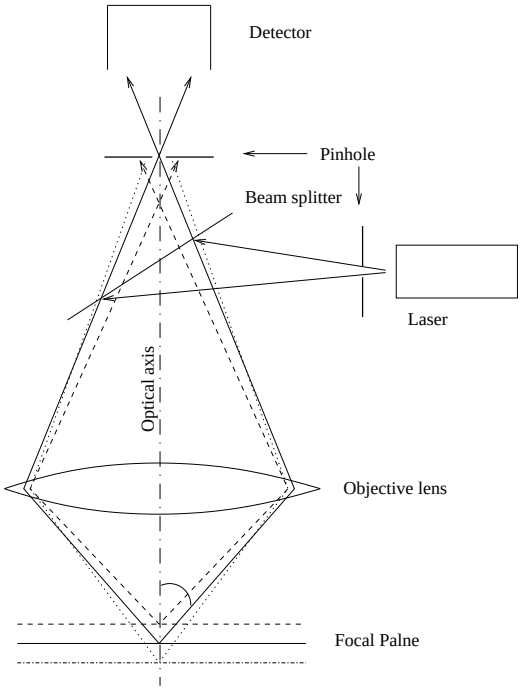


Figure 2.1: Design of the Confocal Microscope.

The basic configuration of confocal microscope is shown in Fig.2.1. A laser beam is reflected by a dichronic mirror and then focused by the objective lens into the focal point. This laser excites the fluorescent dye in the specimen which in turn emits photons at a lower energy level and longer wavelength. The emission light is then collected by the objective and focused into the detector. Before the emission reach the the detector it passes through a pinhole which allows only the in-focus light to reach the detector.

2.2 Optical Components

The optical resolution of a microscope is a measure of its resolving capability. It is defined as the smallest distance between two points that are still distinguished by the microscope as separate points. The optical resolution depends on the objective lens and the wavelength of the laser, in the lateral plane, normal to the optical axis it is calculated as,

$$\text{Lateral resolution} = \frac{0.61NA}{\lambda}. \tag{2.1}$$

Where, λ is the wavelength of the excitation light and NA is the numerical aperture of the objective lens. This optical resolution defines the minimum length of the specimen that can be resolved using the microscope. This can be increased using light with small wavelength or using an objective lens with a higher numerical aperture.

The resolution in the axial direction, along the optical path of the microscope, is much worse than in the lateral direction. For an aberration-free, diffraction-limited optical system, the axial resolution depends on the NA of the objective, wavelength, and the immersion medium of the objective lens,

$$\text{Axial resolution} = \frac{0.88n\lambda}{(n - \sqrt{n^2 - NA^2})}. \tag{2.2}$$

Where, λ is the wavelength of the excitation light, n is the refractive index of the immersion medium, and NA is the numerical aperture of the objective lens.

2.2.1 Objective Lens

A lens is characterized by magnification and *numerical aperture* (NA). Magnification by itself does not determine the resolution, the NA must be known. The NA is a measure of light-gathering capabilities of the lens and depends on the *refractive index* (RI) of the immersion medium of the lens. The NA is calculated as, $NA = n \sin \alpha$, where n is the RI of the immersion medium and α is the half angle of acceptance of the lens (see Fig.2.2). Immersion media used in practice are, water, glycerin, and oil. Although using an objective lens with higher magnification increases the size of the image formed in front of the detector, the overall resolving capability of the microscope is limited by the NA and the wavelength of light used to form the image. Objective lenses with high NA generally have

large magnification factors, 60×-100×, but small *free working distance*. The free working distance, which is the distance between the lens and the cover-slip, is an important factor as it determines the maximum thickness of the specimen that can be imaged using the lens before the lens hits the cover-slip.

Another possibility to increase the number of pixels of an image is through an increase of the sampling frequency, known as *electronic zoom*. However, this is again limited by the minimum resolvable distance determined by the NA of the lens and the wavelength of light.

An important parameter that affects the optical sectioning capabilities of the microscope and hence the ability to acquire images of 3D structures is the *Depth of Field* DOF (see Fig.2.2). The DOF is an indication of the distance above and below the focal point where light is still being collected and is calculated as,

$$\text{DOF} = \pm\frac{n\lambda}{2NA^2}. \tag{2.3}$$

The smaller the DOF, the less suitable a lens is to image specimens that extend in the axial direction.

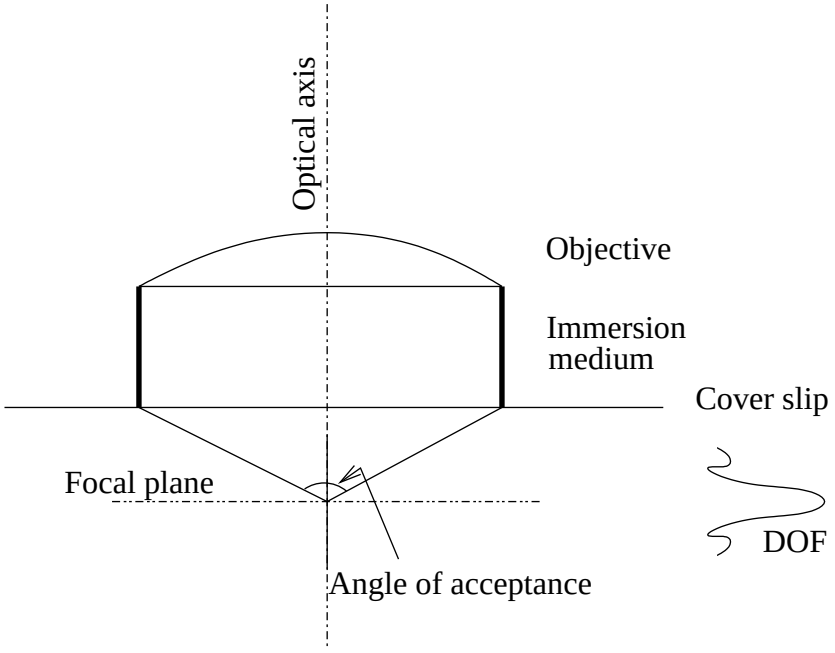


Figure 2.2: Depth of field of an objective lens.

2.2.2 Laser Illumination

The wavelengths of the detection and illumination laser are selected according to the type of the fluorochrome used to stain the specimen. Each type of fluorochrome is excitable by a specific band of wavelengths, and emits light at a different range of wavelengths. Excitation and emission wavelengths are selected to be near the peak of the response curve, excitation and emission, of the fluorochrome to improve efficiency. The power of the laser light should be as small as possible to avoid detrimental effects in the specimen and bleaching of the fluorochrome.

2.2.3 Pinhole

The adjustment of the size of the pinhole is very critical to the quality of the acquired image. Increasing the size of the pinhole, causes more out-of-focus light to be collected and the acquired image is blurry. Decreasing the size of the pinhole, causes less in-focus light to be collected and the acquired image is dark. The size of the pinhole is normally set to be equal to the radius of the Airy disc.

2.3 Point Spread Function

The microscope image of an ideal point source of light is known as the *point spread function* (PSF). Ideally the PSF should be a punctal spot as well, however, due to diffraction and aberration this is never the case. The PSF describes the transfer properties of the optical system. The Fourier transform of the PSF is known as the *optical transfer function* (OTF). For an aberration-free optical system the lateral variation of the PSF has the form of an Airy pattern with its peak at the center and several side lobes,

$$h(r) = \left(\frac{2J_1(\frac{2\pi r NA}{\lambda})}{\frac{2\pi r NA}{\lambda}} \right)^2 . \tag{2.4}$$

Where, r is the distance from the center of the PSF, J_1 is a Bessel function of the first kind of order 1, and λ is the wavelength of the laser light. The first minimum of the $J_1(\zeta)$ function occurs at $\zeta = 3.83$, therefore the first minimum of the Airy pattern is at $r = \frac{0.61\lambda}{NA}$. This explains the value of the optical resolution defined in Eq.2.1. This value of resolution is based on the *Rayleigh's Criterion*, which is

one of the most important principles in microscopy. According to this criterion, two objects are distinguishable, i.e. can be resolved as separate objects, if they are separated by a distance which is not smaller the radius of the central peak of the Airy pattern. The spectrum of visible light ranges from 400 nm for blue light to 700 nm for red light. This means that, using visible light, the maximum achievable resolution is about 200 nm for $NA = 1$. In practice, the optical components of the microscope are not ideal. Several effects, like aberrations, light scattering, focus errors and light absorption worsen the resolution of the microscope and cause the PSF to deviate from its ideal form. In the following section, some sources of distortion are discussed.

2.4 Sources of Distortion

Although the confocal microscope is much better than the conventional wide field microscope and several problems have been solved to improve its performance over that of the conventional microscope, still some effects contribute to the distortion of the acquired image. Some of these effects are outlined in the following.

- Photon noise. This is mainly due to the quantum nature of light. Photons are emitted by the fluorescent dye, collected by the photon multiplier tube (PMT) at randomly distributed time intervals. The intensity of each pixel in the observed image is proportional to the number of photons counted during the time interval corresponding to each pixel, (exposure time). If A represents the mean of the number of photons counted at all the image pixels, then the probability of counting exactly n photons during the time interval corresponding to one pixel is Poisson distributed [2] with

  $$\Pr(n) = \frac{A^n}{n!} exp(-A).$$

  This uncertainty of the number of photons manifests itself in the form of Poisson noise in the observed image. This noise is correlated with the image intensity.

  This noise may be reduced by either increasing the exposure time or increasing the number of emitted photons. However, the exposure time is limited

by the size of the image field and is generally, very small. For example, for a 512 × 378 pixel raster scanned in 1 Sec, the exposure time for one pixel is about 1.6×10^{-12} sec [33]. On the other hand, to avoid bleaching of the dye the number of photons to which the specimen is exposed must be kept as small as possible and this limits the number of emitted photons.

Hence, this type of noise is very hard to avoid and will always be present. It is regarded as an intrinsic [33] noise in the imaging process that cannot be eliminated.

- Dark signal noise. This is due to the generation of electrons in the PMT irrespective of whether it is irradiated or not. This noise can be reduced by cooling the PMT. This noise follows Poisson statistics as well, but is not correlated to the image intensity. This noise adds a background to the observed image.

- Intensity attenuation with depth. Because both the excitation and the emission light pass through the specimen, the images of planes located deeper inside the specimen have lower intensity than those of the planes located above.

- Refractive index mismatch. This is due to the presence of media with different refractive indices along the optical path. The refractive index of the immersion medium is different from the refractive index of the tissue of the specimen or the aqueous medium in which the specimen is mounted. This refractive mismatch makes the PSF depth-variant.

- Spherical aberration. Because the surface of the lens is spherical, light rays near the optical axis, paraxial, and peripheral rays are focused at different planes. This aberration increases with increasing depth inside the specimen. Although it is possible to correct this aberration, using collar rings, this correction is valid only at one depth.

- Chromatic aberration. This arises due to the dispersion of light as it passes through glass because the RI of glass is different for different wavelengths. This causes the lens to focus different wavelengths on different planes. How-

ever, high quality objective lenses, achromats, are corrected to reduce this type of aberration.

- Non-uniform dye distribution. Since the image recorded by the microscope is the image of light emitted by the fluorescence dye, any discontinuities or gaps of dye distribution in the structure of interest is reflected in the recorded image as intermittent structure which is hard to interpret and leads to difficulties in image analysis.

2.5 Preparation of The Specimen

To prepare a specimen for imaging, the structure of interest within the specimen is labeled with some fluorochrome dye. The specimen has to be chemically processed to reduce its undesirable optical characteristics to allow for deeper light penetration. There are several histochemical protocols that were developed and widely used.

In this study two examples of confocal microscope data set are used. The first one is a 3D image stack of a single *Drosophila Melanogaster* optic lobe and the second is a 3D image stack of the inter-neuron A4I1 of the migratory locust (*Locusta migratoria migratorioides*). A brief summary of specimen preparation is given in the sequel.

2.5.1 Drosophila Optic Lobe

A summary of specimen preparation is given here, for complete protocol the reader is referred to [19].

Immunohistochemistry

Fly brains were prepared in Ringer's solution and then immediately fixed in paraformaldehyde (15-45 minutes). The brains were subsequently put into a blocking solution for 30 minutes and incubated with primary antibody for 6-12 hours at 4°C. Several dilutions of antibodies were used. Cy3 and DTAF labeling (Jackson ImmunoResearch) were used as the secondary antibodies. The preparations were embedded in Vectashield (Vector Laboratories, Burlingame, CA).

Imaging Setup

A Leica (Nussloch, Germany) TCS4D confocal microscope equipped with an ArKr laser was used for data acquisition. The excitation and emission wavelengths are 550 and 570 nm, respectively. Series of complete optic lobes were scanned with a 40× objective (NA=1.0) and comprised 182 images of 512×512 pixel resolution at 8 bit color depth.

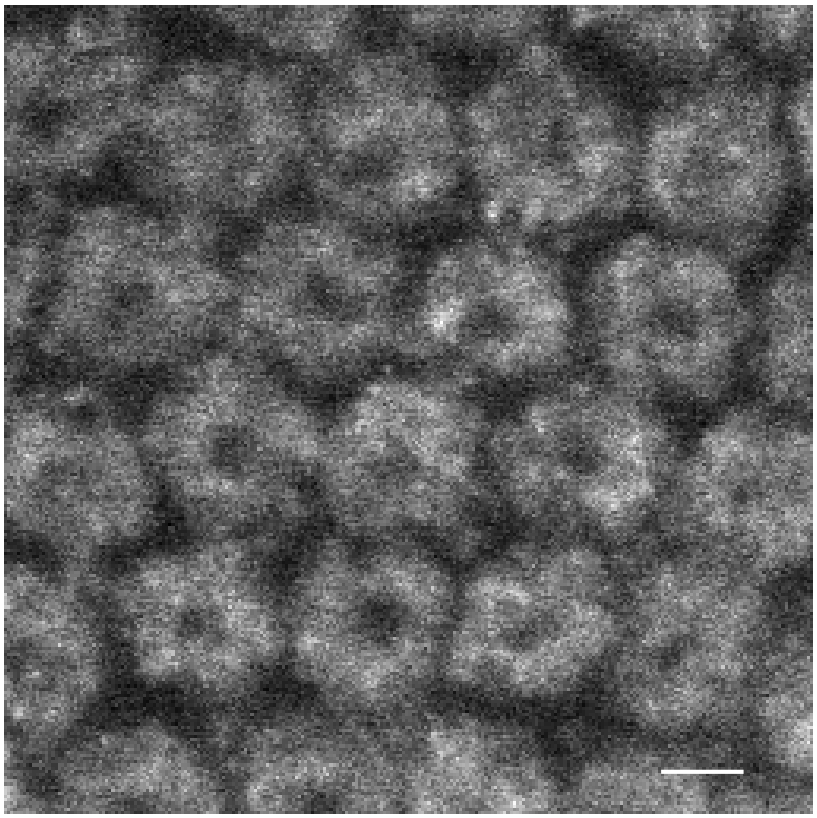


Figure 2.3: An XY section of the confocal microscopy image of the optic lobe of Drosophila. Scale bar = 4 μ. The image stack was scanned using an 40×/1.0 objective lens. The whole stack consists of 182 images with 512×512 voxels.

2.5.2 Locust Inter-neuron A4I1

A summary of the specimen preparation is given here, for the complete protocol the reader is referred to [3].

Immunohistochemistry

Electrodes were inserted into the cell body of the A4I1 neuron. Neurobiotin was injected iontophoretically by de-polarizing current pulses. Axonal diffusion was allowed for 12 to 24 hours, then specimens were fixed in 4% neutral Formalin for 2 to 3 hours, washed and lipids removed in an ethanol series. The specimens were placed in Streptavidin Cy3 (Jackson Immuno Research) overnight at 4°C, dehydrated in an ethanol series and cleared in methyl salicylate.

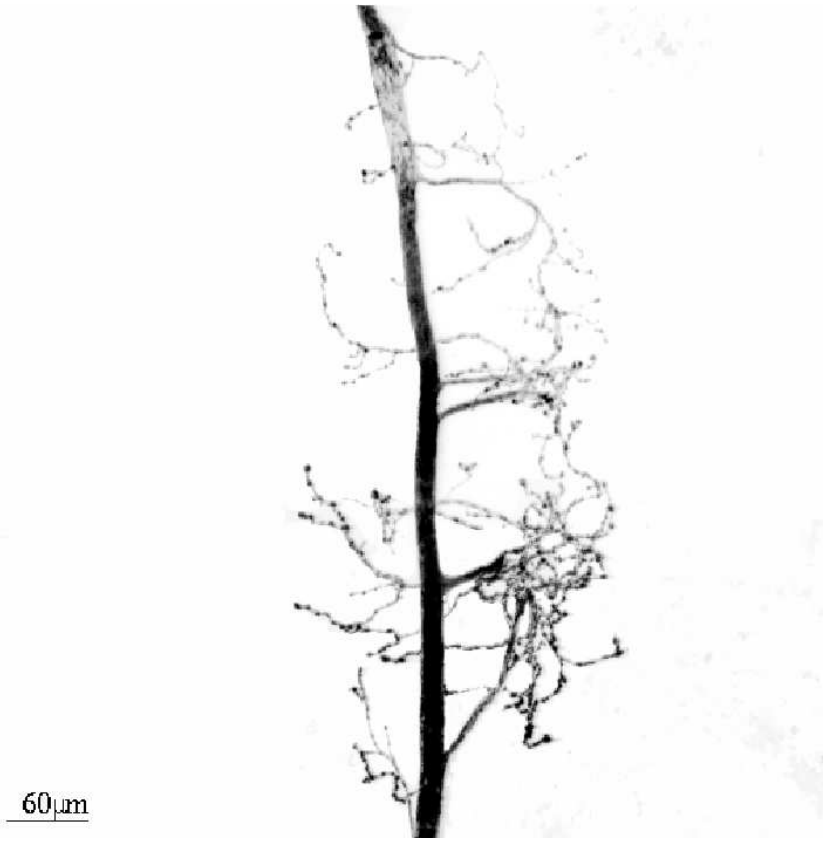


Figure 2.4: Maximum intensity projection of the confocal microscopy image of the mesothoracic part of a CY3 stained A4I1 inter-neuron of Drosophila. The image stack is acquired using a 20×/0.6 objective, and consists of of 135 images with 512×512× voxels.

Imaging Setup

The preparation was imaged with a confocal laser scanning microscope (Leica TCS4D). An 20×0.6 air objective lens was used, which due to its small field of

view enables the scanning of the preparation as a whole. The fluorochrome is excited with 568 nm ArKr laser and the emission signal was detected using an LP590 long-pass filter. The data set is (512×512×135) voxels in size, with a voxel size of 0.98μm in all directions.

Chapter 3

Review of Deconvolution Algorithms

3.1 Image Restoration

The goal of image restoration is to recover an image $\hat{f}$, that is an estimate of the true image f that represents the intensity fucation of the object being imaged. In general, the acquired image is a degraded (blurred), noisy version of the true image, and can be modeled as,

$$g = \mathcal{N}(h \otimes f) = \mathcal{N}(Hf). \tag{3.1}$$

Where $\mathcal{N}$ denotes a noise process, $\otimes$ denotes the convolution, f denotes the true image, and g is the observed image. In confocal microscopy the noise is mainly due to the photon-counting nature of image detection and is known to follow Poisson statistics. Poisson noise is correlated to the degraded image Hf.

Most classical restoration algorithms are developed with the assumption that the noise statistics follow Gaussian distribution. This assumption is not valid in the case of photon-limited imaging, where the noise is mainly due to the photon-counting nature of the image detection process. However, for images with a relatively high SNR and small dynamic range Poisson noise can be reasonably approximated by additive Gauussian noise [39]. In this case the standard deviation of the Gaussian noise is not constant, rather proprotional to the intensity of the image. If the dynamic range of intensity is small the standard deviation can be assumed constant.

If the assumption of Gaussian noise is valid, the model in Eq.3.1, reduces to

$$g = Hf + n. \tag{3.2}$$

Where n, denotes additive white noise of zero mean and known variance. This model of additive white Gaussian noise spans various imaging techniques, where the noise is independent of the intensity of the image. Due to its simplicity, several restoration algorithms are based on this model.

3.2 Linear Methods

These methods are simple and produce fast (possibly crude) solutions. Assuming that the PSF h is known, direct methods seek an estimate of the true image $\hat{f}$ by filtering the acquired image g with some appropriately designed filter W,

$$\hat{f} = Wg. \tag{3.3}$$

The filter W is selected to optimize some criterion (or minimize an error norm). In the following, we review some of the classical linear methods.

3.2.1 Least Squares Filters

The classical direct approach to solve the restoration problem is to find $\hat{f}$ that minimizes the norm

$$||g - H\hat{f}||^2. \tag{3.4}$$

This represents the least squares fit to the data. The linear filter that minimizes Eq.3.4 is given by,

$$W = (HH^T)^{-1}H^T, \tag{3.5}$$

which is the generalized inverse of H. Because the spectral properties of the noise are not considered, the critical issue of this method is noise amplification. This becomes more clear in the frequency domain. In the Fourier domain, the solution can be written as,

$$\hat{\mathcal{F}} = \frac{\mathcal{G}}{\mathcal{H}} = \frac{\mathcal{H}^*\mathcal{G}}{|\mathcal{H}|^2}. \tag{3.6}$$

Where, * denotes the complex conjugate, $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$ are the Fourier transforms of the true image, the observed image and the PSF, respectively. Because the PSF $h(X)$ is usually a low pass filter (LPF), i.e. $\mathcal{H}(\omega) \approx 0$ at high frequencies, the division by $|\mathcal{H}|$ in Eq.3.6 leads to large amplifications at high frequencies, where the noise dominates over the image. This frequency dependent amplification leads to significant errors in the restored image, and amplification of noise. To avoid these problems, a workaround is to use a *pseudo* inverse of $\mathcal{H}$ defined as,

$$\frac{1}{\mathcal{H}} = \begin{cases} 1/\mathcal{H} & \text{if } \mathcal{H} > \epsilon \\ \epsilon & \text{if } \mathcal{H} \leq \epsilon \end{cases}$$

However, the value of ϵ affects the restored image. With no clear objective selection of ϵ, restored images are generally noisy and not suitable for further analysis. We conclude that the least squares solution is very sensitive to noise in the observed image. A small perturbation in g might lead to an un-bounded perturbation in the solution $\hat{f}$. Regularization techniques are used to mitigate the ill-posed-ness of the image restoration problem.

3.2.2 Regularized Linear Methods

The problem of image restoration is ill-posed in the sense of Hadamar. A problem is said to be well-posed if:

1. A solution to the problem exists.
2. The solution is unique.
3. The solution depends continuously on the data.

If a problem fails to satisfy one (or more) of these requirements, the problem is said to be ill-posed. Using regularization an ill posed problem can be converted into an associated well-posed problem. The solution to the well-posed problem provides an approximation to the solution of the ill-posed one.

3.2.3 Wiener Filter

The Wiener filter is optimal in the sense of mean square error. The filter is designed by minimizing the MSE between the restored image $\hat{f}$ and the true image f.

$$\min_f E(||f - \hat{f}||^2) \tag{3.7}$$

The linear filter that minimizes Eq.3.7 is given by,

$$W = (HH^T + R_{nn}/R_{ff})^{-1} H^T. \tag{3.8}$$

Where, $R_{ff} = E(ff^T)$ and $R_{nn} = E(nn^T)$, are the covariance matrices of the true image f and the noise n, respectively. Although, it is optimally derived, the success of the Wiener filter depends on the accurate estimation of the image and noise covariance matrices. In practice, one has only the acquired image g. In this case R_{ff} can be approximated by R_{gg}. The noise variance is assumed to be known, in this case $R_{nn} = \sigma^2 I$, where σ^2 is the variance of noise and I is the identity matrix. It should be noted that the Wiener filter Eq.3.8 is derived under the assumption that the noise n is not correlated to the true image f, i.e. $E(fn^T) = 0$. This assumption is not valid in the case of Poisson noise.

In the frequency domain, the Wiener filter is given by,

$$\hat{\mathcal{F}} = \frac{\mathcal{H}^* \mathcal{G}}{\mathcal{H}\mathcal{H}^* + S_{nn}/S_{ff}}. \tag{3.9}$$

Where, S_{ff} and S_{nn} are the power spectra of the true image and noise, respectively. Wiener filter is a regularized inverse filtering technique, compare Eq.3.6 and Eq.3.9, with S_{nn}/S_{ff} as the regularization term. This regularization term is inversely proportional to the SNR, and provides for adaptive filtering so that $W = 1/\mathcal{H}$, when the signal dominates (high SNR) and limits the gain when the noise dominates.

However, because the filtering is performed in the frequency domain, the smoothness of the image in the spatial domain is not considered. Moreover, the restored image generally exhibits ringing artifacts in the vicinity of edges due the attenuation of high frequency components.

3.2.4 Constrained Least Squares Filters

As we have seen above, direct minimization of the mean square error (MSE) between the acquired image g and the re-blurred estimated solution $H\hat{f}$ leads to undesirable noise amplification. Moreover, the problem exhibits infinitely many solutions. Additional constraints are needed to select only one solution.

Instead of the unconstrained minimization of Eq.3.4, a constrained solution is found by solving,

$$\begin{aligned} \hat{f} &= arg \min_f \|g - Hf\|^2 \\ \text{subject to} \quad & \|Cf\|^2 < \text{constant}. \end{aligned} \tag{3.10}$$

Where C, is a filter used to extract features of $\hat{f}$ that should be penalized and is often selected as a high pass filter. The intuition behind this is that, in general, images are smooth with limited high frequency components. Hence, it is appropriate to penalize solutions with large high-pass energy.

Selecting C as the Laplacian filter, leads to the classical Tikhonov-Miller algorithm. The constrained minimization in Eq.3.10, is found by minimizing the functional,

$$||g - H\hat{f}||^2 + \gamma||C\hat{f}||^2. \tag{3.11}$$

Where γ is the regularization coefficient. This functional consists of two terms: the first one, $||g - H\hat{f}||^2$, measures the fidelity of the estimate $\hat{f}$ to the data and the second term, $||C\hat{f}||^2$, measures features (e.g., the roughness, smoothness) of the solution. The balance between those terms is controlled by the regularization coefficient γ. With no additional constraints, a minimizer of the functional in Eq.3.11 is found analytically, with the resulting filter,

$$W = (HH^T + \gamma CC^T)^{-1}H^T \tag{3.12}$$

Note, the similarity between Eq.3.8 and Eq.3.12.

The constrained least squares solution is equivalent to the Wiener solution when $\gamma C^T C = R_{nn}/R_{ff}$. This technique leads, to adaptive filtering depending on the local spatial activities (e.g. smoothness) of the image.

Relation to previous Methods

This filter is similar to the Wiener filter, Eq.(3.9). While Wiener filter performs adaptive filtering according to the SNR, the Tikhonov-Miller filter performs adaptive filtering according to the roughness of the image.

Popular selections of the filter $C^T C$ are the Laplacian or the identity matrix. Using the identity matrix leads to the standard form of Tikhonov regularization which can be regarded as a *pseudo* inverse, or a Wiener filter with a constant SNR.

Constrained Tikhonov-Miller algorithm

To avoid negative intensities in the restored image, Eq.(3.11) must be minimized under the constraint that $f \geq 0$. This constrained minimization leads to the iterative constrained Tikhonov-Miller algorithm (ICTM) [43], that takes a direct approach for the minimization using the conjugate gradient algorithm. The non-negativity constraint is applied after each iteration by setting the negative pixels to zero. An accelerated version of the algorithm was introduced in [47] to improve the convergence of the algorithm. Despite of improvements in performance, the *ad-hoc* application of the constraints after each iteration is suboptimal. The algorithm is shown ([44],[46]) to perform worse in comparison to other algorithms that employ a strict application of constraint.

3.2.5 Simulation

To show the performance of linear methods introduced above simulations are performed using a synthetic image. The synthetic image that represents the true image is blurred with a PSF of the form $h = e^{-X^2/2\sigma^2}$ with $\sigma = 2$, then a constant background intensity of 16 is added to the blurred image. Finally, a Poisson noise is generated, using the image as the mean for the Poisson process and a photon conversion factor of 1. The true image and the observed image are shown in Fig.(3.1). The restored image using inverse filtering is shown in Fig.(3.2). The main drawback of linear methods, besides negative intensities in the restoration, are the amplification of noise for the inverse filtering and ringing artifacts for the Tikhonov-Miller algorithm.

3.2.6 Conclusion

Although linear methods are simple and can be implemented fast, these methods lead to inappropriate estimates of the true image. These methods suffer from a number of drawbacks,

- The inability to incorporate prior knowledge about the true image.
- Negative intensities might occur in the restoration.
- Ringing artifacts are created in the neighborhood of edges.

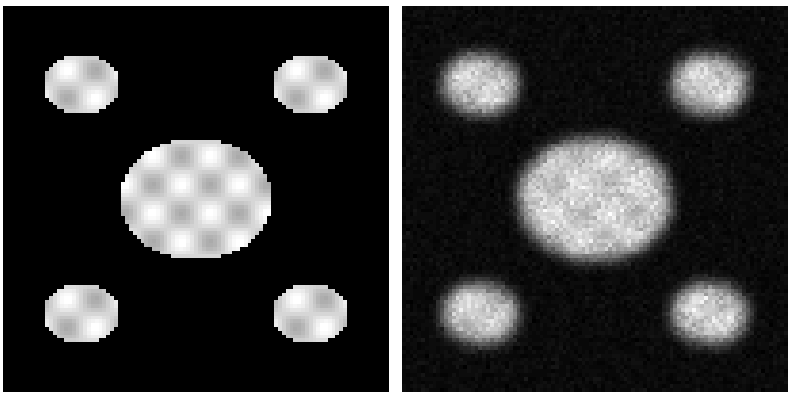


Figure 3.1: Test image used ima the simulation experiments. *Left*: True image. *Right*: Observed image. The true image is blurred with the PSF, then a constant background intensity of 16 is added to the blurred image. Finally, the image is corrupted with Poisson noise, of a photon conversion factor of 1, to create the observed image.

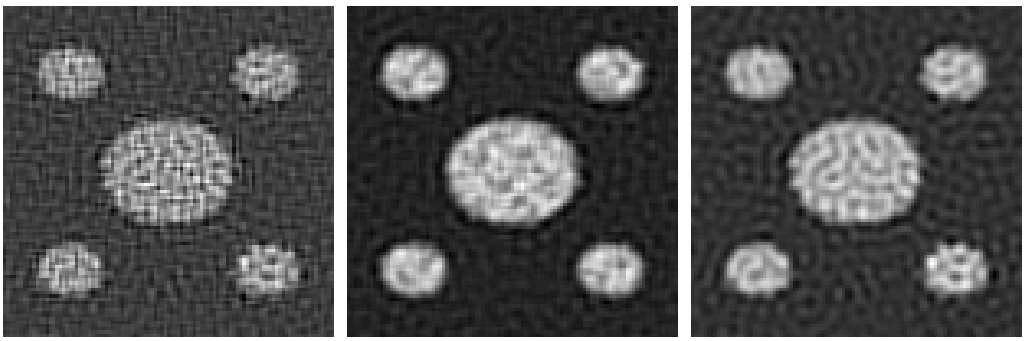


Figure 3.2: Restored images using linear methods. *Left*: Inverse filtering. *Middle*: Wiener filtering. *Right*: Tikhonov-Miller. In the case of inverse filtering the zero coefficients in the Fourier transform of the PSF are set to 0.001 to avoid the division by zero. The results of the Wiener filter are restored the power spectrum of the true image and noise. The regularization coefficient for the Tikhonov-Miller algorithm was 0.003.

Although the regularized linear filtering methods, e.g. Tikhonov-Miller, incorporate some form of prior knowledge using the filter C, the performance of these algorithms needs to be improved.

3.3 Maximum Likelihood Methods

The application of the method of maximum likelihood (ML) estimation to the problem of image restoration is based on knowledge of the random properties of the observed image g. If the PSF is known, the probability density function $\Pr(g|f)$, the likelihood, is function only of the true image f, and the problem of image restoration is to estimate the unknown parameters $f(X), \forall X \in S_f$, where, S_f is the support of f. The ML estimate is the image f which is most likely to give rise to the observed image g. The log-likelihood function is often used for mathematical convenience. The ML solution is found by solving

$$\hat{f} = arg\,\min_f - \log \Pr(g|f,h)$$

In the case of Gaussian noise, the likelihood function is given by,

$$\Pr(g|f) = \frac{1}{(2\pi)^{N/2}\sigma^N} \exp(-||g - Hf||^2/2\sigma^2), \tag{3.13}$$

where, N is the number of voxels in the image and σ^2 is the noise variance. The ML solution is then found by solving,

$$\hat{f} = arg\,\min_f \frac{||g - Hf||^2}{2\sigma^2}. \tag{3.14}$$

This is equivalent to the least squares solution in section (3.2.1). Regularization is needed in order to stabilize the solution. The minimization in Eq.3.14, can be solved using iterative techniques such as the steepest descent method. The solution is found iteratively using the update rule,

$$f^{k+1} = f^k + \eta H^T(g - Hf). \tag{3.15}$$

Where, ^{T} denotes the transpose. This method is usually referred to as the Landweber method [40]. In this case the number of iterations plays the role of regularization, and the algorithm is stopped before complete convergence to avoid noise-over-fitting.

Furthermore, constraints like non-negativity, band-limited constraints can be applied after each iteration of algorithm. This leads to algorithms like projection into convex sets (POC) [22], which we will not consider here.

In the Poisson case

$$\Pr(g|f,h) = \frac{(Hf)^g}{g!}\exp(-Hf). \tag{3.16}$$

The ML solution is found by setting $\frac{\partial \log(\Pr(g|f))}{\partial f}$ to zero, which leads to the Richardson-Lucy algorithm (also known as the expectation maximization maximum likelihood algorithm EM-ML.

Richardson-Lucy Algorithm

A well known algorithm, especially, in the field of astronomy is the Richardson-Lucy (RL) algorithm. Although, the original algorithm was not derived within the Bayesian framework, this algorithm can be regarded as the maximum likelihood algorithm for Poisson noise. This can be shown [26] by the maximization of $\Pr(g|f,h)$ w.r.t. f, which leads to the iterative form of

$$\hat{f}^{(k+1)} = \left(H^T \frac{g}{Hf}\right)\hat{f}^{(k)}. \tag{3.17}$$

Where, k is the iteration number, H^T denotes the transpose of the convolution matrix corresponding to the PSF.

The RL algorithm is constrained but un-regularized. Provided that the initial guess for the true image $\hat{f}^{(0)}$ is non-negative, $\hat{f}^{(k)}$ will remain non-negative. Another important advantage of the algorithm is that, the total intensity is preserved, i.e. $\sum_X \hat{f}(X) = \sum_X g(X)$. The RL algorithm is also related to the maximum aposteriori approach (see subsection (3.5.1)). A major drawback of the ML approach is that it leads to un-regularized algorithms. This leads to solutions that display many artifacts due to noise amplification, or over-fitting. An example convergence characteristics of the RL algorithm are shown in Fig.(3.3). In the beginning of the iterations, MSE (between the estimated image and the true solution) decreases with the number iterations until it reaches a minimum then increases again when noise-over-fitting begins. The algorithm is usually stopped before convergence to avoid noise amplification. This implies using the number of iterations as regularization, a technique known as *truncated-iteration* [31]. Another possibility is to use

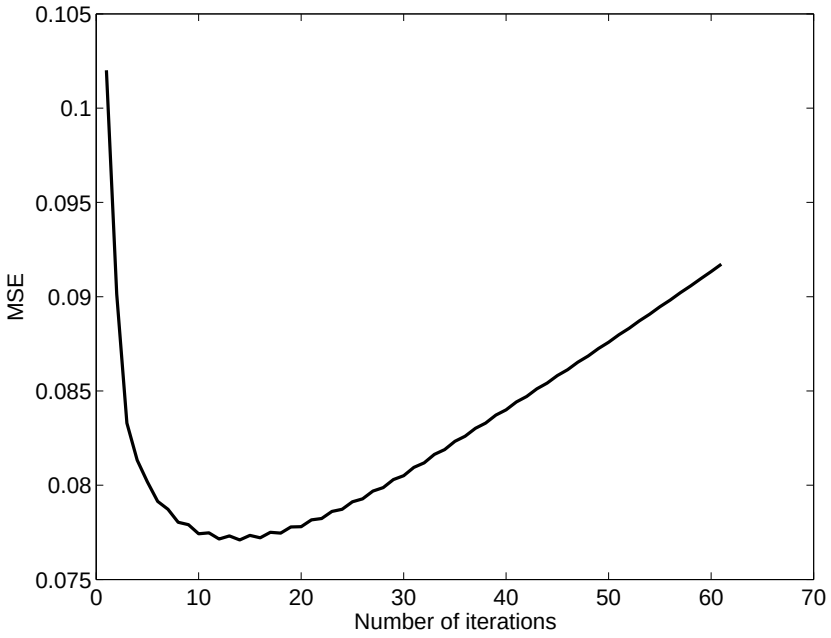


Figure 3.3: Convergence of RL algorithm, see section(3.5.3). The algorithm reaches a minimum value of MSE after a certain number of iterations. However, due to the lack of regularization, the MSE starts to increase again. The algorithm must be stopped before convergence to avoid noise-overfitting.

the statistics of the residual $\|g - Hf\|^2$ to stop the iteration. For a Gaussian likelihood function the optimal number of iterations can be mathematically proven. However, due to the nonlinear constraints (e.g. non-negativity) this no longer applies. This makes the algorithm sensitive to the initial guess that affects the performance of the algorithm to a large extent. A smooth initial solution should be used in order to avoid the amplification of high frequency artifacts. For a large number of iterations the introduction of artifacts still occurs and the algorithm should be terminated before convergence.

Conclusion

As we have seen above, the ML approach leads to un-regularized algorithms. Such algorithms are very sensitive to the noise, and the iterations have to be stopped before convergence. However, the question of the optimal number of iterations is not easy to answer. To solve these problems, regularization is needed. Some authors prefer to introduce regularization as a penalty function added to the likelihood function to alleviate the illposedness of ML method. This leads to penalized

likelihood methods. However, the Bayesian framework provides a means for the introduction of regularization in the form of a prior probability distribution. This leads to the Maximum *a posteriori* (MAP) method. The difference between both approaches, penalized likelihood and MAP, is largely semantic. In the following section, we discuss the Bayesian approach to the image restoration problem.

3.4 Bayesian Approach

Bayesian estimation is a well-established methodology within statistics that has been recently applied in image restoration. Bayesian framework is based on Bayes' theorem due to the English mathematician Thomas Bayes (1702-1761). Restoration algorithms developed within the Bayesian framework are distinguished from other algorithms by the inclusion of prior knowledge about the true image in the form of a prior probability distribution over images. To tackle the problem of image restoration within the Bayesian framework, the first step is to construct the prior probability distribution, $\Pr(f)$. The probability density function $\Pr(f)$ represents our prior knowledge about the true image. $\Pr(f)$ assigns a high probability to solutions that agree with our prior knowledge (or belief) about the true image (solution). The prior distribution is constructed independent of the observed image. Using Bayes theorem, this prior distribution can be modified, based on the observation model, into the *a posterior* distribution. According to Bayes theorem, the posterior can be calculated as,

$$\Pr(f|g) = \frac{\Pr(g|f)\Pr(f)}{\Pr(g)} \tag{3.18}$$

where the evidence $\Pr(g)$, depends on the observed image only and can be regarded as a normalizing constant, the likelihood $\Pr(g|f)$ depends on the observation model and $\Pr(f)$ is the prior. The mode of the posterior distribution is often selected to be the estimated true image. In this case, it is known as the maximum aposteriori solution.

3.5 Maximum *a posteriori* Methods

To avoid the drawbacks of the ML approach, regularization is needed. Regularization is effected through the introduction of an appropriate prior probability density

function $\Pr(f)$. Note that $\Pr(f)$ can be regarded as a penalty function that penalizes undesired features of the solution, in accordance with general knowledge about the true image f. In literature several forms of regularization functions are proposed. The maximization of the penalized likelihood can be interpreted as the maximization of the posterior probability.

$$\hat{f} = arg \max_{f} \Pr(g|f) \Pr(f). \tag{3.19}$$

3.5.1 Prior Models

The selection of the prior probability distribution is still an unsolved problem. It is difficult to declare that a certain prior distribution is suitable for all situations in practice. A prior distribution that performs good for one class of images might not be suitable for another class of images. For a certain class of images, it would be better to learn the form of the prior distribution using several sample images from this class. However, this approach is not always feasible. Several forms of the prior distribution that are based on general knowledge about images have been suggested by researchers from many disciplines of science. The most frequently used priors have the form of Gibbs distribution,

$$\Pr(f) \propto \exp(-\gamma \Omega(f)). \tag{3.20}$$

Where, γ is a parameter that controls the amount of regularization and $\Omega(f)$ is some functional selected to describe prior knowledge about the true image f.

Some prior distributions that are used in image restoration are outlined in the following. We remark that some of the priors introduced in this section, like the Good's roughness penaly, the smoothness, and the total variations are usually considered as penalty function in a penalized ML formulation. However, although the penalilzed ML and the Bayesian approach are cast as different methods, the optimization problem formulated by both methods is similar if the penalty function is, roughly, regarded as a form of prior distribution.

Non-informative Prior

This corresponds to the prior,

$$\Pr(f) = constant. \tag{3.21}$$

This prior assigns equal probability to all images. However, the term 'non-informative' is misleading as it is quite informative to assume that all images are equally probable. using this prior leads again to the maximum likelihood method discussed in (section 3.3). For Poisson noise the non-informative prior leads to the Richardson-Lucy (Eq.3.17) algorithm which is extensively used for astronomical images [31].

3.5.2 Smoothness Prior

$$\Pr(f) \propto \exp\left[-\gamma_f \int ||\nabla f(X)||_2^2 \, dX\right]. \tag{3.22}$$

This formulation of the prior encourages smooth solutions and penalizes rough estimates of the true image. The roughness of the image is measured by the (Eucleadean) norm of the derivative of the image.

Good's Roughness Penalty

$$\Pr(f) \propto \exp\left[-\gamma_f \int f^{-1}(X)||\nabla f(X)||_2^2 \, dX\right]. \tag{3.23}$$

This formulation of the prior leads to the preference of a solution that exhibits smooth characteristics at low intensity sites and allow high gradients (roughness) at sites with large intensity. This smoothness functional was successfully used for the restoration of confocal microscopy images ([48], [25]). In [25], the authors derived an expectation maximization (EM) algorithm regularized by Good's roughness penalty. In [48], quadratic parameterization of the true image was used to derive a more robust restoration algorithm. After a comparison among several image restoration algorithms for confocal microscopy [46], Good's roughness was found to be the best choice for the penalty function according to the MSE criterion.

Total Variation

The total variation (TV) corresponds to the prior,

$$\Pr(f) \propto \exp\left[-\gamma_f \int |\nabla f(X)| \, dX\right]. \tag{3.24}$$

This prior leads to a regularizer of the form,

$$\frac{\partial \Pr(f)}{\partial f} \propto \left[\nabla . \frac{\nabla f}{|\nabla f|}\right]. \tag{3.25}$$

TV was first introduced in [37] as a regularization functional for the de-noising of images corrupted with Gaussian noise. The TV leads to an edge preserving regularization functional that encourages smoothing in the direction tangential to the edges and not in the direction orthogonal to the edges. However, it is argued that TV has the disadvantages that fine details (texture) in the image are not restored and corners are rounded [29].

Entropy

In the absence of any other information about the true image f except its non-negativity, it is possible to formulate the probability $\Pr(f)$ based on the entropy of f. The principle of maximum entropy has its origins in statistical thermodynamics, has been applied to image restoration. Probably, the first application of entropy in image restoration was in [11]. Several researchers have contributed to improve the method into what is currently known as maximum entropy method (MEM). The most commonly used entropy functions [31] are:

- $\Omega(f) = -\log(f).$ (3.26)
- $\Omega(f) = -f\log(f).$ (3.27)
- $\Omega(f, m) = -\left(f - m - f\ln\frac{f}{m}\right).$ (3.28)

Where, m is known as the model. The last definition has the advantage that it has a maximum of zero when f is equal to the model m. The model is used to incorporate knowledge about the spatial intensity distribution of the image. In the absence prior information, the model is assumed to be constant [45]. The MEM is widely used for the restoration of images in astronomy. Despite its capabilities, the MEM has a major defect of ignoring the correlation among image pixels. This leads to the introduction of spurious features, such as ringing, into the restored image. Another difficulty lies in the selection of an appropriate model. An approach to take the correlation among image pixels into account was suggested in [14]. In order to do so, the authors introduced the intrinsic correlation function ICF which encodes the correlation structure of the restored image. In this approach, the restored image is modeled to be the result of the convolution of the ICF with another (hidden) image whose elements (pixels) are uncorrelated and randomly

distributed with unity variance. One deficiency of this approach is that, the ICF introduces a characteristic scale of pixel correlation. In this case the restoration becomes sensitive to the selection of the ICF.

3.5.3 Simulation

In this simulation the effect of the prior on the restored image is examined. The same test image in Fig.(3.2) is used in this simulation. The restored image using the flat, TV, entropy, and Good's roughness priors are shown in Fig.(3.4). Intensity

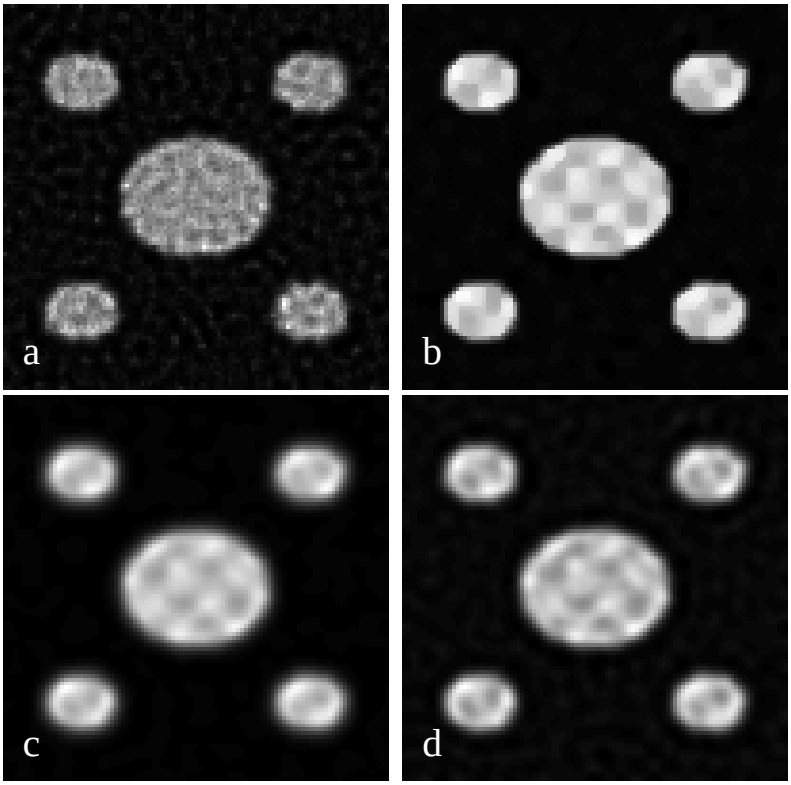


Figure 3.4: Restored images using different priors. (a) Flat prior. (b) Total Variation. (c) Entropy. (d) Good's roughness. The restored image using a flat prios is very noisy due to the absence of regularization.

profiles of a horizontal line through the center of the restored images are shown in Fig.(3.5). The results for the flat prior (RL algorithm), Fig.(3.4a), are noisy and characterized by wild oscillations due to noise amplification. This is due to the lack of regularization. This can be avoided by truncated iteration (see section(3.3)). In Fig.(3.4 b), the restored image using the TV prior. This image exhibits

sharp edges (see Fig.(3.5)) and is less noisy. However, fine details (texture) within the white regions are not well-restored and the smooth peaks are flattened. This type of prior is not suitable for the restoration of fine details. The results for the

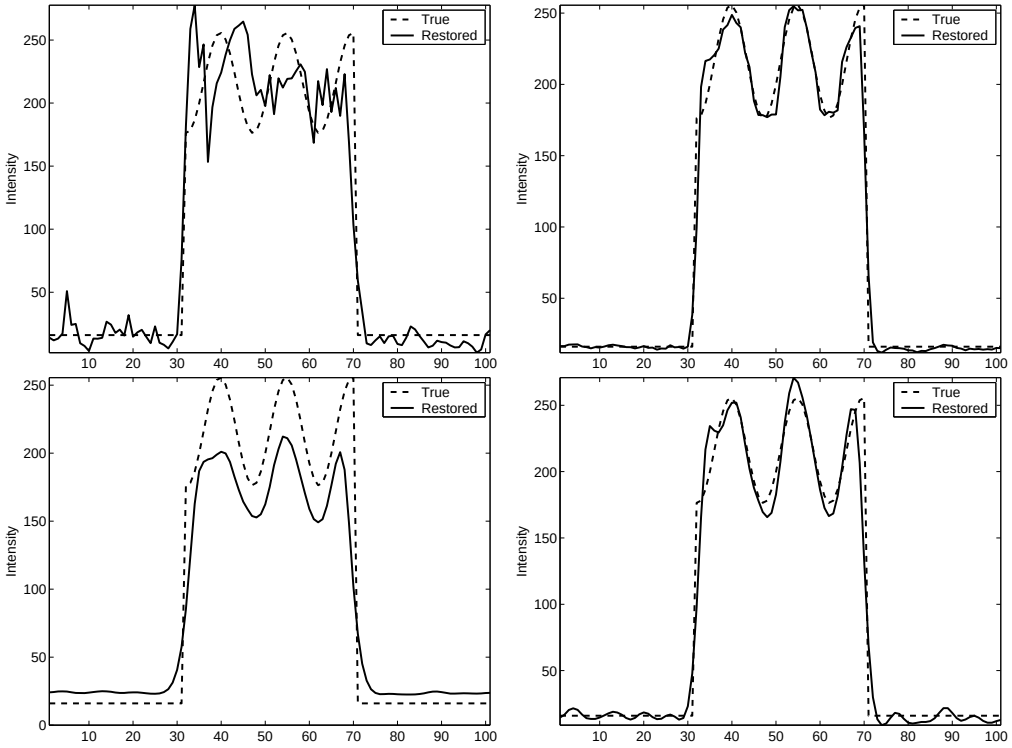


Figure 3.5: Intensity profiles of a horizontal line through the center of the restored images (in Fig(3.4)) using different priors. *Top left:* Flat prior. *Top right:* Total Variation. *Bottom left:* Entropy. *Bottom left:* Good's roughness. Overfitting occurs using a flat prior due to the lack of regularization. Using TV, the restored image exhibits sharp edges, but smooth peaks are flattened. The image restored using an entropic prior (with uniform m) shows a small dynamic range of intensity. Using the Good's roughness prior the edges are relatively smoothed in comparison to the restoration using TV but the smooth peaks are not flattend.

entropy prior are shown in Fig.(3.4c). In this simulation a flat model is used. From this result, it is clear that the entropic prior leads to restored images with narrow intensity range. Another disadvantage of this prior is the difficulty to restore both high and low intensities at the same time using a flat prior. In order to do so, the model should be selected to reflect knowledge about the (local) intensity distribution of the image. Moreover, because no information about the spatial structure of the image is considered, the edges are still blurred. The result for the Good's roughness prior is shown in Fig.(3.4d). This result is smoother than

the results corresponding to other priors. Although, the edges are not as sharp as those in the result of TV prior, the fine details inside the white regions are better restored and not flattened like in the case of TV prior.

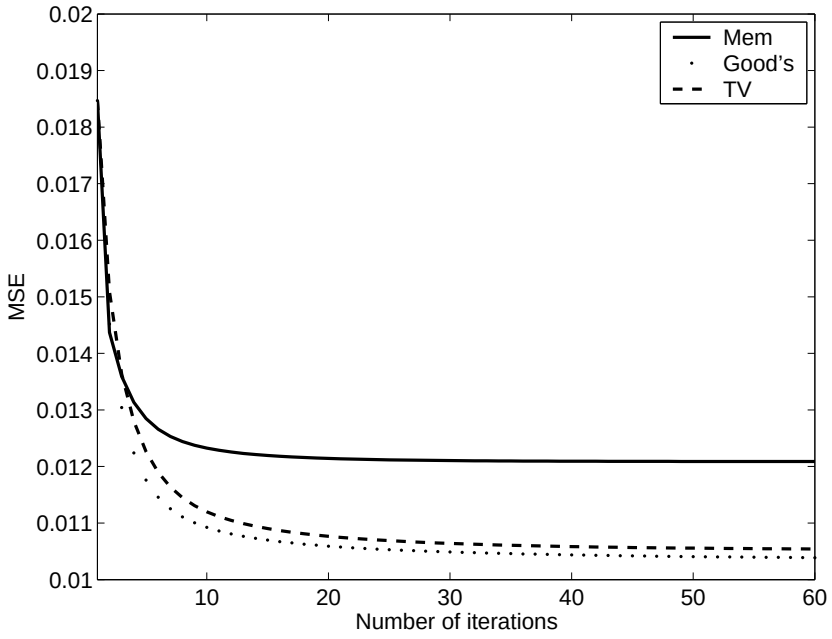


Figure 3.6: Convergence of MAP deconvolution algorithms with different priors: maximum entropy (Mem), Good's roughness, and TV. Using a uniform model (m) for entropic prior leads to a solution with large MSE at convergence. The Good's roughness and the TV priors leads to a better solution in terms of the MSE.

3.5.4 Conclusion

As we have seen above, each prior model has it advantages and disadvantages. The smoothness prior leads to smoothed edges in the restored image. On the other hand, the entropy prior depends on the selection of the model, and ignores the correlation among neighboring image pixels. Even the introduction of ICF imposes a characteristic scale on the restored image. Although the TV prior leads to an edge-preserving regularization, fine-details are not restored. The Good's roughness is better suited to application involving Poisson noise. Because Poisson noise is correlated to the image intensity, adaptive regularization based on the local intensity values is expected to perform better than any isotropic form of regularization. This adaptation is inherent in the Good's roughness penalty due

to the term in the denominator (Eq.3.23). This type of prior has been used for the restoration of confocal microscopy images and shown to have good performance [46].

3.6 Discussion

In this chapter a brief review of deconvolution algorithms was introduced. Linear algorithms estimate the original image as a filtered version of the observed image. These algorithms are simple, fast and might be useful for some simple applications. However, the performance of such algorithms is far from adequate for applications in medical image processing. More complicated algorithms that utilize prior knowledge about the original image and apply imaging constraints (e.g. non-negativity) are needed for the restoration of medical and biological images. Some of these algorithms were introduced in this chapter, e.g. ML and MAP methods. Such algorithms are more successful than the linear algorithms. This improvement in performance is due the incorporation of more knowledge into the algorithm and the consideration of imaging constraints.

Chapter 4

A Constrained Blind Deconvolution Algorithm

4.1 Blind Deconvolution

The problem of blind deconvolution is under-determined, given the observed image g it is required to determine both the true image f and the point spread function h. This lack of information can be alleviated if the PSF, h, is known. The PSF can be calculated using theoretical models, such as the model of Gibson and Lanni [13]. However, this theoretical PSF describe the imaging properties of the microscope under ideal operating conditions. The PSF can be measured by imaging small beads of size below the resolution limit of the microscope. However, this measured PSF does not portray the transfer characteristics of the imaging system under all conditions [8] [35]. During image acquisition the light passes through the specimen. Hence, an accurate PSF should be measured under conditions that exactly replicate the conditions under which the image was acquired. This approach is impractical as each specimen has its own optical characteristics that are not easy to replicate provided that they can be measured to a reasonable accuracy. The blind deconvolution problem is ill-posed in the sense that it admits many solutions. For example, solutions like $f = \delta(X),\ h = g$ and $h = \delta(X),\ f = g$; where $\delta(X)$ is the delta Dirac function, represent two trivial solutions to the problem that should be avoided. Moreover, due to the low pass characteristics of the PSF, spatial frequencies that are outside the passband of the PSF are either not imaged or attenuated

and the addition of such components to one solution creates a new solution that is also admissible. To alleviate this problem, constraints on the feasible solutions have to be incorporated into the blind deconvolution algorithm. In the Bayesian approach to the deconvolution problem, an estimate of the original image is found by the minimization of some risk function. This risk function is derived from the imaging model and (possibly) prior knowledge. In this case the intensity at each pixel in the image are parameters that have to be estimated by the deconvolution algorithm. Due to the large number of parameters, the risk function becomes less sensitive to changes in each single parameter. Moreover, as the number of parameters increases, the risk function exhibits an irregular shape and might become oscillatory [27]. This increases the chance that an iterative algorithm gets stuck in a local minimum.

Hence, there is a need to ameliorate this under-determinacy of the blind deconvolution problem, either by increasing the data or decreasing the number of unknowns. Increasing the data can be achieved through the incorporation of more prior information into the deconvolution algorithm. Decreasing the number of parameters can be achieved through the use of a parametric model. Using a parametric model reduces the number of parameters to those of the parametric model. A model that has a small number of parameters would, intuitively, leads to a more accurate estimation. The use of a parametric model excludes the solutions that are not captured by the model and penalizes infeasible solutions. This may be effected by enforcing constraints on the solutions to the problem and leads to constrained deconvolution algorithms.

Holmes [21] suggested constraining the PSF to be circular symmetric and band-limited. In his implementation, those constraints are applied after each iteration of the algorithm. In another context, Thiébaut and Conan [42] proposed a constrained deconvolution algorithm for speckle noise in astronomical images. This algorithm uses quadratic parameterization to enforce non-negativity on f and a PSF parameterization based on phase aberrations in the pupil plane for h. Recently, Conchello [30] developed a parametric blind deconvolution algorithm that is based on a mathematical model of the PSF. In this algorithm a model based on phase aberrations is used for the PSF.

Motivated by the success of those methods, a parametric deconvolution algorithm is proposed here.

4.2 Parametric Model of the PSF

To solve the blind deconvolution problem one starts by assuming a certain support of the PSF, S_h. If the relation among the PSF coefficients is not known, the deconvolution algorithm is required to estimate all the coefficient inside the support of the PSF. This case corresponds to a parametric model of the form, $h = \sum_i h_i \delta(X - X_i)$. Where, $\delta(.)$ is the canonical delta Dirac function. This model has a large descriptive capacity so that it can capture all possible combinations of the PSF coefficients. However, only quite a small set of these combinations represents feasible estimates of the PSF. This means that, a great deal of the descriptive power of the model will be used to explain infeasible solutions. This approach is unrealistic.

Designing a parametric model, we are actually looking for some function, defined as $h : S_h \mapsto \mathcal{R}$ that approximates the set of points inside the support of the PSF. Specifically, we want to construct an expansion of h of the form $h = \sum_{i=1}^{N_h} a_i \Phi_i(X - X_n)$, where Φ_i denotes a preselected set of basis functions, a_i denotes the corresponding weights, and N_h denotes the number of basis functions. The case of no parametric model corresponds to basis functions that are delta Dirac functions, $\delta(X - X_n)$, located at each point inside the PSF support. We are now looking for a different basis that leads to the reduction of the number of parameters while retaining a reasonable fit to every feasible PSF. The basis functions should be selected in such a way that the model can capture all feasible solutions. For example, linear polynomials are not a reasonable choice as the model may lead to non-smooth functional forms of the PSF. If the variation of the PSF over the image volume is perfectly known, a model can be selected to describe the system without any ambiguity. A good understanding of the physics involved in the image formation process would enable the construction of a model that encompasses the main optical effects. However, in practice this is never the case and the PSF exhibits variations that are not known apriori and hence are hard to model. On the other hand, these variations are dependent of the specimen being imaged and are not unique. A proper model should be as simple as possible in order to reduce the number of unknown parameters. On the other hand, the model should be able to describe the main features of the PSF. Models that are too simple may be not adequate to describe the PSF, while too complicated models may not lead to a

significant improvement in the convergence of the algorithm. Needless to say, if the wrong model is selected the deconvolution algorithm will fail or at least lead to estimates that are of little use for subsequent analysis.

Constraints and prior knowledge about the PSF

In order to design a proper parametric model of the PSF, one has to consider the general constraints on the PSF so that every plausible PSF can be well approximated by the model. Under normal operating conditions the PSF of the confocal microscope follows a mathematical model and exhibits a shape known as the Airy disc. In the radial direction, normal to the optical axis of the microscope, the PSF is circularly symmetric, and has the form of concentric rings of decreasing intensity with maximum intensity at its center,see Fig.4.1. In the axial direction, the PSF extends above and below its center, and exhibits the same form of repeating rings at each slice but with different intensity. Another constraint is the

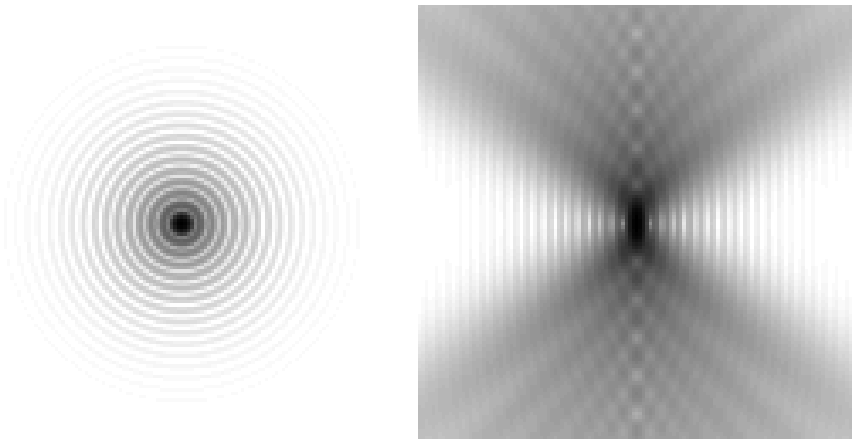


Figure 4.1: Sections through the center of the confocal Microscope PSF. *left*: XY section. *right* YZ section. Shades of gray indicate the intensity on a logarithmic scale (large → small, black → white). The PSF exhibits circular symmetry in the radial (XY) plane. In the axial direction (YZ), above and below the center of the PSF the circular symmetry is still valid although the intensity is different.

band-limitedness of the PSF. Sections through the center of the Fourier transform of the PSF (OTF) are shown in Fig.4.2. The PSF is band-limited in both the radial and the axial direction.

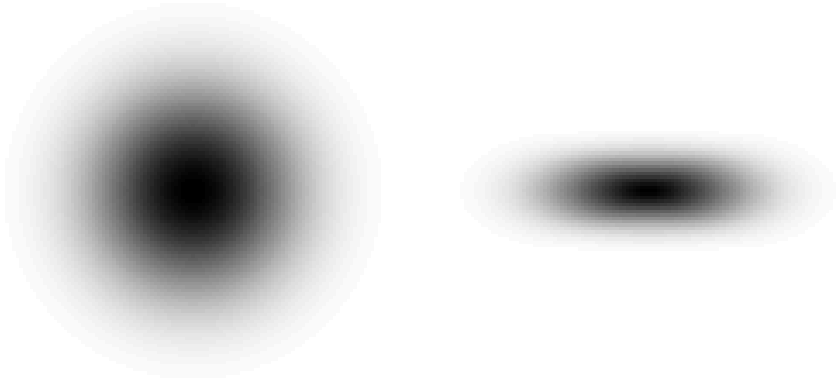


Figure 4.2: Sections through the center of the Fourier transform of the PSF (OTF). *Left*: XY section. *Right* YZ section. Shades of gray indicate the magnitudes on a logarithmic scale (large → small, black → white).

4.2.1 RBF model of PSF

Our goal is to design a parametric model that implicitly incorporates the general constraints on the PSF, like circular symmetry, non-negativity, band-limited-ness and unity sum. Starting from the parametric form,

$$h(X) = \sum_{n=1}^{N_h} a_n \Phi_n(X), \tag{4.1}$$

where, N_h, is the number of basis functions. We proceed by incorporating all the constraints in a sequential manner. The first constraint is that the PSF has a unity sum. Normally, imaging systems neither absorb nor generate energy, and the blurring preserves the mean of the input image. To enforce this constraint on the PSF function, the model is normalized to have unity sum,

$$h(X) = \frac{\sum_{n=1}^{N_h} a_n \Phi_n(X)}{\sum_n a_n (\sum_X \Phi_n(X))}. \tag{4.2}$$

Since negative values of $h(X)$ have no physical interpretation, the PSF function must be constrained to be non-negative. This is ensured using quadratic parameterization of the weights, i.e. $a_n = \beta_n^2$. Now the PSF model is,

$$h(X) = \frac{\sum_{n=1}^{N_h} \beta_n^2 \Phi_n(X)}{\sum_{n=1}^{N_h} \beta_n^2 (\sum_X \Phi_n(X))}, \tag{4.3}$$

where, β_n^2 are the weights, and Φ_n are the basis functions.

This parameterization ensures that the PSF is non-negative provided that the basis functions $\Phi_n(X)$ are non-negative. This condition is satisfied by the proper selection of the basis functions.

Selection of the basis functions

The parametric model proposed here, is based on Gaussian radial basis functions (RBF). The selection of RBF is motivated by several advantages of these functions that are suitable for the re-parameterization of the PSF. These functions are smooth, band-limited and circular-symmetric. Using RBF as the basis functions of the parametric model enforces the features of these functions on the reconstructed PSF function, and hence improves the performance of the deconvolution algorithm. To enforce circular symmetry on the PSF, the basis functions are selected as,

$$\Phi_{n_r}(\rho) = \exp[-2\frac{(\|\rho\|_2 - n_r\lambda_r)^2}{\lambda_r^2}] + \exp[-2\frac{(\|\rho\|_2 + n_r\lambda_r)^2}{\lambda_r^2}], \tag{4.4}$$

where, $\rho = \sqrt{x^2+y^2}, \{x, y\} \in \mathrm{R}^2$, and λ_r controls the width of the basis function. This set of basis functions are concentric rings, and any linear combination of these functions will also be circular symmetric. Moreover, since the Fourier transform of a Gaussian function is also a Gaussian function,

$$\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-X^2}{2\sigma^2}} \quad \overset{\mathrm{F}}{\rightarrow} \quad e^{\frac{-\omega^2\sigma^2}{2}}, \tag{4.5}$$

this choice of Φ_i as a low pass filter (LPF) with a cutoff frequency $\frac{1}{2\lambda_r}$ gives a PSF which is band-limited to frequencies up to $\frac{1}{2\lambda_r}$. On the other hand, the non-negativity constraint is enforced on the model, because $\Phi_{n_r}(\rho) \geq 0 \ \forall \ \rho$. Now, the proposed model assures that all the constraints of the PSF function are always satisfied and there is no need for the *ad-hoc* enforcement of the constraints after each iteration of the algorithm. In the three dimensional

$$\Phi_{n_r,n_z}(\rho, z) = \Phi_{n_r}(\rho)\Phi_{n_z}(z), \tag{4.6}$$

where,

$$\Phi_{n_z}(z) = \exp\left[-2\frac{(\|z - n_z\lambda_z\|_2)}{\lambda_z^2}\right]. \tag{4.7}$$

This choice of basis functions enforces the constraint of band-limitedness in all directions on the PSF. Each basis function is like a shell with the thickness of the shell determined by λ_r, λ_z.

Figure 4.3: Basis functions of the parametric model in 2D. The first three plots on the left show 3 basis functions created using Eq.4.4. The rightmost plot shows the constructed PSF. The basis function are circular symmetric in the radial (XY) plane to ensure that the estimated PSF is also circular symmetric.

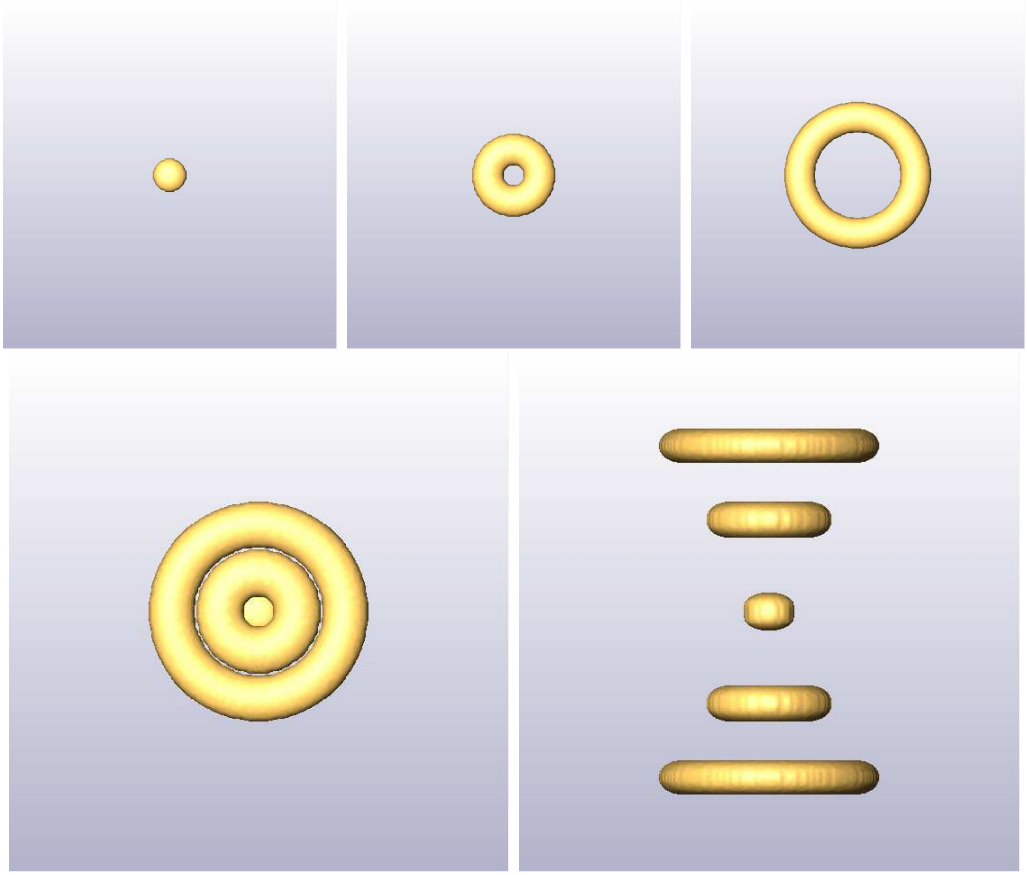


Figure 4.4: Construction of the PSF using the 3D basis functions. *Top:* Basis functions of the parametric model. Each basis is an elliptic (or circular) shell with the radius of the shell and the thickness controlled by the parameters of the model, Eq.4.6 . *Bottom*: The linear combination of the three basis functions shown above. The left panel shows the XY view and the right panel shows the YZ view. Using the parametric model, it is possible to synthesize every physically plausible PSF.

How to select the centers of the basis functions

The model parameters n_r, λ_r, n_z, and λ_z are selected such that the basis functions are uniformly distributed over the whole support of the PSF with an overlap of 50%. The basis functions have constant width in the radial and axial direction, selected by λ_r,λ_z and the radius of the basis functions increase by multiples of λ_r, λ_z. Such a choice of the parameters insures that the fitted function is smooth.

4.3 Derivation of the Parametric Deconvolution Algorithm

In confocal microscope the noise predominantly follows Poisson distribution. For this type of noise the risk function is,

$$R(f,h) = \sum_X [(f \otimes h)(X) - g \ln(f \otimes h)(X)] + \gamma_f \Omega_f(X), \tag{4.8}$$

where γ_f is a regularization coefficient that controls the tradeoff between the fidelity to the data and the smoothness of the solution, and $\Omega_f(X)$ is the negative logarithm of the prior probability $\Pr(f)$ and can be interpreted as a regularization term. The Good's roughness penalty ([25], [48]) is used for regularization, which corresponds to assuming a prior of the form

$$\Pr(f) \propto \exp\left[-\gamma_f \int f^{-1}(X) ||\nabla f(X)||_2^2 \, dX\right]. \tag{4.9}$$

To ensure the non-negativity of the estimate f, we use quadratic parameterization $f(X) = \alpha^2(X)$. It can be shown [48] that for this parameterization the Good's roughness penalty function is $\Omega_f = -4\langle \alpha, L(\alpha) \rangle$, where L denotes the Laplacian.

The blind restoration problem now reduces to finding the image $\hat{f}$ and $\hat{h}$ that minimize the risk function Eq.(4.8).

$$\{\hat{f}, \hat{h}\} = \min_{f,h} R(f,h) \tag{4.10}$$

This minimization is performed iteratively using a gradient descent method. Instead of the simultaneous estimation of both $\hat{f}$ and $\hat{h}$, alternating minimization (AM) is used. In each iteration of the algorithm, we update the image f estimate while keeping the PSF function constant and then update the PSF while keeping the image f constant.

For the image $f = \alpha^2$ the update rule is:

$$\alpha^{k+1} = \alpha^k - \eta \left(\frac{\partial R(\alpha^2, h)}{\partial \alpha} \right), \tag{4.11}$$

where k is the iteration number, η is the step size, and where

$$\frac{\partial R(\alpha^2, h)}{\partial \alpha} = -2\alpha \left(\frac{g}{f \otimes h} - 1 \right) \otimes h(-X) - 4\gamma_f L(\alpha). \tag{4.12}$$

For the weights of the PSF model, the gradient of the risk function is,

$$\frac{\partial R(f,h)}{\partial \beta_n} = \frac{\partial R(f,h)}{\partial h(X)} \frac{\partial h(X)}{\partial \beta_n} \tag{4.13}$$

$$\frac{\partial R(f,h)}{\partial h} = -\left(\frac{g}{f \otimes h} - 1 \right) \otimes f(-X) \tag{4.14}$$

$$\frac{\partial h(X)}{\partial \beta_n} = 2\beta_n \left(\Phi_n - (\sum_X \Phi_n(X)) h(X) \right) / \sum_i \beta_n^2 (\sum_X \Phi_n(X)). \tag{4.15}$$

$$\begin{aligned} \frac{\partial R(f,h)}{\partial \beta_n} &= \frac{-2\beta_n}{\sum_i \beta_n^2 (\sum_X \Phi_n(X))} \\ &\quad \left[\left(\frac{g}{f \otimes h} - 1 \right) \odot \left(\Phi_n \otimes f - (\sum_X \Phi_n(X)) h \otimes f \right) \right]. \end{aligned} \tag{4.16}$$

Where, $\odot$ denotes the correlation operation such that $A(X) \odot B(X) = A(X) \otimes B(-X)$. Now the update rule for the weights is

$$\beta_n^{k+1} = \beta_n^k - \eta \frac{\partial R(f,h)}{\partial \beta_n} \tag{4.17}$$

4.4 Experiments and Results

4.4.1 Quantification of the Results

To quantify the quality of the deconvolution results, the mean square error (MSE) criterion is used. The MSE is calculated as

$$MSE(f, \hat{f}) = \sum_X (f(X) - \hat{f}(X))^2. \tag{4.18}$$

Working with simulated images, one has the advantage of having the true image available and it is easy to compare the result of the deconvolution algorithm to

the true image. The MSE is calculated between the final result of deconvolution and the unblurred true image of the object used for simulation.

For real data, one has the problem that the exact image is unknown. One can resort to the MSE between the observed (blurred and noisy) image, and the estimated image after deconvolution re-blurred again using the estimated PSF i.e. $MSE(g, H\hat{f})$. Ideally, the deconvolution algorithm should generate an image that is a de-noised and less blurry version of the observed image.

4.4.2 Benchmark Algorithm

To compare the performance of the constrained parametric deconvolution algorithm to other deconvolution algorithms, a blind version of the deconvolution algorithm proposed in [48] is used. This algorithm was selected due to its improved performance compared to the existing deconvolution algorithms for optical imaging [46]. In this algorithm quadratic parameterization is used for for the image $f = \alpha^2$, and no parameterization is used for the PSF. In the following this algorithm is referred to as the non-parametric deconvolution algorithm.

The benchmark algorithm consists of the alternating minimization of the risk function Eq.4.8. For the PSF, the update rule is

$$h^{k+1} = h^k - \eta \frac{\partial R(f,h)}{\partial h}. \tag{4.19}$$

Where, $\frac{\partial R(f,h)}{\partial h}$ is given by Eq.4.14. In this algorithm quadratic re-parameterization is used to assure the non-negativity of the image, i.e. $f = \alpha^2$. The update rule for the image is the same as in Eq.4.12. Because no parameterization is used for the PSF, this algorithm must estimate all the coefficients inside the support of the PSF.

4.4.3 Experiments with Artificial Data

To evaluate the performance of the algorithm a test image is created. The image is convolved with a confocal PSF, calculated using the model of Gibson and Lanni [13] with the parameters: Objective 60×/1.4, illumination wavelength of 530 nm. Then Poisson noise is generated in the image (using the blurred image as the mean for the Poisson process), the maximum number of photons/voxel is 2040

photons/voxel to set the SNR =31.5 dB. The number of basis functions used is N_h =6. The parametric algorithm was initialized as follows:

- initial guess for the true image is the observed image.

- initial guess for the PSF, uniform weights.

For the non-parametric (benchmark) algorithm, the same initializations are used. The initial guess of the PSF and the confocal PSF (the true PSF) are shown in Fig.4.5. The algorithms are run until convergence. We indicate convergence when $\frac{||f^{k+1} - f^k||}{||f^k||} < 10^{-4}$.

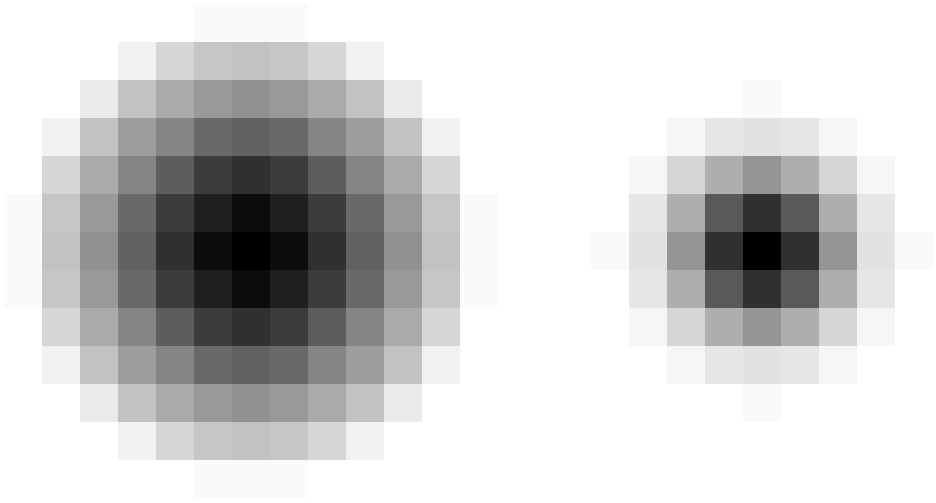


Figure 4.5: The initial guess of the PSF (*left*), and the true PSF (*right*). The true PSF is the XY section of a confocal PSF corresponding to an 60×/1.4 objective lens, illumination wavelength of 530 nm, calculated using the model of Gibson and Lanni [13]

The simulated confocal microscope image is restored using both the proposed algorithm and the benchmark algorithm. Fig.4.6, shows the MSE plotted against the number of iterations for both the parametric algorithm and the non-parametric algorithm. As the figure reveals, the parametric deconvolution algorithm exhibits a faster convergence rate than the benchmark algorithm. Even more, the result of the parametric algorithm exhibits a smaller value of MSE after convergence. The above experiment is repeated using a different value of the photon conversion factor to set the signal to noise ratio SNR at 22.5 dB. This corresponds to a maximum number

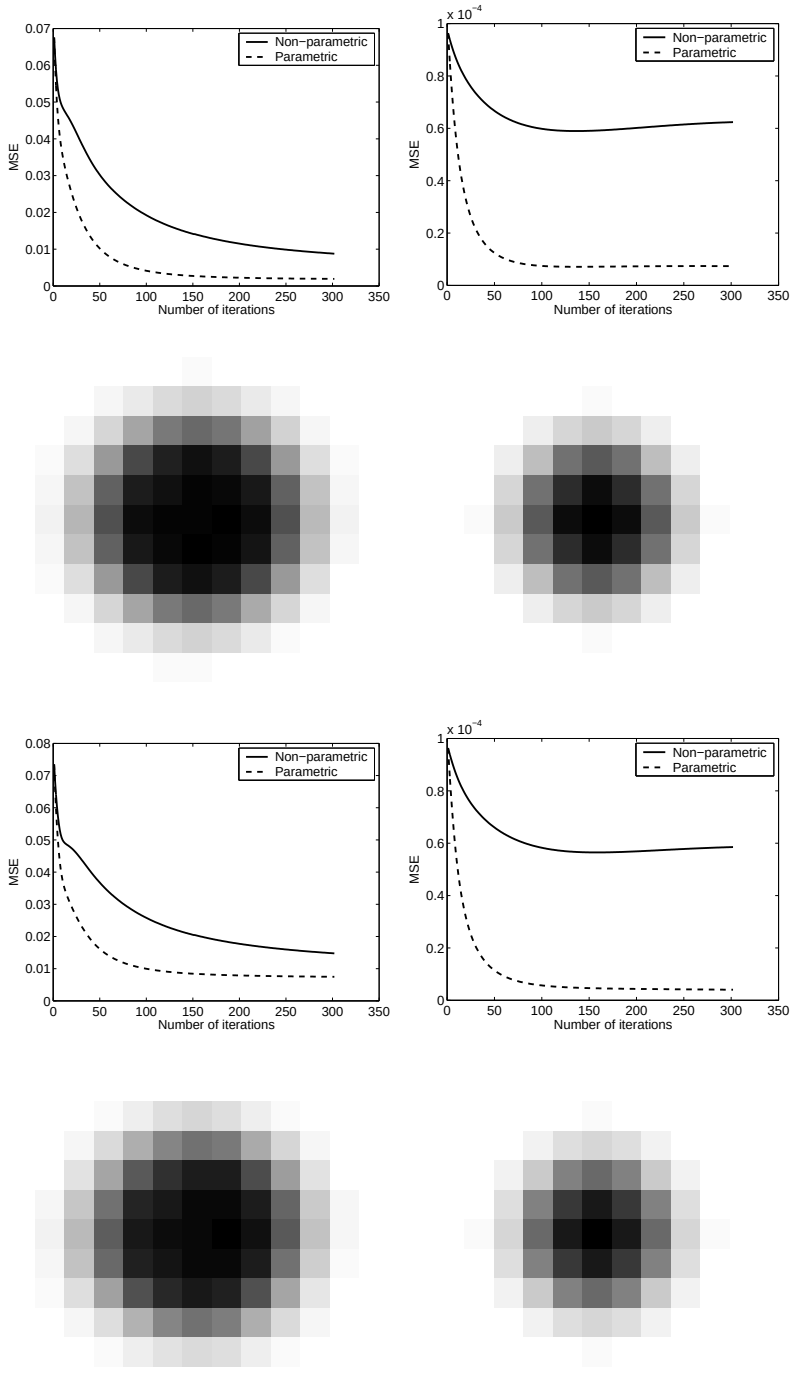


Figure 4.6: Comparison of convergence of both the non-parametric and parametric deconvolution algorithms. *Top* row: results corresponding to a maximum number of photons/voxel=2040, SNR=31.5 dB. *Left*: the restoration MSE. *Right*: the MSE between the estimated PSF and the true PSF. The *second* row shows images of the estimated PSF using the non-parametric (*left*) and the parametric (*right*) algorithms. The *third* and the *fourth* rows show the results corresponding to a maximum number of photons/voxel=255, SNR=22.8 dB.

of photons of 255 photons/voxel to set the SNR=22.8 dB. Fig.4.6, shows the MSE plotted against the number of iterations for both the parametric algorithm and the non-parametric algorithm. As we can see in the figure, the parametric algorithm still exhibit a faster convergence rate and a smaller value of MSE at convergence.

4.4.4 Restoration of Confocal Images

The proposed constrained parametric deconvolution algorithm is applied to an image stack of the optic neuropils of *Drosophila*. This image represents a challenge to deconvolution algorithms. Due to the small size of the cell compartments, electron microscopic studies are the necessary choice. The visualization of such small cells using confocal microscopy was explored in [20]. In this experiment a confocal image of these cells is used to test the performance of the proposed deconvolution algorithm. This image data set has been introduced in chapter 2 along with the immunohistochemical preparation and imaging setup.

Working with real data, one faces the problem of how to evaluate the results. Normally, a deconvolved image would show more intensified, less-blurred structures in comparison to the observed image. However, to assess the performance of the deconvolution algorithm, one needs the ground truth to judge the restored image. This is very important as in some situations artifacts might arise in the restored image. These artifacts would be hard to distinguish from biological structures unless some knowledge about the true image is available. On the other hand, the image formed by the microscope is due to the light emitted from the fluorochrome and any non-uniformities in the staining will be reflected in the observed image. Fortunately, an electron microscopy image of the neuropils was available [20]. Although this image is not of the same preparation whose image is acquired by the confocal microscope, the image reveals the general ultrastructure of the lamina cross section (see Fig.4.7) and provides the ground truth for a qualitative comparison of the deconvolution results of both algorithms.

The restoration results of both the parametric and non-parametric deconvolution algorithms are shown in Fig.(4.8). For this experiment the regularization coefficient used was: 15×10^{-3}. An initial guess of the PSF size $16 \times 16 \times 16$, was used. For the RBF model of the PSF the number of basis functions was three basis functions both in the radial and axial direction. Examining Fig.4.8, we can see

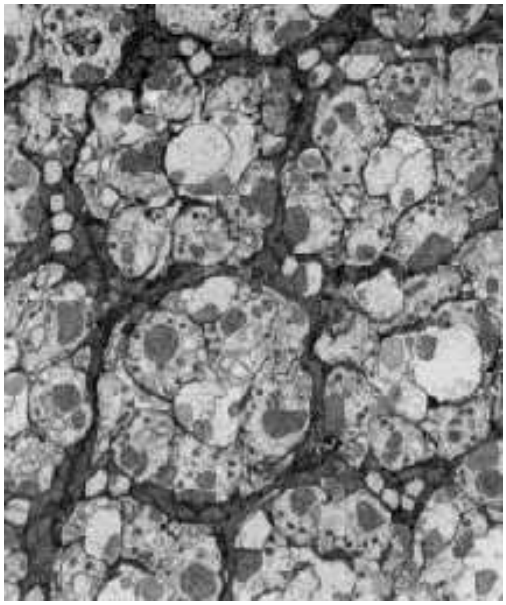


Figure 4.7: Electron microscope image of the optic neuropils of *Drosophila* The image reveals the ultrastructure of the lamina and is used for a qualitative judgment of the deconvolution results. The lamina shows an orderly organization of cartridges. Each cartridge consists of six gray photoreceptors that encircle two or three lighter axons.

that the results of the constrained parametric deconvolution algorithm are better than those of the non-parametric algorithm. The results of the proposed algorithm show more details of the cells, while the results of the non-parametric algorithm are more noisy and lack several fine details.

4.4.5 Execution times

The non-parametric algorithm is computationally cheap compared to the parametric algorithm. The average execution time per iteration was 3 minutes for the non-parametric algorithm and 5 minutes for the parametric algorithm. The simulations were done on Silicon Graphics Octane (SGI) dual processor workstation (R10K, 195 MHZ; 896 MB and 1,024 MB RAM). The parametric (constrained) algorithm generally much slower than the non-parametric algorithms due to the increased computational effort needed. The main computational load comes from the PSF update step which needs several convolutions to be computed.

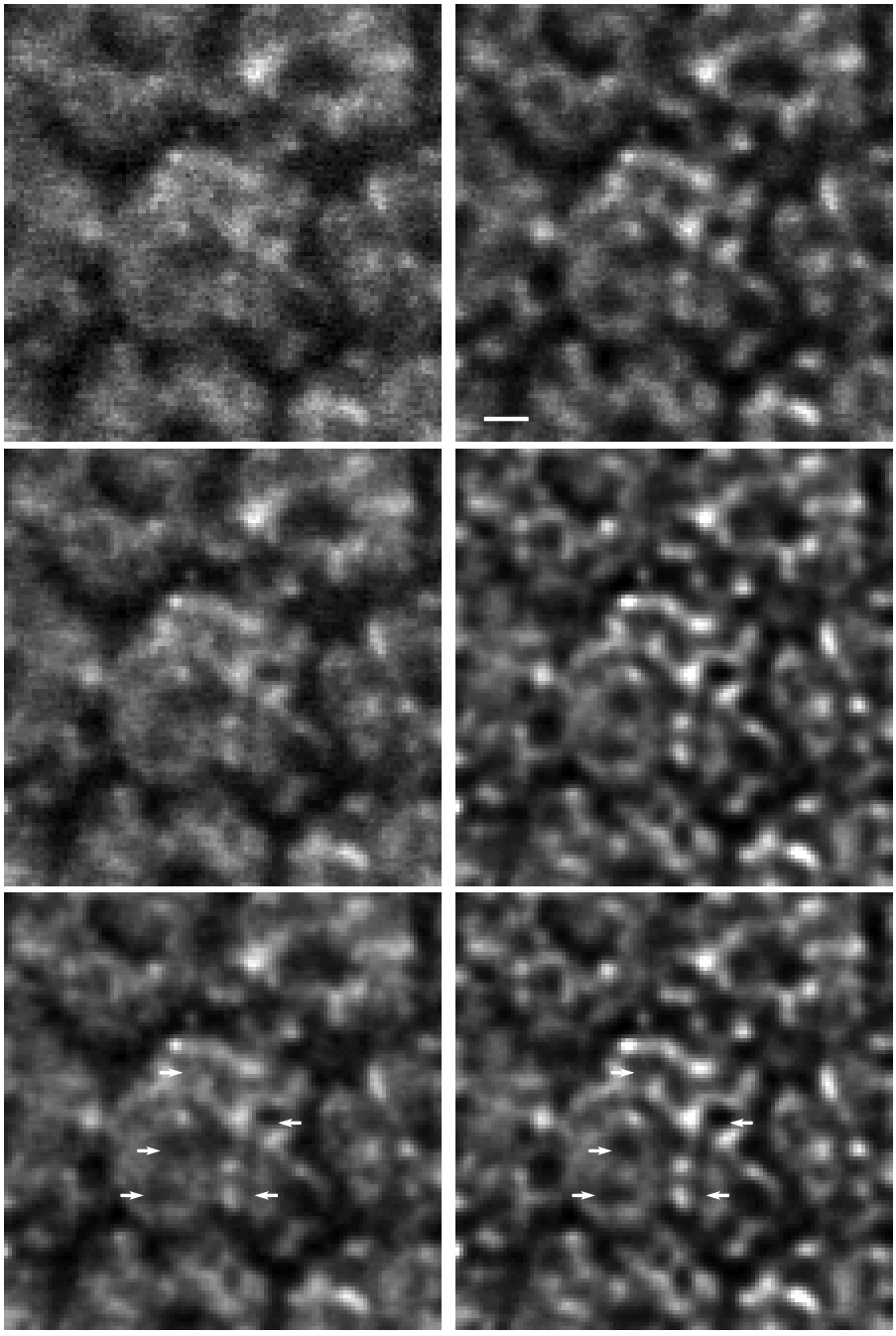


Figure 4.8: Axial sections (Z-slices) of restoration results of the confocal microscope image stack. *Left*: the results of the non-parametric deconvolution algorithm. *Right*: the results of the parametric deconvolution algorithm. From *top* to *bottom*, results after 8, 20, 40 iterations, respectively. The results of the parametric deconvolution algorithm reveals more details of the structure of the lamina. The compartments of the central cartridge are clearly recognizable in the restored images. The results of the non-parametric (benchmark) algorithm are deblured and smoothed but small compartments are still not adequately resolved.

4.5 Discussion

As the experiments showed, the parametric deconvolution algorithm convergence faster in comparison to the non-parametric algorithm. This is mainly attributed to the reduction of the number of unknowns using the RBF model of the PSF. For the non-parametric algorithms all the PSF coefficients inside the support of the PSF have to be estimated, which is quite a large number of unknowns in comparison to the number of weights of the PSF model of the parametric algorithm.

Intuitively, for any estimator the rate of convergence improves with the reduction of the number of unknowns. For the model proposed here, the number of basis function should not be too small in order to assure that the parametric model is able to fit the true PSF. For example, if the true PSF is smaller than the size of the basis function, the model will not be able to produce a good estimate of the PSF and will always converge to an estimate that is the same as the smallest basis function located at the center of the support of the PSF. Because, The parametric

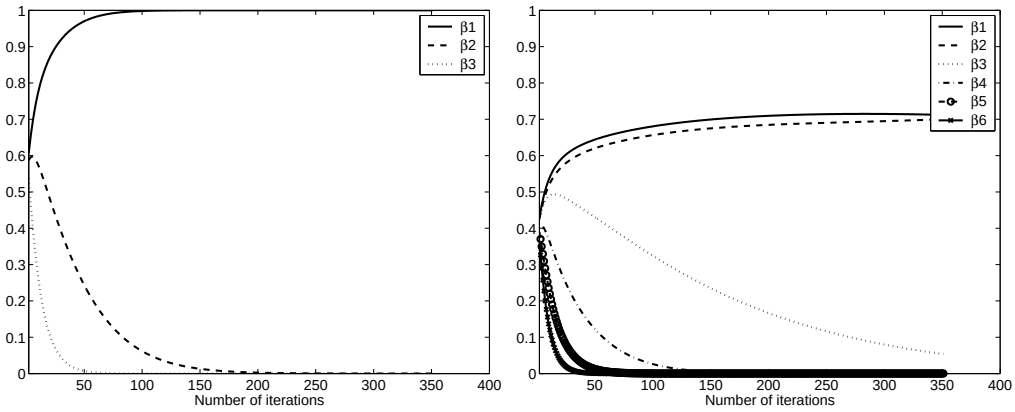


Figure 4.9: Effect of the number of basis functions on the convergence of the weights of the parametric model of the PSF. *Right*) Three basis functions. *Left*) Six basis functions. For large number of basis functions the convergence is slow, but the estimation of the PSF is more accurate. For a small number of basis function, the convergence is fast, but the accuracy might be worse.

model cannot converge to an estimate of the PSF which is smaller than the size of the basis function, the size of the basis function have to be smaller than the size of the true PSF. In practice, the size of the true PSF might not be known, one should start with a large number of basis function.

In our experiments, we have found that a number of basis functions of the

model equal to half the size of the PSF in each direction (radial and axial) leads to good results.

Number of Basis Functions

An important question is how to select the number of basis functions used in the model. Intuitively, the number of basis functions has to be as small as possible in order to increase the rate of convergence. However, a too small number of basis functions might not be enough to reach a reasonable estimate of the PSF. On the other hand, the size of the basis functions should not be larger than the size of the true PSF, otherwise the algorithms will always converge to an estimate of the PSF which is always the same as the smallest basis function. Of course, an objective selection of the number of basis functions would be better. The problem is that the actual PSF is not known. In most cases, only the size (support) of the PSF function can be guessed. Based on the size of the PSF, a safer choice is to use the maximum number of basis functions i.e. half the support of the PSF in both the radial and the axial directions. In our experiments, we have developed an empirical criterion that a number of basis functions equal to half the support of the PSF is enough and usually leads to better results.

4.6 Conclusion

In this chapter a novel parametric constrained blind deconvolution is proposed. The proposed algorithm is based on a parametric model of the PSF using radial basis functions. The experiments show that the algorithm performs better than existing algorithms [46][49]. The results of the experiments presented here clearly support the feasibility of parametric blind deconvolution. We found that the parametric blind deconvolution algorithm accurately estimates the parameters that define the PSF and produces good estimates of both the true image and the point spread function. The computational load of the algorithm is higher than the non-parametric algorithm. However, the algorithm exhibits a faster rate of convergence and a smaller number of iterations would be enough to achieve comparable accuracy.

Chapter 5

Restoration of Neuron 3D Images

In this chapter, we present a deconvolution method to correct the effect of anisotropic depth-variant blur introduced by the imaging device (confocal microscopy). When objects of tubular-like structures like neurons are imaged, the acquired images are degraded and the extraction of the accurate morphology of neurons is hampered due to these anisotropic deformations. To mitigate the lack of information of the deconvolution problem, we propose a method to estimate the support of the point spread function from the acquired image without any prior knowledge of the imaging system. This method which is based on the estimation of the original object, is suitable in cases for which the objects being imaged have a known geometry, like neurons. Using the proposed deconvolution method, geometric distortions are eliminated and the restored images are more suitable for further analysis.

5.1 Introduction

Computational optical sectioning microscopy is an important tool that enables the three-dimensional (3D) visualization of fluorescently-stained biological specimens without the need to physically slice the specimen into thin sections [30]. Using a confocal microscope a series of two dimensional images is collected by focusing the microscope at different depths through the specimen. Only a small portion of the specimen is illuminated at a given time and light from other points in the specimen is rejected by a small aperture in front of the detector [33]. One important application area of confocal microscopy is neurobiology. Neurons exhibit a

tree-like structure with an enormous variety in configuration and size that range from hundreds of micrometers to well below the resolution limit of conventional microscopy. Proper visualization and analysis of these cells requires high resolution and quality of the acquired images, especially in neuromorphometric studies, where the accurate description of neuron structures plays an important role in analyzing the relation between the functional properties and the morphological features of the neuron. Although the resolution of the confocal microscope is superior to that of the conventional light microscope, there are effects that distort the image and make it insufficient for precise visualization and the precise assessment of the exact geometry of small biological structures. Examples are: shot noise, detector noise, blurring, and background noise. Moreover, the presence of the specimen in the optical path of the microscope changes the overall characteristics of the imaging system and gives rise to specimen-induced artifacts in the observed image. Typical mounting media, for example, have a refractive index that is different from that of the immersion medium of the objective lens of the microscope. This mismatch in refractive index causes spherical aberration that increases with increasing depth into the specimen. Although some objectives can be adjusted to eliminate spherical aberration at a given plane inside the specimen, they are aberration-free only for the depth for which they were adjusted [35]. The effect of these depth-dependent degradations is an anisotropic point spread function (PSF) which results in geometric distortions and gives rise to anisotropic artifacts.

As accurate neuron morphology is often important subsequent processing is needed to correct for the blurring artifacts and geometric distortion to improve the quality of the acquired images. In order to do so, several deconvolution algorithms were developed to solve this problem (e.g. [48], [21], [30]). Non-blind deconvolution algorithms require knowledge of the PSF. The PSF of the imaging system can, in principle, be either calculated or measured by imaging beads of small size near to the resolution limit of the microscope. However, following the above discussion, the PSF is actually a function of the specimen and the imaging depth, and this would require calculating the PSF for each specimen and at all depths inside the specimen. A more robust approach is to extract the PSF from the observed image. Although blind deconvolution algorithms try to estimate both the original image and the PSF from the observed image, prior information about the PSF size is necessary otherwise the problem is ill-defined. Without proper initialization, these

algorithms may lead to sub-optimal or incorrect results due to the ill-posed-ness of the blind deconvolution problem. In this chapter, a new method to estimate the PSF based on the geometric distortions in the confocal microscope images of neurons is proposed. This method is based on the biologically plausible assumption that the cross-sectional shape of neuronal processes (dendrites and axons) can be approximated, on average and over a long distance, by a circle. However, due to the geometric distortion introduced by the imaging device, this shape is deformed into an ellipse in the observed image. Based on the energy conservation assumption of the blurring process, the true shape of the cross-section of the neuron can be estimated. A similar constraint is used for neuron tracing algorithms [1], that fit generalized cylinders to the neuronal processes in the microscope image. By comparing the observed cross-sections to the estimated ones, the support (size) of the PSF can be estimated. This support gives a measure of the blur and its aspect ratio provides information about the anisotropy of the PSF. Finally, we apply non-blind deconvolution algorithm to improve the quality of the observed image. The chapter is organized as follows: After a brief description of the data set used in this study (section 5.2.1), we explain the details of the algorithm for estimating the support of the PSF in section 5.2.2. In section 5.2.3 we describe the deconvolution algorithm. The results are shown in section 5.3. Finally, conclusion and discussion are given in section 5.4.

5.2 Materials and Methods

5.2.1 Data Set

In this work, a 3D image stack of the inter-neuron A4I1 of the migratory locust (*Locusta migratoria migratorioides*) (courtesy of Dirk Bucher, Department of Biology, Brandeis University, USA) is used. The immunohistochemistry techniques and imaging setup are described in section (2.5.2). To view the structure of the neuron, surface rendering techniques are employed and iso-surfaces of the data set are constructed for different threshold of image intensities. For proper visualization a compromise is made between a high threshold value that is necessary to suppresses the background noise and a small threshold value that is necessary to keep the fine, low-intensity, details of the data set. An iso-surface of the image

details of a part of the stack (128×128×128), with the visually optimal intensity threshold of 25, constructed using the AMIRA software (Indeed - Visual Concepts GmbH) is shown in Fig.5.1. This value of threshold is chosen as the smallest intensity value above noise level. The geometric distortion in the form of elongation in the axial (Z) direction can be clearly seen. Without correction, any further analysis or absolute measurements (volume, surface area, etc.) will lead to incorrect values. Although using a larger threshold to construct the iso-surface of the

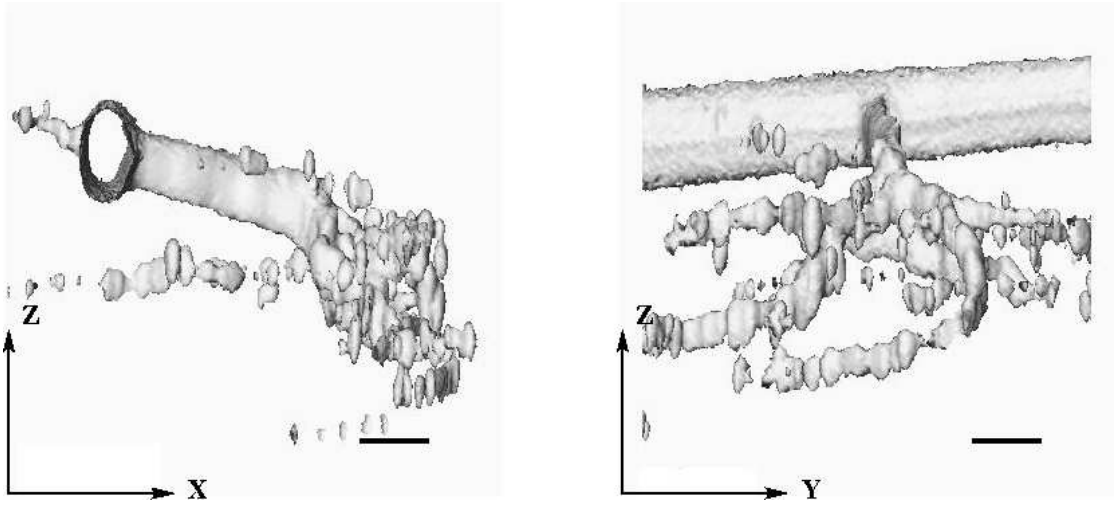


Figure 5.1: The Iso-surface constructed from the raw data set (using the visually optimal threshold value of 25). *Left*: XZ view. *Right*: YZ view. The anisotropic deformation causes the elongation of objects in the axial direction, hence the cross-section of the neuronal branches are elliptic. Scale bar: 16μm

neuron would help to obtain a more circular cross-section of some of the branches, fine details are lost and the geometric distortion is still not completely corrected, as can be seen in Fig.5.2. Hence, an adaptive method is needed to correct for the anisotropic distortion artifacts without loss of fine image details.

5.2.2 Estimation of the Radial and Axial Blur

The blur introduced by the imaging system is characterized by its PSF, that describes the optical characteristics of the imaging system including the specimen. As previously discussed in [18], different refractive indices along the optical path in the microscope result in a change in the shape of the PSF and this change increases dramatically with the depth inside the specimen. This can be seen in Fig.5.3. The

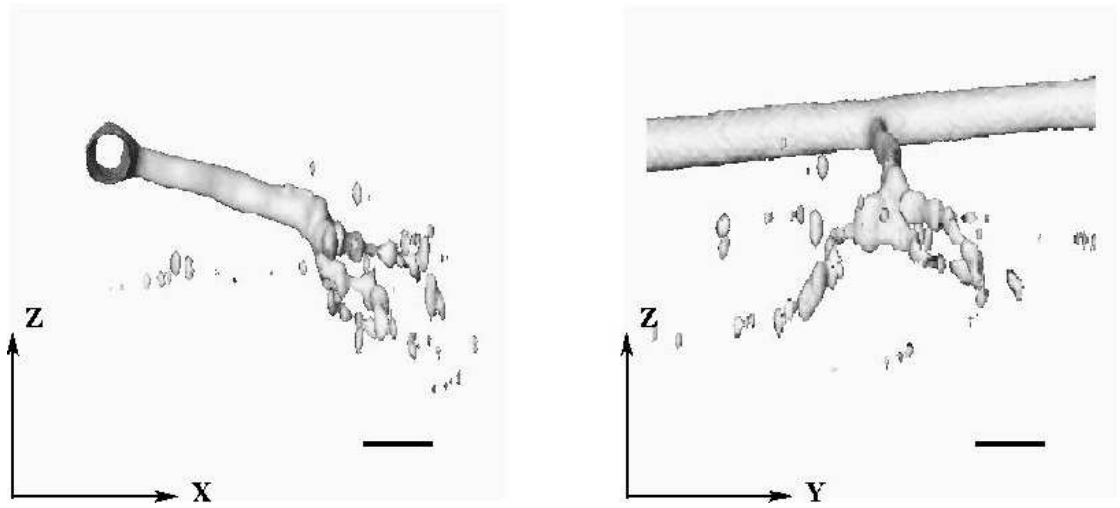


Figure 5.2: Left: XZ view of the iso-surface constructed using a threshold value of 55, right: YZ view. Fine details are lost when compared to Fig.5.1, and the cross-section is still elliptic though with less eccentricity. Scale bar: 16μm

figure shows XZ-slices of an image of a small bead (0.45μm in diameter) at three different depths, calculated using the software package XCOSM [6], [5]. As can be seen, the PSF changes with depth, exhibits halos and rays, and the PSF becomes more anisotropic with depth. Even worse its centroid is not aligned with its maximum (the maximum is marked by the white dot inside the central elliptical region). If we neglect intensity values smaller than 10% of the maximum value, the PSF has approximately the shape of an ellipsoid whose major axis is parallel to the optical axis (Z-direction) and whose aspect ratio increases as we focus deeper into the specimen. The aspect ratio of the point spread function causes a strong blur in the axial direction than in the radial direction.

In the case of confocal microscopy images of neuron, we can assume that the neuron's branches have, on average, circular cross-sections. If we additionally assume that the blur does not change the signal energy, i.e., that the sum of intensities of all the pixels that belong to the unblurred object is equal to the sum of the intensities of the pixels in the blurred image it is possible to estimate the average spatial extent of the PSF.

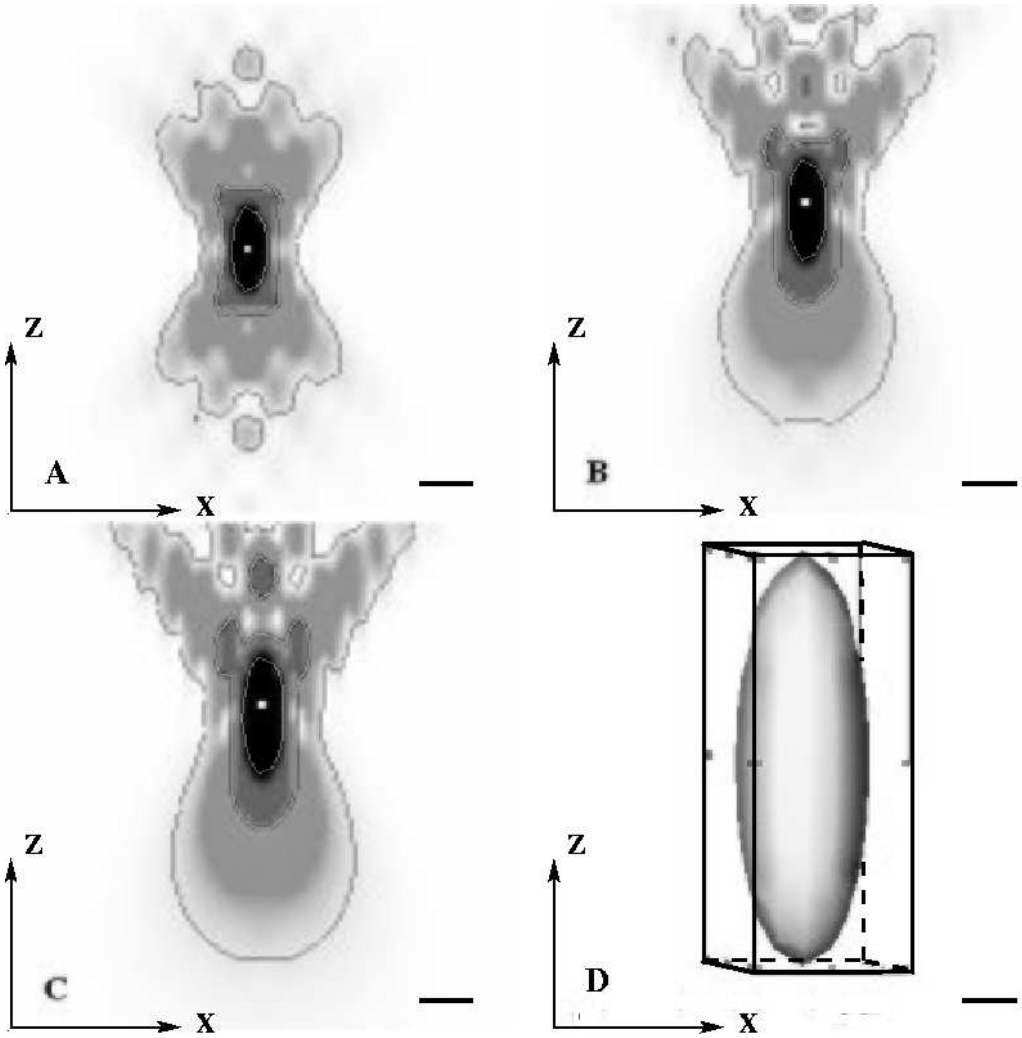


Figure 5.3: Bead images at different depths. **A, B, C:** Calculated image of an XZ-slice through the center of a bead (0.45μm in diameter) at different depths (**A,B,C**: 0.0μm, 0.90μm, 1.80μm). The bead is assumed to be immersed in a medium with an average homogeneous refractive index of 1.33. Simulated microscope parameters: 60×/1.4 objective immersed in oil, refractive index of (1.515) and illumination wavelength of 530 nm. Gray values denote image intensities (large → small: black → white), and the contours indicate iso-intensity lines at 0.1%, 5%, 10% of the maximum value. The white dot indicates the centroid of the image. The iso-line at 10% of the maximum value, which encloses the PSF coefficients with high values, can be well approximated by an ellipse whose aspect ratio increases with depth. **D:** The iso-surface (threshold 10% of the maximum) is shown together with its bounding box, which is an approximation of the support of the PSF. Scale bar: 0.6μ for **A, B, C**; and 0.2μ for **D**.

PSF Support Estimation method

The proposed method is based on the prior knowledge of the geometry of the true object being imaged, a cylindrical object in the case studied here. If the true object is convolved with an anisotropic PSF during image acquisition, the geometry of the (boundaries of the) object changes and the imaged object is deformed depending on the aspect ratio of the blurring PSF. The confocal PSF exhibits a larger support along the optical axis of the microscope (Z-direction) in comparison to its support in the radial direction, see Fig.5.3, and this causes more deformation along the axial direction. A graphical illustration of the proposed blur support estimation method is shown in Fig.5.4. In order to calculate the support of the PSF, an estimate of the true object is needed. This estimated object is found by summing the intensity of all the pixels belonging to the imaged object and mapping this sum of intensity into a circle of constant intensity. The diameter of the circle is determined such that the sum of intensity of all the pixels inside the circle is equal to the sum of the intensity of all the pixels belonging to the imaged object divided by the maximum intensity value. An estimate of the support of the PSF is calculated as the difference between the major and minor axes lengths of the boundary of the imaged object and the diameter of the circle.

Consistency Analysis

The boundaries of both the blurred object and the reblurred (the convolution of the estimated PSF and the estimated true object) are compared. If both boundaries are not the same, the estimated PSF is corrected untill the two boundaries are congruent, see Fig.5.5.

Simulation

In this section, we simuate the propsed method for the estimation of the support of the PSF using artifical data. The true image is shown in Fig.5.6. The support of the PSF is varied from 2 to 20 pixels, and the blurred image is corrupted with noise to emulate the observed image. The proposed method is then applied to the observed image to estimated the support of the PSF. The estimated support plotted against the true support of the PSF is shown in Fig.5.6. As the figure reveals, the estimation accuracy is within ± 1 pixels over the whole range.

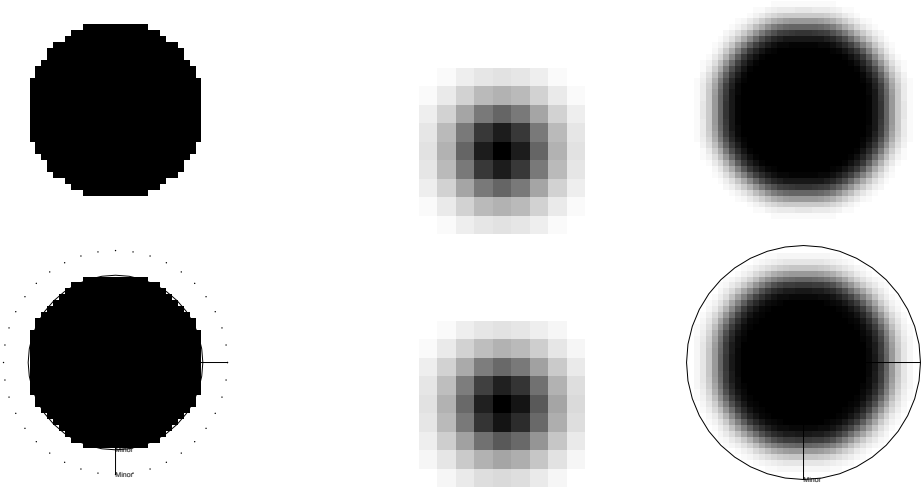


Figure 5.4: Illustration of the PSF support estimation method. *Top row:* blurring during the image acquisition; the test image in the left panel is blurred by the PSF in the center panel. The blurred object is shown in the right panel. *Bottom row:* estimation of the PSF support and the true object; the estimted object is shown in the left panel, and the estimated PSF is shown in the center panel. The reblurred object (the convolution of the estimated object and the estimated PSF) is shown in the right panel.

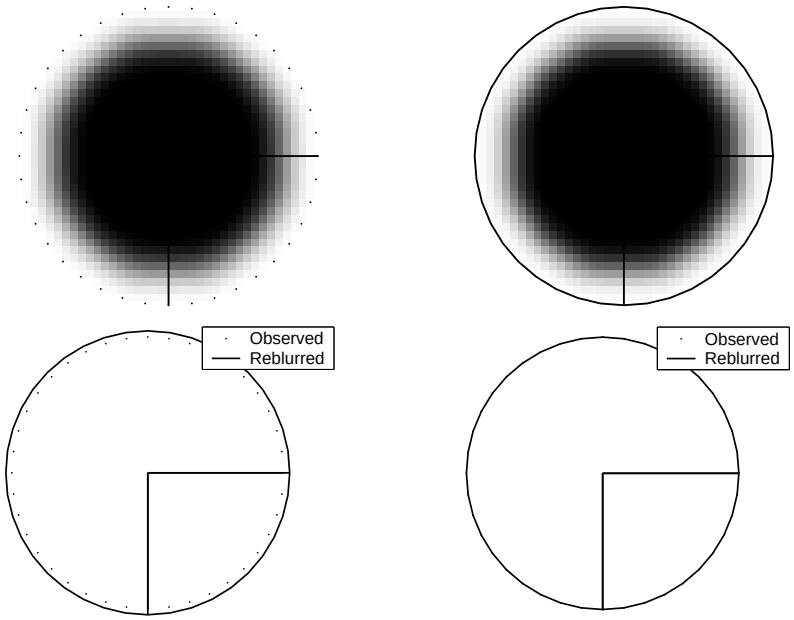


Figure 5.5: Consistency test. The boundaries of the reblurred object are compared to the boundaries of the blurred object. The estimated PSF is corrected such that the two boundaries are congruent.

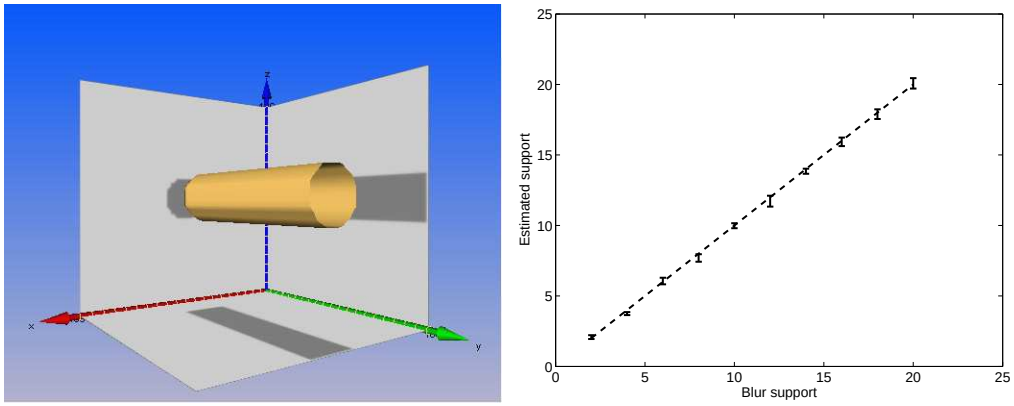


Figure 5.6: Simulation of the PSF support estimatio method. The left panel shows the test image used in the simulation. The test image is blurred with a PSF of support ranging from 2 to 20 pixels. Afterwards, the blurred image is corrupted with noise to emulate the observed image. The support of the PSF is then estimated using the proposed method. The right panel shows the estimation results.

Some problems face the application of this method in the case of confocal images. First, neuron processes are oriented in random directions and not necessarily parallel to the coordinate axes. Second, background noise might be present in the acquired image, which hampers the extraction of the boundaries of neuron processes. In the following, the method is extended into a general algorithm to handle these problems. The algorithm is outlined as follows:

1. Select a slice that is parallel to the Z axis (either XZ- or YZ-slice), and suppress the background noise. This can be done by setting all image intensities which are below a chosen threshold to zero, see Fig.5.7. The threshold is selected as the minimum intensity value that removes the back ground noise. This value of threshold can be determined from the histogram of the image.

2. Find the boundaries of each object in the thresholded image. This is done using the contour following algorithm described in [24].

3. For each object, fit an ellipse to the boundary (for the details of ellipse-fitting see [15]; [10]), find the centroid of the area and calculate the sum of intensity values of all pixels inside the object.

4. Correct for the tilt of the neuronal branches. The cross sectional shape of

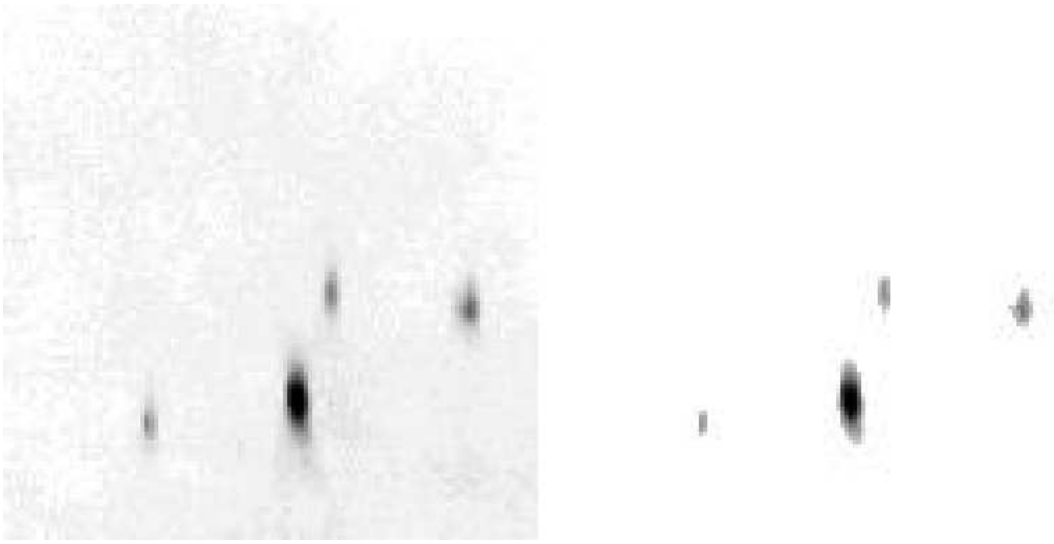


Figure 5.7: An YZ slice of the image stack, before (left) and after (right) thresholding. A threshold of 25, which is about 10% of the maximum gray value is used.

the neuron is assumed to be a circle only in the plane orthogonal to the trace of the neuron. For tilted branches, a geometrical correction is needed (see Fig.5.8). The correction done is as follows:

- Calculate the trace of the neuron, i.e. the vector which connects the-centroid of the object in the current slice and that of the object in the neighboring slice. This vector is the normal to the cross-sectional area of the neuron, see Fig.5.8.
- This normal vector, together with the centroid of the object, defines a plane. The shape of the cross-sectional area of the neuron in this plane is assumed to be a circle. Hence, a dummy circle (with certain radius) is drawn in this plane.
- Project the points of this circle into the plane of the current slice (XZ or YZ plane).
- Fit an ellipse to the projected points and calculate the aspect ratio. This ellipse represents the corrected cross-sectional shape.

5. Apply the constraint of the conservation of image intensity. We assume that the true object has a circular cross-section, and that blurring causes

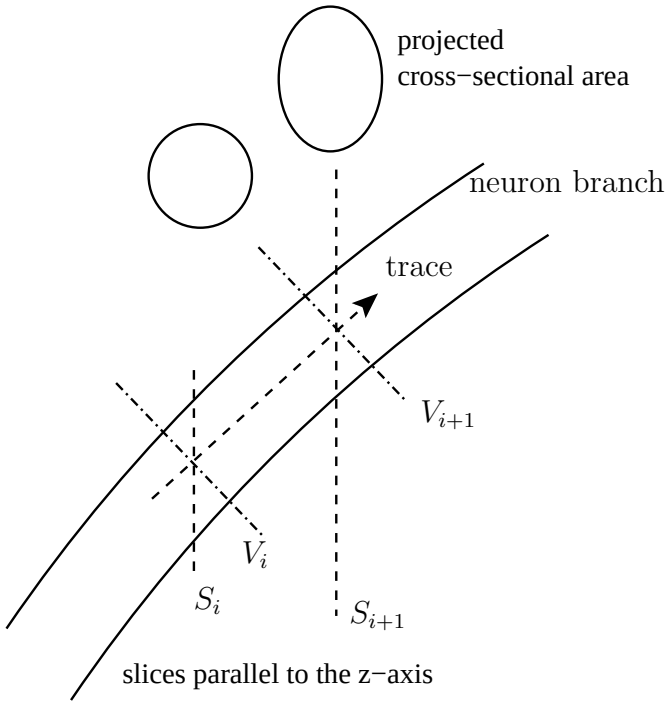


Figure 5.8: Illustration of the geometrical correction. The figure shows a contour of a branch of the neuron (solid line), whose trace is not normal to the plane of the slices S_i and S_{i+1} (dashed lines labeled S_i and S_{i+1}). The cross-section of the neuron is circular in the planes normal to the trace of the neuron (dashed lines labeled V_i and V_{i+1}).

a redistribution of the total intensity over a larger number of voxels that constitute the blurred object. The ellipse calculated in step 4 is then scaled, without changing its aspect ratio, so that the area of this ellipse (number of voxels) is equal to the total intensity of the blurred object in this slice divided by the maximum intensity in the acquired image. This gives the estimated object.

6. Compare the images of the estimated and the blurred (observed) object within the slice (Fig.5.9). The difference between the extent of these two ellipses along the coordinate axes gives an estimation of the support of the PSF.

Fig.5.10 shows the results obtained with the above method for the data set described in section 5.2.1. The algorithm described in the previous subsection is repeated using branches with different sizes to measure the variation of the PSF

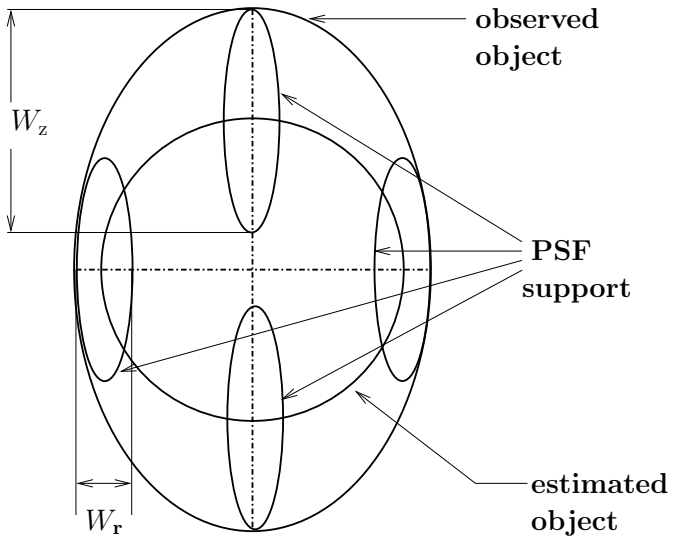


Figure 5.9: Fitted ellipses for the observed and the estimated (unblurred) objects within a slice parallel to the Z-axis. The elongation of the object along the coordinate axes gives an estimation of the support of the PSF. The size of the PSF support in the radial and axial direction is measured by W_r and W_z, respectively.

size with depth. The range (support) of the PSF in the axial and the radial direction is plotted as a function of depth. As can be seen, the PSF support in the radial direction is much smaller than that in the axial direction and the PSF support in the axial direction changes rapidly with depth, by almost a factor of 2.

5.2.3 Deconvolution Algorithm

The blur varies with depth and this means that most of the existing deconvolution algorithms that are based on a depth-invariant PSF model cannot correct the anisotropic distortion over the whole image stack. The usual approach is to divide the image stack into sub-blocks along the Z axis, then each block is separately deconvolved using different PSF. These sub-blocks can then be recombined to form the restored image [35]. We suggest using a space-variant deconvolution method, which can be implemented in a computationally efficient way for the case of a separable PSF. Our approach to blind deconvolution thus consists of estimating the range of the PSF as described in section 5.2.2, constructing a simplified space-variant model of the PSF, and then applying depth-variant deconvolution to restore the image stack as whole. In the remainder of this section, we describe the whole

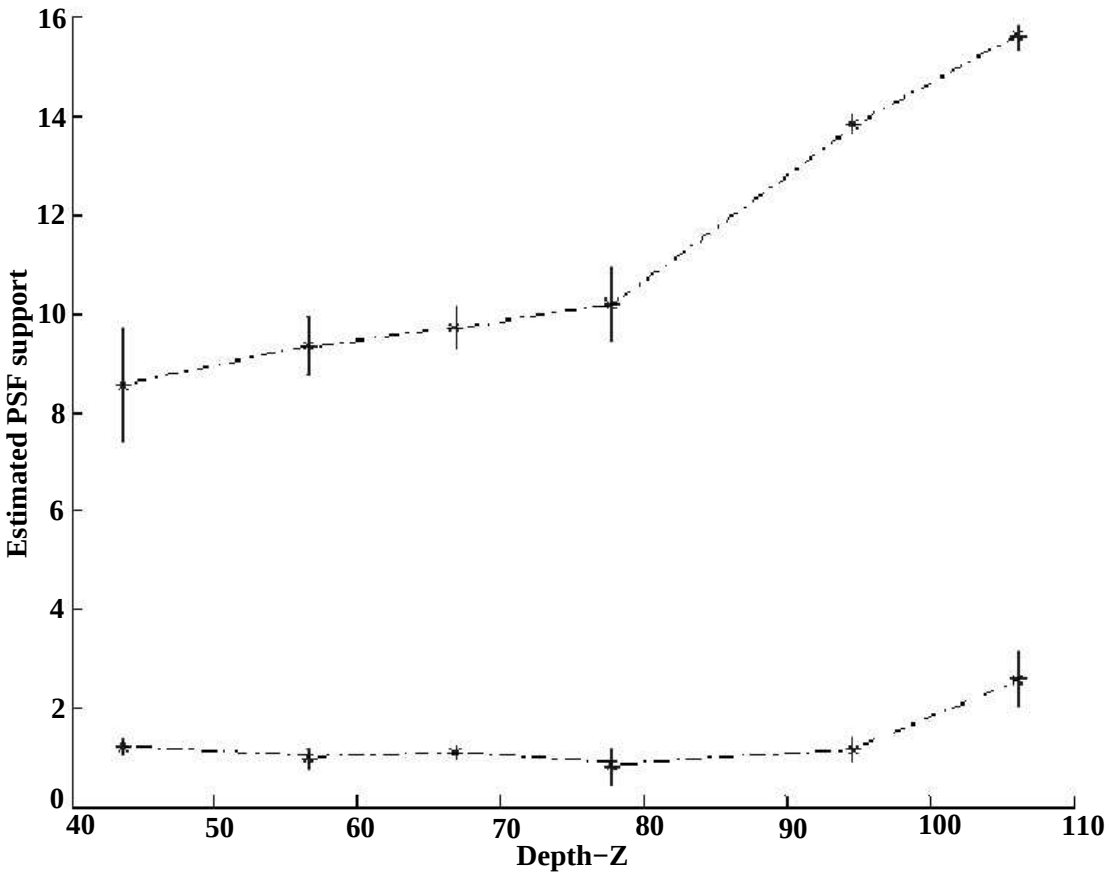


Figure 5.10: The estimated support of the PSF (measured in voxels) as a function of depth (in voxels). Upper curve: axial support; lower curve: radial support. The voxel size is 0.98μm. The number of measurement points is not the same for each error bar. From right to left, the number of points are, 75, 38, 13, 6, 5, 3. The mean is calculated as the average of the measurement points around the specific depth, and the standard deviation is marked by the size of the error bar.

deconvolution algorithm.

Imaging Model

The observed image is modeled as a blurred noisy version of the true object,

$$g(X) = \mathcal{N}[(f * h)(X)], \tag{5.1}$$

where g is the observed image, f is the original object, h is the PSF, $\mathcal{N}[.]$ denotes the noise function, and X is the 3D position vector $X = (x, y, z)$. The symbol $*$

denotes the space-variant convolution operation,

$$(f * h)(X) = \sum_{X'} f(X')h(X, X'). \tag{5.2}$$

The noise is assumed to follow the Poisson statistics:

$$\Pr(g|f, h) = \frac{(f * h)^g}{g!} \exp[-(f * h)]. \tag{5.3}$$

Mathematical Model of the PSF

In section 5.2.2 we have estimated the support of the PSF. In order to apply deconvolution we now need a model for the PSF in order to calculate its values at each point inside this support. The significant PSF coefficients ($\geq 10\%$ of the maximum coefficient) are contained in an elliptic contour (Fig.5.3 A, B and C), and in an ellipsoid (Fig.5.3 D). This motivates the following model for the PSF,

$$h(X) = \begin{cases} \frac{\Phi(X)}{\sum_X \Phi(X)} & , X \in W \\ 0 & , \text{elsewhere,} \end{cases}$$

where W denotes the support of the PSF and

$$\begin{aligned} \Phi(X) &= \Phi_r(X)\Phi_z(X) \\ &= e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma_r^2}} e^{-\frac{1}{2}\frac{z^2}{\sigma_z^2}}. \end{aligned} \tag{5.4}$$

The parameters σ_r, σ_z control the size of the PSF in the radial and axial directions. They are related to the estimated values of the radial and axial support W_r, W_z of the PSF via

$$\sigma_r = W_r/4, \sigma_z = W_z/4.$$

This simplified model of the PSF also satisfies the constraints and assumptions on the PSF function, in particular the non-negativity, the circular symmetry in the radial plane and the unity sum of the PSF. Finally, this simplified model is a good candidate for an efficient implementation of the space-variant deconvolution algorithm.

Although, the RBF model proposed in the previous chapter has an improved performance, due to the reduction of the number of unknowns, in comparison to the existing deconvolution methods, however, in the space-variant case, all the

parameters of the model (RBF) must be estimated at every pixel in the image which is quite a large number. In the following, we use the simplified model of the PSF defined above and extend the deconvolution method outlined in the previous chapter into a space-variant deconvoluion method. This simplified model of the PSF corresponds to an RBF model that consists of only one basis function with the parameters chosen as, $\lambda_r = \sigma_r/2$ and $\lambda_z = \sigma_z/2$.

Deconvolution Method

Given the model Eq.5.4 of the PSF, the image is deconvolved using the MAP-method. The posterior probability of f is given by

$$\Pr(f|g) = \frac{\Pr(g|f)\ \Pr(f)}{\Pr(g)}. \tag{5.5}$$

The maximum *a posteriori* (MAP) approach is to find the image f which maximizes the numerator of Eq.5.5. Taking the negative of the logarithm of $\Pr(f|g)$ leads to the risk function $R(f,h)$ which is minimized to obtain a good solution. The risk function is given by

$$R(f,h) = \sum_X [(f*h)(X) - g\ln(f*h)(X)] + \gamma_f \Omega_f(X), \tag{5.6}$$

where γ_f is a regularization coefficient that controls the tradeoff between the fidelity to the observation and the smoothness of the solution, and $\Omega_f(X)$ is the negative logarithm of the prior probability $\Pr(f)$ and can be interpreted as a regularization term. This risk function is the same as the risk function defined in Eq.4.8 but extended into the space-variant case here. The Good's roughness penalty [25], [48] is used for regularization, which corresponds to assuming a prior of the form of Eq.4.9. To ensure the non-negativity of the estimate f, we use a square parameterization $f(X) = \alpha^2(X)$. For this parameterization the roughness penalty function is $\Omega_f = -4\langle\alpha, L(\alpha)\rangle$, where L denotes the Laplacian. The restoration problem now reduces to finding $\hat{f}$ that minimizes the risk function. This minimization is performed iteratively using a gradient descent method with the update rule given by Eq.4.11, with

$$\frac{\partial R(f,h)}{\partial \alpha} = -2\alpha\left(\frac{g}{f*h} - 1\right) * h(-X) - 4\gamma_f L(\alpha). \tag{5.7}$$

Choice of the Regularization Parameter

Various methods exist in literature for the selection of the regularization parameter (see [41]). The most well studied methods are the cross-validation and the L-curve method. One problem of cross-validation is that it does not perform well with correlated noise [41]. The L-curve method is more robust for correlated noise, but some convergence problems were reported in [50]. However, although both methods are computationally complex, they may lead to non-optimal value of the regularization parameter γ, especially for large data sets. These methods are useful to indicate a suitable range for γ. Finally, one resorts to visual inspection of the restored image to choose the optimal value of the regularization parameter. The PSF is calculated according to the simplified mathematical model discussed above, and the estimated PSF support from section 5.2.2. The observed image stack is used as the initial guess for f. The learning step size and regularization coefficient are fine-tuned experimentally using previous experience in a similar work [20] to find the range of suitable values of these parameters.

Stopping Rule and Validation of Results

From the neuroscientist's standpoint, it is important to determine the morphology of all dendrites and axons, their lengths, surface area and volumes. Clearly quantitative evaluation of the quality of the restoration should be based on these characteristics. However, ground truth of these characteristics is not available. The best that we can do is to follow the assumption that the cross-section of neuronal branches is, on average, circular in shape. Hence, the algorithm is run until this assumption is valid, see Fig.5.11. A quality criterion could be a measure of the circularity of the cross-sectional shape of the neuronal branches. Restoring a nearly circular shape of the cross-section means that, we have actually improved the axial resolution to a level comparable to the radial one, which is a good improvement. Another issue is how to automatically stop the restoration iteration. A measure of circularity (aspect ratio = major axis length/minor axis length) is calculated at different test points in the image stack, and can be used to stop the deconvolution iteration. For the raw data set, the average aspect ratio was 1.85, after 10 iterations of the restoration algorithm the aspect ratio was 1.18 and after 20 iterations it was 1.11. The aspect ratio improves with increasing number

of iterations. Ideally, the iteration should proceed until a circularity measure of 1.00 (perfect circle) is obtained. However, with increasing number of iterations the improvement in the aspect ratio is small (it reaches 1.08 after 30 iterations) and some artifacts appear in the restored image stack (see, e.g., arrowheads in Fig.5.14). These artifacts are not captured by the circularity measure and might resort to visual inspection to stop the restoration iteration. However, the prob-

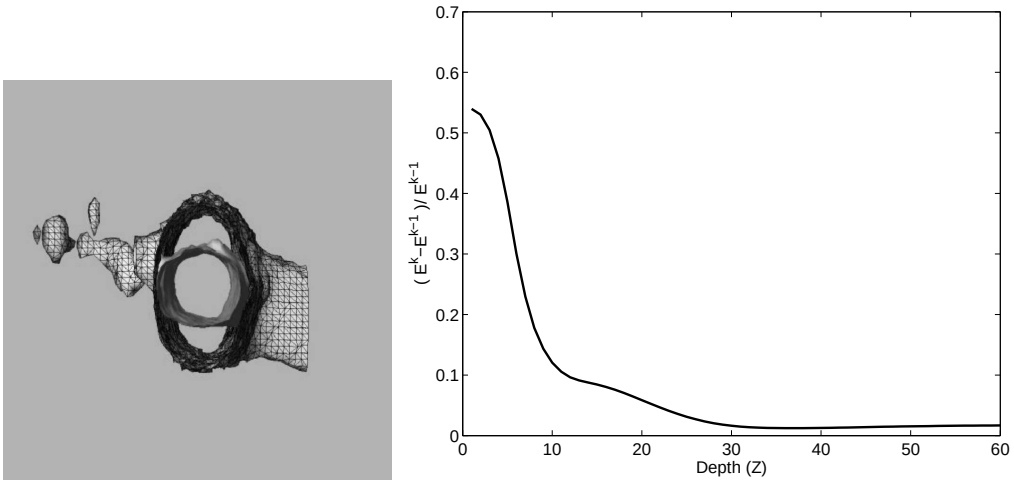


Figure 5.11: Illustration of the stopping rule. The iso-surface of the raw (hatched) and the restored (shaded) image of one neuron branch. The algorithm is stopped when the cross-section is, on average, circular in shape.Alternatively, the algorithm can be run until convergence. If the estimated PSF is close to the truth, then the restoration algorithm leads to an estimated image which is as close as possible to the true object.

lem of this approach is that, the circularity of the true object is not known. An alternative approach is to let the algorithm run until convergence. We used the criterion,

$$(E^k - E^{k-1})/E^k, \tag{5.8}$$

where, $E^k = ||g - H\hat{f}||$ is the MSE at iteraion k, to denote convergence.

5.3 Results

In this section, results are presented to illustrate the effectiveness of the deconvolution algorithm. The results show that the anisotropic distortion can be removed

with a few number of iterations of the suggested deconvolution algorithm. The deconvolution algorithm is implemented in C++ and compiled for a Silicon Graphics Octane dual processor workstation (R10K, 195MHz; 896 MB and 1,024 MB RAM). The execution time is about 6 minutes for one iteration of the algorithm for the data set described in Fig.5.1. We tested several values for the parameters of the new method including learning step size, number of iterations, and regularization coefficient. For all the results shown in this chapter we have used a regularization coefficient $\gamma = 0.0015$ and a learning step size $\eta = 0.15$. Fig.5.12 shows the projection view of the restored image stack after 20 iterations. It can be seen in this figure that the blur in the axial direction is reduced to a great extent. The

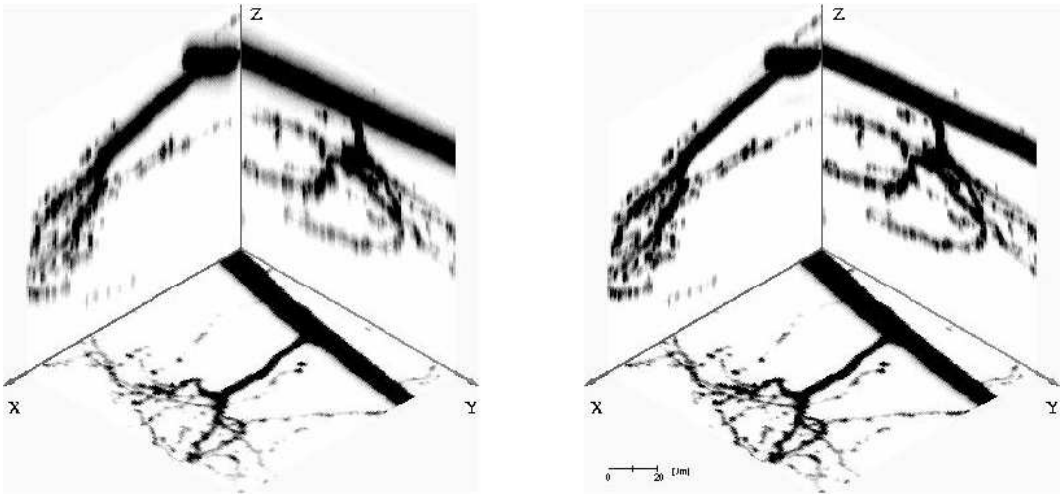


Figure 5.12: Maximum intensity projection view of the observed (left) and the restored (right) image stack. Gray values denote image intensity (large → small, black → white)

rendered iso-surfaces, using a threshold value of 25 (the minimum threshold above the noise level), are shown in Fig.5.13 for both the restored and the observed image stacks. The restored image is more suitable for subsequent analysis than the raw one. Moreover, the fine details of the neuron are enhanced in the restored stack (cf. arrowheads in Fig.5.13), which is desirable for neuron tracing algorithms and proper visualization. Fig.5.14 illustrates the restoration results after 2, 5, 10, 30 iterations of the deconvolution algorithm. As can be seen, the aspect ratio of the neuron branches improves with increasing number of iterations. The required number of iterations to obtain a reasonably circular cross-section of the neuron branches depends on the step size η (for the gradient descent algorithm used for

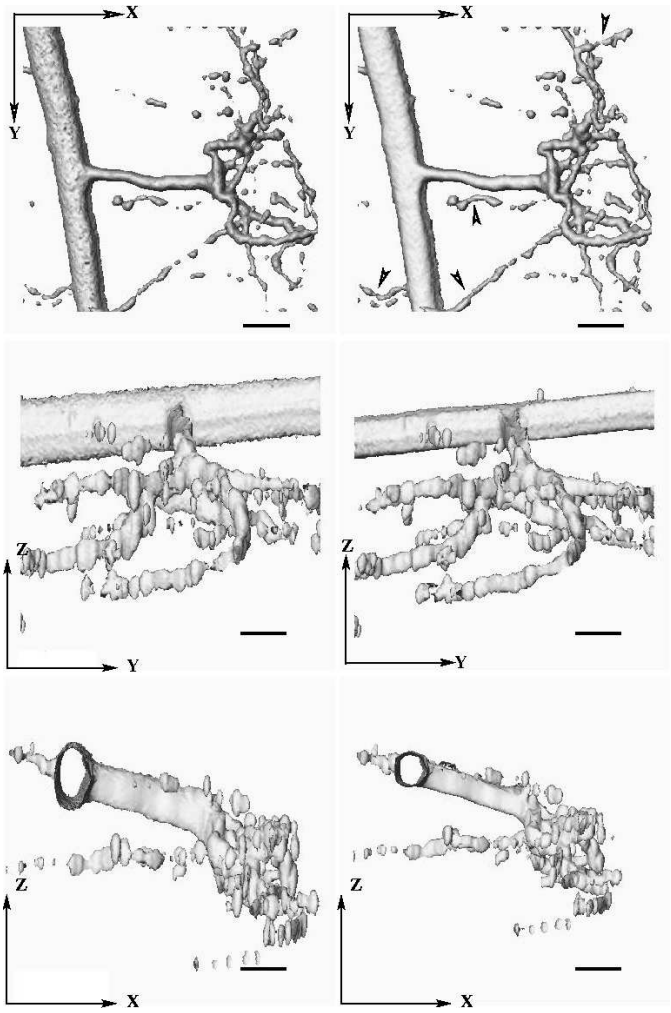


Figure 5.13: The rendered iso-surface of the observed (left column) and the restored (right column) image stack after 20 iterations. The figure shows the XY view (first row), the YZ view (center row), and the XZ view (bottom row). The iso-surfaces are constructed using a threshold value of 25. The geometric distortion is reduced and the fine details of the neuron structure are enhanced (cf. arrowheads in the right panel at the top). Scale bar: 16μm.

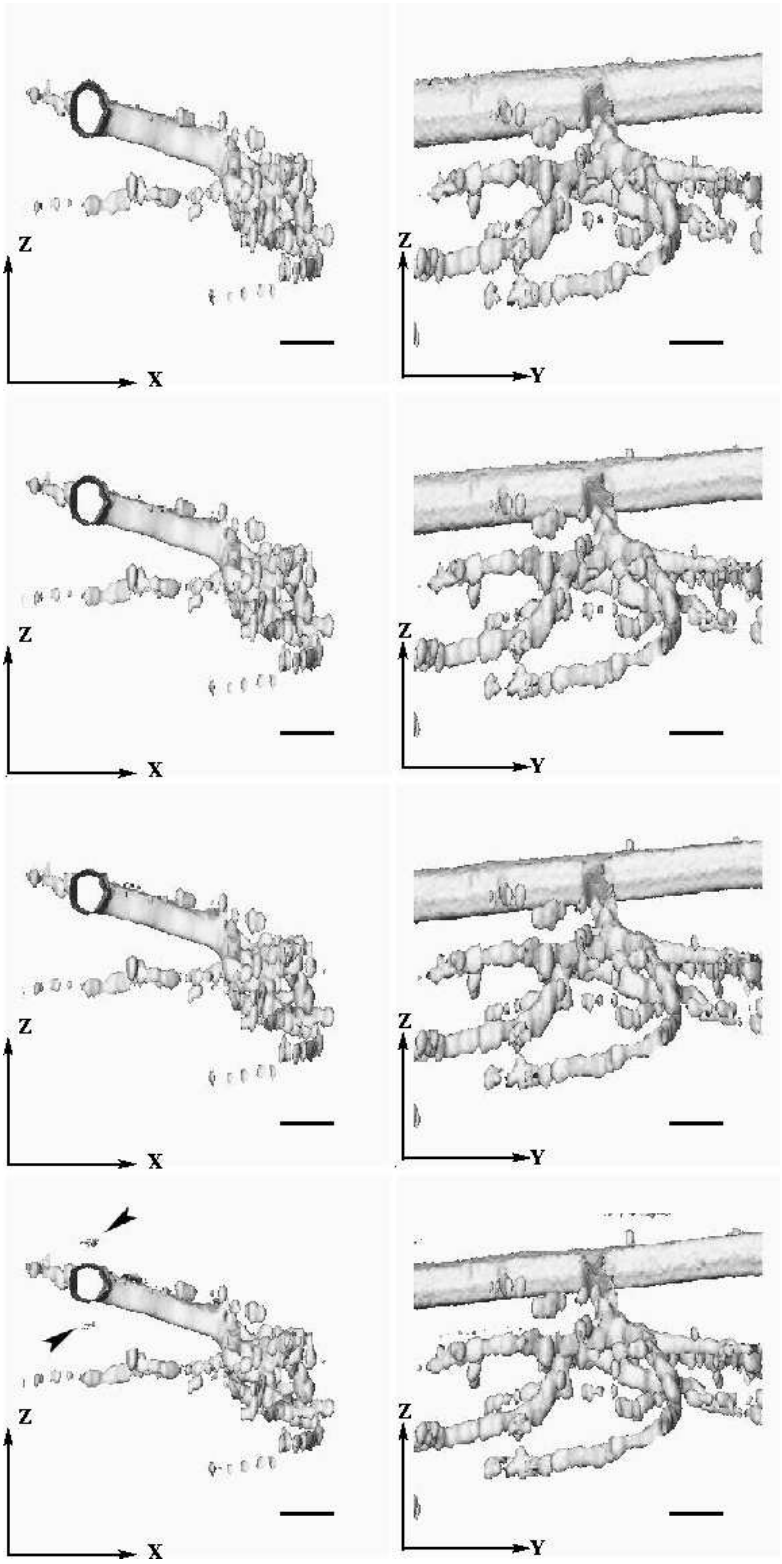


Figure 5.14: Illustration of the convergence properties of the algorithm. *Left column*: XZ view of the iso-surface; *Right column*: YZ view. From *top* to *bottom*: results after 2, 5, 10, and 30 iterations. Scale bar : 16μm. The average circularity improves from 1.85 for the raw image into 1.11 for the restored image.

the minimization of the risk function) and the value of the regularization coefficient. Without proper choice of the regularization coefficient, the algorithm may introduce artifact (cf. arrowheads in Fig.5.14). Using a large η leads to the rapid introduction of artifacts which will not be corrected in the following iterations.

5.3.1 Validation

In this experiment, the PSF estimated in section 5.2.2 is used to apply the deconvolution algorithm to other parts (cuts) from the same data set. Those cuts are different from the part (which is given the name Pow2) used to estimate the PSF. The whole data set is of size (512×512×138). The part of the data set used so far is of size (128×128×128). The results of the first cut (which is given the name Pow3, is of size 180×119×135) are shown in Fig.5.16. The results of the second cut (which is given the name Pow4, is of size 217×160×135) are shown in Fig.5.17.

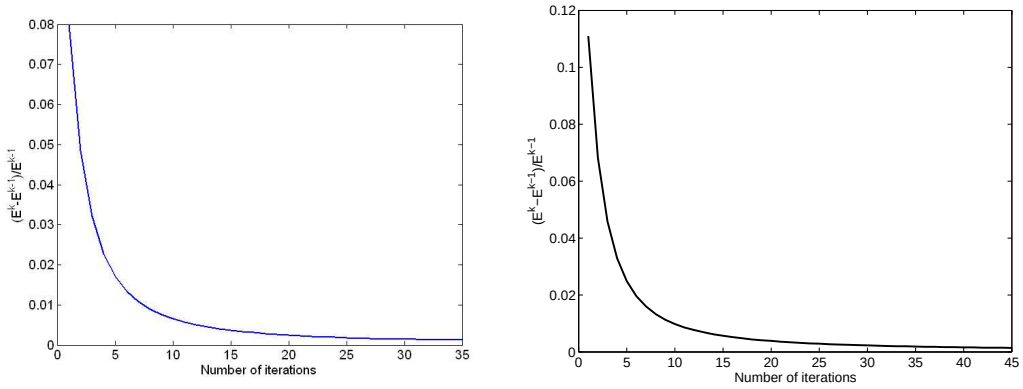


Figure 5.15: Convergence rate of the restoration method for the Pow2 (*left*) and the Pow4 (*right*) data set.

5.3.2 Experiments with other data sets

So far, we have used a single confocal data set to evaluate the proposed methods. In this section the proposed restoration method is evaluated using data sets collected under different settings. All the preparations used in this section were neurobiotin/streptavidine staining [3].

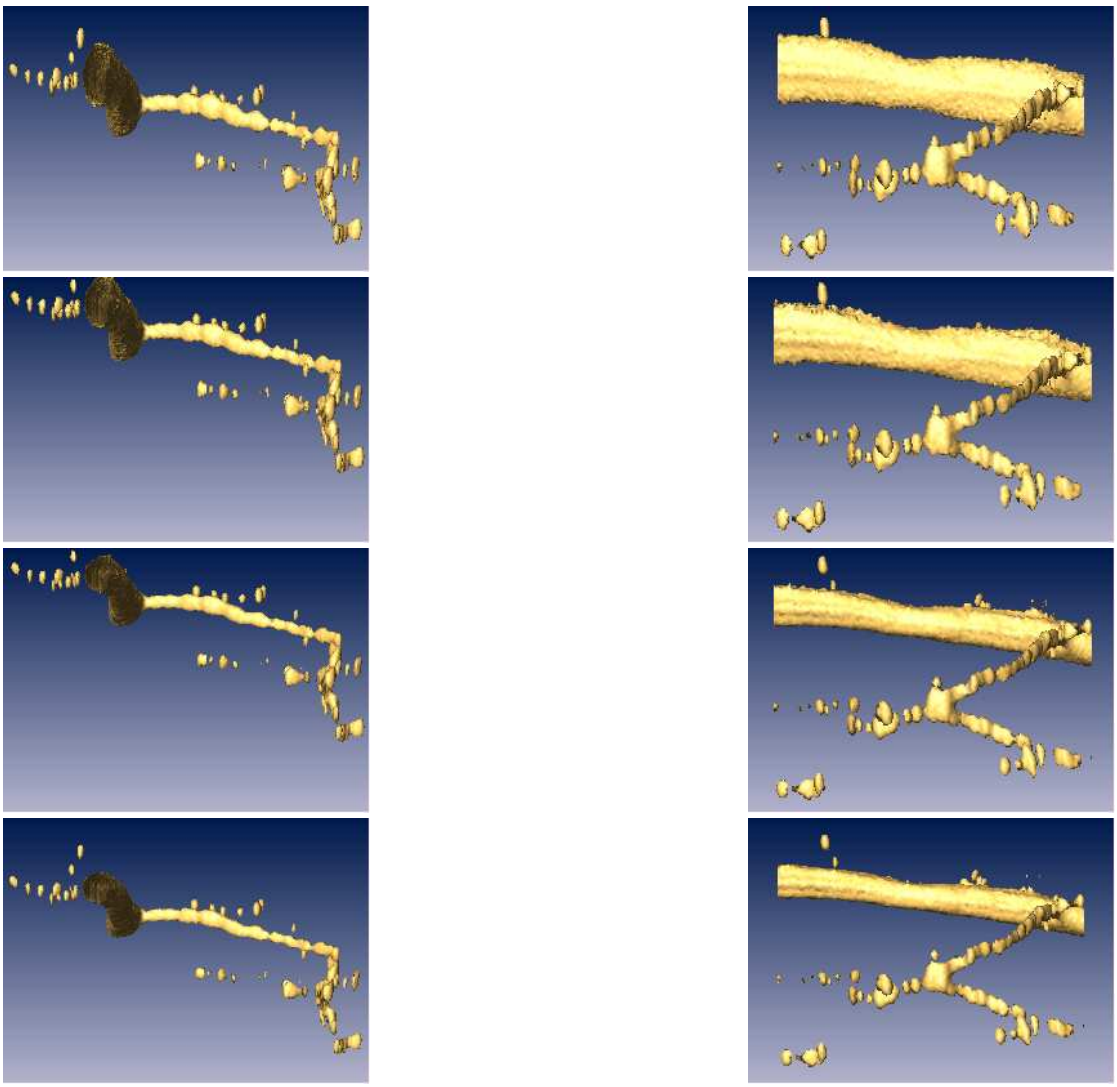


Figure 5.16: Convergence of the algorithm for the (Pow3) data set. *Left column*: XZ view of the iso-surface; *Right column*: YZ view. From *top* to *bottom*: Original, results after 5, 15, and 35 iterations, respectively. The restored image exhibits an improved average circularity of 1.1223 in comparison to an average circularity of 2.12 for the raw image.

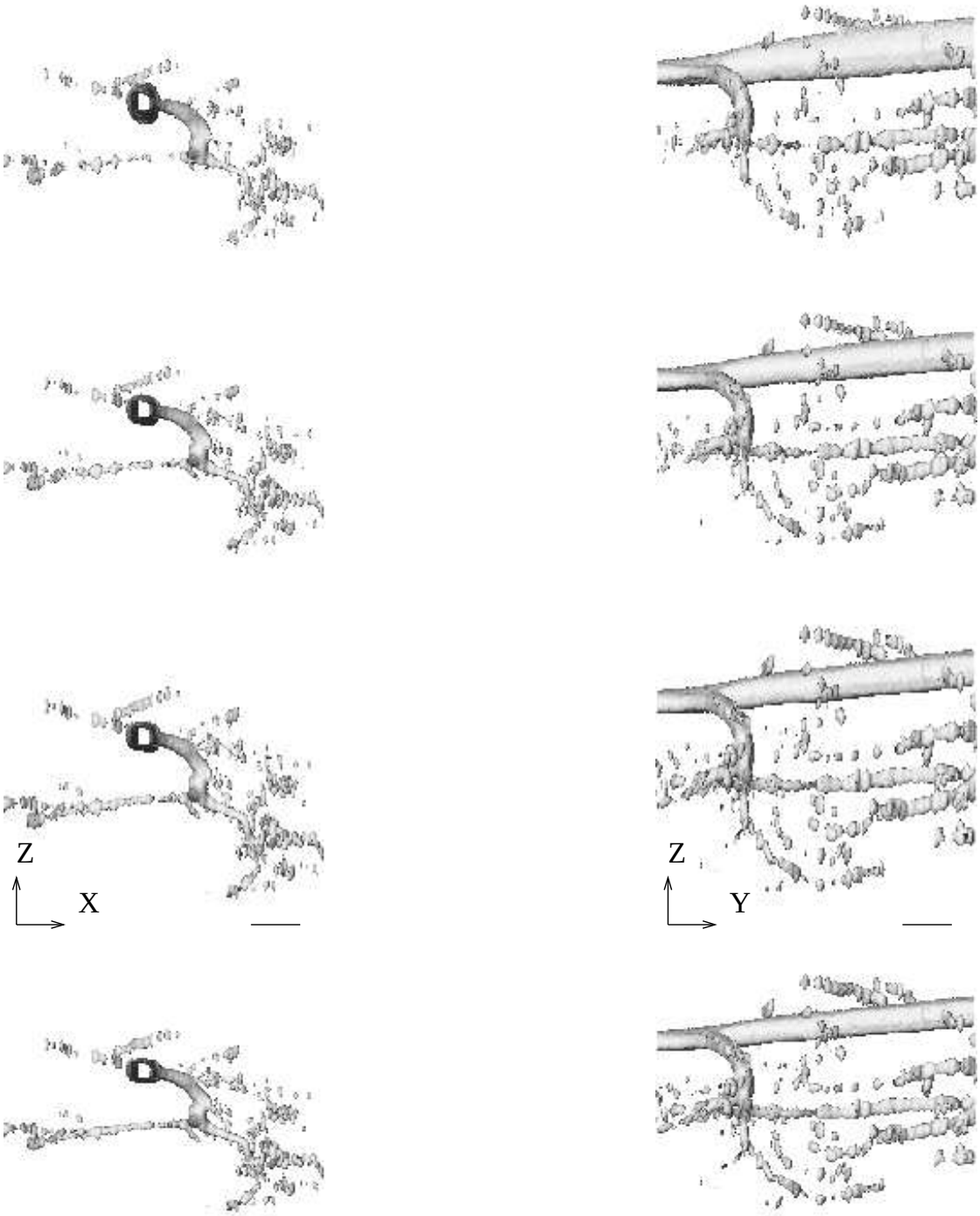


Figure 5.17: Convergence of the algorithm for the (Pow4) data set. *Left column*: XZ view of the iso-surface; *Right column*: YZ view. From *top* to *bottom*: Original, results after 5, 15, and 25 iterations, respectively. The restored image exhibits an improved average circularity of 1.1269 in comparison to an average circularity of 1.923 for the raw image.

The Trio data set

This data set is of a sensory neuron of the locust (512×512×130 voxels in size, corresponding to 0.38×0.38×0.76 μm voxel size). This data set, given the name Trio, is acquired using an objective 50× oil immersion lens. There is very little background noise in this image in comparison to the other data sets. The iso-surface of the raw data set is shown in Fig5.18. As can be seen the neuron processes are not significantly elongated in the direction of the optical axis of the microscope, because of using an oil lens that reduces the refractive index mismatch. The PSF support is estimated using the propsed method, the estimated support is shown in Fig5.18. The estimated support reflects the less severe anisotropy of the PSF in comparison to the case of an air lens. The raw confocal image is restored using the restoration method and the iso-surface of the restored image is shown in Fig.5.18. As the figure shows, in the restored image the geometric distortion is significantly reduced. We emphasis that the restoration algorithm performs only cautious corrections, and that none of the neuron processes is mapped into a cylinder, rather than the distortion is removed to obtain a reasonable estimate of the true object.

The Ww42Cut2 data set

Finally, another neuron data set, given the name Ww42Cut2, is used to evaluate the methods proposed here. The data set consists of 512×512×20 voxels with a voxel size of 1.465×1.465×1.994 μm and acquired with an 10×0.4 objective lens. The iso-surface of the raw and the restored data set is shown in Fig.5.19. This data set includes part of the soma of the neuron and only one dendrites. The support of the PSF is estimated only at the depth of the dendrites. However, all the structures in the image are nearly at the same depth and space invariant restoration is used in this experiment. The iso-surface of the restored image is shown in Fig.5.19. Contour plots of an XZ section through both the raw and the restored images are shown in the same figure. These plots show how the structrues in the restored image are deblurred (reflected as dense isolines in the plots of the restored image compared to those of the raw image.

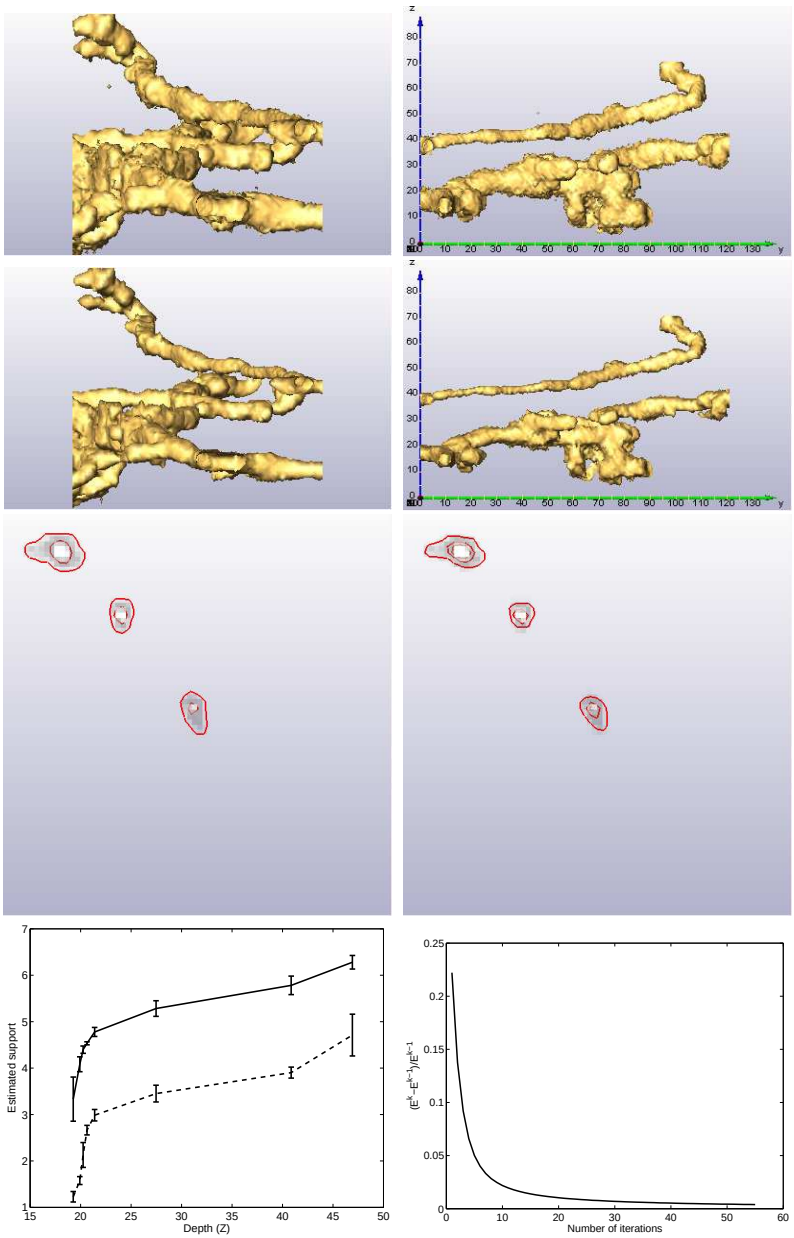


Figure 5.18: Restoration of the Trio data set. This data set is scan of a sensory neuron of the locust, using an 50× oil immersion lens. The consists of 130 image of 512×512 pixels, with a voxel size of 0.38×0.38×0.76 μm. *Top row:* the iso-surface of the raw confocal image constructed using the minimum threshold above the nois elevel; *left:* the XZ view and *right:* the YZ view. *Second row:* the iso-surface of the restored image using the same threshold; *left:* the XZ view and *right:* the YZ view. *Third row:* contour plots of an XZ section of (*left*) the raw, and (*right*) the restored data sets. *Fourth row:* the estimated support of the PSF (*left*), the solid line is the axial support and the dashed line is the radial support; and the rate of convergence (*right*).

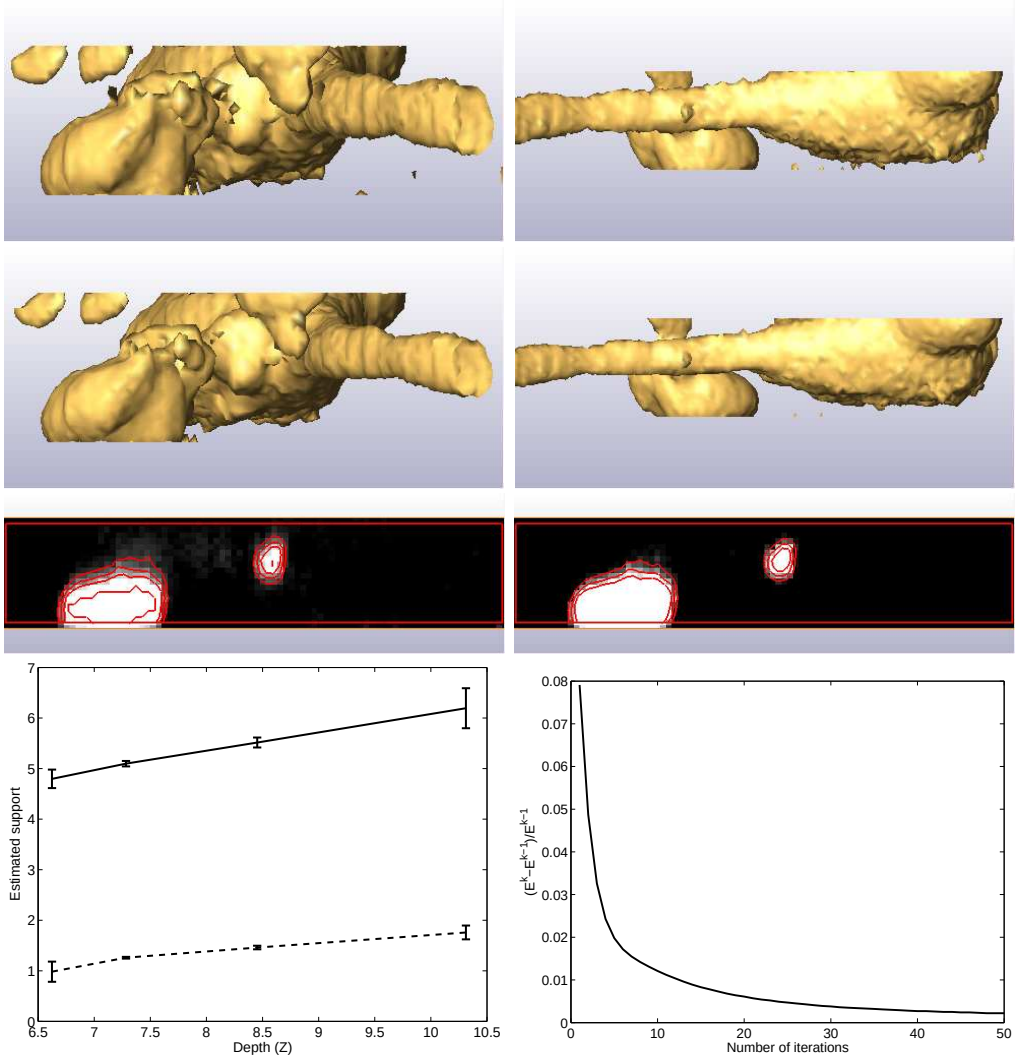


Figure 5.19: Restoration of the Ww42Cut2 data set. This data set is acquired using an 10×0.4 lens, and consists of 512×512×20 voxels in size, with a voxel size of 1.465×1.465×1.994 μm). *Top row:* the iso-surface of the raw confocal image constructed using the minimum threshold above the noise level; *left:* the XZ view, *right:* the YZ view. *Second row:* the iso-surface of the restored data set using the same threshold; *left:* the XZ view, *right:* the YZ view. *Third row:* contour plots of an XZ section of (*left*) the raw, and (*right*) the restored data sets. *Fourth row:* the estimated support of the PSF (*left*), the solid line is the axial support and the dashed line is the radial support; and the rate of convergence (*right*).

5.4 Discussion

In this chapter, we proposed a restoration method to correct the geometric distortion artifacts that are prominent in confocal microscopy images of neurons. We also proposed a new estimation method for the support of the PSF, that is based on information extracted from the image stack itself. There are other approaches that have been suggested in literature to extract a measure of the blur from the observed image. Most of these approaches (e.g., [9], [36]) are based on the image of an ideal edge. To estimate the PSF differential operators are used which are very sensitive to noise in the observed image. However, for a stable evaluation of these operators, a smoothing filter of an appropriate scale is usually used. This scale should be adaptively selected according to some objective [9]. Although some methods were proposed to select the optimum scale [32], [28], these methods either require the knowledge of the amount of noise, or a proper edge detection formalism. Our approach for PSF estimation, however, is based only on the shape information. Although it is specifically developed for neuron images, it can be applied in other cases where information about the shape of the object being imaged is available.

The simultaneous estimation of the PSF and the specimen function is computationally expensive and requires care to get a reliable estimate and to avoid suboptimal solutions. In this chapter we followed an efficient approach to apply blind deconvolution. We have opted to estimate the support of the PSF *apriori* from the observed image. Then, using the estimated support of the PSF as a constraint, we used a simplified mathematical model for the PSF function. This model is not based on microscope parameters, and is more suited to the blind deconvolution approach. The model is separable and this enables the efficient implementation of depth-variant deconvolution of the whole image stack. Thus one avoids the likely introduction of artifacts and the additional post-processing of dividing the image stacks into sub-blocks and the application of space-invariant deconvolution using a different PSF for each block.

5.4.1 Circumreasoning and PSF Estimation

Is the cross-section of different neuron branshes a perfect circle? Of course this is not always true, different neuron processes exhibit a smooth shape which changes along the length of the branch. However, for the blur estimation method we assume that the cross-sectional shape of the neuron processes is, on the average, circular. On the average means that, if subsequent cross-sections are averageed together, the resulting cross-sectional shape is well approximated by a circle. This assumption is biologically plausible and is used for neuron tracing methods that fits a generalized cylinder model into the structure of the neuron (excluding the soma). Moreover, the literature of neurbiology use the word "diameter" to describe the size of dendrites and aother neuorn processes. If the cross-sectional shape is not well approximated by a circle, then the word diameter would be inappropriate. In the experiment shown in this work, we processed several data sets, all the imaged structures are elongated in the direction of the optical axis of the microscope. If it is inappropriate to assume that the cross-section of the denrites is, on the average, circular; it is also inappropriate to claim that all neuron processes are elongated in the direction of the optical axis of the microscope. There is no biological explanation for this as well. The only reasonable explanation is that, this deformation is due to the anisotropy of the PSF of the microscope. This anisotropy increases in the presence of different refractive indices along the optical path of the microscope. This refractive index mismatch is unavoidable in the case of neuron images because of the presence of the tissue containing the cells in the optical path of the microscope during the process of image acquisition. This results in a depth-variant anisotropic PSF that incurs geometric distrotion of the imaged objects. For the restoration of the acquired images, accurate knowlede of parameters of the PSF, most important is its support (size), is needed. Because the optical characteristics of the tissue are unknown and probably not uniform over the whole specimen, it is more appropriate to estimate the support of the PSF from the acquired image. In this work, we introduced a new method for estimating the size of the PSF, based on the knowledge of the geometry of the object being imaged (the neurons). To estimate the PSF, we assumed the original cross-sectional shape of the neuron is circular. Since, this assumption is not perfectly true everywhere along the trace of neuron, several measurement points are used for each branch

and averaged together. Measurement sites that are located near to bifurcation points, were avoided. Moreover, the estimation is repeated using several branches of different size. The estimated PSF exhibits a larger size in the axial direction than in the radial direction, which is in agreement with the calculated PSF. The variation of the PSF size is relatively small for depths $< 60\mu$m, and when focusing deep inside the specimen, the PSF size shows significant dependence on depth (Z) of the structure being imaged. Although, thick specimens can be cut into smaller ones that are scanned separately and the resulting image stacks are then combined together, however, this is not always possible for all specimens and some specimens must be scanned as a whole to avoid the damage of the structure of interest. The results of the experiments presented in the current work show that it is possible to correct for the depth-dependent anisotropic deformation using restoration techniques, provided that the PSF can be estimated to a reasonable accuracy. Finally, we stress that the method introduced in this work are not a kind of a mapping method rather a restoration method. If the cross-section of the neuron is circular everywhere, then a mapping method that converts the ellipse (the deformed cross-section) into a circle (the true cross-section) would be a perfect agent. However, none of the restored images exhibits a perfect circle as the shape of the cross-section, rather a slowly varying shape which is approximately circular, on the average.

In summary, we have introduced a blind deconvolution method for the correction of geometric distortion in confocal microscope images of neurons. The method is based on the estimation of the blur from the geometric deformation that occurs during image acquisition. However, direct analysis of the acquired images without prior correction of these distortions could lead to incorrect conclusions. Hence, the proposed deconvolution method serves as an important pre-processing step for the proper visualization and extraction of accurate morphological features of neurons.

Chapter 6

Adaptive Regularization

Regularization is necessary to mitigate the illposedness of the image restoration problem. A less studied aspect of restoration algorithms is how to select the form of the regularization functional. The regularization functional imposes prior assumptions on the restored image. In the literature, both Tikhonov and TV regularization functional are widely used. Both functionals impose a smoothness constraint on the restored image. In this chapter, the impact of these regularization functionals on the characteristics of the restored image is studied. The geometrical behavior of regularization in different regions in the image is analyzed. Based on this analysis, an adaptive regularization technique is introduced. The performance of the adaptive technique is evaluated using both simulated and real confocal microscopy images.

6.1 Introduction

Imaging systems fail to transmit all the information about the true image into the observed image. Image restoration aims at recovering the underlying (true) image from a noisy blurred observation. The information provided by the observed image is, in general, not enough for an accurate estimation of the true image. In order to solve the restoration problem more information or assumptions about the structure of the true image is needed.

In the Bayesian approach, assumptions about the statistics of the image are incorporated as a prior probability density. In the (penalized) ML approach, prior knowledge is embedded in the penalty function used for regularization. Regardless of the approach used to formulate the restoration problem, the solution is found by constructing some functional $\mathcal{J}$ which upon minimized leads to the the solution of the problem. An estimate of the true image $\hat{f}$ is found by solving,

$$\hat{f} = arg \ \min_{f} \mathcal{J}(f, g) \tag{6.1}$$

The functional $\mathcal{J}$ is the sum of two terms,

$$\mathcal{J} = \mathcal{L} + \gamma V$$

Where, the first term measures the faithfulness of the estimate to the observed image, the second term is the regularization term, and γ is the regularization coefficient which controls the balance between the two terms. The functional $\mathcal{L} \propto -\log \Pr(g|f)$, and is dependent on the characteristics of the observation noise. In the case of Gaussian noise $\mathcal{L} = \int_{S_f}(g - h \otimes f)^2$. In the case of Poisson noise $\mathcal{L} = \int_{S_f}[h \otimes f - g \ln((h \otimes f)/g)]$ which is the I-divergence, (Kullback-Leibler measure) [7] between g and $h \otimes f$. In the case, where the information about the true image is qualitative in nature, e.g. smoothness, the regularization term V is formulated as some functional which upon being optimized achieves the stated goal. Among the well-known functionals are the Dirichlet integral and the total variation integral,

$$V_{TV}(f) = \int_{S_f} |\nabla f| \, dX \qquad V_{Dirichlet} = \int_{S_f} (|\nabla f|)^2 dX. \tag{6.2}$$

Where, S_f is the support of the image, and $|\nabla f|$ denotes the gradient of the image. The Dirichlet integral leads to the classical form of Tikhonov regularization. A

generalization of these functionals can be written as,

$$V(f) = \int_{S_f} \psi(|\nabla f(X)|) dX, \tag{6.3}$$

where $\psi : \mathcal{R}^+ \mapsto \mathcal{R}$, is a potential function (an even function, and $\psi(0) = 0$). With $\psi(s) = s$, $V(f)$ is the the total variation functional, and with $\psi(s) = s^2$ is the Dirichlet functional.

The minimization in Eq.6.1 can be solved by a variety of iterative methods, like the gradient descent. The gradient of $\mathcal{J}$ w.r.t. f is,

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial f} &= \frac{\partial \mathcal{L}}{\partial f} + \gamma \frac{\partial V}{\partial f} \\ \nabla \mathcal{J} &= \nabla \mathcal{L} + \gamma \nabla V \\ &= h \odot \left(1 - \frac{g}{h \otimes f}\right) - \gamma \, div \left(\frac{\psi'(|\nabla f|)}{|\nabla f|} \nabla f\right) \end{aligned} \tag{6.4}$$

Where, *div* denotes the divergence operator. This leads to an *adaptive regularization-based restoration algorithm*, depending on the form of the function ψ. The derivation of Eq.6.4 is given in Appendix A. In the following section, the geometrical behavior of both V_{TV} and $V_{Dirichlet}$ is studied, to pave the way for an adaptive form of regularization based on the formulation of a new function ψ.

6.2 Geometrical Behavior of Regularization

Now, lets have a closer look on the regularization term,

$$\nabla V(f) = - \, div \left(\frac{\psi'(|\nabla f|)}{|\nabla f|} \nabla f\right). \tag{6.5}$$

First, it should be noted that, the above formulation encompasses other techniques in image processing. For example, this regularization term is similar to the right hand side of the partial differential equation used in the formulation of anisotropic diffusion [34] [52],

$$\frac{\partial f}{\partial t} = div \left(c(|\nabla f|) \nabla f\right). \tag{6.6}$$

Where, c(.) is known as the diffusivity [34]. Note that if the diffusivity function is defined as,

$$c(|\nabla f|) = \frac{\psi'(|\nabla f|)}{|\nabla f|}, \tag{6.7}$$

the right hand side of Eq:6.6 becomes the same as $-\nabla V$.

To gain more insight into the behavior of the regularization term, Eq.6.5 needs to be further simplified. A decomposition of ∇V can be obtained by expanding the *div* operator, see Appendix B for a detailed derivation,

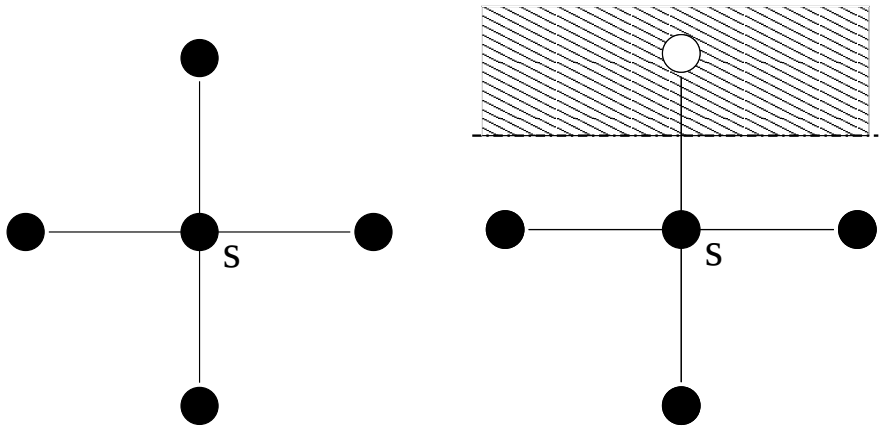
$$div\left(\psi'(|\nabla f|)\frac{\nabla f}{|\nabla f|}\right) = \left(\frac{\psi'(|\nabla f|)}{|\nabla f|}f_{\zeta\zeta} + \psi''(|\nabla f|)f_{\eta\eta}\right). \tag{6.8}$$

Where, η, ζ are unit vectors that are perpendicular to each other, ζ is the unit vector in the direction normal to the gradient, η is the unit vector in the direction of the gradient, and $f_{\eta\eta}$, $f_{\zeta\zeta}$ are the second derivatives in the direction of η and ζ, respectively. Since the regularization aims at imposing a smoothness constraint on the estimated solution $\hat{f}$, it is now clear that the selection of ψ affects the direction of the smoothing incurred by the regularization functional. Consider the case when $\psi(s) = s^2$. In this case, the regularization term reduces to

$$\nabla V(f) = div(\nabla f) = f_{\zeta\zeta} + f_{\eta\eta} = \nabla^2 f. \tag{6.9}$$

Where, ∇^2 is the second derivative (Laplacian) of f. This corresponds to the classical isotropic and space-invariant form of Tikhonov regularization. The isotropy can be seen in Eq.6.9, where the orthogonal decomposition reveals that the smoothness constraint is always enforced both in the direction of the gradient (across the edges) and in the direction normal to the gradient (along the edges). It should be noted that, imposing a smoothness constraint in the direction normal to the edges hampers the adequate restoration of (sharp) edges in the image. Note that in this case, Eq.6.6 reduces to the isotropic heat equation, $\frac{\partial f}{\partial t} = \nabla^2 f$. Let us now investigate the behavior of the Tikhonov regularization (Laplacian) in different regions in the image. In a homogeneous region in the image, the difference in intensity between the neighboring pixels is small and the Laplacian $\nabla^2 f$ is well estimated by the mean of the intensity differences between the pixel at the center s and its neighbors. This situation is depicted in Fig.6.1a, where each circle represents a pixel in a 4-connected neighborhood centered at s. Now consider the case when the pixel s is located near an edge Fig.6.1b. In this case, one (or more) neighbor of pixel s belongs to a different region and the sum of the intensity difference will be biased. Moreover, in the presence of correlated noise (like Poisson noise) the bias of the sum of intensity differences will increase. Hence, depending on the intensities of the two regions, either over- or under-regularization will be effected on pixels at

the boundary (edge) between them. This sudden change of regularization leads to the undesirable ringing artifacts near the edges in the restored image.



(a) (b)

Figure 6.1: A neighborhood of pixels. a) inside a homogeneous region. b) near a boundary between different regions.

Now consider the case where $\psi(s) = s$, corresponding to the TV regularization. The regularization term reduces to

$$\nabla V(f) = div\left(\frac{\nabla f}{|\nabla f|}\right) = \frac{1}{|\nabla f|} f_{\zeta\zeta}. \tag{6.10}$$

Note that in this case $\Psi''(s) = 0$, which means that the smoothing is always oriented in the direction normal to the gradient and no smoothing is effected in the direction of the gradient. This is a desirable feature in order to enhance edges in the restored image. However, problems arise from the term in the denominator of ∇V. In homogeneous regions, where all the pixels in the neighborhood around s have similar intensities, the magnitude of the gradient, $|\nabla f| \to 0$. In this situation, over-regularization occurs, which hampers the adequate restoration of fine details in the image.

In the following section, a new regularization functional is introduced to overcome the shortcomings of the (TV and Tikhonov) regularization schemes discussed in this section.

6.3 Adaptive Regularization Approach

In the previous section, two cases for the regularization functional $V(f)$ are discussed. The Dirichlet functional leads to isotropic smoothing, which hampers the adequate restoration of edges. On the other hand, the TV functional leads to anisotropic smoothing, which has the advantage of preserving edges but has the drawback of attenuating fine details in the restored image. In this section, an adaptive functional, that has the advantages of encouraging the adequate restoration of both edges and fine details, is proposed. In order to achieve the desired adaptability, a combination of the two regularization techniques (TV and Tikhonov) is proposed; using the following form for the function ψ,

$$\psi(s) = \begin{cases} s^2/2\sigma_e, & |\nabla f| < \sigma_e \\ s, & |\nabla f| \geq \sigma_e \end{cases} \tag{6.11}$$

where, σ_e is a threshold that has to be selected. In regions where $|\nabla f| < \sigma_e$, Tikhonov regularization is used in order to preserve fine details and avoid the problems associated with TV regularization. Near the boundaries of different regions in the image, where the gradient is large $\nabla f \geq \sigma_e$, the TV regularization is used in order to encourage the adequate restoration of edges. The value of σ_e determines the value of $|\nabla f|$ at which the type of regularization (either TV or Tikhonov) is changed. The proposed function ψ is plotted in Fig.6.2, along with those functions associated with both Tikhonov regularization and the TV regularization.

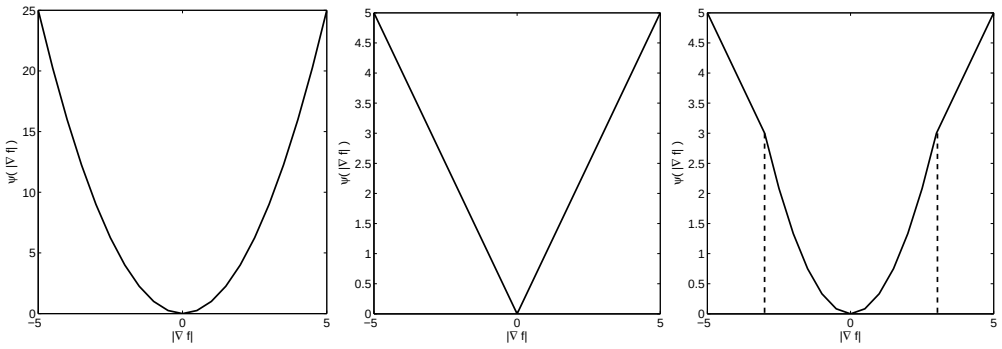


Figure 6.2: Potential functions $\psi(.)$ of Tikhonov, TV, and adaptive regularization (the value of σ_e, is marked by the dotted lines). When $\sigma_e \to 0$, $\psi(s) \to |s|$ and when $\sigma_e \to \infty$, $\psi(s) \to s^2/2$.

6.3.1 How to select the value of the threshold σ_e

The tuning parameter σ_e is a positive constant that has to be carefully selected. It should be noted that, if $\sigma_e \to 0$, then $\psi \to |s|$ and as $\sigma_e \to \infty$, $\psi \to s^2/2$. A proper value of σ_e should be larger than the gradient of the noise, in order to avoid the enhancement of weak edges that arise due to noise. On the other hand, σ_e determines how large the (magnitude of) gradient can be before it is considered to denote the presence of a true (not caused by noise) edge and should be selected as small as possible to avoid the smoothing of edges. Hence, σ_e should be selected just above the noise level. If the noise level is not known, it can be estimated using robust estimation techniques [23].

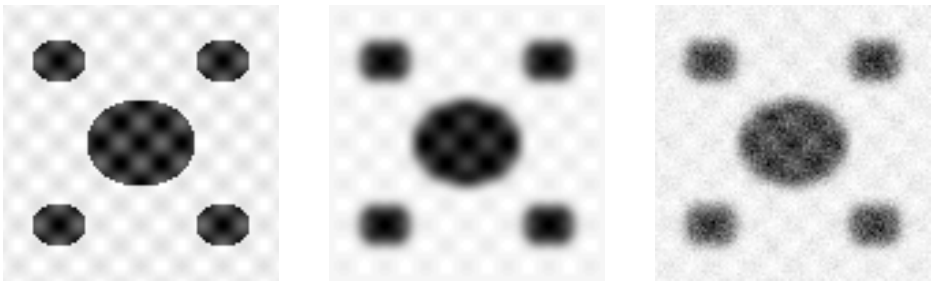
In the following section, the performance of the proposed adaptive regularization scheme is evaluated and its performance is compared to both the classical Tikhonov regularization and the TV regularization techniques.

6.4 Experimental Results

In this section, experiments are performed to illustrate the performance of the proposed adaptive regularization technique. A synthetic image that simulate the true image, is blurred with a PSF calculated using the model of Gibson and Lanni [13] with the parameters: Objective 60×/1.4, illumination wavelength of 530 nm. Poisson noise is generated in the blurred image, using the blurred image as the mean of the Poisson process and a photon ratio of 100 photons/pixel, to set the SNR at 18 db. The true, blurred, and the observed image are shown in Fig.6.3. The degraded image is restored using the three regularization techniques, TV, classical Tikhonov form of regularization and the proposed adaptive regularization technique. The initial guess for f is selected as the observed image. For a fair comparison between different algorithms, the L-curve method [17] is used to calculate the regularization coefficient for each algorithm. In all the experiments, the value of σ_e is calculated using the robust estimator [23] [16]

$$\sigma_e = \text{MAD}(\nabla f)/0.6745,$$

where, MAD denotes the median absolute deviation, to select the threshold σ_e. The number 0.6745, in the denominator is the median absolute deviation of a Gaussian



true blurred observed

Figure 6.3: Test image. The true image is convolved with a calculated confocal PSF [13] to generate the blurred image. The blurred image is then used as the mean of a poisson process to simulate the detection noise. The photon conversion factor is 100 photons/pixel. This corresponds to a SNR of 18dB.

distribution of unity variance. The algorithms are run until convergence. The convergence criterion used is $(E^k - E^{k-1})/E^k \leq 10^{-4}$, where, k is the number of iterations, and $E^i = ||g - h \otimes f^{(i)}||_2^2,\ i = k-1,\ k$. The restoration results are shown in Fig.6.4, along with a plot of the intensity profile of a horizontal line through the center of each image.

Examining these images shows that both the TV and the adaptive regularization produce images that are visually superior to the classical Tikhonov regularization technique (the Laplacian). In the line plots, it can be seen that, the image restored using the classical Tikhonov regularization exhibits edges that are not adequately restored in comparison to the results of other algorithms. Although, both the adaptive regularization and the TV regularization produced similar results, the fine details are better preserved (restored) in the results using the adaptive regularization technique.

In order to quantitatively compare the restoration results, the following criteria are used: the improvement in SNR, defined as ISNR$=||h \otimes f_T - g||/||h \otimes f_T - h \otimes f^T||$, where, f_T denotes the true image, and f^T is the restored image after convergence. To measure the sharpness of edges in the restoration, the gradient of the restored image is calculated, $||\nabla f|| = \sum_{S_f} |(\nabla_x f)^2 + (\nabla_y f)^2|$. Finally, the entropy $(\sum_{S_f} f \log_2 f)$, of the restored image is also calculated. The entropy is a measure of the information content of the image. The values of these measures calculated for the restored images using the TV, Tikhonov, and adaptive regularization are shown in Table 6.1. The values in the table show that, using the TV regularization, the edges in the restored image are well restored as indicated by largest value

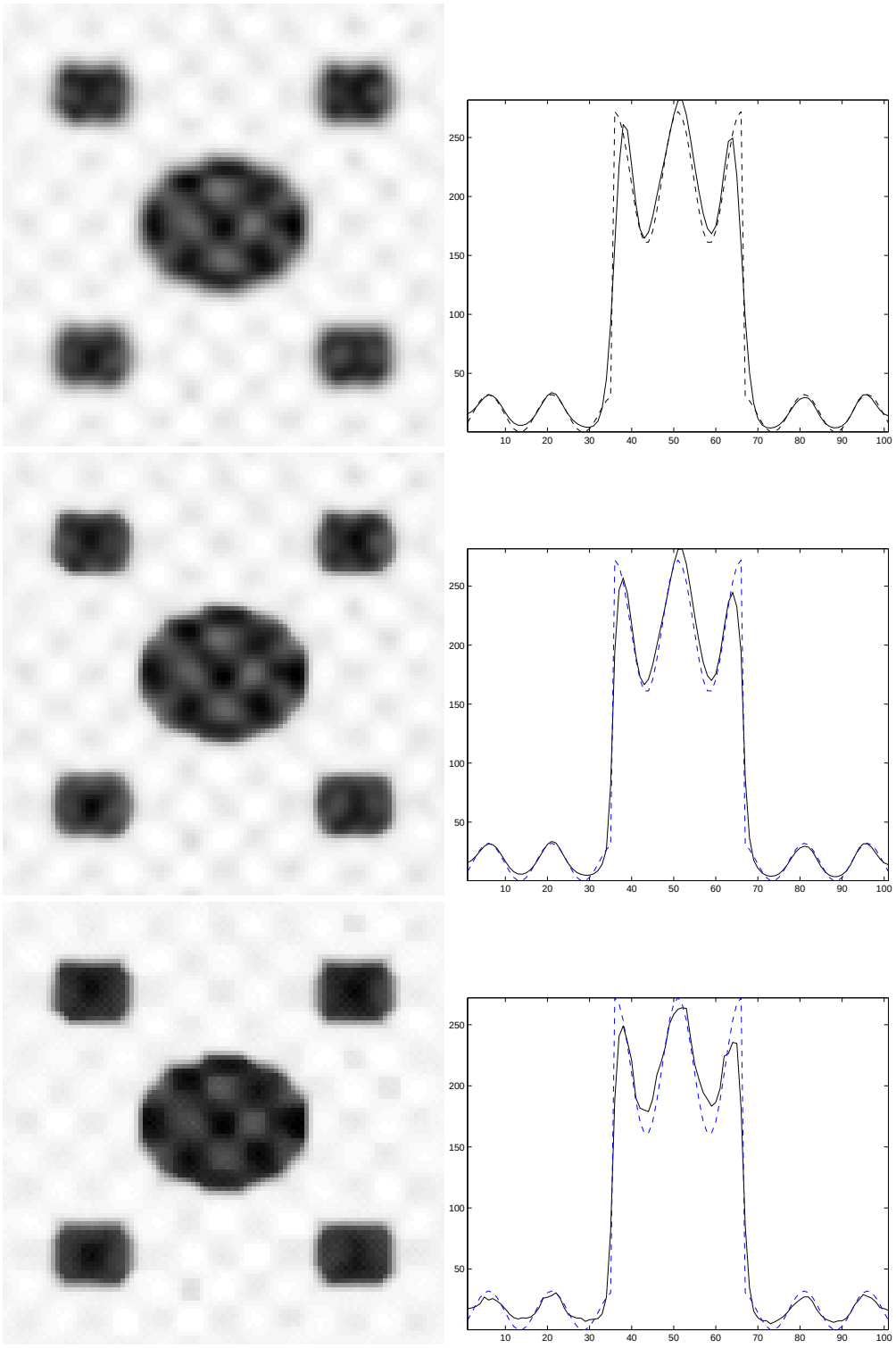


Figure 6.4: Restoration results. In each row the restored image is shown on the *left* and the intensity profile of a line through the center of the image is shown to the *right*. *Top*: using the classical Tikhonov regularization. *Middle*: using the adaptive regularization technique. *Bottom*: using the TV regularization.

	Adaptive	TV	Tikhonov
$\|\|\nabla \hat{f}\|\|$	1.6250	1.8960	1.5221
Entropy	253.029	252.745	252.235
ISNR	1.987	1.8960	1.4714

Table 6.1: Quality measures of the restored test images.

of $||\nabla \hat{f}||$. However, the attenuation of fine details in this image is reflected as a smaller value of both the entropy and the ISNR. For the image restored using the Tikhonov regularization, the sharpness of the edges is the worst which reflects the over-smoothing of edges using this technique. This over-smoothing is also reflected as a smaller value of entropy (which is interpreted as a loss of information due to the smoothing of edges) and a smaller value of ISNR in comparison to the results of both the TV and the adaptive regularization. The image restored using the adaptive technique has the highest value of both entropy (indicating a less severe loss of information) and ISNR. However, the sharpness $||\nabla \hat{f}||$ is inferior in comparison to the results of the TV regularization. This might be attributed to the stair-casing of smooth transitions in intensity in the results of the TV regularization.

In the next experiment, the adaptive regularization is tested using a confocal microscopy image stack. This data set is an image of the medula staining of inactive TeTxLC expressing optic lobes (of Drosophila) with anti-Irrec [19]. The inactive TeTxLC produces a highly organized staining pattern of the distal and proximal medula. The image stack is collected using a TCS4D (Leica Gmbh) confocal microscope equipped with ArKr laser and an 40×1.0 objective. The data set consists of 183 image with 512×512 pixels. An XY-slice of the observed image stack is shown in Fig.6.5 along with the restored image using each regularization technique, for visual comparison of the results. Since the true image is not known, it is not easy to quantitatively evaluate the result. However, in Fig.6.5, it can be seen that the boundaries between different structures in the image are more enhanced in the image restored by the adaptive regularization in comparison to the result of the Tikhonov regularization. On the other hand, the fine details are better restored in comparison to the TV regularization, compare the sites marked by arrows in Fig.6.5.

In order to quantify the performance of the three regularization techniques, following measures are used: MSE=$||h \otimes f - g||^2$ and the contrast defined [8] as

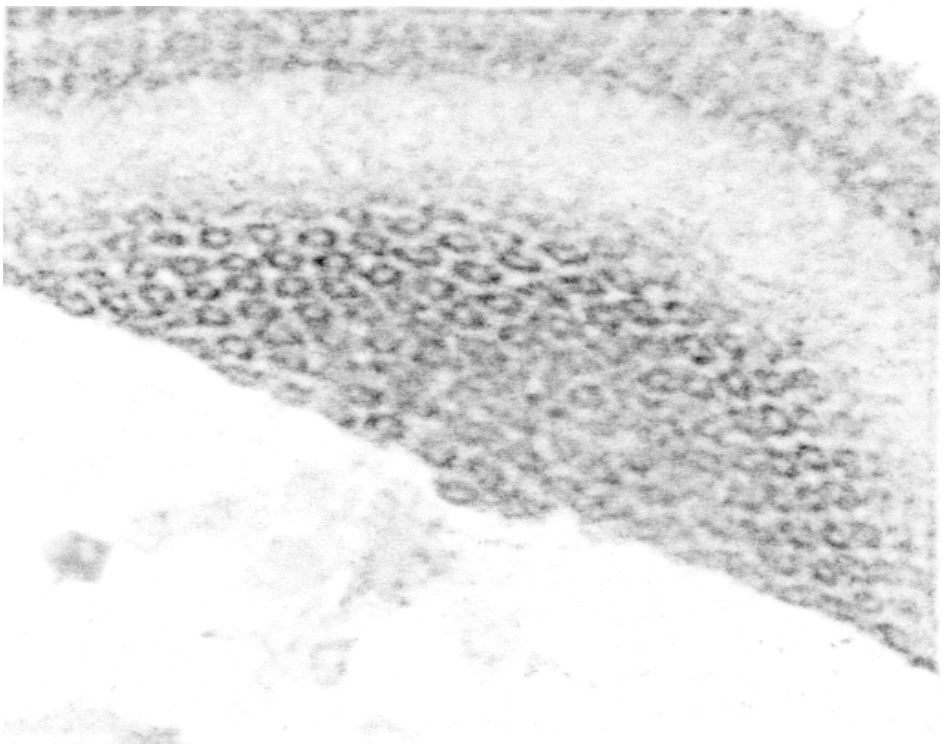
the ratio between the standard deviation of the restored image and the standard deviation of noise. In current case, the true image is not available and the residual noise $(h \otimes f - g)$ is used to calculate the standard deviation of the noise in the restored image. The values of each measure calculated for the restored images are shown in table 6.2. The image restored using the proposed algorithm exhibits a smaller MSE and higher contrast in comparison to other results, in agreement with the visual comparison.

	TV	Adaptive	Tikhonov
MSE	8.34×10^{-5}	7.3513×10^{-5}	7.4739×10^{-5}
Contrast	0.9635	1.1288	1.0974

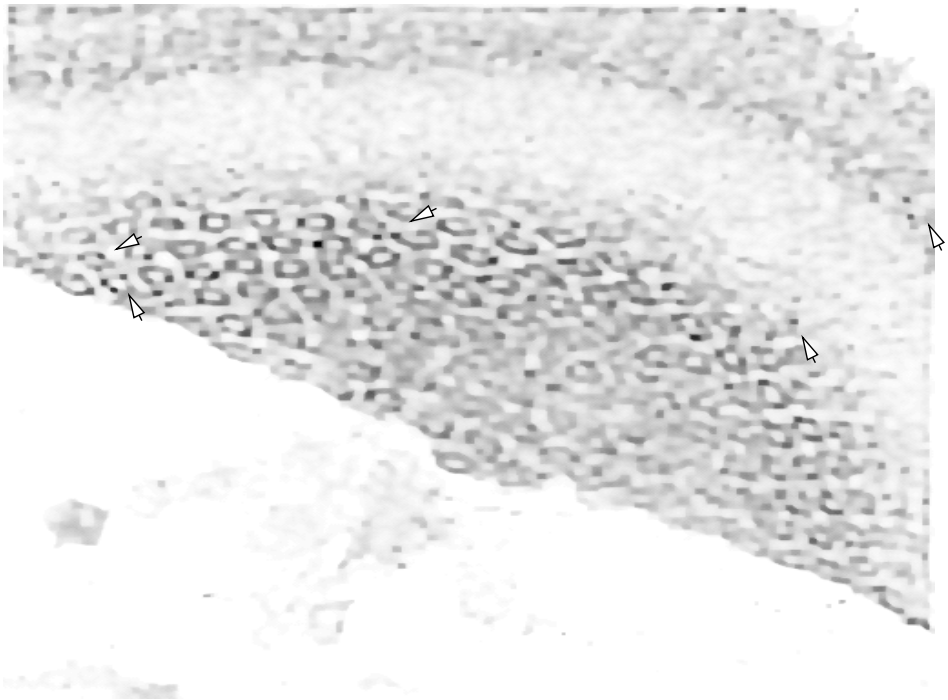
Table 6.2: Quality measures of the restored confocal microscopy image.

6.5 Conclusion

In this chapter, a new regularization technique for deconvolution algorithms is proposed. The proposed regularization technique is indeed a compromise between both TV and the classical Tikhonov regularization schemes. The proposed technique is compared to both the classical Tikhonov regularization and the TV regularization technique. The proposed technique is tested on both synthetic images and on real confocal microscopy images. The experiments show that the images restored using the proposed algorithm exhibit more contrast, adequately restored edges, and fine details in comparison to images restored by other algorithms.

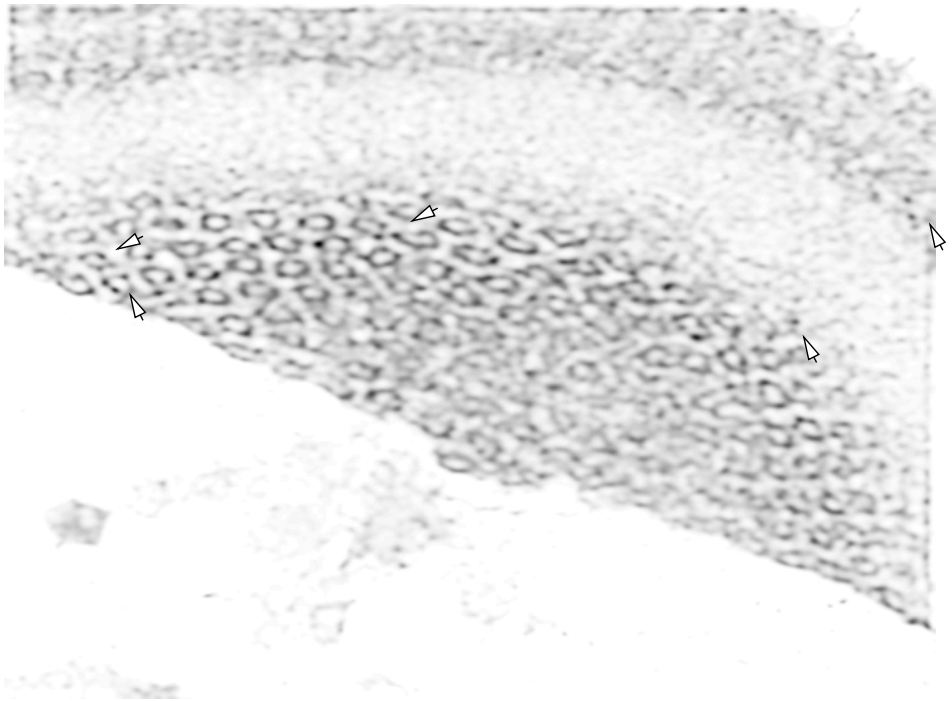


(Raw confocal image)

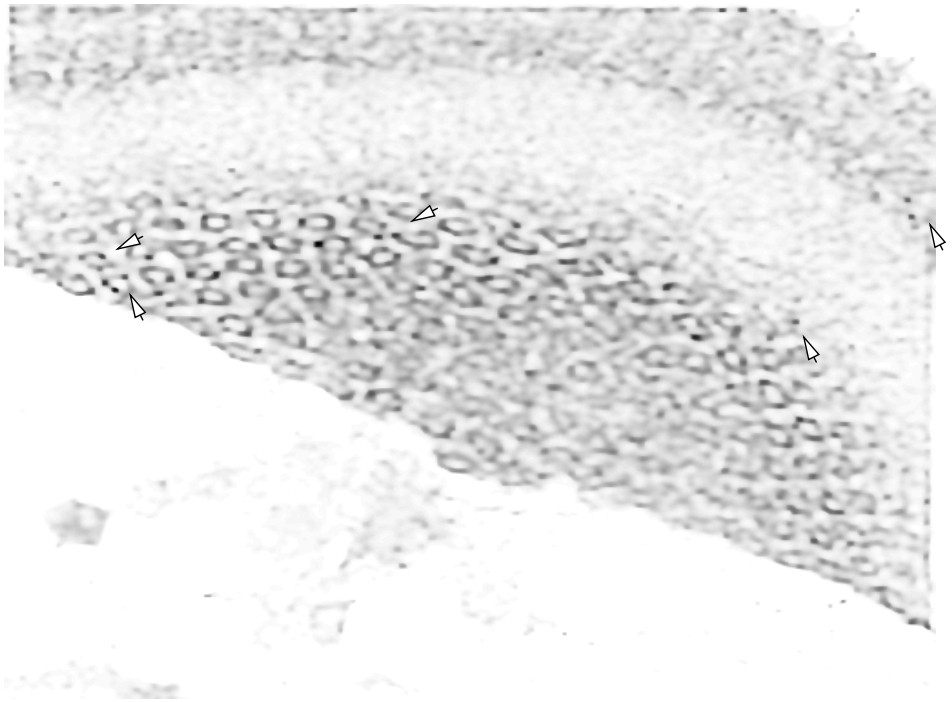


(Restoration result using the TV regularization)

Figure 6.5: *continued on the next page.*



(Restoration result using Tikhonov regularization)



(Restoration result using the adaptive regularization)

Figure 6.5: Visual comparison of the restoration results of a confocal microscopy image using the Tikhonov, TV, and adaptive regularization functionals. Gray values denote image intensities (large → small: black → white). The sites marked by the arrows show the improvement of the image contrast using the adaptive regularization.

Chapter 7

Conclusions

In this thesis, we have developed new methods to solve the problem of image restoration in confocal microscopy. We have investigated linear and Bayesian restoration methods. We have studied the influence of the selection of the prior distribution in MAP methods on the restoration results. We have shown that the performance of blind deconvolution algorithms could be improved if more constraints are incorporated into the algorithm. We have also developed a new method for the restoration of confocal microscopy images of neurons. Our results show that the accuracy of morphological measurements are improved when performed after the restoration of the acquired images. Finally, we have studied the effect of the regularization functional on the restored image. A new adaptive regularization technique is introduced and tested on both artificial and confocal microscope images.

7.1 Confocal Microscopy

In chapter 2, we started by a description of the confocal microscope and its imaging properties. We have briefly discussed the components of the microscope, and the effect of each component on the overall performance of the microscope. We have outlined the different sources of degradation and their impact on the observed image.

7.2 Comparison of Deconvolution Algorithms

In chapter 3, we reviewed several deconvolution algorithms. we have discussed the Bayesian approach to image restoration, the maximum likelihood ML method and the maximum *a posteriori* MAP method. The ML method leads to unregularized algorithms e.g. Richardson-Lucy algorithm. Unregularized algorithms must be stopped before complete convergence to avoid noise-amplification. We have investigated different priors: entropy, Good's roughness penalty, total variation. The entropy prior was found to be dependent on the model m. A flat model is not the best selection, a better model should include more information about the spatial distribution of intensity in the image. Such prior knowledge is rarely available in practice. However, this prior ignores the correlations among neighboring image pixels. The TV prior tolerates discontinuities in the intensity and leads to sharper edges in the restored image. However, this prior is not suitable for images with fine details or texture. Among all the priors we tested, the Goods roughness prior was the most favorable for images corrupted with Poisson noise. This prior leads to a spatially-adaptive regularization depending on the local intensity of image. This prior is suited to images with smooth transitions in intensity. We have used this prior in the restoration algorithms developed in chapter 5. However, a promising alternative might be a compromise between the Good's roughness prior and the total variation prior.

7.3 A Constrained Blind Deconvolution Algorithm

The blind deconvolution problem is known to be ill-posed and admits infinitely many solutions. Moreover, the problem is under-determined, given the observed image it is required to estimate both the true image and the PSF.

In chapter 4, We have introduced a constrained blind deconvolution algorithm. In this algorithm, we have used a parametric model for the PSF. This model is based on Gaussian radial basis functions. To limit the set of admissible solutions and avoid unfeasible solutions, we have incorporated the most general constraints about the PSF, circular symmetry, unity sum, band-limitedness and non-negativity, into the parametric model. For the image, we have used quadratic parameterization to impose the non-negativity constraint. Proceeding this way, the *ad-hoc* enforcement of the constraints after each iteration could be avoided. To compare the performance of the proposed algorithm against other algorithms, we used a blind version of the algorithm developed by Verveer and Jovin [48] as a benchmark. This algorithm uses quadratic re-parameterization for the image to assure non-negativity and no parametric model for the PSF. To the best of our knowledge, this is the best algorithm for the restoration of confocal microscopy images [46]. Our experiments showed that, the proposed algorithm has an improved performance in comparison to the non-parametric algorithm. Experiments with noisy test images, showed that the proposed algorithm is more robust to noise. The algorithm is also tested on real confocal microscopy data set, and the results of the proposed algorithm were better than those of the benchmark algorithm.

7.3.1 Further Research Directions

Further improvements in performance might be possible through a further reduction of the number of unknowns. A parametric model for the image might decrease the number of unknowns drastically. However, we are not aware of a basis that can be generally used for all kinds of images. Moreover, the initialization of the locations of the centers of the basis functions is another major problem that face this approach. Our experiments showed that blind deconvolution algorithms are sensitive to the noise level in the observed image. Hence an alternative research av-

enue is to test methods for reducing the noise in the observed image (pre-filtering) before the application of the restoration algorithms. Methods like Gaussian filtering are known to blur the image being filtered. However, the added blur might be taken into account by using a larger PSF. Wavelet de-noising methods [38] are a better candidate for the pre-filtering task.

7.4 Restoration of 3D Neuron Images

In chapter 5, a new method for the restoration of confocal microscopy images of neurons was introduced. In this method the PSF is estimated from the observed image *a priori* and a non-blind deconvolution algorithm is used for the restoration of the true image. To estimate the PSF, we have studied the deformation of the cross-sectional shape of the neuron. The estimated PSF exhibits a strong dependence on the depth (Z) inside the specimen, this necessitates the application of a depth-variant deconvolution algorithm. To avoid the artifacts associated with the depth-variant approach, we have used a separable model for the PSF that enabled the restoration of the neuron image stack as a whole.

Our experiments show that using the proposed restoration algorithm it is possible to remove the distortions introduced during the image acquisition process. We have also investigated the improvement in the circularity of the cross-section of the neurons after the application of the restoration algorithm. The conventional approach to the blind deconvolution problem is based on the simultaneous estimation of both the image and the PSF from the observed image using an alternating minimization technique that consists of two non-blind steps. The image estimation step estimates the true image assuming that the current estimate of the PSF is correct, and the PSF estimation step uses the thus estimated image as the true image in order to obtain an estimate for the PSF. To the contrary, we have opted to split the blind deconvolution problem into two separate tasks. The first task is to estimate the PSF *a priori* form the observed image. The second task is to use a non-blind deconvolution algorithm to restore the true image based on the estimated PSF. We have obtained an estimated PSF based on the distortions of the cross-sectional shape of the neurons in the observed image. Using this estimated PSF we were able to apply a non-blind deconvolution algorithm to restore the true

image.

7.4.1 Further Research Directions

This technique of splitting the blind deconvolution problem avoids the computational load associated with the PSF estimation step of the conventional alternating minimization approach. Such an approach can be generalized into a new generation of blind deconvolution algorithms. However, this necessitates the development of new methods for the estimation of the PSF from the observed images. A promising method is introduced in [9]. This method estimates the PSF from edges in the observed image, based on a model of an ideal edge. However, in the same paper the method is tested on a synthetic image and the errors in the estimated blur scale were found to increase roughly linearly with the scale of the blur. To the time of writing this these, we are not aware of a better method for the estimation of the PSF. Most other approaches for the estimation of the PSF are based on using two images acquired under different conditions [12].

7.5 Adaptive Regularization

In chapter 6, an adaptive regularization functional is proposed. The proposed functional is a combination of the total variation and the classical Tikhonov regularization functionals. The proposed regularization technique is spatially adaptive. In regions where the image is smooth, as reflected by a small value of the magnitude of the gradient, the Tikhonov form of regularization is used in order to avoid the attenuation of fine image details, which would have been the case if the TV regularization were used. In regions where, the magnitude of the gradient is large, the TV regularization is used to avoid oversmoothing the edges. We have tested and compared the adaptive regularization technique to both the TV and the classical Tikhonov regularization. The results of the experiments agree with the predicted performance. We have also evaluated the proposed regularization technique using confocal microscopy images. The algorithm is shown to improve the contrast of the restored image.

7.5.1 Further Research Directions

The adaptive regularization technique introduced here is based on an analysis of the behavior of regularization in different regions in the image. There are other possibilities to achieve the adaptability of regularization. An alternative approach might be to formulate the regularization functional using another basis, e.g. wavelets. In [29], an approach to combine TV and wavelets for image restoration is proposed. Such techniques are promising, the only problem is the high computational load especially for 3D image.

7.6 Summary

In this thesis we have proposed new methods to solve the problem of image restoration in the biological and medical fields. Using an RBF model for the PSF, we introduced a blind restoration method that has an improved estimation accuracy and an accelerated convergence performance. However, in the space-variant case, there is a severe lack of information; one has to estimate both the true image and the PSF at each pixel given only the observed image. A new method for the estimation of the support of the PSF, based on the prior knowledge of the geometry of the true object (neurons), is introduced in chapter 5. The PSF is approximated by a simplified model that corresponds to only one basis function of the RBF model proposed in chapter 4. The parameters of the model are estimated from the observed image. Using the estimated PSF, we were able to reduce the computational load associated with the simultaneous estimation of both the image and the PSF in the conventional approach for the solution of the blind deconvolution problem. Finally, an adaptive regularization technique is proposed in chapter 6 to avoid the disadvantages of existing techniques. Using this regularization technique, the quality of the restored image (contrast, edge sharpness) is much improved.

However, research will never come to an end. We hope that we have provided a good starting point for future improvements on the methods introduced in this thesis.

Appendix

A The Adaptive Regularization-Based Restoration Algorithm

The goal of this appendix is to derive Eq.6.4

$$\nabla \mathcal{J} = h \odot (1 - \frac{g}{h \otimes f}) - div \left\{ \psi'(|\nabla f|).\frac{\nabla f}{|\nabla f|} \right\}. \quad \text{(A-1)}$$

We start by a simplified derivation of the Euler-Lagrange equation which is then used to find derive the above equation.

Euler-Lagrange Equation

The calculus of variation is used to minimize a functional (function of functions) in a way that is similar to the minimization of functions in calculus. Our goal is to minimize some functional $\mathcal{J}$, which can generally be a function of f, f', X, where, f is the image (function), f' is the derivative of f, and X is the coordinates vector.

Let $\hat{f}$ denote an image (function) that minimizes the functional $\mathcal{J}$. Assume an infinitesimal variation of magnitude ϵ around $\hat{f}$ in the direction of the function $\zeta(X)$. We have,

$$\mathcal{J}(\hat{f}, \nabla\hat{f}, X) \leq \mathcal{J}(\hat{f} + \epsilon\zeta(X), \hat{f}' + \epsilon\zeta'(X), X) \quad \text{(A-2)}$$

where $\zeta(X)$, is a smooth function that satisfies the boundary conditions on the region S_f, i.e $\zeta(X) = 0, \quad \forall X \notin S_f$.

To find the condition that must be satisfied by any minimizer of $\mathcal{J}$, take the derivative of $\mathcal{J}(\hat{f} + \epsilon\zeta(X), \hat{f}' + \epsilon\zeta'(X), X)$ w.r.t ϵ, and set it to zero as $\epsilon \to 0$, i.e.,

$$\frac{d\mathcal{J}(\hat{f} + \epsilon\zeta(X), \hat{f}' + \epsilon\zeta'(X), X)}{d\epsilon}|_{\epsilon=0} = \int_{S_f} \left[\frac{\partial \mathcal{J}}{\partial f}\zeta(X) + \frac{\partial \mathcal{J}}{\partial f'}\zeta(X) \right] dX \quad \text{(A-3)}$$

Integrating the second term by parts, we get

$$= \int_{S_f} \left[\frac{\partial \mathcal{J}}{\partial \hat{f}} - \frac{d}{dX} \left\{ \frac{\partial \mathcal{J}}{\partial \hat{f}'} \right\} \right] \zeta(X) dX = 0 \tag{A-4}$$

For this to hold for any function $\zeta(X)$, it is obvious that the remaining part of the integrand must be zero,

$$\left[\frac{\partial \mathcal{J}}{\partial f} - \frac{d}{dX} \left\{ \frac{\partial \mathcal{J}}{\partial f'} \right\} \right] = 0 \tag{A-5}$$

This is the Euler-Lagrange equation. Any minimizer of $\mathcal{J}$ must satisfy this equation.

Based on this background, we now turn to our main objective. The solution to the deconvolution problem is found by the minimization of a risk function (or functional) $\mathcal{J}$, which consists of two terms,

$$\mathcal{J} = \mathcal{L} + \gamma V. \tag{A-6}$$

Where, $\mathcal{L} = \int_{S_f}((Hf - g) - g \ln(Hf/g))dX$, and $V(f) = \int_{S_f} \psi(|\nabla f|)dX$. Applying the Euler-Lagrange equation, Eq.A-5 to the functional $\mathcal{J}$,the condition that must be satisfied by any minimizer of $\mathcal{J}$ is that,

$$(h \otimes f - g \ln(h \otimes f)) - \frac{d}{dX} \left\{ \psi'(|\nabla f|). \frac{\nabla f}{|\nabla f|} \right\} \tag{A-7}$$

$$= h \odot (1 - \frac{g}{h \otimes f}) - div \left\{ \psi'(|\nabla f|). \frac{\nabla f}{|\nabla f|} \right\}. \tag{A-8}$$

Where, $div = \sum_{i=1}^{d} \frac{\partial}{\partial x_i}$, d is the dimension of the image (d =2 for 2D and d=3 for 3D), and $\odot$ denotes the correlation operation.

This proves Eq.6.4.

B Decomposition of The Generalized Regularization Functional

The goal of this appendix is to show that,

$$div\left\{\psi'(|\nabla f|).\frac{\nabla f}{|\nabla f|}\right\} = \frac{\psi'(|\nabla f|)}{|\nabla f|}f_{\zeta\zeta} + \psi''(|\nabla f|)f_{\eta\eta} \tag{B-1}$$

Where, $div = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$, $|\nabla f| = \sqrt{f_x^2 + f_y^2}$, and $\nabla f = (f_x, f_y)$, we start by developing the div operator,

$$div\left\{\psi'(|\nabla f|).\frac{\nabla f}{|\nabla f|}\right\} = \frac{\partial}{\partial x}\left(\frac{\psi'(\sqrt{f_x^2+f_y^2})}{\sqrt{f_x^2+f_y^2}}f_x\right) + \frac{\partial}{\partial y}\left(\frac{\psi'(\sqrt{f_x^2+f_y^2})}{\sqrt{f_x^2+f_y^2}}f_y\right) \tag{B-2}$$

The first term can be further simplified as,

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\psi'(\sqrt{f_x^2+f_y^2})}{\sqrt{f_x^2+f_y^2}}f_x\right) &= \frac{\partial\psi'(\sqrt{f_x^2+f_y^2})}{\partial x}\frac{f_x}{\sqrt{f_x^2+f_y^2}} \\ &+ \psi'(\sqrt{f_x^2+f_y^2})\frac{\partial}{\partial x}\frac{f_x}{\sqrt{f_x^2+f_y^2}}\end{aligned} \tag{B-3}$$

Where,

$$\frac{\partial\psi'(\sqrt{f_x^2+f_y^2})}{\partial x} = \psi''(\sqrt{f_x^2+f_y^2})\frac{f_xf_{xx}+f_yf_{yy}}{\sqrt{f_x^2+f_y^2}} \tag{B-4}$$

$$\frac{\partial}{\partial x}\frac{f_x}{\sqrt{f_x^2+f_y^2}} = \frac{f_y^2f_{xx}-f_xf_yf_{xy}}{(f_x^2+f_y^2)^{3/2}} \tag{B-5}$$

Finally, we get,

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\psi'(\sqrt{f_x^2+f_y^2})}{\sqrt{f_x^2+f_y^2}}f_x\right) &= \psi''(\sqrt{f_x^2+f_y^2})\frac{f_x^2f_{xx}+f_xf_yf_{xy}}{f_x^2+f_y^2} \\ &+\psi'(\sqrt{f_x^2+f_y^2})\frac{f_y^2f_{xx}-f_xf_yf_{xy}}{(f_x^2+f_y^2)^{3/2}}\end{aligned} \tag{B-6}$$

Similarly, the second term in Eq.B-2 can be simplified as,

$$\begin{aligned}\frac{\partial}{\partial y}\left(\frac{\psi'(\sqrt{f_x^2+f_y^2})}{\sqrt{f_x^2+f_y^2}}f_y\right) &= \psi''(\sqrt{f_x^2+f_y^2})\frac{f_y^2f_{yy}+f_xf_yf_{xy}}{f_x^2+f_y^2} \\ &+\psi'(\sqrt{f_x^2+f_y^2})\frac{f_x^2f_{xx}-f_xf_yf_{xy}}{(f_x^2+f_y^2)^{3/2}}\end{aligned} \tag{B-7}$$

Hence,

$$div\left(\frac{\psi'(|\nabla f|)}{|\nabla f|}\nabla f\right) = \frac{\psi'(|\nabla f|)}{|\nabla f|}\left(\frac{f_y^2 f_{xx} - 2f_x f_y f_{xy} + f_x^2 f_{yy}}{(f_x^2 + f_y^2)}\right) + \psi''(|\nabla f|)\left(\frac{f_x^2 f_{xx} + 2f_x f_y f_{xy} + f_y^2 f_{yy}}{(f_x^2 + f_y^2)}\right) \quad \text{(B-8)}$$

To simplify further, we use a local orthogonal coordinate system defined by the gradient and the vector normal to the gradient. The orientation of the gradient is defined as,

$$\cos\theta = \frac{f_x}{\sqrt{f_x^2 + f_y^2}}, \sin\theta = \frac{f_y}{\sqrt{f_x^2 + f_y^2}} \quad \text{(B-9)}$$

The second term in Eq.B-8 can be written as,

$$\left(\frac{f_x^2 f_{xx} + 2f_x f_y f_{xy} + f_y^2 f_{yy}}{(f_x^2 + f_y^2)}\right) = f_{xx}\cos^2\theta + 2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta \quad \text{(B-10)}$$

$$= [\cos\theta \;\; \sin\theta]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

Which is the second derivative in the direction of the gradient.

Similarly,

$$\left(\frac{f_y^2 f_{xx} - 2f_x f_y f_{xy} + f_x^2 f_{yy}}{(f_x^2 + f_y^2)}\right) = f_{xx}\sin^2\theta - 2f_{xy}\cos\theta\sin\theta + f_{yy}\cos^2\theta \quad \text{(B-11)}$$

$$= [-\sin\theta \;\; \cos\theta]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$= [\cos(\theta+\pi/2) \;\; \sin(\theta+\pi/2)]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}\begin{bmatrix} \cos(\theta+\pi/2) \\ \sin(\theta+\pi/2) \end{bmatrix}$$

Which is the second derivative in the direction normal to the gradient.

Hence

$$div\left\{\psi'(|\nabla f|).\frac{\nabla f}{|\nabla f|}\right\} = \frac{\psi'(|\nabla f|)}{|\nabla f|}f_{\zeta\zeta} + \psi''(|\nabla f|)f_{\eta\eta} \quad \text{(B-12)}$$

This proves Eq.6.8.

Bibliography

[1] Khalid A. Al-Kofahi, Sharie Lasek, Donald H. Szarowski, Christopher J. Pace, George Nagy, James N. Turner, and Badrnath Roysam. Rapid automated three-dimensional tracing of neuron from confocal image stacks. *IEEE Transaction on Information Technology in Biomedicine*, 6(2), 2002.

[2] Peter L. Becker. *Quantitative fluorescence measurements*. Fluorescence imaging spectroscopy and microscopy. Wiley N.Y., 1996.

[3] Dirk Bucher and Hans-Joachim Pflüger. Directional sensetivity of an identified wind-sensetive interneuron during the postembryonic development of the locust. *Journal of Insect Physiology*, 46:1545–1556, 2000.

[4] T.F. Chan and C.K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Process.*, 7:370–375, 1998.

[5] José-Angel Conchello, J. J. Kim, and E. W. Hansen. Enhanced three-dimensional reconstruction from confocal scanning microscope images. depth discrimination versus signal-to-noise ratio in partially confocal images. *Applied Optics*, 33(17):3740–3750, 1994.

[6] José-Angel Conchello and J. W. Lichtman. Theoretical analysis of a rotating-disk partially confocal focal scanning microscope. *Applied Optics*, 33(4):585–596, 1994.

[7] I. Csiszár. Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems. *The Annals of Statitics*, 19(4):2032–2066, 1991.

[8] Jaques Boutet de Monvel, Sophie Le Calvez, and Mats Ulfendahl. Image restoration for confocal microscopy: Improving the limits of deconvolution,

with application to the visualization of the mammalian hearing organ. *Biophisical Journal*, 80(5):2455–2470, 2001.

[9] James H. Elder and Steven W. Zucker. Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7):699–716, 1998.

[10] Andrew W. Fitzgibbon, Maurizio Pilu, and Robert B. Fisher. Direct least squares fitting of ellipses. In *Proceeding of the 13th International Conference on Pattern Recognition, Vienna*, pages 253–257, September 1996.

[11] B. R. Frieden. Restoring with maximum likelihood and maximum entropy. *Journal of Optical Society of America*, 62:511–518, 1972.

[12] G. B. Giannakis, Heath, RW, and Jr. Blind identification of multichannel FIR blurs and perfect image restoration. *IEEE Tr. Im. Proc.*, 9(11):1877–1896, November 2000.

[13] F. S. Gibson and F. Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *J. Opt. Soc. Amer. A*, 8:1601–1613, 1991.

[14] S. F. Gull. Development in maximum entropy data analysis. In J. Skilling, editor, *Maximum entropy and Bayesian methids*, pages 53–71, Dordrecht, 1988. Kluwer academic paublishers.

[15] Radim Halíř and Jan Flusser. Numerically stable direct least squares fitting of ellipses. In *Proceedings of the 6th International Conference in Central Europe on Computer Graphics and Visualization. WSCG'98, CZ, Plzeň*, pages 125–132, 1998.

[16] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuv, and W. A. Stahel. *Robust Staitistics: the approach based on Influence Functions.* John Wiley and sons., New York, N. Y., 1986.

[17] P. C. Hansen. Analysis of the L-curve method in the regularization of ill-posed problems. *SIAM J. Sci. Comput.*, 14(6):1487–1503, 1993.

[18] S. Hell, G. Reiner, C. Cremer, and E. H. Stelzer. Aberrations in confocal fluorescance microscopy induced by mismatches in refractive index. *Journal of Microscopy*, 169:391–405, 1993.

[19] Peter Robin Hiesinger, Christian Reiter, Harald Schau, and Karl-Friedrich Fischbach. Neuropil pattern formation and regulation of cell adhesion molecules in drosophila optic lobe development depend on synaptobrevin. *Journal of Neuroscience*, 17(19):7548–7556, 1999.

[20] Peter Robin Hiesinger, Michael Scholz, Ian A. Meinertzhagen, Karl-Friedrich Fischbach, and Klaus Obermayer. Visualization of synaptic markers in the optic neuropils of drosophila using new constrained deconvolution method. *The Journal of Comparative Neurobiology*, 429:277–288, 2001.

[21] Timothy J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum likelihood approach. *J. Opt. Soc. Am. A*, 9:1052–1061, 1992.

[22] M.R.P. Homem, N.D.A. Mascarenhas, L.F. Costa, and C. Preza. Biological image restoration in optical sectioning microscopy using prototype image constraints. *Real-Time Imaging*, 8:475–490, 2002.

[23] P. J. Huber. *Robust Statistics*. New York: Wiley, 1981.

[24] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall Inc., 1989.

[25] S. Joshi and M. I. Miller. Maximum a posteriori estimation with Good's roughness for optical-sectioning microscopy. *Journal of the Optical Society of America A*, 10(5), 1993.

[26] G. M. P. Kempen. *Image restoration in fluorescence microscopy*. PhD thesis, Department of imaging science and technology, faculty of applied sciences, Delft university of technology, 1999.

[27] Reginald L. Lagendijk, Jan Biemond, and Dick E. Boekee. Identification and restoration of noisy blurred image using the expectation-maximization algorithm. *IEEE Transaction on Acoustics. Speech. and Signal Processing.*, 38(7), 1990.

[28] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–154, 1998.

[29] F. Malgouyres. Mathematical analysis of a model which combines total variation and wavelet for image restoration. *Journal of information processes*, 2:1–10, 2002.

[30] J. Markham and José-Angel Conchello. Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *Journal of the Optical Society of America A*, 16(10), 1999.

[31] R. Molina, J. Núñez, F.J. Cortijo, and J. Mateos. Image restoration in astronomy. a bayesian perspective. *IEEE Signal Processing Magazine*, 18(2):11–29, 2001.

[32] Clark F. Olson. Adaptive-scale filtering and feature detection using range data. *IEEE Trassaction on Pattern Analysis and Machine Inteligince*, 22(9), 2000.

[33] J. B. Pawley. *Handbook of Biological Confocal Microscopy*. New York: Plenum Press, 1995.

[34] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(7):629–639, jul 1990.

[35] Chrysante Preza and José-Angel Conchello. Image estimation accounting for point-spread function depth variation in thee-dimensional fluorescence microscopy. *3D and Multidimensional Microscopy: Image Aquisition and Processing X, Proc. SPIE*, 4964(27), 2003.

[36] F. Rooms, M. Ronsse, A Pizurica, and W. Philips. PSF estimation with application in autofocus and image restoration. In *3rd IEEE Benelux Signal Processing Symposium (SPS-2002), Leuven, Belgium*, pages 13–16, 2002.

[37] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total veriation based noise removal algorithm. *Physica D*, 60:259–268, 1992.

[38] Paul Scheunders. Wavelet thresholding of multivalued images. *IEEE Transaction on Image Processing*, 13(4):475–438, 2004.

[39] D. L. Snyder and M. I. Miller. *Random Point Processes in Time and Space*. Springer Verlag, Berlin, 1991.

[40] J.-L Starck, E. Pantin, and F. Murtagh. Deconvolution in astronomy: a review. *Pub. Astron. Soc. Pac.*, 114:1051–1069, 2002.

[41] Luis Tenorio. Statistical regularization of inverse problems. *Society for Industrial Applied Mathematics, SIAM REVIEW*, 43(2):347–366, 2001.

[42] E. Thiébaut and J.-M. Conan. Strict a priori constraints for maximum-likelihood blind deconvolution. *J. Opt. Soc. Am. A*, 12:485–492, 1995.

[43] H.T.M. van der Voort and K.C. Strasters. Restoration of confocal images for quantitative image analysis. *J. Microsc.*, 178:165–181, 1988.

[44] G. M. P. van Kempen and H.T.M. van der Voort. A quantitative comparison of two restoration methods as applied to confocal microscopy. *ASCI'96*, pages 196–201, June 1996.

[45] Peter J. Verveer. *Computational and optical methods for improving resolution and signal quallity in fluorescence microscopy*. PhD thesis, Department of imaging science and technology, faculty of applied sciences, Delft university of technology, 1998.

[46] Peter J. Verveer, M. J. Gemkow, and Thomas M. Jovin. A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy. *J. Microsc.*, 193:50–61, 1999.

[47] Peter J. Verveer and Thomas M. Jovin. Acceleration of the ICTM image restoration algorithm. *J. Microsc.*, 188:191–195, 1997.

[48] Peter J. Verveer and Thomas M. Jovin. Image restoration based on Good's roughness penalty with application to fluorescence microscopy. *Journal of Optical Society of America*, A(15):1077–1083, 1998.

[49] Peter J. Verveer and Thomas M. Jovin. Image restoration based on good's roughness penalty with application to fluorescence microscopy. *J. Opt. Soc. Am. A*, 15:1077–1083, 1998.

[50] C. R. Vogel. Nonconvergence of the L-curve regularization parameter selection method. *Inverse Problems*, 12:535–547, 1996.

[51] Yu-Li You and Mostafa Kaveh. A regularization approach to joint blur identification and image restoration. *IEEE Transaction on Image Processing*, 5(3), 1996.

[52] Yu-Li You, Wenyuan Xu, Allen Tannenbaum, and Mostafa Kaveh. Behavioral analysis of anisotropic diffusion in image processing. *IEEE Transaction on Image Processing*, 5(11), 1996.