Stellingen
behorend bij het proefschrift

Image Restoration in Fluorescence Microscopy
Geert van Kempen
11 januari 1999

1. De Generalized Cross Validation methode is een efficiënte manier om een waarde voor de regularisatie parameter te bepalen die de optimale waarde benadert. (Hoofdstuk 4)

2. Constrained beeld restauratie algoritmen hebben een gedegen achtergrondkennis nodig. (Hoofdstuk 5)

3. Waar twee honden vechten om een been uit de cache, loopt niemand meer. (Hoofdstuk 6)

4. Indien de puntspreidingsfunctie slecht bepaald is, reduceren niet-lineaire beeldrestauratie-algoritmen zich tot een dure vorm van image enhancement.

5. Er is geen betere intervaltraining mogelijk dan op een zonnige zondagmiddag hard te willen fietsen langs de Vliet.

6. Een oplossing is pas goed als ie ook mooi is.

7. Een bezoek aan een National Park is te beschouwen als een quantum-mechanische meting: de menselijke aanwezigheid verstoort de te bewonderen natuur.

8. Alleen van je eigen data weet je precies hoe het gemeten is.

9. Een overzicht van het onderwerp van een wetenschappelijke publicatie wordt het snelst verkregen door de referentielijst te lezen.

10. Een artificieel neuraal netwerk dient gebruikt te worden als laatste redmiddel, niet als een generieke zwarte doos ter vervanging van domeinspecifieke voorkennis.

11. De evolutie zit in een lokaal optimum genaamd de mens.

12. He who does not travel does not know the value of men. (Moorish proverb, Bruce Chatwin, The Songlines)

Stellingen
Together with the thesis

Image Restoration in Fluorescence Microscopy
Geert van Kempen
11 January 1999

1. The method of Generalized Cross Validation is an efficient way for producing a value of the regularization parameter which is close to its optimal value. *(Chapter 4)*

2. Constrained image restoration algorithms require thorough background knowledge. *(Chapter 5)*

3. When two dogs fight for a bone from the cache, nothing moves. *(Chapter 6)*

4. A poorly determined point spread function reduces non-linear image restoration algorithms to an expensive form of image enhancement.

5. There is no better interval training than trying to bike as fast as possible along the Vliet on a sunny Sunday afternoon.

6. A solution becomes good when it is beautiful as well.

7. A visit to a National Park can be regarded as a quantum measurement: the human presence distorts the Nature to be admired.

8. Only the measurer knows exactly how the data has been measured.

9. The quickest way to get an overview of the subject of a scientific publication is to read the list of references.

10. An artificial neural network should be used as a last resort, not as a generic blackbox replacing domain-specific, a priori knowledge.

11. Evolution is in a local optimum called mankind.

12. He who does not travel does not know the value of men. *(Moorish proverb, Bruce Chatwin, The Songlines)*

13. Bugs are found on the way home.
Image Restoration

in

Fluorescence Microscopy
Image Restoration
in
Fluorescence Microscopy

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,

op maandag 11 januari 1999 te 16.00 uur
door Gijsbert Michiel Peter VAN KEMPEN

natuurkundig ingenieur
geboren te Haarlem
Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. I.T. Young

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. dr. I.T. Young, Technische Universiteit Delft, promotor
Dr. ir. L.J. van Vliet, Technische Universiteit Delft, toegevoegd promotor
Prof. dr. T. Wilson, Oxford University, Groot Brittannië
Prof. dr. ir. J. Biemond, Technische Universiteit Delft
Prof. dr. ir. A. van den Bos, Technische Universiteit Delft
Prof. dr. ir. A.W.M. Smeulers, Universiteit van Amsterdam
Dr. H.T.M. van der Voort, Scientific Volume Imaging, Hilversum

Advanced School for Computing and Imaging

This work was carried out in graduate school ASCI.
ASCI dissertation series number 39.

Published and distributed by:

Delft University Press
Mekelweg 4
2628 CD Delft
The Netherlands
Telephone: +31 15 2783254
fax: +31 15 2781661
E-mail: DUP@DUP.TUDelft.NL


Copyright © 1998 by G.M.P. van Kempen

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without the permission from the publisher: Delft University Press.

Printed in the Netherlands
Aan mijn ouders

Aan Lise
# Table of Contents

1. **Image Restoration in Fluorescence Microscopy** ...................................................... 1  
   1.1 Fluorescence Microscopy ...................................................................................... 2  
   1.2 Image Restoration ............................................................................................... 5  
   1.3 Scope of this Thesis ............................................................................................ 11  

2. **Image Formation** ...................................................................................................... 13  
   2.1 Wave Description of Light .................................................................................. 14  
   2.2 Fluorescence Microscopy .................................................................................... 18  
   2.3 Image Sampling ................................................................................................... 28  
   2.4 Particle Description of Light .............................................................................. 33  
   2.5 Image Formation in a Fluorescence Microscope ............................................... 34  

3. **Methods for Image Restoration** ............................................................................. 35  
   3.1 Introduction ........................................................................................................ 36  
   3.2 Classical Image Restoration ............................................................................... 37  
   3.3 Constrained Tikhonov Restoration .................................................................... 40  
   3.4 Maximum Likelihood Restoration ..................................................................... 43  
   3.5 Software Implementation .................................................................................... 47  

4. **Iterative Image Restoration Algorithms** ............................................................... 49  
   4.1 Performance Measures ...................................................................................... 50  
   4.2 Generation of Simulated Microscope Images .................................................... 52  
   4.3 Regularization Parameter ................................................................................. 57  
   4.4 First Estimates ................................................................................................... 67  
   4.5 Stop Criterion ..................................................................................................... 74  

5. **Applications of Image Restoration in Microscopy** ............................................. 77  
   5.1 Background Estimation ...................................................................................... 78  
   5.2 Prefiltering: Reducing the Noise Sensitivity ....................................................... 88  
   5.3 Improvement of Quantitative Measurements .................................................... 96
## Table of Contents

6. **Optimal Cache Usage for Separable Filters** .................................................. 105
6.1 Prologue ........................................................................................................... 106
6.2 Introduction ..................................................................................................... 108
6.3 Cache Organization ......................................................................................... 109
6.4 Optimization of Cache Utilization ................................................................. 111
6.5 Experiments .................................................................................................... 115
6.6 Conclusions .................................................................................................... 119

7. **Conclusions** .................................................................................................. 121
7.1 Image Formation ............................................................................................. 122
7.2 Image Restoration Algorithms ......................................................................... 122
7.3 Restoration Parameters ................................................................................... 123
7.4 Application of Image Restoration ................................................................... 125
7.5 Cache-optimized Separable Filters ................................................................. 126

Appendix A ............................................................................................................ 129
References ............................................................................................................ 131
Summary ................................................................................................................ 139
Samenvatting ......................................................................................................... 143
Dankwoord ........................................................................................................... 147
Curriculum Vitae & Bibliography ......................................................................... 149
Chapter 1

Image Restoration in Fluorescence Microscopy

This thesis presents image restoration techniques for applications in (confocal) fluorescence microscopy. We have gained a better understanding of the behavior of non-linear image restoration algorithms and we have developed novel methods to improve their performance in such a way that more accurate measurements can be performed on three-dimensional fluorescence images.

In this first chapter we introduce the principles of fluorescence microscopy and discuss the properties of the three-dimensional image formation in a fluorescence microscope. An image is blurred and distorted by noise during its formation and acquisition. These distortions hide fine details in the image hampering both the visual and the quantitative analysis of the image. The purpose of image restoration is to invert this and to suppress the noise restoring the fine details in the image which results in an improved analysis of the image.

The principles of image restoration are discussed in the second part of this chapter. We give an overview of various restoration techniques used in fluorescence microscopy. We discuss the influence of regularization and the background on the performance of non-linear image restoration algorithms.
1.1 Fluorescence Microscopy

Since the invention of the first compound light microscope by Zacharias Jansen in 1595 (Jones, 1995), the instrument has evolved enormously. The modern light microscope is a versatile instrument for microscopic analysis. The construction of the first epi-illuminated fluorescence microscope by Ploem (Ploem, 1967), made the light microscope a useful instrument for fluorescence microscopy. In epi-illumination, the illumination of the sample and the detection of its emitted fluorescence light are done using the same objective lens (Figure 1.1). This strongly reduces the penetration of illumination light in the detection light path, which makes the detection of the weak fluorescence light feasible\(^1\).

![Diagram of epi-illumination fluorescence microscope]

**Figure 1.1** Schematic diagram of a conventional wide-field fluorescence microscope showing its inability of to discriminate the out-of-focus light from the in-focus light.

A strong characteristic of the epi-fluorescence microscope is its wide-field illumination, which enables the simultaneous imaging of the entire focal plane. Modern scientific grade fluorescence microscopes are excellent tools for acquiring microscopic images of two-dimensional samples with a discriminating power of well below one micrometer.

The performance of a wide-field microscope in acquiring three-dimensional data is not as good as one would like. The wide-field illumination turns out to be a major drawback of the microscope. Since the whole sample is illuminated simultaneously, it will not only

\(^1\) The emitted fluorescence light is typically a thousand times less intense than the illumination light
excite fluorophores in the focal plane but also in the out-of-focus regions of the sample as well.

When a fluorescence sample is illuminated with light of the proper wavelength (in the absorption spectrum of the fluorescence molecules), it will emit light of a longer wavelength. This emitted light can then be detected using, for example, a CCD camera. A camera will acquire a two dimensional image of the emitted light intensity. The light emitted from out-of-focus regions cannot however be distinguished from the light emitted from the in-focus light as illustrated in Figure 1.1. Therefore the image will be a combination of a sharp, focused image of the in-focus plane with a blurred, unsharp image of the out-of-focus light. (It is for this reason that the primary application of a wide-field microscope is in the imaging of thin samples.)

The acquisition of both the in-focus and out-of-focus light results in the very poor resolution of a conventional wide-field fluorescence microscope along the optical axis. To illustrate this further, we have acquired an image of a spherical fluorescence bead (diameter 6.42 μm) with such a microscope. Both the lateral plane and the axial plane through the center of the bead are shown in Figure 1.2 together with a circle indicating the nominal diameter of the bead.

![Figure 1.2 Center x-y (left) and x-z (right) slices of a fluorescence bead imaged by a conventional fluorescence microscope. The circle indicates the nominal diameter of the bead (6.42 μm). The image is sampled laterally at 0.23 μm, and at 0.225 μm in the axial direction.](image)

### 1.1.1 Confocal Microscopy

A confocal microscope (Pawley, 1995) yields improved axial resolution by sacrificing the wide-field illumination in favor of a point illumination and point detection. By illuminating a single point at the time, one can discriminate the in-focus light from the out-of-focus light. In a confocal microscope this is implemented by placing a pinhole in front of the detector (see Figure 1.3). Although the axial resolution is greatly improved by preventing the detection of the out-of-focus light, fundamental laws of optics still restrict
the axial resolution of a confocal microscope to less then one third of the lateral resolution of that microscope.

![Confocal Microscope Diagram](image)

**Figure 1.3** Schematic diagram of confocal microscope showing the blocking of the out-of-focus light by the detection pinhole.

### 1.1.2 Image Formation

The resolution of a microscope determines the limit for the detection of fine details in the sample. The way the resolution of the microscope influences the image can be modeled by the point spread function of the microscope. The point spread function determines how a single point in the sample is being imaged. Representing the sample as a collection of points, the image of the sample can be composed by replacing each point in the sample with the point spread function weighted by the intensity of each point. This operation is defined as a convolution. We model the formation of an image in a fluorescence microscope as a convolution of the sample with the point spread function of the microscope.

A large point spread function will impose significant blurring of the sample. This reduces the resolution in the image. The point spread function of a conventional wide-field fluorescence microscope is relatively large whereas the full-width-at-half-maximum of the confocal point spread function is considerably smaller (see Figure 1.4).
conventional PSF

confocal PSF

Figure 1.4 Center x-z slices of a conventional and confocal fluorescence point spread function. Each black-white transition in the image represents a decay in intensity of a factor of ten, with the highest intensity in the center of the image.

Besides blurring by the point spread function, two other factors influence the image formation of a fluorescence microscope. Not only does the illuminated sample contribute to the intensity acquired in the image, but background intensity originating from autofluorescence, background fluorescence, scattering and offsets in the detector gain contribute as well. Furthermore, noise will distort the image. Noise in a fluorescence microscope originates from different sources. Not only will electronic components of the detector contribute noise, but the inherent uncertainty in photon counting yields a Poisson distributed signal as well. In modern, scientific grade microscope systems Poisson noise, caused by photon counting, is the dominant source of noise. An illustration of the image formation of a fluorescence microscope modeled in this way is given in Figure 1.5.

Object    PSF    Blurred Object    Image

Figure 1.5 Flow diagram of the image formation in a fluorescence microscope. The symbol ⊗ denotes the convolution operator.

1.2 Image Restoration

Blurring of the image by the point-spread-function of a fluorescence microscope can hamper quantitative analysis of the image by decreasing the accuracy of the measurements performed.

The goal of image restoration is to invert the degradations that the microscope imposes on the image. This requires an accurate model of the image formation.
A fluorescence microscope imposes two types of distortions on the image, a deterministic blurring by the point spread function and a stochastic distortion (by noise). We can therefore formulate the goal of image restoration as the reconstruction of the original sample from the acquired image using the point spread function and a model for the type of noise the image is disturbed with as a priori knowledge (see Figure 1.6). Poisson noise will be the main source of noise in an image acquired with a scientific grade microscope system. Most image restoration methods require a priori knowledge about the point spread functions. Blind image restoration algorithms, however, do not require knowledge about the point spread function. Instead they aim to restore the original image from the original image and estimate, as a by-product, the point spread function. (Ayers & Dainty, 1988; Holmes, 1992; Fish et al., 1995; Krishnamurthi et al., 1995; Thiebaut & Conan, 1995; Conchello & Yu, 1996; Kundur & Hatzinakos, 1996).

1.2.1 Linear Image Restoration
We model the deterministic blurring as a convolution of the image with the point spread function. In the frequency domain a convolution transforms into a multiplication of the Fourier transform of the sample with the optical transfer function. The optical transfer function (OTF) is the Fourier transform of the point spread function. Therefore a naive form of image restoration is to divide the Fourier transform of the image by the OTF. This procedure is known as inverse filtering (Jain, 1989). Figure 1.7 shows the result of the
Inverse filter on a synthetic image of a spherical object blurred with a confocal point spread function and disturbed with Poisson noise.

*Figure 1.7* Restoration result of the inverse filter\(^2\) of an image of a syntactic spherical object which has been blurred by a confocal point spread function.

It is clear from Figure 1.7 that the result of inverse filtering is not usable due to large amplification of noise. This amplification is caused by the reconstruction by the inverse filter of the high frequencies in the Fourier spectrum of the image. Not only are these frequencies dominated by noise, but the optical transfer function has low intensities at these frequencies as well. Therefore the inverse filter divides noise-dominated frequencies by low transfer values which result in a strong amplification of the noise in the restoration result.

The inverse filter takes only the deterministic distortion, the blurring, into account, which results in an unacceptable performance on noisy images. Therefore we need more sophisticated algorithms that take both the deterministic and the stochastic distortions into account.

A common way to restore blurring in the presence of noise is to incorporate regularization in the restoration procedure. In the first part of chapter 3 we discuss the Wiener filter and the Tikhonov-Miller filter. Both are linear, regularized restoration filters. They regularize their results by restoring the frequencies, which are dominated by the object and to suppress those frequencies that are dominated by noise. This way the problems that arise when using the inverse filter are avoided.

Both the Wiener filter and the Tikhonov-Miller filter are linear space-invariant filters (Andrews & Hunt, 1977; Jain, 1989). Therefore they cannot restrict the intensities in the restoration result to only positive values. This is an unacceptable property of linear image restoration since the intensities in a fluorescence image represent photons of which only a positive number can exist. Furthermore these linear filters restore the image by

---

\(^2\) The result of the inverse filter is obtained with a pseudo-inverse filter which sets those frequencies in the restoration result to zero at which the transfer function is zero as well.
convolving it with a restoration filter. These restoration filters approximate the inverse filter at frequencies dominated by the object, and have a low or zero response at frequencies dominated by noise and at those frequencies at which the OTF has low or zero values. Therefore these linear restoration filters will only restore frequencies inside the bandwidth of the OTF and will thus not restore frequencies beyond the bandwidth of the OTF. Algorithms that do restore frequencies beyond the bandwidth of the OTF are said to produce images at "superresolution" (Gerchberg, 1979; Jain, 1989; Carrington et al., 1995). Clearly linear restoration algorithms lack this property.

1.2.2 Non-linear Image Restoration

Agard and Sedat (Agard, 1984; Agard et al., 1989; Agard et al., 1989) were among the first to apply image restoration in three-dimensional fluorescence microscopy. In their work they have used three restoration approaches, nearest neighbor deconvolution, Wiener filtering and a modified Jansson-van Cittert technique (Jansson et al., 1976; Castleman, 1996). Realizing that normally most of the blurring will arise from neighboring planes, a considerable removal of the image blur can be obtained using the intensities from only these planes. Such a nearest neighbor algorithm does not require the computation of the three-dimensional Fourier transform, which results in a quick restoration algorithm. The van-Cittert algorithm is a deterministic deconvolution technique. Agard and Sedat modified the Jansson-van Cittert algorithm to incorporate a non-negativity constraint, which they implemented by clipping the negative intensities of each iteration to zero.

Over the years the computational power and memory capacity of computers has increased enormously. At the present time iterative restoration algorithms that compute a few three-dimensional Fourier transforms in each iteration can be applied on images of a useable size (for example 256 by 256 by 128 pixels). These algorithms have been shown to outperform the above mentioned Jansson-van Cittert algorithm (Conchello & Hansen, 1990).

Image Restoration based on the Richardson-Lucy algorithm

Holmes (Holmes, 1988; Holmes, 1989; Holmes & Liu, 1989; Holmes & Liu, 1991; Willis et al., 1993; Holmes et al., 1995) has introduced the EM-MLE algorithm to fluorescence microscopy. This algorithm has been developed as a reconstruction algorithm for computer tomography (Shepp & Vardi, 1982) using the Expectation-Minimization Algorithm (EM) (Dempster et al., 1977). The algorithm iteratively finds the maximum likelihood estimator (MLE) when the image is distorted by Poisson noise.
The EM-MLE had been derived earlier by Richardson-Lucy (Richardson, 1972; Lucy, 1974), using different principles, for image restoration in astronomy. In this thesis we refer to the EM-MLE algorithm as the Richardson-Lucy algorithm. The algorithm is a non-linear, iterative algorithm that will produce a positive constrained restoration result. The Richardson-Lucy algorithm computes the unregularized maximum likelihood estimator. It is therefore sensitive to the noise realizations present in the acquired image (van Kempen et al., 1997). To reduce this sensitivity, several authors have proposed various methods to regularize the Richardson-Lucy algorithm. Snyder (Snyder & Miller, 1991) investigated the use of sieves for regularizing the EM-MLE algorithm. Good's roughness regularization has been incorporated in the Richardson-Lucy algorithm by Joshi and Miller (Joshi & Miller, 1993). Conchello derived a modified Richardson-Lucy algorithm (Conchello & McNally, 1996) which incorporates Tikhonov regularization (Tikhonov & Arsenin, 1977). We will refer to this algorithm as the RL-Conchello algorithm. In the implementation of the Richardson-Lucy and RL-Conchello algorithms we have used the incorporation of the background as proposed by Snyder et al. (Snyder et al., 1993). Finally, based on the EM-MLE algorithm, several authors (Holmes, 1992; Fish et al., 1995; Krishnamurthi et al., 1995; Thiebaut & Conan, 1995) have derived blind deconvolution algorithms which do not require a priori knowledge about the point spread function. Instead these algorithms restore both the original object and the point spread function from the acquired image.

**Non-linear Tikhonov image restoration**

The Tikhonov-Miller algorithm is the linear restoration filter found when minimizing the Tikhonov functional (Tikhonov & Arsenin, 1977). The Tikhonov functional consists of a mean-square-error term, which measures the distance of the restoration result, blurred by the point spread function, to the acquired image and of a Tikhonov regularization term. Both the iterative constrained Tikhonov-Miller (ICTM) algorithm (Lagendijk & Biemond, 1991; van der Voort & Strasters, 1995) and the Carrington algorithm (Carrington & Fogarty, 1987; Carrington, 1990; Carrington et al., 1995) are non-linear algorithms that iteratively minimize the Tikhonov functional. The ICTM algorithm imposes the non-negativity constraint by clipping the negative intensities of each iteration to zero. Like the ICTM algorithm, the Carrington algorithm uses the conjugate gradient descent algorithm to numerically find the minimum of the Tikhonov functional. However, it constrains the solution by transformation. The Carrington and ICTM algorithm as well as the Richardson-Lucy algorithm is discussed in detail in the second part of Chapter 3. Recently Verveer (Verveer & Jovin, 1997; Verveer & Jovin, 1997; Verveer et al., 1997; Verveer et al., 1998) proposed an alternative method for minimizing the Tikhonov
functional. He used a quadratic transformation to incorporate the non-negativity constraint and used the conjugate gradient descent algorithm for minimizing the Tikhonov functional. Furthermore, Verveer has shown how to optimize the conjugate gradient descent algorithm for the proposed transformation. Using this transformation, he (Verveer & Jovin, 1997; Verveer et al., 1997; Verveer et al., 1998) has derived a family of restoration algorithms which are suitable for different types of noise (Poisson noise and additive Gaussian noise) as well as various types of regularization (Tikhonov, entropy and Good's roughness).

1.2.3 Regularization

Most of the restoration algorithms studied in this thesis incorporate regularization. These algorithms balance the fit of their restoration result to the acquired image with an a priori model of the restoration result (Andrews & Hunt, 1977). These two terms are balanced by the regularization parameter. A large value of the regularization parameter results in a stronger influence of the regularization on the restoration result, whereas a low value of the regularization parameter will make the restoration algorithms more sensitive to the noise in the acquired image. The regularization parameter, therefore, has a large influence on the result produced by the restoration algorithm.

![Figure 1.8 Restoration results of the ICTM algorithm for a high value (0.1) (left), a low value (0.00001) (right) and a well-chosen value (0.006) for the regularization parameter.](image)

Figure 1.8 shows the restoration result produced by the ICTM algorithm for a low, a high and a well-chosen value for the regularization parameter. In chapter 4 we investigate the influence of the regularization parameter on the performance of various restoration algorithms and compare various methods for determining the regularization parameter for the ICTM and Carrington algorithm. Furthermore it is shown how the regularization parameter for the RL-Conchello algorithm can be estimated.
1.2.4 The Influence of the Background

The ICTM, Carrington, Richardson-Lucy and RL-Conchello algorithms are non-linear algorithms that minimize the Tikhonov functional to produce non-negative restoration results. The incorporation of the non-negativity makes these algorithms non-linear and are said to be capable of producing superresolution results (Holmes, 1988; Carrington et al., 1995). The ICTM algorithm, for example, implements this non-negativity by clipping the negative values of each iteration to zero. The clipping however will only be in effect when negative intensities are computed by the iteration. Therefore parameters that influence these intensities will influence the effectiveness of the clipping. One of these parameters, the estimated amount of background, is an input parameter of the non-linear restoration algorithms. In chapter 5 we study extensively the influence of the estimated background on the performance of the non-linear restoration algorithms and propose a method for the estimation of the background.

1.3 Scope of this Thesis

In this thesis we investigate the application of non-linear image restoration algorithms in fluorescence microscopy. In our research we use three popular and often applied image restoration algorithms: the ICTM algorithm, The Carrington algorithm and the Richardson-Lucy algorithm. Furthermore we have included the Tikhonov regularized Richardson-Lucy algorithm proposed by Conchello. We study the influence of various parameters, such as the regularization parameter, first estimate, and background estimation on the performance of these algorithms and show how values for these parameters can be determined. We propose novel methods to improve the performance of the algorithms by reducing the influence of the noise on the restoration. Finally, we study the improvement on quantitative analysis when image restoration is applied prior to the measurement.

Chapter 2 is devoted to the image formation in a fluorescence microscope. We derive a mathematical description of the point spread function assuming diffraction limited optics. We discuss the conditions for modeling the image as a convolution of the sample with the point spread function and show how this model can be used to model the properties of both a conventional fluorescence microscope and a confocal fluorescence microscope. Finally we use the particle description of light to discuss the noise influence on the image formation in a fluorescence microscope.
Chapter 1  Image Restoration in Fluorescence Microscopy

In Chapter 3 we derive the image restoration algorithms used in this thesis. In the first part of this chapter an introduction into classical image restoration is given and two linear image restoration algorithms, the Wiener filter and the Tikhonov-Miller filter are derived. In the second part of this chapter we discuss iterative, non-linear restoration algorithms. We derive the iterative constrained Tikhonov-Miller algorithm, the Carrington algorithm, the Richardson-Lucy algorithm and a Tikhonov regularized Richardson-Lucy algorithm.

In the first part of Chapter 4 we discuss the setup of the simulation experiments performed. We discuss several measures for measuring the performance of the restoration algorithms and the generation of synthetic bandlimited objects. In the second part of this chapter we investigate the influence of the regularization parameter and the first estimate on the performance of the non-linear iterative image restoration algorithms. Finally we discuss a criterion for stopping the iteration process of these algorithms.

Chapter 5 discusses the application of image restoration in fluorescence microscopy. We start by showing the dependency of the performance of the non-linear restoration algorithms on the background estimation and show how the optimal background value can be determined. We continue with discussing a method, which we call prefiltering, that reduces the influence of the noise in the acquired image on the restoration result. In a third section we present a feasibility study of applying confocal microscopy in the analysis of mitotic cells. We show how image restoration prior to a quantitative measurement can improve the accuracy of such a measurement.

In Chapter 6 we present methods to improve the cache performance of separable filters on modern workstations. The most time-consuming part of iterative restoration algorithms is the fast Fourier transform used to compute the convolution. The fast Fourier transform is an example of a separable image operation. We show that a straightforward implementation of separable image operations will lead to a worst case use of the cache. We present methods, based on transposing the image that optimize cache usage.

In Chapter 7 we draw conclusions and summarize the research presented in this thesis.
Chapter 2

Image Formation

This chapter describes a model for the image formation in a fluorescence microscope based on the wave description and the quantum nature of light. Using the wave description of light, the finite resolution of an image obtained with a microscope is derived. We discuss the conditions under which a fluorescence microscope can be modeled as a linear translation invariant system. Diffraction theory is used to model the field of incident light near focus. Using this model, we derive the image formation in a general fluorescence microscope, having both a finite sized illumination and detection aperture. The image formation in a confocal fluorescence microscope, a wide-field fluorescence microscope and a scanning fluorescence microscope are derived as special cases.

We discuss sampling theory to formulate the conditions for an error free conversion of an analog image into a digital representation.

Using the quantum nature of light, we describe the noise properties of a light detection system. Both intrinsic and extrinsic noise sources are treated, as well as the photon-limited characteristics of scientific grade light detectors.
2.1 Wave Description of Light

In this section we use the wave description of light to describe the diffraction properties of an optical system. In the next paragraphs we discuss the conditions which allow an optical system to be regarded as linear translation invariant. Diffraction theory is used to model the imaging properties of incoherent optical systems. A comprehensive description of these subjects can be found in (Born & Wolf, 1980; Goodman, 1996).

2.1.1 Light as a wave disturbance

Light can be described as an electromagnetic wave, the propagation of which is governed by the Maxwell equations. Although the properties of an optical system are described in terms of electromagnetic waves no sensor exists for directly measuring the amplitude and phase of optical radiation. Light detectors and sensors are sensitive to the intensity of the incident radiation. Intensity is the time average of the radiation energy which crosses a unit area in a unit time,

\[ I(x) = \langle |U(x,t)|^2 \rangle \]

(2.1)

with \( U \) the wave disturbance, \( I \) the resulting intensity, and the brackets \( \langle \rangle \) to indicate the temporal averaging process.

2.1.2 Linear translation invariant system

In order to use linear system theory we have to show that a high quality optical system is, to a good approximation, a linear translation invariant (LTI) system.

An optical system is translation-invariant if the image of a point source changes only in location, not in form, when the point source is translated over the object field. The point spread function (or impulse response) \( h_{\xi, \eta, \zeta}(x, y, z) \) denotes the response at position \((x, y, z)\) to an unit impulse response at position \((\xi, \eta, \zeta)\). The unit impulse response \( \delta \) is defined as

\[ \delta(x, y, z) = \delta(x) \delta(y) \delta(z) \]

(2.2)

with

\[ \delta(x) = 0 \quad \forall x \neq 0 \]

\[ \int dx \delta(x) = 1 \]  

(2.3)

The impulse response of a translation-invariant system depends only on the distances \((x - \xi), (y - \eta)\) and \((z - \zeta)\). For a translation-invariant system we can therefore write

\[ h_{\xi, \eta, \zeta}(x, y, z) = h(x - \xi, y - \eta, z - \zeta) \]

(2.4)

The image \( g \) of an object \( f \) obtained by a translation-invariant system is therefore written as a convolution.
\[
g(x,y,z) = \int \int \int d\xi d\eta d\zeta h(x-\xi, y-\eta, z-\zeta) f(\xi, \eta, \zeta)
\]
\[= h(x, y, z) \otimes f(x, y, z)
\]

with \( \otimes \) denoting the convolution operator.

In practice optical systems are seldom translation-invariant or *isoplanatic* over the entire object field. However it is usually possible to divide the object field into regions (isoplanatic patches) in which the system is approximately translation-invariant.

So far we have ignored the magnification properties. The magnification makes it more difficult to model an optical system as a translation invariant system. We will not discuss it here, but instead refer to section 2.1.4.

A system is said to be linear if a weighted combination of inputs yields a weighted combination of outputs. Linear superposition requires that the response to multiple inputs at one time is equal to the sum of the individual responses. For an optical system this results in the following requirement

\[
\left\langle aU(x_1) + bU(x_2) \right\rangle = |a|^2 \left\langle U(x_1) \right\rangle + |b|^2 \left\langle U(x_2) \right\rangle \quad a, b \in C
\]

(2.6)

From equation (2.6) follows that

\[
\left\langle U(x_1)^* U(x_2) \right\rangle = 0
\]

(2.7)

which is the definition of incoherence. If wavefronts from different sources exhibit a fixed phase relationship, they are said to be coherent. The complete absence of this dependency is called incoherence.

In this thesis we focus on incoherent imaging as it occurs in fluorescence microscopy. No fixed phase relationship exists among the fluorescence from the molecules that compose the object making the fluorescence light incoherent (Lakowicz, 1983). In wide-field microscopy Köhler illumination provides approximately incoherent wide-field illumination of the object if the condenser is wide enough to fill the entire entrance pupil of the objective lens (Martin, 1966).

### 2.1.3 Diffraction theory

In this section we discuss scalar diffraction theory following the derivations found in (Born & Wolf, 1980). A comprehensive description of diffraction theory can be found in (Born & Wolf, 1980; Stammes, 1986; Williams & Becklund, 1989; Goodman, 1996).

The Kirchhoff diffraction integral describes the complex amplitude \( U \) at an observation point \( P \) at a distance \( s \) from a point \( Q \) on the aperture caused by a point source at a distance \( f \) from the aperture (see Figure 2.1),
Chapter 2  Image Formation

\[ U(P) = \frac{i}{\lambda} A e^{-i\chi} \int \frac{dS e^{ikq}}{s} \]  \hspace{2cm} (2.8)

where \( Af \) is the amplitude of the incident wave at \( Q \) (Born & Wolf, 1980).

![Diagram of diffraction](image)

**Figure 2.1** Diffraction of a converging spherical wave at a circular aperture

If \( q \) denotes the unit vector in the direction of \( OQ \), we have to a good approximation,

\[ s - l = -q \cdot R \]  \hspace{2cm} (2.9)

We define \( (x, y, z) \) to be the Cartesian coordinates of \( P \) and \( (\xi, \eta, \zeta) \) of \( Q \). Changing to the cylindrical coordinates gives (Born & Wolf, 1980)

\[ \begin{align*}
\xi &= ap \cos \psi, \\
\eta &= ap \sin \theta, \\
x &= R \cos \psi, \\
y &= R \sin \psi, \\
dS &= a^2 \rho d\rho d\theta / f^2,
\end{align*} \]  \hspace{2cm} (2.10)

with \( a \) the radius of the objective aperture. Since \( Q \) lies on the spherical wave-front \( S \),

\[ \zeta = -\sqrt{f^2 - a^2 \rho^2} = -f \left(1 - \frac{a^2 \rho^2}{f^2} + \ldots\right) \]  \hspace{2cm} (2.11)

and

\[ q \cdot R = \frac{x \xi + y \eta + z \zeta}{f} = \frac{a p \cos (6 - \psi)}{f} = \zeta \left(1 - \frac{a^2 \rho^2}{f^2} + \ldots\right) \]  \hspace{2cm} (2.12)

It is useful, at this stage, to introduce the dimensionless variables \( u \) and \( v \),

16
\[ u = \frac{2\pi}{\lambda} \left( \frac{a}{f} \right)^2 z \]  
(2.13)

\[ v = \frac{2\pi}{\lambda} \left( \frac{a}{f} \right) \frac{x}{\sqrt{x^2 + y^2}} \]

Substituting (2.10) and (2.12) in (2.8) yields

\[ U(u, v) = -\frac{ika^2 A e}{z_2^2} e^{\frac{iz_2}{a}} \left[ \int_0^1 \rho d \rho J_0(\rho v) e^{-\frac{i}{2} \rho^2} \right]^2 \]  
(2.14)

In this derivation the following relation has been used

\[ \cos \theta \cos \psi + \sin \theta \sin \psi = \cos(\theta - \psi) \quad \text{and} \quad \int_0^{2\pi} d \theta e^{-ix \cos(\theta - \phi)} = 2\pi J_0(x) \]  
(2.15)

The incoherent diffraction point spread function is given by

\[ h(u, v) = |U(u, v)|^2 = \left( \frac{ka^2 A e}{z_2^2} \right)^2 \left[ \int_0^1 \rho d \rho J_0(\rho v) e^{-\frac{i}{2} \rho^2} \right]^2 \]  
(2.16)

This expression can be approximated by a finite series of Bessel function (Born & Wolf, 1980) or evaluated numerically using numerical integration techniques (Press et al., 1992).

The derivations given in this section result in only one approximation for the three dimensional point spread function. A large collection of papers has been published deriving alternative approximations for the three dimensional point spread function based on scalar diffraction theory (Hopkins, 1955; Frieden, 1967; Born & Wolf, 1980; Sheppard, 1986; Gibson & Lanni, 1989; Wilson, 1990) as well as on vectorial diffraction theory (Richards & Wolf, 1959; van der Voort & Brakenhoff, 1990).

2.1.4 Magnification

So far, we have not discussed the magnification properties of a lens. From geometrical optics, the relation between an image \( g \) and the object \( f \) is given by (Young, 1989; Tao & Nicholson, 1995; Goodman, 1996),

\[ g(x, y) = \frac{1}{|M|} f \left( \frac{x}{M}, \frac{y}{M} \right) \]  
(2.17)

This system is characterized by the following geometrical impulse response,

\[ h_{x, y, z}(x, y, z) = \frac{1}{|M|^3} \delta \left( \frac{\xi - x}{M^3}, \eta - \frac{y}{M^3}, \zeta - \frac{z}{M^3} \right) \]  
(2.18)

This impulse response is not translation-invariant. Geometrical optics does not include the effects of diffraction. A more complete description of the image system can be obtained if
both are combined. To do this we have to reduce the geometrical object-image relation to a convolution equation. This can be obtained by normalizing the object coordinates,

$$\tilde{x} = M \xi \quad \tilde{\eta} = M \eta \quad \tilde{\zeta} = M^2 \zeta$$

(2.19)

The image obtained with a diffraction limited magnifying lens can thus be modeled as

$$g(x, y, z) = h(\tilde{x}, \tilde{y}, \tilde{z}) \otimes f(\tilde{x}, \tilde{y}, \tilde{z})$$

(2.20)

with $h$ the diffraction PSF as derived in the previous section and $\tilde{x}, \tilde{y}, \tilde{z}$ normalized coordinates defined by (2.21).

In the following sections we will not consider the magnification in our derivations, but assume a magnification of one. The results can be used for magnifying lenses using the coordinate normalization as defined by (2.20).

### 2.2 Fluorescence Microscopy

In fluorescence microscopy the fluorescent molecules in the illuminated object are excited by incident light of wavelength $\lambda_{ex}$. The excited molecules emit light of wavelength $\lambda_{em}$ which is collected by the microscope forming a fluorescent image. The difference $\Delta \lambda$ ($\Delta \lambda = \lambda_{em} - \lambda_{ex} > 0$) between $\lambda_{em}$ and $\lambda_{ex}$ is called the Stokes shift of the fluorescent molecule (Lakowicz, 1983). In this thesis we will assume in our derivations that both the excitation and emission light is monochromatic,

$$I_{ex}(\lambda) = I_{ex} \delta(\lambda - \lambda_{ex})$$

$$I_{em}(\lambda) = I_{em} \delta(\lambda - \lambda_{em})$$

(2.22)

Although most fluorescent molecules have broad excitation and emission spectra, this assumption can be fulfilled by placing narrow bandpass filters in the illumination and detection light paths. In the absence of these filters the formulae we derive in the following sections have to be integrated over the excitation and emission spectra.

The illumination intensity $I_{ill}$ is the amount of light that is projected from the source onto the sample. The illumination intensity determines the probability that an excitation photon "hits" a fluorescent molecule at a certain point in the object. Similarly, the detection intensity quantifies the probability that a fluorescent photon emitted from a point in the object is able to propagate to the detector. The detection intensity is therefore dependent on the intensity of both the excitation and emission light.
Figure 2.2 Model of a general epi-fluorescence microscope.

We use a general model of an epi-fluorescence microscope (see Figure 2.2) to model the illumination intensity and the detection intensity. In epi-fluorescence, the illumination and detection light path share the same objective which strongly reduces the amount of illumination light penetrating the detection light path (Ploem, 1967). A system diagram of the general epi-fluorescence microscope, as shown in Figure 2.3, makes the “flow of data” in the microscope clearer. This diagram also depicts the symbols we use to refer to the various components of the microscope. This model assumes that the illumination and the detection are limited by finite sized apertures, $A_{\text{ill}}$ and $A_{\text{det}}$, respectively. In a wide-field fluorescence microscope, the illumination aperture is determined by the size of diaphragm, in a confocal microscope by the illumination pinhole. The detection aperture is determined by the size of the detection pinhole in the case of a confocal microscope, and in the case of a wide-field microscope by the camera.

In our general epi-fluorescence microscope a point in the object is illuminated through the illumination aperture $A_{\text{ill}}$ of the light source. The light source can be modeled as a collection of point sources spread over the aperture. Each point source is projected by the objective on the sample (Jovin et al., 1990). The illumination intensity can therefore be written as the convolution between the illumination aperture and the diffraction point spread function of the objective at the excitation wavelength,

$$I_{\text{ill}}(x) = \int d\xi h_{\text{ill}}(x - \xi) A_{\text{ill}}(\xi)$$

where $h_{\text{ill}}$ represents the diffraction PSF at the illumination wavelength $\lambda_{\text{ill}}$. 

19
In (2.23) we have used a single lateral coordinate \( x \). We will continue to do so as it makes the derivations easier to follow. At the end of this section we will reformulate the most important results for the three dimensional case.

\[
h_{ir}(x) = h(x, \lambda_{ex}) \tag{2.24}
\]

Figure 2.3 System diagram of a general epi-fluorescence microscope, showing the “flow of data” in the microscope.

In fluorescence imaging, the intensity of the emission light is determined by the intensity of the illumination light and the amount of fluorophores. For a low concentration of fluorophores, the emitted light \( I_{em} \) is the product of the fluorophore distribution (the object) \( O(x) \) and the illumination intensity,

\[
I_{em}(x, \lambda_{em}) = O(x, \lambda_{ex}) \lambda_{em} h_{il}(x, \lambda_{ex}) \tag{2.25}
\]

The emitted light is imaged by the objective onto a photo-sensitive sensor. Behind the objective, the detection aperture limits the emission light with respect to the sensor. The intensity behind this aperture can be modeled as,

\[
I_{det}(x) = A_{det}(x) \int d\xi h_{em}(x - \xi) I_{em}(\xi) \tag{2.26}
\]

with \( h_{em} \) the diffraction point spread function of the objective at the emission wavelength

\[
h_{em}(x) = h(x, \lambda_{em}) \tag{2.27}
\]

The point spread function of a general fluorescence microscope can now be found by substituting an impulse function \( \delta(x) \) for \( O(x) \),

\[
h_{GFM}(x) = A_{det}(x) h_{em}(x) \int d\xi h_{ex}(-\xi) \lambda_{il}(\xi) \tag{2.28}
\]
In the next two sections we will derive the detection intensity of a confocal and a wide-field fluorescence microscope as limiting cases of the presented general fluorescence microscope.

2.2.1 Confocal fluorescence microscopy

A confocal fluorescence microscope illuminates the object through a pinhole aperture. This is often implemented by illuminating the focal plane of the sample with a focused laser beam. The emission light is collected with a light detector (typically a photomultiplier tube (PMT)) placed behind a detection pinhole. The detection pinhole blocks the out-of-focus light. Figure 2.4 shows the schematic setup of an epi-fluorescence confocal microscope. Its system diagram is shown in Figure 2.5.

![Schematic setup of an epi-fluorescence confocal microscope](image)

**Figure 2.4** Schematic setup of an epi-fluorescence confocal microscope.

The confocal setup as shown in Figure 2.4 detects the light at a single position. A lateral scan of the sample yields a two dimensional image. A three dimensional image is obtained when the sample is scanned both laterally and axially. The lateral scanning of the object is performed in a confocal laser scan microscope (CLSM) by scanning the laser beam over the objective. In the less frequently used confocal scanning laser microscope (CSLM) the scanning is performed by the x-y and z stages of the microscope.
An expression for the signal produced by the detector $I_{nur}$ can be derived from our model of the general fluorescence microscope. Assuming that both pinholes are infinitely small, we model them by an impulse function $\delta(x)$. In a confocal microscope, the light that falls through the detection pinhole is collected by a single detector. Therefore $I_{nur}$ yields at a single point

$$I_{PMT}(0) = \int dx \, \delta(x) \int d\xi h_{em}(x, -\xi)O(\xi) \int dX \, h_{ex}(\xi - X) \delta(X)$$

$$= \int dx \, \delta(x) \int d\xi h_{em}(x, -\xi)O(\xi)h_{ex}(\xi)$$

$$= \int d\xi h_{em}(x, -\xi)O(\xi)h_{ex}(\xi)$$  \hspace{1cm} (2.29)

A three-dimensional confocal image is acquired by scanning the object in three dimensions. The detection intensity of each point of the object can be derived from (2.29) by shifting both the illumination and detection over the object,

$$I_{PMT}(x) = \int dx' \delta(x - x') \int d\xi h_{em}(x', -\xi)O(\xi) \int dX \, h_{ex}(\xi - X) \delta(X - x)$$

$$= \int d\xi \int dx' \delta(x - x')h_{em}(x', -\xi)O(\xi)h_{ex}(\xi - x)$$

$$= \int d\xi h_{em}(x - \xi)O(\xi)h_{ex}(\xi - x)$$ \hspace{1cm} (2.30)

$$= h_{em}(x)h_{ex}(-x) \otimes O(x)$$

In practice the detection pinhole is not infinitely small, but finite; its size is a compromise between resolution and the signal-to-noise ratio of the detected signal (Pawley, 1995, chapters 2, 11, and 22). The size of the illumination pinhole in the object space is
determined by the spot size of the focused laser source and the magnification of the objective. In practice these parameters are such that the illumination pinhole can be approximated by a point. For a finite sized detection pinhole $A_{de}$ we can derive from (2.30),

$$I_{PMT}(x) = \int dx' A_{de}(x-x') \int d\xi h_{em}(x'-\xi) O(\xi) h_{ex}(\xi-x)$$

$$= \int d\xi \int dx' A_{de}(x-x') h_{em}(x'-\xi) O(\xi) h_{ex}(\xi-x)$$

$$= \{A_{de}(x) \otimes h_{em}(x)\} \otimes O(x)$$

(2.31)

![PSF and OTF images](image)

Figure 2.6 The center XY (left) and XZ (right) slices of the confocal PSF (top) and confocal OTF (bottom). The images are sampled at the Nyquist rate in both the lateral (50 nm) and axial direction (170 nm). The images are displayed using a logarithmic contrast stretch.

For an infinitely large detection pinhole in equation (2.31), the convolution between the pinhole aperture and the detection PSF becomes a constant yielding,

$$A_{de}(x) = 1 \rightarrow I_{PMT}(x) \rightarrow h_{ex}(-x) \otimes O(x)$$

Finally, for a small Stokes shift ($\lambda_{ex} = \lambda_{em}$) and an infinitely small detection pinhole the confocal PSF can be found by replacing the object function in (2.30) with an impulse function.
Chapter 2  Image Formation

\[ h_{c}\beta(x, y, z) = h(x, y, z)^3 \]  \hspace{1cm} (2.32)

Images of the confocal PSF and the corresponding OTF are shown in Figure 2.6. These images are computed numerically from a theoretical description of the three dimensional PSF (van der Voort & Brakenhoff, 1990).

2.2.2 Wide-field fluorescence microscopy

In wide-field microscopy the object is illuminated by flat field illumination. The emission light is recorded by a camera. Figure 2.7 shows a schematic model of a wide-field epi-fluorescence microscope. The corresponding system diagram is shown in Figure 2.8.

![Model of a wide-field epi-fluorescence microscope.](image)

Using our model of a general fluorescence microscope, we can derive the detection intensity for a wide-field epi-fluorescence microscope. By enlarging the illumination aperture to an infinite size the illumination intensity (2.23) becomes a constant,

\[ I_{in}(x) = \int d\xi h_{in}(x - \xi) = C_{in} \]  \hspace{1cm} (2.33)

The detection intensity of a wide-field fluorescence microscope is found by enlarging the detection aperture to an infinite size,
\[ I_{\text{det}}(x) = A_{\text{det}}(x) \int d\xi h_{\text{em}}(x - \xi) O(\xi) I_{\text{ill}}(\xi) \]
\[ = \int d\xi h_{\text{em}}(x - \xi) O(\xi) C_{\text{ill}} \]
\[ = h_{\text{em}}(x) \otimes O(\xi) C_{\text{ill}} \]  \hspace{1cm} (2.34)

For a wide-field fluorescence microscope the aperture function \( A_{\text{det}} \) is defined as
\[ A_{\text{det}}(x, y) = \begin{cases} 1 & \text{x, y inside the aperture} \\ 0 & \text{x, y outside the aperture} \end{cases} \]

**Figure 2.8** System diagram of a wide-field epi-fluorescence microscope.

Often a CCD camera is used to record the detection intensity as modeled in Figure 2.8. A CCD camera consists of an array of detectors (sometimes referred to as pixels) (van Vliet et al., 1997). Each detector simultaneously acquires the light that reaches it. The number of acquired photons is converted to a discrete brightness value. This allows us to model the acquisition of an image with a CCD camera as a convolution of the image with the response function of a CCD element \( h_{\text{camera}} \):

\[ I_{\text{camera}}[n] = (I_{\text{det}}(x) \otimes h_{\text{camera}}(x)) \sum_{n=-\infty}^{\infty} \delta(x - n X_o) \]
\[ = (I_{\text{det}} \otimes h_{\text{camera}})(n X_o) \]
\[ = (I_{\text{det}} \otimes h_{\text{camera}})[n] \]  \hspace{1cm} (2.35)

where \( X_o \) is the center-to-center distance between the CCD elements and \( n \) is a discrete variable.
For square shaped CCD elements \( h_{\text{camera}} \) is a two-dimensional block function with its size determined by the size of the light-sensitive part of the CCD element. Substituting (2.34) in (2.35) we obtain,

\[
I_{\text{camera}}[n] = (h_{\text{camera}} \otimes h_{\text{em}} \otimes O)[n] \cdot C_{\text{ill}}
\]  

(2.36)

where we have used an infinitely large detection aperture. The PSF \( h_{\text{wfm}} \) of a wide field fluorescence microscope is found by substituting an impulse function for the object function, \( O(x) = \delta(x) \) and setting \( C_{\text{ill}} \) to one:

\[
I_{\text{wfm}}[n] = (h_{\text{camera}} \otimes h_{\text{em}})[n]
\]  

(2.37)

Given the size of a CCD pixel element (typically 6 - 23 μm (van Vliet et al., 1997)), and a typical magnification of 40 - 100 times, the CCD elements are small enough to allow the PSF to be approximated by (Young, 1989):

\[
I_{\text{wfm}}[n] = h_{\text{em}}[n]
\]  

(2.38)

**Figure 2.9** The center XY (left) and XZ (right) slices of the wide-field PSF (top) and wide-field OTF (bottom). The images are sampled and are displayed in the same way as Figure 2.6.
A three-dimensional image of an object can be obtained with a wide field fluorescence microscope by acquiring a stack of two-dimensional images of different \( x \cdot y \) planes of the object.

Figure 2.9 shows the PSF and the OTF of a wide-field fluorescence microscope. The image of the \( XZ \) slice of the OTF clearly shows the “missing cone” of a three-dimensional wide-field OTF (see also section 2.1.3).

In a scanning microscope the role of the apertures is reversed in comparison with a wide-field microscope: the object is illuminated with a point source and the emitted light is acquired by a light detector without a detection aperture. The resulting point spread function of scanning fluorescence microscope is therefore,

\[
h_{\text{SprM}}(x) = h_{\text{a}}(x) \tag{2.39}
\]

Although a \textit{single} photon scanning microscope is rarely used, a \textit{two-photon} scanning fluorescence microscope is becoming increasingly popular. In a two-photon setup a fluorescence molecule is excited by two photons of half the excitation energy (double the excitation wavelength). Only the presence of two photons at the same time and place enables the excitation of the fluorescent molecule by providing a sufficient amount of energy. Realizing that the point spread function can be seen as a spatial probability function (Denk et al., 1995), the chance of a two photon excitation is dependent on the product of the PSFs of both photons, resulting in the following PSF for a two photon fluorescence microscope,

\[
h_{\text{2PFM}}(x) = h_{\text{a}}(x)^2 \tag{2.40}
\]

where \( h_{\text{a}} \) is the diffraction PSF at the “two-photon” excitation wavelength. Although two photon excitation requires expensive, high powered, pulsed lasers, its improved depth penetration, reduced scattering and reduced bleaching characteristics make it an interesting alternative for a single photon confocal microscope, especially when combined with confocal detection (Pawley, 1995, chapters 26 and 28).

### 2.2.3 Microscopic image formation in three dimensions

In the previous sections we have used a single lateral coordinate \( x \) to describe the image formation in fluorescence microscopes. In this section we will reformulate the most important results for the three dimensional case. It will be convenient to use the following notation for the convolution of a two dimensional function with a three dimensional function,

\[
g(x,y,z) = \int d\xi d\eta h(x-\xi, y-\eta) f(\xi, \eta, z) = h(x,y) \otimes f(x,y,z) \tag{2.41}
\]
Chapter 2  Image Formation

To express the detector intensities in three dimensions it is necessary to realize that the aperture functions are two-dimensional lateral functions. The one-dimensional convolution in (2.28) should therefore be expressed as a two dimensional convolution as defined by (2.41). Using this definition the sensor intensity of a confocal fluorescence microscope (equation (2.31)) in three dimensions yields,

\[
I_{PMT}(x, y, z) = \left( A_{\text{pinhole}}(x, y) \otimes h_{em}(x, y, z) \right) \otimes h_{ex}(-x, -y, -z) \otimes O(x, y, z) \quad (2.42)
\]

The PSF of a confocal fluorescence microscope in three dimensions can be found by substituting an impulse for the object function,

\[
h_{CFM}(x, y, z) = \left( A_{\text{pinhole}}(x, y) \otimes h_{em}(x, y, z) \right) h_{ex}(-x, -y, -z) \quad (2.43)
\]

The sensor intensity of a wide-field fluorescence microscope is in three dimensions the two dimensional convolution of the camera function with the diffraction PSF at the detection wave length,

\[
I_{\text{camera}}(x, y, z) = h_{\text{camera}}(x, y, z) \otimes h_{em}(x, y, z) \otimes O(x, y, z)) \quad (2.44)
\]

The PSF of wide-field fluorescence microscope yields

\[
h_{\text{WFM}}(x, y, z) = h_{\text{camera}}(x, y, z) \otimes h_{em}(x, y, z) \quad (2.45)
\]

2.3 Image Sampling

A digital image can be obtained by sampling an analog image. The sampling theorem states that a digital image is an aliasing free representation of the analog image if the analog image is sampled with a sampling frequency higher than twice the highest frequency \( f_C \) (the cutoff frequency) of the analog image. The analog image can then be reconstructed without any errors from the samples by interpolation if the sampling density satisfies this Nyquist criterion.

The cutoff frequencies or bandwidths of a microscope system are determined by diffraction and the illumination and detection apertures. In the following we start by giving the bandwidth of the diffraction PSF. Then we derive the bandwidth of the entire microscope.

2.3.1 The bandwidth of the diffraction PSF

The frequency response of a point spread function is described by its Fourier transform, the optical transfer function or OTF. The bandwidth and the region of support of a diffraction limited OTF have been derived by (Wilson, 1990; Wilson & Tan, 1993; Goodman, 1996). The frequency response of the OTF can be determined geometrically as
the overlap between two displaced pupil functions (Goodman, 1996). The size of these pupil functions is determined by the numerical aperture of the objective lens. The region in which the frequency response of the OTF is non-zero (the footprint) can be found using the overlap of these pupil functions (Wilson & Tan, 1993). The footprint of a diffraction limited OTF is in good approximation given by (Wilson & Tan, 1993),

\[
\text{Re}\left\{\sqrt{|l^2 - r^2 - \frac{1}{2}d^2|}\right\} \geq 0 \quad \text{with} \quad |l| \leq 1, |d| \leq 1
\]  

(2.46)

where \( l = f_{\text{lateral}}/f_{c,\text{lateral}} \) and \( a = f_{\text{axial}}/f_{c,\text{axial}} \) are normalized spatial frequencies. The lateral and axial cutoff frequencies for a diffraction OTF are given by (Wilson & Tan, 1993)

\[
f_{c,\text{lateral}} = \frac{2\sin \theta}{\lambda} = \frac{2NA}{\lambda} \quad f_{c,\text{axial}} = \frac{n(1 - \cos \theta)}{\lambda} = \frac{2n\sin^2(\theta/2)}{\lambda}
\]

(2.47)

with \( NA \) the numerical aperture of the objective and \( \theta \) the half-aperture angle as defined by Figure 2.1. The diffraction limited confocal OTF is non-zero if (Wilson & Tan, 1993),

\[
|l| \leq 0.5 \quad \text{and} \quad |d| \leq 1.0
\]

(2.48)

\[
\text{Re}\left\{\sqrt{|l^2 - r^2 - \frac{1}{2}d^2|}\right\} \geq 0 \quad \text{with} \quad |l| \leq 1, |d| \leq 1
\]

Figure 2.10 Schematic representation of the footprints of a wide-field (dark grey) and a confocal (light grey) fluorescence OTF.

In the confocal case, the lateral and axial cutoff frequencies are twice those given by (2.47). Figure 2.10 shows the footprints of the wide-field and confocal fluorescence OTFs. This figure clearly shows the region around the axial frequency axis where the wide-field OTF has no support. This region is known as the missing cone of the wide-field OTF (Streibl, 1984; Streibl, 1985). This region is filled in the confocal case, which explains the optical sectioning properties of a confocal fluorescence microscope.
2.3.2 The bandwidth of linear systems

To derive the bandwidth of a confocal and a wide-field fluorescence microscope it is convenient to derive the bandwidth of the convolutions defined by (2.5) and (2.41) and the bandwidth of a signal obtained by the multiplication of two signals.

When $F$ and $H$ are the Fourier transforms of respectively $f$ and $h$, the Fourier transform of the convolution between $f$ and $h$ yields,

$$h(x,y,z) \otimes f(x,y,z) \xrightarrow{T_{sp}^{-1}} H(\omega_x, \omega_y, \omega_z)F(\omega_x, \omega_y, \omega_z)$$

(2.49)

Using the property that the Fourier transform of a convolution is equivalent to the multiplication of the Fourier transforms of the two signals, it is easy to derive that the bandwidth of the resulting signal is equal to the minimum of the bandwidth of the two signals,

$$h(x,y,z) \otimes f(x,y,z) \rightarrow f_c = \min\left( f_c^h, f_c^f \right)$$

(2.50)

Similarly the multiplication of two signals is equivalent to the convolution of their Fourier transforms,

$$h(x,y,z)f(x,y,z) \xrightarrow{T_{sp}^{-1}} \frac{1}{2\pi} \mathcal{F}^{-1} H(\omega_x, \omega_y, \omega_z) \otimes F(\omega_x, \omega_y, \omega_z)$$

(2.51)

where we have used the Fourier transform as defined by (Oppenheim et al., 1983). The bandwidth after the multiplication of two bandlimited signals is the sum of the bandwidths of the signals,

$$h(x,y,z)f(x,y,z) \rightarrow f_c = f_c^h + f_c^f$$

(2.52)

Finally the bandwidth of the two dimensional convolution is described as follows. This convolution is equivalent to a multiplication in every lateral slice of $F$ with $H$,

$$g(x,y,z) = h(x,y) \otimes f(x,y,z)$$

(2.53)

$$g(\omega_x, \omega_y, \omega_z) = \mathcal{F}_{2D} \{ g(x,y,z) \} = H(\omega_x, \omega_y)F(\omega_x, \omega_y, \omega_z)$$

In other words, $H$ is repeated without change in the axial direction yielding an infinite axial bandwidth. If the axial bandwidth of $f$ is a function of the lateral frequency, the axial bandwidth of the resulting signal $g$ is equal to the highest axial frequency of $f$ within the lateral bandwidth of $g$. The bandwidth therefore yields,

$$f_{c_{lateral}} = \min \left[ f_{c_{lateral}}^h, f_{c_{lateral}}^f \right]$$

(2.54)

$$f_{c_{axial}} = \max \left[ f_{c_{axial}} \mid F(f_{c_{lateral}}, f_{c_{axial}}) > 0 \right] \quad \forall f_{c_{lateral}} \leq f_{c_{lateral}}$$

A graphical representation of this result is given in Figure 2.11.
2.3.3 The bandwidth of fluorescence microscopes

The bandwidth of the confocal fluorescence microscope and the wide-field fluorescence microscope can easily be derived using the results of the previous sections. In the confocal case it is convenient to derive the "diffraction" bandwidth first and then combine this with the bandwidth of the detection pinhole.

As expressed by (2.43) the confocal PSF is the two dimensional convolution of the detection pinhole with the product of two diffraction PSFs. Using (2.52), the diffraction bandwidth of a confocal microscope yields,

\[ f_{c,\text{diffraction}} = f_{c,\text{illumination}} + f_{c,\text{detection}} \]  \hspace{1cm} (2.55)

where \( f_{c,\text{detection}} \) and \( f_{c,\text{illumination}} \) are the cutoff frequencies of the diffraction OTF (2.47) at the emission wavelength and excitation wavelength respectively. The bandwidths of the confocal microscope can be found using (2.54),

\[ f_{c,\text{lateral}} = \min \left[ f_{\text{pinhole}}, f_{c,\text{lateral}}^{\text{diffraction}} \right] \] \hspace{1cm} (2.56)
\[ f_{c,\text{axial}} = \max \left[ f_{\text{axial}}, |H_{\text{CRM}}(f_{\text{lateral}}, f_{\text{axial}})| > 0 \right] \forall f_{\text{lateral}} \leq f_{c,\text{lateral}} \]

where \( H_{\text{CRM}} \) is the confocal OTF. The bandwidth of a wide-field fluorescence microscope can be derived in a similar way. Using (2.54) the bandwidths of the wide-field PSF (2.45) are,

\[ f_{c,\text{lateral}}^{\text{conventional}} = \min \left[ f_{c,\text{lateral}}^{\text{camera}}, f_{c,\text{lateral}}^{\text{detection}} \right] \] \hspace{1cm} (2.57)
\[ f_{c,\text{axial}}^{\text{conventional}} = \max \left[ f_{\text{axial}}, |H_{\text{WFM}}(f_{\text{lateral}}, f_{\text{axial}})| > 0 \right] \forall f_{\text{lateral}} \leq f_{c,\text{lateral}}^{\text{conventional}} \]

If diffraction is determining the bandwidth in (2.56) and (2.57) and the Stokes shift is negligible, the bandwidth of a confocal microscope is twice that of a wide-field fluorescence microscope.

The influence of the pinhole on the frequency response of a confocal OTF is shown in Figure 2.12. For an increasing pinhole size the region of support of the confocal OTF decreases and approaches the shape of the wide-field OTF (as shown in Figure 2.9).
Chapter 2  Image Formation

For typical working conditions of a confocal microscope with an excitation wavelength of 488 nm, an emission wavelength of 520 nm, a backprojected pinhole size of 280 nm, and an oil-immersion objective with a NA of 1.3, the bandwidths are 1/102 nm$^{-1}$ in the lateral direction and 1/341 nm$^{-1}$ in the axial direction. This corresponds to a lateral Nyquist frequency of 19.6 pixels/μm and an axial Nyquist frequency of 5.86 pixels/μm.

![Figure 2.12](image.png) The center XZ slice of a simulated confocal OTF ($\lambda_e = 488$ nm, $\lambda_o = 520$ nm, NA = 1.3) for different pinhole sizes. The diameter of the pinhole is set to 0.1, 0.3, and 0.5 μm (top row, left to right), 0.8, 1.0 and 1.5 μm (middle row) and 2.0, 3.0 and 5.0 μm (bottom row). All images are sampled with the same sampling frequency as the top-left image, which is sampled at the Nyquist frequency. The images are displayed with a logarithmic contrast stretch.
2.4 Particle Description of Light

The quantum nature of light can be used to explain the noisy images that often occur in low light-level situations. Light can also be considered as a series of particles called photons. Each photon carries a certain amount of energy $E$,

$$E = h\frac{c}{\lambda}$$

where $c$ is the speed of light and $h$ is Planck’s constant. Photon production by any light source is a statistical process governed by the laws of quantum physics. The source emits photons at random time intervals. The number of photons in a fixed observation interval will result in a number that obeys Poisson statistics. Each observation will measure a number $p$ with a probability given by the Poisson distribution

$$P(p|\rho T) = \frac{\rho T)^p e^{-\rho T}}{p!}$$ \hspace{1cm} (2.58)

with $T$ the observation interval (or exposure time) and $\rho$ the photon flux (the probability distribution for $\rho T = 8$ is shown in Figure 2.13). The average of a large number of observations will approximate the expected photon production $\rho T$.

![Figure 2.13 Poisson distribution for an expected number of photons $\rho T = 8$.](image)

This photon detection induced Poisson noise is sometimes referred to as *intrinsic* noise (Pawley, 1995 chapter 2), and is unavoidable when acquiring an image. However, light detectors and sensors contribute in general extra *extrinsic* noise to the detected signal. This extrinsic noise can be caused by various sources such as dark current (Poisson distributed), electronic noise, detector readout (both Gaussian distributed), and quantization noise (uniformly distributed) (Castleman, 1996).

A light detector is said to be *photon-limited* if the extrinsic noise is negligibly small compared to the amount of intrinsic noise induced by the detection of photons. Measurements show that scientific CCD cameras and PMTs can be regarded as photon-limited (Mullikin et al., 1994; Pawley, 1995; van Vliet et al., 1997).
2.5 Image Formation in a Fluorescence Microscope

Using both the wave description and quantum nature of light, as discussed in the previous sections, we can build a model for the image formation of a fluorescence microscope. The incoherent nature of the fluorescence light allows us, given the assumption of isoplanatism, to model the fluorescence microscope as a linear translation invariant system. The image is in that case a convolution of the object with the point spread function of the fluorescence microscope.

The detection of photons in a finite time interval distorts the observed image with Poisson noise. Extrinsic noise sources can further hamper the image.

Finally it is common in fluorescence microscopy to measure a non-zero background level arising from auto-fluorescence, inadequate removal of fluorescent staining material, glare and reflections, and/or offset levels associated with the gain of the detector or other electronic sources.

This leads to the following general image formation model of a fluorescence microscope,

$$m(x, y, z) = N(h(x, y, z) \otimes f(x, y, z) + b(x, y, z))$$  \hspace{1cm} (2.59)

with $m$ being the observed image, $h$ the point spread function of the fluorescence microscope, $f$ the object, $b$ the background and $N$ the noise distortion function.
Chapter 3

Methods for Image Restoration

In this chapter several methods for image restoration are discussed. We start with two classical image restoration filters, the Wiener filter and the Tikhonov-Miller filter.

In this chapter we will assume that the acquired image can be modeled as the original image blurred by a translation-invariant point spread function and distorted by noise. The Wiener filter is the linear filter that minimizes the mean square error between the original image and its restored estimate. The Wiener filter assumes that the acquired image is distorted by additive Gaussian noise.

The Tikhonov-Miller filter is the linear filter found when minimizing the Tikhonov functional. This functional is the squared difference between the acquired image and a blurred estimate of the original object regularized by a Tikhonov energy bound.

Both the Wiener filter and the Tikhonov-Miller filter are linear operations on the recorded image. Therefore they cannot restrict the domain in which the solution is to be found nor can they restore information at frequencies that are set to zero by the image formation process.

These restrictions are tackled by algorithms discussed in the second part of this chapter. Both the iterative constrained Tikhonov-Miller algorithm and the Carrington algorithm iteratively minimize the Tikhonov functional. They differ however in the way the non-negativity constraint is incorporated.

In the last section we discuss the Richardson-Lucy algorithm. This iterative algorithm finds the maximum likelihood solution, using the EM algorithm, when the acquired image is distorted by Poisson noise.
3.1 Introduction

The field of image restoration studies methods that can be used to recover an original signal from degraded observations. Many image restoration algorithms have their roots in well-developed areas of mathematics such as estimation theory, the solution of ill-posed problems, linear algebra and numerical analysis. Image restoration techniques model the degradations, usually blur and noise, and apply an inverse procedure to obtain an estimate of the original signal. These techniques have been studied for some time in the fields of signal processing, astronomy, optics, medical imaging and microscopy. Some excellent treatments and overviews on image restoration can be found in (Andrews & Hunt, 1977; Demoment, 1989; Jain, 1989; Lagendijk & Biemond, 1991; Banham & Katsaggelos, 1997; den Dekker & van den Bos, 1997).

Several fields of image processing, such as image enhancement, image reconstruction, and parameter estimation are related to image restoration. Image restoration differs from image enhancement in that the latter is designed to emphasize features of the image that make the image more pleasing to the observer. In general these techniques do not employ knowledge about the image formation process.

Image restoration is usually treated separately from image reconstruction techniques since image reconstruction operates on a set of projections, not on a full image. Restoration and reconstruction have in common that they try to recover the original signal. Therefore they end up solving the same mathematical problem: finding a solution to a set of (non-) linear equations.

Image restoration can also be distinguished from parameter estimation. Parameter estimation fits a model of the object and the image formation to the data by optimizing a small set of model parameters, whereas in image restoration the intensity of every pixel of the object has to be determined. However image restoration can be seen as a parameter estimation problem. If every pixel of the recorded image is regarded as a parameter of the object model one can think of image restoration as a parameter estimation problem in which a (very) large set of parameters has to estimated.

Both deconvolution and blind-deconvolution can be regarded as subfields of image restoration. Deconvolution is sometimes used to refer to the entire field of image restoration. It is, however, more common to refer to deconvolution as techniques that invert the blurring process in a deterministic way. Often these techniques ignore the effects of noise on the inverse procedure. In applications where the knowledge about the
degradations is very limited, blind deconvolution techniques (Ayers & Dainty, 1988; Holmes, 1992; Fish et al., 1995; Krishnamurthi et al., 1995; Thiebaut & Conan, 1995; Conchello & Yu, 1996; Kundur & Hatzinakos, 1996) can be applied to produce both an estimate of the object as well as an estimate of the (translation-invariant) degradation.

In general, image restoration methods yield an estimate of the original image given an imaging model, a model of the noise, and some additional criteria. These criteria can depend on the imposed regularization or the constraints implied on the solution found by the restoration algorithm. In this chapter we investigate restoration methods that use the imaging model that we discussed in the previous chapter. These methods differ significantly due to the different modeling of the noise distortion on the image and the imposed constraints and regularization.

### 3.2 Classical Image Restoration

In this section we review classical image restoration by introducing minimum mean square error restoration and least squares filters. A comprehensive treatment of this subject can be found in (Andrews & Hunt, 1977; Jain, 1989). Classical image restoration methods are linear filters that correct for the degradations of the image formation. These methods model the image formation by a translation-invariant blurring distorted by additive Gaussian noise.

In the previous chapter we have shown that the image formation of a fluorescence microscope can be modeled in the following general form

$$m(x, y, z) = N(h(x, y, z) \otimes f(x, y, z) + b(x, y, z))$$  \hspace{1cm} (3.1)

In this equation $N$ represents a general noise distortion function, $b$ a background signal, and $m$ the recorded fluorescence image. For scientific-grade light detectors $N$ is dominated by Poisson noise.

For images with a relatively high signal-to-noise ratio and small dynamic range the additive Gaussian noise assumption made by classical image restoration methods can be motivated by the Central Limit theorem (Snyder & Miller, 1991): Under these circumstances the distribution of a Poisson process can be approximated by a Gaussian distribution. This approximation of an inhomogeneous Poisson process will in general lead to a different sigma of the Gaussian distribution for each recorded data point, since the variance of a Poisson process is equal to its intensity. However, for images with a relatively high signal-to-noise ratio and small dynamic range we can approximate this
with a constant sigma, making the amount of Gaussian noise independent from the pixel intensity. Using this additive Gaussian noise model for \( N \) we rewrite (3.1) as
\[
g(x, y, z) = m(x, y, z) - b(x, y, z) = h(x, y, z) \otimes f(x, y, z) + n(x, y, z)
\] (3.2)

After sampling equation (3.2) becomes
\[
g[x, y, z] = \sum_{i=1}^{M_x} \sum_{j=1}^{M_y} \sum_{k=1}^{M_z} h[x-i, y-j, z-k] f[i, j, k] + n[x, y, z]
\] (3.3)

with \( M_x, M_y \) and \( M_z \) the number of sampling points in respectively the \( x, y \) and \( z \) dimension. For conveniences we will adopt a matrix notation
\[
g = Hf + n
\] (3.4)

where the vectors \( f, g \) and \( n \) of length \( M \) (\( M = M_x M_y M_z \)) denote respectively the object, its image and the additive Gaussian noise. The \( MxM \) matrix \( H \) is the blurring matrix representing the point spread function of the microscope.

The two classical restoration filters, the Wiener filter and the Tikhonov-Miller filter, that will be discussed in the following sections are both convolution filters operating on the measured image. They can be written as
\[
\hat{f} = Wg
\] (3.5)

with \( W \) the linear restoration filter and \( \hat{f} \) its result.

### 3.2.1 Minimum mean square error restoration: the Wiener filter

Minimum mean square error image restoration finds the restoration result by minimizing the mean square error between the restored image \( \hat{f} \) and the object \( f \). The mean square error between \( \hat{f} \) and \( f \) is defined as
\[
MSE(\hat{f}, f) = E\left[ (\hat{f} - f)^2 \right] = E\left[ (\hat{f} - f)^T (\hat{f} - f) \right]
\] (3.6)

When the solution \( \hat{f} \) is required to be a linear estimate in the form (3.5) the minimum of (3.6) can be found by solving
\[
\frac{\partial MSE(\hat{f}, f)}{\partial W} = E\left[ 2g^T (Wg - f) \right] = 0
\] (3.7)

The solution of (3.7),
\[
W_{Wiener} = H^T \left( H^T H + E\left[ n n^T \right] + E\left[ f f^T \right] \right)^{-1} \left( H^T H + \frac{S_n}{S_f} \right)
\] (3.8)

is called the Wiener filter. \( S_n \) and \( S_f \) are the noise and object power spectra. In this derivation we have used the assumption that the noise is signal-independent
\[
E\left[ fn^T \right] = E\left[ fn^T \right] = 0
\] (3.9)
Note that from (3.8) it follows that \( \hat{f} \) will be zero for those frequencies for which \( H \) is zero. Therefore, being a linear filter, a Wiener filter cannot restore information beyond the bandwidth of the point spread function.

### 3.2.2 Least squares filters

Besides image restoration based on minimizing the mean square error, the least squares approach is also a well developed method for image restoration. This approach is based on minimizing the squared difference between the acquired image and a blurred estimate of the original object,

\[
\|H\hat{f} - g\|^2
\]  
(3.10)

However a direct minimization of (3.10) will produce undesired results since it does not take into account the (high) frequency components of \( \hat{f} \) that are set to zero by the convolution with \( H \).

Finding an estimate \( \hat{f} \) from (3.10) is known as an ill-posed problem (Tikhonov & Arsenin, 1977). To address this issue Tikhonov defined the regularized solution \( \hat{f} \) of (3.2) as the one that minimizes the well-known Tikhonov functional (Tikhonov & Arsenin, 1977)

\[
\Phi(f) = \|H\hat{f} - g\|^2 + \lambda \|C\hat{f}\|^2
\]  
(3.11)

with \( \|\cdot\|^2 \) the Euclidean norm. In image restoration \( \lambda \) is known as the regularization parameter and \( C \) as the regularization matrix. The Tikhonov functional consists of a mean square error fitting criterion and a stabilizing energy bound which penalizes solutions of \( \hat{f} \) that oscillate wildly due to spectral components which are dominated by noise. The minimum of \( \Phi \) can be found by solving

\[
\nabla_{\hat{f}} \Phi(\hat{f}) = 2H^T(H\hat{f} - g) + 2\lambda C^T \hat{f} = 0
\]  
(3.12)

which yields the well-known Tikhonov-Miller (TM) solution \( W_{TM} \)

\[
W_{TM} = \frac{H^T}{H^TH + \lambda C^TC}
\]  
(3.13)

By comparing (3.13) with (3.8) it can be seen that the Tikhonov-Miller solution is equal to the Wiener filter when \( \lambda C^TC = S_{nn}/S_{ff} \).

### 3.2.3 Limitations of linear image restoration algorithms

The Wiener filter and the Tikhonov-Miller restoration filter are both convolution filters. Their linear nature makes them incapable of restoring frequencies for which the PSF has a zero response. In particular the PSF of a 3-D conventional fluorescence microscope has large regions with zero response known as the missing cone (as discussed in the previous chapter).
Furthermore, the linear methods cannot restrict the domain in which the solution should be found. This property is a major drawback since the intensity of an imaged object represents light energy, which is non-negative. Finally, van der Voort (van der Voort & Strasters, 1995) showed that the Tikhonov-Miller filter is very sensitive to errors in the estimation of the PSF causing ringing artifacts.

In the next two sections we discuss three iterative non-linear algorithms that we will use in the experiments presented in this thesis. These algorithms, the iterative constrained Tikhonov-Miller algorithm, the Carrington algorithm and the Richardson-Lucy algorithm are frequently used in fluorescence microscopy. These iterative and non-linear algorithms tackled the above mentioned problems in exchange for a considerable increase in the computational complexity. These algorithms involve many iterations each with a complexity comparable to that of the discussed linear restoration filters.

3.3 Constrained Tikhonov Restoration

3.3.1 The iterative constrained Tikhonov-Miller algorithm

The iterative constrained Tikhonov-Miller (ICTM) (Lagendijk & Biemond, 1991; van der Voort & Strasters, 1995) finds the minimum of (3.11) using the method of conjugate gradients (Press et al., 1992). The non-negativity constraint is incorporated by setting the negative intensities after each iteration to zero. The conjugate gradient direction \( d \) of (3.11) is given by

\[
d^k = r^k + \gamma^k d^{k-1}, \quad \gamma^k = \frac{\|r^k\|^2}{\|r^{k-1}\|^2}
\]  

with \( r \), denoting the steepest descent direction

\[
r^k = -\frac{1}{2} \nabla r^k \phi(\hat{f}) = A\hat{f} - H^Tg, \quad A = H^TH + \lambda C^TC
\]

A new conjugate gradient estimate is now found with

\[
\hat{f}^{k+1} = P(\hat{f}^k + \beta^k d^k)
\]

where \( P \) is the projection function

\[
P(f) = \begin{cases} f, & f \geq 0 \\ 0, & otherwise \end{cases}
\]
In the absence of the projection function $P$ the optimal step size $\beta^k$ can be found analytically (Press et al., 1992).

$$\beta = \frac{d^T r}{d^T A d} \quad (3.17)$$

In this equation we have dropped the superscript $k$. In the presence of the projection constraint Verveer (Verveer & Jovin, 1997) showed that (3.17) will not always produce correct values for $\beta$. This can result in a poor performance of the ICTM algorithm. The values for $\beta$ can be found using an iterative one-dimensional minimization algorithm such as the golden section rule (Press et al., 1992) as used by (Lagendijk & Biemond, 1991; van der Voort & Strasters, 1995). Alternatively $\beta$ can be found using a first order Taylor approximation of (3.16) with respect to $\beta$. This yields (Verveer & Jovin, 1997),

$$\beta = -\frac{d^T T(f)r}{d^T T(f)AT(f)d} \quad (3.18)$$

with the diagonal threshold matrix $T$ defined as

$$T(f)_{ij} = \begin{cases} 1 & f_i(x) \geq 0 \text{ and } i = j \\ 0 & \text{otherwise} \end{cases} \quad (3.19)$$
Both methods produce good values for $\beta$ (Verveer & Jovin, 1997). However, to calculate $\beta$ the Taylor approximation needs only two Fourier transforms\(^1\) whereas the golden section rule is an iterative procedure which needs one transform per iteration. For a typical number of five iterations this results in a substantial increase in computational complexity. A schematic diagram of the ICTM algorithm is shown in Figure 3.1.

### 3.3.2 The Carrington algorithm

Like the ICTM algorithm the Carrington algorithm (Carrington, 1990; Carrington et al., 1995) minimizes the Tikhonov functional under the constraint of non-negativity. However, the algorithm is based on a more solid mathematical foundation. The derivation of the Carrington algorithm as given in this section is based on (Verveer, 1995).

Given the Tikhonov functional (3.11) (without the regularization matrix) and its derivative

\[
\frac{1}{2} \nabla_\widehat{f} \Phi (\widehat{f}) = H^T (H\widehat{f} - g) - \lambda \widehat{f}
\]

(3.20)

we want to find the non-negative solution of $\widehat{f}$ that minimizes this functional. Then the Kuhn-Tucker conditions apply (Carrington, 1990):

\[
\nabla_\widehat{f} \Phi = 0 \text{ and } \widehat{f}_i > 0 \quad \text{or} \quad \nabla_\widehat{f} \Phi_i \geq 0 \text{ and } \widehat{f}_i = 0
\]

(3.21)

Then from (3.21) we find

\[
\nabla_\widehat{f} \Phi = 0 \quad \rightarrow \quad \widehat{f} = \frac{1}{\lambda} H^T (g - H\widehat{f}) = H^T c
\]

(3.22)

From the Kuhn-Tucker conditions we see that:

\[
\widehat{f} = H^T c \text{ where } H^T c > 0 \quad \text{ and } \quad \widehat{f} = 0 \text{ where } H^T c \leq 0
\]

(3.23)

This can be written as

\[
\widehat{f} = \max(0, H^T c) = P(H^T c)
\]

(3.24)

On the set where $H^T c > 0$ we then obtain after inserting (3.24) into (3.20) that

\[
\nabla_\widehat{f} \Phi = H^T (H P (H^T c) - g + \lambda c) = 0
\]

(3.25)

Solving (2.23) is equivalent to minimizing the functional $\Psi$ (on the set $H^T c > 0$)

\[
\Psi(c) = \frac{1}{2} \| P(H^T c) - c \|_2^2 + \frac{1}{2} \lambda \| c \|_2^2
\]

(3.26)

Since $\Psi$ is strictly convex\(^4\) and twice continuously differentiable, a conjugate gradient algorithm similar to (3.14) and (3.16) can be used to minimize $\Psi$. In this case the optimal

---

\(^1\) The multiplication of matrix $A$ with $T(f)d$ can be written as a convolution which can efficiently be calculated using the Fourier transform.

\(^4\) $\Psi$ is strictly convex since its second derivative is positive definite (Carrington, 1990).
step size can be found using Newton's rule. No blurring operations are necessary for the sub-iterations resulting in a faster line-search than in the case of ICTM. A schematic representation of the Carrington algorithm is given in Figure 3.2.

Figure 3.2 Diagram of the algorithm of Carrington.

3.4 Maximum Likelihood Restoration

In contrast with the two algorithms discussed previously the Richardson-Lucy is not derived from the image formation model (3.2) which assumes additive Gaussian noise. Instead the general noise distortion function $N$ is assumed to be dominated by Poisson noise.

A fluorescence object can be modeled as a spatially inhomogeneous Poisson process $F$ with an intensity function $f$ (Snyder & Miller, 1991),

$$P(F_i | f_i) = \frac{F_i^{f_i} e^{-f_i}}{f_i!}$$

The image formation of such an object by a fluorescence microscope can be modeled as a translated Poisson process (Snyder & Miller, 1991). This process models the transformation of $F$ into a Poisson process $m$ subjected to a conditional probability density function $H$,

$$E[m] = Hf + b$$  \hspace{1cm} (3.27)
with $b$ the mean of an independent (background) Poisson process. The conditional probability density function is in our case the point spread function of the fluorescence microscope. The log likelihood function of such a Poisson process is given by (Snyder & Miller, 1991)

$$L(f) = -\sum Hf + m^T \ln(Hf + b)$$  \hspace{1cm} (3.28)

where we have dropped all terms that are not dependent on $f$. The maximum of the likelihood function $L$ can be found iteratively using the EM algorithm as described by (Dempster et al., 1977). This iterative algorithm was first used by (Vardi et al., 1985) in emission tomography. Holmes (Holmes, 1988) introduced the algorithm to microscopy.

3.4.1 The Expectation-Maximization algorithm
In this section we review the Expectation-Maximization (EM) algorithm in a general format. We will discuss how the algorithm can be used to compute maximum likelihood (ML) parameter estimates. In the next section we will focus on the application of the EM algorithm to find intensity estimates that maximize (3.28). Finding maximum likelihood parameter estimates can be very difficult. The EM algorithm is a general iterative algorithm for finding the maximum likelihood estimator if the observed data can be regarded as "incomplete".

As stated by (Dempster et al., 1977) the term "incomplete data" implies the existence of two data spaces $X$ and $Y$ and a many-to-one mapping $h$ from $X$ to $Y$. The observed data $y$ is a realization from $Y$. The corresponding $x$ in $X$ cannot be observed directly but is only observed indirectly via the non-invertible transformation (see Figure 3.3)

$$y = h(x)$$  \hspace{1cm} (3.29)

![Figure 3.3 Translation of points on the space X to points on space Y with translation density h.](image)

Let $p(y; \theta)$ be the probability density function of the incomplete data with $\theta$ the parameters to be estimated. The ML estimator of the incomplete data loglikelihood function $L_{ad}(\theta)$ is given by
\[ \hat{\theta}_{ML}(y) = \arg \max_{\theta \in \Theta} \{ L_{ad}(\theta) \} = \arg \max_{\theta \in \Theta} \{ \log p(y; \theta) \} \]  

(3.30)

Solving (3.30) is in general a complicated problem. Instead of solving it directly the EM algorithm computes the ML estimator of the complete data loglikelihood function
\[ \hat{\theta}_{ML}(x) = \arg \max_{\theta \in \Theta} \{ L_{cd}(\theta) \} = \arg \max_{\theta \in \Theta} \{ \log p(x; \theta) \} \]  

(3.31)

which is assumed to be significantly simpler. In fact the complete data space is contrived to make the EM algorithm analytically tractable and computationally feasible.

Starting with an estimate of the parameters \( \hat{\theta}^k \) the EM algorithm finds the conditional expectation of the complete data, given \( y \) and \( \hat{\theta}^k \)
\[ Q(\theta|\hat{\theta}^k) = E \left[ L_{cd}(\theta) | y; \hat{\theta}^k \right] \]  

(3.32)

This is called the expectation (E) step of the EM algorithm. In the maximization (M) step, the condition expectation is maximized with respect to \( \theta \). This leads to a new parameter estimate
\[ \hat{\theta}^{k+1} = \arg \max_{\theta} \{ Q(\theta|\hat{\theta}^k) \} \]  

(3.33)

By repeating the E and M step, the iterative EM algorithm is obtained which converges to a stationary point of \( L_{ad}(\theta) \).

3.4.2 The Richardson-Lucy algorithm
The EM algorithm can be used to find the ML estimator of (3.28). By regarding the recorded image \( m \) as the incomplete data the complete data \( f \) is mapped to \( m \) by the PSF \( H \). The intensity of all pixels in \( f \) are now the parameters to be estimated. The incomplete-data log-likelihood \( L_{ad} \) is given by (2.30). The complete-data loglikelihood function is simply the log-likelihood on the space of \( F \) and is given by
\[ L_{cd}(f) = -\sum f + F^T \ln f \]  

(3.34)
In the $E$-step the conditional expectation $Q$ of the complete-data log-likelihood given the measured data $\mathbf{m}$ and the $k$-th estimate $\hat{f}^k$ of $f$ is

$$Q(\hat{f}^k | \hat{f}^k) = E \left[ L_{cd}(f) | \mathbf{m}, \hat{f}^k \right] = -\sum \mathbf{f} + \sum E \left[ \mathbf{F} | \mathbf{m}, \hat{f}^k \right] \ln f$$

(3.35)

The conditional expectation of a Poisson process $\mathbf{F}$ given the measured data $\mathbf{m}$ is given by (Appendix A: The conditional expectation of a translated Poisson process),

$$E[\mathbf{F} | \mathbf{m}] = \left( \frac{H^T \mathbf{f}}{H \mathbf{f} + b} \right)^T \mathbf{m}$$

(3.36)

Using (3.36) we obtain

$$Q(\hat{f}^k) = -\sum \mathbf{f} + \sum \left( \frac{H^T \hat{f}^k}{H \hat{f}^k + b} \right)^T \mathbf{m} \ln f$$

(3.37)

The $M$-step maximizes $Q$ using $\nabla f Q = 0$ for $\mathbf{f} = \hat{f}^{k+1}$ which yields (Snyder et al., 1993)

$$\hat{f}^{k+1} = \hat{f}^k \left( \frac{H^T}{H \hat{f}^k + b} \right)$$

(3.38)

The EM algorithm ensures a non-negative solution when a non-negative initial guess $\hat{f}^0$ is used. Furthermore the likelihood of each iteration of the EM algorithm will strictly increase to a global maximum (Snyder & Miller, 1991). The EM algorithm (2.47) for finding the maximum likelihood estimator of a translated Poisson process (often referred to as EM-MLE) is identical to the Richardson-Lucy algorithm (Richardson, 1972). A
schematic representation of the Richardson-Lucy algorithm is shown in Figure 3.4. This figure also includes optional regularization of the Richardson-Lucy algorithm as discussed in the next section.

The convergence of the Richardson-Lucy algorithm has been observed to be slow. Therefore several authors have suggested methods to accelerate the Richardson-Lucy algorithm (Lewitt & Muehlehner, 1986; Kaufman, 1987; Lange et al., 1987; Holmes & Liu, 1991; Nuyts et al., 1993; Biggs & Andrews, 1995; Zaccheo & Gonsalves, 1996).

### 3.4.3 Regularization of the Richardson-Lucy algorithm

The Richardson-Lucy algorithm is a constrained but unregularized iterative image restoration algorithm. The ICTM and Carrington algorithms, however, incorporate Tikhonov regularization to suppress undesired solutions. Conchello has derived an algorithm that incorporates Tikhonov regularization into the Richardson-Lucy algorithm (Conchello & McNally, 1996).

Incorporation of Tikhonov regularization in (3.32) yields

\[
Q(\theta | \hat{\theta}^k) = E\left[ L_{cd}(\theta) - \alpha \| \theta \|^2 | y, \hat{\theta}^k \right]
\]

(3.39)

The conditional expectation \( Q \) for estimating the intensity of a translated Poisson process now becomes

\[
Q(f | \hat{f}^k) = -\sum f + E\left[ F|m, \hat{f}^k \right]^T \ln f - \alpha \| f \|^2
\]

(3.40)

The maximization step now yields

\[
\frac{E\left[ F|m, \hat{f}^k \right]}{\hat{f}_{\text{regularized}}^{k+1}} - 2\alpha \hat{f}_{\text{regularized}}^{k+1} - 1 = 0
\]

(3.41)

Solving this quadratic equation in \( \hat{f}_{\text{regularized}}^{k+1} \) yields

\[
\hat{f}_{\text{regularized}}^{k+1} = \frac{-1 + \sqrt{1 + 2 \lambda \hat{f}_{\text{regularized}}^{k+1}}}{\lambda}
\]

(3.42)

with regularization parameter \( \lambda = 4\alpha \) and \( \hat{f}_{\text{regularized}}^{k+1} \) given by (3.38). Using l’Hospital’s rule it is easy to show that (2.29) becomes (2.45) when \( \hat{\lambda} \rightarrow 0 \). We will refer to this algorithm as the RL-Conchello algorithm.

### 3.5 Software Implementation

We have used a software implementation of the ICTM, Carrington and Richardson-Lucy algorithms as part of a software library written by Peter J. Verveer in standard C. Since
Chapter 3  Methods for Image Restoration

this software library was provided to us in source, we could base our implementation of the RL-Conchello algorithm on his implementation of the Richardson-Lucy algorithm. The provided software uses a real-valued fast Fourier transform to compute convolutions. The implemented real-valued fast Fourier transform uses the FFTW 1.2 library (Frigo & Johnson, 1997) which has been compiled in floating point single precision.
Chapter 4

Iterative Image Restoration Algorithms

This chapter discusses various issues that play a role when testing and comparing image restoration algorithms.

We start by defining two performance measures, the mean-square-error and L-divergence, that we use for measuring and comparing the performance of image restoration algorithms.

Many of the tests we present have been performed on simulated images. We discuss the properties of objects generated using an analytical description of their Fourier transform. Images are created by distorting the objects with Poisson noise at a predetermined signal-to-noise ratio.

We continue with a comparison of different methods to determine the regularization parameter for the Tikhonov functional. These methods use different criteria to balance the fit of the restored data to the measured image with the imposed regularization on the restored data.

The next two sections deal with the iterative character of the discussed algorithms. In the first we compare different choices for a first estimate. In the second section stop criteria for iterative optimizations are discussed.
4.1 Performance Measures

In this and the following chapter we present tests and comparisons of the iterative image restoration algorithms discussed in the previous chapter. Many of these experiments are performed on simulated data. This section discusses two measures used to judge the performance of these iterative image restoration algorithms. The next section discusses the generation of the synthetic data used in these experiments.

The main reason for using simulated data instead of real data for testing image restoration algorithms is that the undistorted object data is available as the "ground truth". This allows objective comparisons of the results produced by these algorithms. A difference measure between the undistorted object data and the restored data can then be used to measure the performance of the algorithm. In our experiments we have used two difference measures, the mean-square-error and the I-divergence.

4.1.1 The mean square error

The mean-square-error (MSE) is defined as

$$MSE(f, \hat{f}) = \int |f(x) - \hat{f}(x)|^2 \, dx$$  \hspace{1cm} (4.1)

where $f$ is the original image, and $\hat{f}$ the restored image. The mean-square-error measures the difference in energy between the compared images. The MSE can be calculated both in the spatial domain and in the frequency domain

$$\int |f(x) - \hat{f}(x)|^2 \, dx = \frac{1}{2\pi} \int |F(\omega) - \hat{F}(\omega)|^2 \, d\omega$$  \hspace{1cm} (4.2)

using Parseval’s relation (Oppenheim et al., 1983).

With (4.2), the mean-square-error can be measured in a specific regions of the frequency domain using a mask. For example the region outside the footprint of the OTF can be selected to measure an algorithm’s performance to restore information beyond the resolution of the microscope. This way one can test the claim that iterative constrained image restoration algorithms are capable of partially restoring information outside the bandwidth of a microscope (Holmes, 1988). This capability is sometimes referred to as superresolution (Gerchberg, 1979; Jain, 1989; Carrington et al., 1995).

4.1.2 The I-divergence

Image restoration is an example of the class of inverse problems in which a function has to be inferred from insufficient information that specifies only a feasible set of functions. Often the practical solution to such a problem is to select an element from the feasible set
by a more or less ad hoc selection rule, like minimizing the mean-square-error or the negative entropy.

Csiszár has postulated a set of axioms (consistency, distinctness, continuity, locality, and composition consistency) that a selection rule should satisfy (Csiszár, 1991). It is beyond the scope of this thesis to discuss these axioms in full detail but a short description is appropriate (Snyder et al., 1992):

- **Consistency** If the function \( c \) is selected to satisfy some constraints for one selection rule and \( c \) also satisfies some stronger constraints for another rule, then combining these constraints do not motivate changing \( c \).
- **Distinctness** Two pieces of information should lead to different conclusions about \( c \) unless they are irrelevant.
- **Continuity** The restriction of a selection rule to subspaces is continuous.
- **Locality** If the domain of the functions of interest is partitioned into disjoint subsets on each of which constraints are imposed that only affect that subset, then the function to be selected can be constructed piecewise using on each subset the constraints for that subset.
- **Composition consistency** If information specifies that individual components in a system are composed of two components, but says nothing about their interaction, then the selection rule should lead to the correct result if the components do not interact.

Csiszár concluded that for real-valued functions, having both negative and positive values, only the mean-square-error is consistent with these axioms. For functions that are required to be non-negative the I-divergence,

\[
I(f, \hat{f}) = \int dx \frac{f(x)}{\hat{f}(x)} \ln \frac{f(x)}{\hat{f}(x)} - f(x) - \hat{f}(x)
\]  

(4.3)

is the only selection rule that satisfies all five axioms (Note that \( I(a,b) \neq I(b,a) \)).

In the previous chapter we discussed the Richardson-Lucy algorithm. This algorithm iteratively finds the maximum of the log likelihood function of Poisson noise distorted data. Snyder (Snyder et al., 1992) has shown that maximizing the mean of this log likelihood function is equal to minimizing Csiszár's I-divergence,

\[
L(f) - E[L(\hat{f})] = \int dx \left[ \ln \frac{g(x)}{\hat{g}(x)} - g(x) + \hat{g}(x) \right]
\]  

(4.4)

with

\[
g(x) = h(x) \otimes f(x) \quad \text{and} \quad \hat{g}(x) = h(x) \otimes \hat{f}(x)
\]  

(4.5)
This can also be derived from a Bayesian derivation of maximum entropy using assumptions similar to Csiszár’s axioms (Skilling, 1989).

4.2 Generation of Simulated Microscope Images

We will frequently use simulated data to test or compare the performance of image restoration algorithms. The validity of the results obtained by simulation experiments is determined by the validity of the model of the microscope system used to generate the data and by the choice of objects used to test the algorithms.

We have devoted the second chapter of this thesis to the development of a model of the image formation in a fluorescence microscope. In the last section of that chapter the image formation is modeled as a convolution between the undistorted object f and the microscope’s point spread function H. Noise (N()) - dependent on the intensity of the blurred object and an added background b - disturbs the image g. This gives the following model of the image formation in a fluorescence microscope,

\[ g = N(Hf + b) \]  

(4.6)

where we have used the matrix notation introduced in the previous chapter. We use this model to generate simulated microscope images. This section discusses the generation of objects and the distortion of the data with noise at a predetermined signal-to-noise ratio.

4.2.1 Object generation

In our experiments we use solid spheres as objects. Spheres have a simple description both in the spatial domain and in the frequency domain. This makes the interpretation of obtained results easier. Furthermore spheres serve as a good model for body-stained fluorescence beads. This makes it possible to validate simulation experiments with spheres in real data. Finally spheres can be regarded as a (very) crude model for body-stained biological cells. Therefore such simulations can provide insight into the ability of image restoration algorithms to perform on data relevant to the biological community.

We have generated spheres using an analytical description of their Fourier transform. In spherical coordinates \( r, \varphi, \theta \) the Fourier transform of a sphere is given by (van Vliet, 1993),

\[ S_{\text{sphere}}(r, \varphi, \theta) = \frac{-6\pi \cos(2\pi rs) + 3\sin(2\pi rs)}{(2\pi rs)^3} \]  

(4.7)

with \( s \) the radius of the sphere. To ensure bandlimitation the Fourier transform of the sphere has been multiplied with a Gaussian transfer function, (we used a sigma of 1 pixel
in the spatial domain). Generated in this way, the spheres are free from aliasing effects that arise from sampling non-bandlimited analytical objects below the Nyquist rate. Measurements performed on digital representations of analog objects only produce reliable information about the analog object when the digital representation is free from aliasing (van Vliet, 1993).

We have chosen to generate spheres with the same oversampling factor in the lateral and the axial direction. As derived in chapter 2 the Nyquist frequency along the optical axis is about three times lower than the lateral Nyquist frequency. Our sampling at a fixed oversampling ratio therefore results in an anisotropic spatial sampling. To generate spheres in the spatial domain we therefore have to generate the Fourier transform of ellipsoids in the frequency domain. These ellipsoids can be generated using the Fourier description of spheres and scaling the frequency axes by the ratio of the spatial sampling distances.

The microscope's point spread function has been computed using a theoretical model of the microscopic image formation which is based on vectorial diffraction theory (van der Voort & Brakenhoff, 1990). This model takes important microscopic parameters such as the finite-size pinhole, high numerical apertures, and polarization effects into account; lens aberrations are not modeled.

4.2.2 Noise distortion
In scientific grade microscopes the signal-to-noise ratio is determined by the amount of collected photons (Art, 1995; van Vliet et al., 1997). This is determined both by the light intensity and the exposure time. A shorter exposure time will give a better time resolution but will produce noisier images. Furthermore physical constraints like photobleaching and saturation of the fluorophores can impose limits on the number of collected photons. The signal-to-noise ratio is therefore an interesting parameter when measuring the performance of image processing algorithms.

We define the signal-to-noise ratio (SNR) as

\[
SNR = \frac{E}{\varepsilon} = \frac{\sum_{N} (Hf)^2}{\sum_{N} (g - (Hf + b))^2}
\]  

with \(\varepsilon\) the total power of the noise and \(E\) the total power of the object in the image. This is by no means the only possible definition of the SNR. Examples of different definitions of the SNR are given by Young (Young et al., 1998).
We have distorted the images of the simulated spheres with Poisson noise. The noise is generated using the intensity of the spheres convolved with the PSF as averages of a spatially variant Poisson process (Press et al., 1992). The conversion factor $c$ determines the number of photons needed to increase the grey-value of a pixel with a single level. Denoting the grey-value of a pixel, measured in analog-to-digital units (ADU), with $I$ and the average of the number of photons acquired at that pixel with $P$, then

$$I = P/c$$

The collection of photons can be described with a Poisson process with mean $P$ and standard deviation $\sqrt{P}$. The noise power $\varepsilon$ (ADU$^2$) is then given by

$$\varepsilon = \sigma^2 = c(I_0 + VI_b)$$

with $V$ the image volume ($\mu m^3$), $I_b$ the constant background intensity (ADU), and

$$I_0 = \sum_N Hf$$

The object energy is given by

$$E = I_o^2 = \sum_N (Hf)^2$$

In our simulations we want to disturb the image with noise at a predetermined signal-to-noise ratio. We have implemented this by determining that value of conversion factor that will disturb the image with Poisson noise at the desired SNR. Using (4.8), (4.10) and (4.12) the desired value for the conversion factor can be found

$$c = \text{SNR} \cdot (I_0 + VI_b) / I_o^2$$

### 4.2.3 Generation of test images

The previous sections describe the generation of bandlimited spheres and noise. Using these methods we have generated test images containing spheres that are blurred with the PSF of a fluorescence microscope and distorted by Poisson noise at a predetermined signal-to-noise ratio. Objects generated with different signal-to-noise ratios are shown in Figure 4.1. A diagram of this procedure is shown in Figure 4.2.
Figure 4.1 Center x-y slice and center line of the x-z slice of simulated spheres with a SNR of respectively 1.0 (top), 16.0 (middle) and 256.0 (bottom).
Figure 4.2 Diagram of the object generation procedure.
4.2.4 Implementation of the Simulation Experiments
The simulation experiments described in this thesis have been implemented using the scientific image processing library 3D-Pro (Pattern Recognition Group, The Netherlands) and the image processing package SCIL-Image (TNO-TPD, The Netherlands). All experiments have been performed on images of which the intensities are represented in floating point single precision. The generation of Poisson noise has been implemented using the algorithm described by Knuth (Knuth, 1981). We have ensured the use of a different seed for this pseudo-random noise generator in each experiment by using the time at which the noise was generated as seed.

We generate the synthetic objects, as discussed in this section, with a random subpixel shift to ensure that errors caused by the specific digitization of each object canceled out when averages were taken over multiple repetitions of the same experiment.

4.3 Regularization Parameter
In the previous chapter we introduced the ICTM restoration algorithm and the Carrington algorithm. Both algorithms minimize the Tikhonov functional. The Tikhonov functional consists of a mean-square-error criterion and a regularization term. These two terms are "balanced" by the regularization parameter. Restoration results will be smoother for higher values of the regularization parameter $\lambda$ (more regularization). Lower values of $\lambda$ will result in "crisper" results that are in general more sensitive to the noise in the image.

Previous work (van Kempen et al., 1996) showed that when the regularization parameter $\lambda$ is equal to $1/$SNR, the result of the ICTM algorithm is too smooth. Both an oversmoothed result as well as a result which is too adapted to the noise hides the actual information present in the images. It is therefore of great importance to have a reliable method for determining the regularization parameter.

4.3.1 The SNR method
The signal-to-noise ratio (SNR) method sets $\lambda$ equal to the inverse of the signal-to-noise ratio

$$\lambda_{SNR} = \frac{\epsilon}{E} = \frac{\epsilon}{\sum |g - b|^2} \quad (4.14)$$

4.3.2 The method of constrained least squares
In (Galatsanos & Katsaggelos, 1992) the methods of constrained least squares, generalized cross-validation and maximum likelihood are described to determine $\lambda$ for the Tikhonov-Miller algorithm (see section 3.2.2). These methods define different criteria to
balance the difference between the recorded data and the reblurred restoration result given the amount of noise present in the image. The method of constrained least squares (CLS) finds a $\lambda_{\text{cls}}$ such that the mean-square-error between the recorded data and the blurred restored data equals the noise power,

$$\sum |x - \mathbf{H}\hat{\mathbf{x}}(\lambda)|^2 = \sum |(\mathbf{I} - \mathbf{H}\mathbf{A}(\lambda))\mathbf{x}|^2 = \varepsilon$$  \hspace{1cm} (4.15)

with $\mathbf{I}$ denoting the unity matrix and $\varepsilon$ the total noise power. $\mathbf{A}$ is the linear Tikhonov-Miller filter (see chapter 3),

$$\mathbf{A}(\lambda) = \frac{\mathbf{H}^T}{\mathbf{H}^T\mathbf{H} + \lambda \mathbf{C}^T\mathbf{C}}$$  \hspace{1cm} (4.16)

We determine $\lambda_{\text{cls}}$ numerically using Brent’s method (Press et al., 1992) to find the zero crossing of the function

$$\sum |(\mathbf{I} - \mathbf{H}\mathbf{A}(\lambda))\mathbf{x}|^2 - \varepsilon$$  \hspace{1cm} (4.17)

4.3.3 The method of generalized cross-validation

The method of generalized cross-validation (GCV) is derived from the leave-one-out principle (Galatsanos & Katsaggelos, 1992). For every pixel the Tikhonov-Miller restoration result is calculated using all but the pixel under consideration. The cross-validation function is derived from the mean-square-error between original data and the restoration result derived by filtering each pixel with its associated Tikhonov-Miller filter. The generalized cross-validation function (CV),

$$CV(\lambda) = \frac{\sum |(\mathbf{I} - \mathbf{H}\mathbf{A}(\lambda))\mathbf{x}|^2}{\text{trace}((\mathbf{I} - \mathbf{H}\mathbf{A}(\lambda))^T)}$$  \hspace{1cm} (4.18)

can be efficiently evaluated in the discrete Fourier domain (Galatsanos & Katsaggelos, 1992),

$$CV(\lambda) = \sum_{\omega} \frac{\lambda^2 |C(\omega)|^4 |G(\omega)|^2}{\left( \sum_{\omega} \frac{\lambda^2 |C(\omega)|^2}{|H(\omega)|^2 + \lambda^2 |C(\omega)|^2} \right)^2}$$  \hspace{1cm} (4.19)

See (Golub et al., 1979) for a step-by-step derivation of the generalized cross validation function. The minimum of this function can be found without prior knowledge of the noise variance (Reeves & Mersereau, 1992). We have determined the minimum of the cross-validation function numerically with Brent’s one-dimensional search algorithm (Press et al., 1992).
4.3.4 The method of maximum likelihood

An alternative method to determine $\lambda$ without prior knowledge of the noise variance has been named the maximum likelihood method (ML) by (Galatsanos & Katsaggelos, 1992). It is based on a stochastic approach which assumes $\sqrt{\lambda} C_f$ and the noise to be Gaussian distributed. The former assumption can be met with a proper choice of the regularization matrix $C$ (Galatsanos & Katsaggelos, 1992). The derived maximum likelihood function,

$$ML(\lambda) = \frac{g^T (I - HA(\lambda))g}{(\det[I - HA(\lambda)])^{1/2}}$$

can be evaluated in the discrete Fourier domain (Galatsanos & Katsaggelos, 1992)

$$ML(\lambda) = \frac{\sum_{\omega} \lambda |C(\omega)|^2 |G(\omega)|^2}{\prod_{\omega} \frac{\lambda |C(\omega)|^2}{|H(\omega)|^2 + \lambda |C(\omega)|^2}}^{1/\Omega}$$  (4.20)

with $\Omega$ the number of Fourier coefficients. We used Brent’s one-dimensional search algorithm to find the minimum of (2.54).

4.3.5 The Golden search method

We compare these four methods for determining $\lambda$ (SNR, CLS, GCV and ML) with the “optimal” value for $\lambda$. We define the optimal value of $\lambda$ as the value for which the image restoration algorithm (either the ICTM algorithm or Carrington’s algorithm) produces the smallest mean-square-error. We have used the golden search algorithm to find the optimal value numerically (Press et al., 1992).

4.3.6 Estimation of the noise variance

Both the CLS method and the SNR method use the total noise power $e$. Therefore they require prior knowledge about the noise variance. Often this knowledge is not available. It is possible however to estimate the noise variance using the GCV method (Galatsanos & Katsaggelos, 1992). Denoting the value of the regularization parameter determined by the GCV method with $\lambda_{GCV}$, the noise variance can be estimated by

$$\sigma^2 = \frac{\left\| (I - HA(\lambda_{GCV}))g \right\|^2}{\text{trace}[(I - HA(\lambda_{GCV}))]}$$  (4.21)

Although this estimation of the noise variance assume that the data is distorted by additive Gaussian noise, the results presented in this section shows that (4.21) can also be used to estimate the average noise variance in a fluorescence microscopic image distorted by Poisson noise.
Figure 4.3 The top left graph shows the relative error of the estimated variance using GCV. The top right graph shows the relative error of the regularization parameter determined with CLS using the variance estimated by GCV. The bottom graph shows the relative error of the mean-square-error performance of the ICTM with the CLS regularization parameter determined with an estimated or true value for the variance.

Figure 4.3 shows the relative error (the estimated value minus the true value divided by the true value) of the estimated noise variance and the value of the regularization parameter determined with the CLS method. The relative difference between the mean-square-error performance of the ICTM algorithm with the regularization parameter determined with CLS method using an estimated variance and the true value of the variance is shown as well. We conclude from Figure 4.3 that using the estimated variance produces values of the regularization parameter very close to those found by the CLS method using the true value for the variance. Furthermore it is shown that the mean-
square-error performance of the ICTM algorithm is hardly influenced by using the estimated variance.

![Graph showing the value of the regularization parameter as a function of the signal-to-noise ratio.](image)

**Figure 4.4.** The value of the regularization parameter, using SNR, CLS, GCV and ML, as a function of the signal-to-noise ratio. The optimal values (Golden) of λ for the ICTM and Carrington’s algorithm are plotted as well.

The values of the regularization parameter as determined by SNR, CLS, GCV and ML are plotted in Figure 4.4 as a function of the signal-to-noise ratio. We define the optimal value of the regularization parameter as the value that minimizes the mean-square-error between the original image and the restored image. We have determined the optimal values of the regularization parameter for both the ICTM algorithm and Carrington algorithm determined with a Golden search algorithm (Press et al., 1992). These values are plotted in Figure 4.4 as well.

4.3.7 Comparison of the regularization parameter methods

The simulation experiments described in this section are performed on images with a size of 128×128x64 pixels. We used a two times oversampled confocal point spread function, an NA of 1.3, an excitation wavelength of 480 nm, an emission wavelength of 530 nm, a

---

' These excitation and emission wavelength correspond with the excitation and emission spectra of popular fluorophores like FITC and Nile-Red.
refractive index of 1.515 and a backprojected pinhole size of 300 nm. Spheres with a diameter of 1000 nm were used as objects. The sampling distance was 25.0 nm lateral and 85.0 nm axial. We varied the signal-to-noise ratio (as defined by (4.8)) from 1.0 to 256.0 (0.0 dB to 24.1 dB), which corresponds to a conversion factor ranging from 0.14 to 35.8, with an object intensity of 200.0 ADU and a background of 20.0 ADU. This results in an average number of photons in the object ranging from 28.0 photons ($SNR = 1.0$) up to 7000.0 photons/pixel ($SNR = 256.0$) and 2.8 up to 700.0 photons/pixel in the background. We have based our implementation of the various methods for determining the regularization parameter on sources provided by Peter J. Verveer. We reimplemented these algorithms in 3D-Pro to incorporate the methods to estimate the signal-to-noise ratio as presented in section 4.3.6.

Figure 4.5 shows the mean-square-error performances of the ICTM and Carrington algorithms for the various methods used to estimate the regularization parameter. We have also included the performance of the ICTM with no regularization ($\lambda$ set to zero). Due to numerical instability (see also the low performance of the Carrington algorithm for low values of $\lambda$ as estimated with ML), we were not able to include these results for the Carrington algorithm. All the results shown in Figure 4.5 are averages over eight experiments. All data points have a coefficient-of-variation of less than 5%.

Figure 4.5 clearly shows that high values of the regularization parameter as determined by the SNR method result in a low performance caused by an oversmoothed restoration result (see Figure 4.6).

The CLS method is also known to overestimate the value of the regularization parameter (Galatsanos & Katsaggelos, 1992). Therefore the CLS method will produce somewhat smooth (restored) images resulting in a suboptimal performance.

As described in Chapter 3, the regularization of the ICTM algorithm includes a regularization matrix $C$. The Carrington algorithm does not include this regularization matrix. We have therefore chosen to make $C$ equal to the unity matrix $I$ and not to use a specific regularization matrix $C$ that would make $Cf$ Gaussian distributed, as required by the ML algorithm. This explains the weak performance of the ML method in combination with the ICTM algorithm. In our experiments the ML method produces very small values for the regularization parameter which could not be handled properly by the Carrington algorithm. We attribute the low performance of the Carrington algorithm in this case to numerical instability of the algorithm.
Figure 4.5. The value of the mean-square-error of the restoration result of the Carrington (left) and ICTM algorithm (right). The regularization parameter was determined with the CLS, GCV, SNR and ML method and a Golden search. The performance of an unregularized ICTM are shown as well.

The performances of the GCV and the Golden search method are almost equal. The GCV method produces values of the regularization parameter with a very small positive bias (this can be seen in Figure 4.4 where the values of $\lambda_{GCV}$ are slightly larger than those of $\lambda_{Golden}$). This can explained as follows: the GCV method uses the Tikhonov-Miller algorithm to produce an estimate of $f$. The Tikhonov-Miller algorithm is a linear unconstrained algorithm that can produce ringing artifacts. These artifacts can be suppressed by increasing the regularization. The Golden search method uses the constrained ICTM algorithm. The constraints of the ICTM algorithm will prevent ringing artifacts to some extent reducing the need for an increased regularization. Clearly the GCV method is our method of choice for estimating the regularization parameter and will be used in the experiments presented in this and the following chapter.

Figure 4.6 shows center x-y and x-z slices of the restoration results obtained with the ICTM and Carrington algorithms using the different methods to determine the regularization parameter. At first, the result obtained with the SNR method might seem the most appealing since it lacks the structures, or texture present in the other results that are not present in the original object (see Figure 4.2). These structures are caused by the noise realization in the acquired image. The SNR method imposed the strongest
regularization at the SNR chosen in this figure (see Figure 4.4), which smooths not only the noise-induced texture but the restoration result as well.

\[ ICTM \]

\[
\begin{array}{cccc}
\text{SNR} & \text{CLS} & \text{GCV} & \text{ML} & \text{Golden} \\
\end{array}
\]

\[ Carrington \]

Figure 4.6. Restoration results of the ICTM and Carrington algorithms with the regularization parameter determined by the SNR, CLS, GCV, ML, and Golden search method. The pictures show the center x-y an x-z slices of the center (80x80 pixels x-y, 80x40 pixels x-z) of an image of a sphere with a 0.50 \( \mu \text{m} \) radius and a SNR of 8.0.

4.3.8 The regularization parameter of the RL-Conchello algorithm

In the previous sections we have shown that the GCV method is a good method for determining the regularization parameter for the ICTM algorithm and the Carrington algorithm. The method is computationally feasible since we use the Tikhonov-Miller algorithm to compute an estimate for \( f \).

The RL-Conchello algorithm is based on a regularized maximum likelihood function for images distorted by Poisson noise. No linear solution exists for this functional. We have therefore approximated the functional such that it can be solved linearly. This linear
solution can then be used to estimate the regularization parameter with a feasible computational complexity.

The RL-Conchello maximizes the regularized loglikelihood functional of a translated Poisson process

$$ L(\hat{f}) = -\sum H\hat{f} + m^T \ln(H\hat{f} + b) - \lambda\|\hat{f}\|^2 $$

(4.22)

Maximizing (4.22) is equal to minimizing

$$ c + \sum H\hat{f} - m^T \ln(H\hat{f} + b) + \lambda\|\hat{f}\|^2 $$

(4.23)

were $c$ is a scalar constant. If we choose $c$ to be the loglikelihood of the recorded data (4.23) becomes

$$ \sum g - m^T \ln m + \sum H\hat{f} - m^T \ln(H\hat{f} + b) + \lambda\|\hat{f}\|^2 $$

(4.24)

A second order Taylor expansion of $H\hat{f}$ around $g$ yields

$$ \sum \frac{(H\hat{f} - g)^2}{2m} + \lambda\|\hat{f}\|^2 = \sum (H\hat{f} - g)^2 + 2\lambda\sum m\hat{f}^2 $$

(4.25)

To be able to find a linear solution of (4.25) in the form of a Tikhonov-Miller filter, we need to approximate $m$ by a constant. Using the mean of $m$ (4.25) becomes

$$ \sum (H\hat{f} - g)^2 + 2\lambda m\|\hat{f}\|^2 $$

(4.26)

Using one of the methods to determine the regularization parameter for the Tikhonov-Miller filter, we can determine the regularization parameter for the RL-Conchello algorithm

$$ \lambda_{RL-Conchello} = \lambda_{TM}/2\bar{m} $$

(4.27)

Thompson (Thompson, 1989) has used similar approximations for determining the regularization parameter for a mean-square-error functional regularized by an entropy measure.

Figure 4.7 shows the values of the regularization value for the RL-Conchello algorithm determined by the CLS, GCV, SNR and ML algorithms. Again we used the Golden search algorithm to find the optimal value of the regularization parameter which is included as well.
Figure 4.7 The value of the regularization parameter for the RL-Conchello algorithm determined with SNR, CLS, GCV and ML, as a function of the signal-to-noise ratio. The optimal values (Golden) are plotted as well.

Figure 4.8 The value of the mean-square-error of the restoration result of the RL-Conchello algorithm. The regularization parameter is determined with the CLS, GCV, SNR and ML method and a Golden search. The performance of the unregularized Richardson-Lucy algorithm are shown as well.

The mean-square-error performance of the RL-Conchello algorithm is shown in Figure 4.8 with the regularization parameter determined by the five different methods as a
function of the signal-to-noise ratio. We have also included the performance of the Richardson-Lucy algorithm, the unregularized version of the RL-Conchello algorithm.

It is clear from Figure 4.7 that the optimal value of the regularization value is determined considerably less accurately by any of the four proposed methods than in case of the ICTM and Carrington algorithms. However, the GCV method still produces results that are considerably better in the mean-square-error sense than the results of the unregularized Richardson-Lucy algorithm.

4.4 First Estimates

The investigated iterative restoration algorithms need a first estimate to start their iterations. In this section we compare the performance of the ICTM, Carrington and Richardson-Lucy algorithms using different first estimates. We have tested the following first estimates:

- **the recorded image**
  An obvious choice for the first estimate is to use the image \( m \) acquired by the microscope. For the ICTM and Carrington algorithms we have subtracted the (constant) background \( b \) from \( m \) to make the summed intensity of \( (m - b) \) equal to that of the original object. In chapter 3 we have defined \( m - b \) as \( g \). We used \( m \) for the RL-Conchello and Richardson-Lucy algorithms to guarantee the positivity of the first estimate.

- **the mean of the recorded image**
  Kaufman (Kaufman, 1987) has argued that starting from a uniform first guess is a good approach if no other reasonable guess as a first estimate can be obtained. We have used the mean of \( g \) for the ICTM and Carrington algorithms and the mean of \( m \) for the RL-Conchello and Richardson-Lucy algorithms as a uniform first estimate.

- **a smoothed version of the recorded image**
  The recorded image contains the smoothed object as well as noise. To prevent that the restoration algorithm will adapt to noise realizations in the recorded image one could suppress the noise in \( g \) (\( m \) for the RL-Conchello and Richardson-Lucy algorithms) by means of local smoothing. We have smoothed \( g \) with a Gaussian filter with a sigma of 2.0 pixels (this imposes a smoothing comparable to the smoothing of the point spread function used in this experiment).

- **a homogeneous noise image**
  The mean of \( g \) (or \( m \)) is a constant image in the spatial domain and an impulse in the Fourier domain. In that case the restoration algorithms have to restore the complete
Fourier domain starting from an impulse. It is interesting to see if it is beneficial to start with a first estimate that fills the Fourier domain. We have generated such an image by generating a constant image distorted by Poisson noise. The mean of this image has been set to the mean of $g$.

- **the result of Tikhonov-Miller restoration**
  The iterative algorithms based on the Tikhonov functional should converge more quickly to their solutions if a first estimate close to the final solution is provided. We used the restoration result produced by the linear Tikhonov-Miller filter to test this hypothesis. For the RL-Conchello and Richardson-Lucy algorithms we have clipped this result at zero to guarantee the positivity of the first estimate.

- **the original object**
  Ideally the restoration algorithms should produce the original object. Using the original object as a first estimate enables us to test whether this is the solution found by these algorithms under realistic circumstances.

- **a zero image for the Carrington algorithm only**
  The Carrington algorithm finds the estimate of $f$ by optimizing a functional as function of a transformed parameter $c$ (see chapter 3 for a derivation of the Carrington algorithm). Therefore none of the first estimates proposed above is a really good estimate for the Carrington algorithm. Carrington has proposed to use an image set to zero as the first estimate (Carrington, 1990).

We have measured the mean-square-error and the I-divergence of the ICTM algorithm, the Carrington algorithm, the RL-Conchello algorithm and the Richardson-Lucy algorithm as a function of the number of iterations for the various first estimates. The measured mean-square-error and I-divergence values of the first two hundred iterations are shown in Figure 4.9 for the ICTM algorithm, in Figure 4.10 for the Carrington algorithm, in Figure 4.11 for the RL-Conchello algorithm, and in Figure 4.12 for the Richardson-Lucy algorithm. The data shows the averages over eight repetitions of the experiment using different noise realizations. Note that we have used different scales in these graphs to account for the differences in the range of the data.

From these figures we conclude that the ICTM and RL-Conchello algorithms using the various first estimates converge to the same end result in both mean-square-error and I-divergence sense. Using the relatively noisy first estimates (the recorded image and homogeneous noise image) the Carrington and Richardson-Lucy algorithms produce a
result with a somewhat poorer performance. Also the zero image first estimate for the Carrington algorithm produces a slightly poorer performance.

The ICTM algorithm shows the fastest convergence of the four algorithms tested. The results also show that none of the first estimates is a really good first estimate for the Carrington algorithm. The convergence is slow and rough when compared to the other results.

The three regularized algorithms (ICTM, Carrington and RL-Conchello) converge to a result that is very similar in both the mean-square-error and the I-divergence sense. The performance of the (unregularized) Richardson-Lucy algorithm is lower than that of the other three and does not converge to a minimum in both mean-square-error and I-divergence sense. The performance even decreases somewhat for a large number of iterations. This can be explained by the lack of regularization in the Richardson-Lucy algorithm which allow the restoration result to adapt to the noise present in the acquired image.

Only the ICTM and Carrington algorithms benefit from using the smoothed image as a first estimate. The ICTM algorithm converges quickly for this first estimate, whereas the Carrington algorithm produces a (somewhat) better restoration result. The convergence of this first estimate is worse than that of the recorded image for both the RL-Conchello and the Richardson-Lucy algorithm. Using the restoration result of the Tikhonov-Miller filter as a first estimate gives by far the fastest convergence for all algorithms except for the Carrington algorithm.

Even when the original object is used as a first estimate, the various algorithms converge to the same solution in the mean-square-error sense as obtained with the other first estimates. This can be explained by the regularization. The Tikhonov regularization can be regarded as a Gaussian a priori model of the original image (Andrews & Hunt, 1977; Verveer & Jovin, 1997). Since the intensities of our original object are not Gaussian distributed, the algorithms will find a solution, which is a balance between the distance to the original object and the Gaussian model imposed by the regularization.

Although the result was a surprise, we have not investigated why the ICTM algorithm produces a maximum mean square error at about ten iterations when the original object is used as a first estimate.
Figure 4.9 The mean-square-error and l-divergence for different first estimates as a function of the number of iterations performed by the ICTM algorithm.
Figure 4.10 The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the Carrington algorithm.
Figure 4.11 The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the RL-Conchello algorithm.
Figure 4.12 The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the Richardson-Lucy algorithm.
4.5 Stop Criterion

Starting from a first estimate \( \hat{f}_0 \), the ICTM, Carrington and Richardson-Lucy algorithms iterate to their solution. In principle one can continue to generate new estimates of \( \hat{f} \) until the restoration algorithm finds the optimum of its functional. In practice this procedure is undesirable. Experiments (Holmes & Liu, 1991) show that the likelihood of an Richardson-Lucy estimate increases logarithmically as a function of the number of iterations. This growth makes the search for the maximum of the likelihood function computationally very expensive. We have therefore used a threshold on the change of the functional

\[
\left( \hat{f}^{k+1} - \hat{f}^k \right) / \hat{f}^k
\]  

(4.28)

to stop the iteration. We have used \( 10^{-6} \) as threshold. This criterion can be seen as a speed-of-convergence criterion since it is a threshold on the slope of the change of the functional as a function of the number of iterations. The functionals of the restoration methods considered have been shown to converge as an exponential function of the number of iterations to their optimum (Holmes & Liu, 1991; Lagendijk & Biemond, 1991). Therefore the threshold determines how close to the optimum the algorithms are stopped.

4.5.1 Stopping the Richardson-Lucy algorithm

For the unregularized Richardson-Lucy algorithm the threshold stop criterion also serves as a means of regularization. A complete convergence of the Richardson-Lucy algorithm on noisy images will lead to an adaptation in the restoration results to the noise present in the acquired image (see results presented in the previous section). This will hamper the restoration result, decreasing the performance of the Richardson-Lucy algorithm. This can be observed in the performance measurements of the Richardson-Lucy algorithm as a function of the number of iterations presented in the previous section (see Figure 4.12). Stopping the iteration before a complete convergence will therefore limit this adaptation effectively regularizing the restoration in this sense.

In our simulation experiments, we could have measured the mean-square-error between the restoration result and the original object and stop when a minimum is reached as used by Verveer (Andrews & Hunt, 1977; Verveer & Jovin, 1997). However this cannot be done when restoring real data. Therefore we used for the Richardson-Lucy algorithm a threshold stop criterion and stopped when the change in functional dropped below \( 10^4 \). This choice will be very dependent on the type of image to be restored. We found that for our simulation experiments, performed on similar objects, this choice for the threshold produced results near the optimum in the mean-square-error sense.
An alternative to using a threshold stop criterion, Perry has derived the randomized GCV method to determine the optimal number of iterations for the Richardson-Lucy algorithm (Perry & Reeves, 1994). As a drawback, this method adds a considerable computational overhead (three convolutions) to each iteration of the Richardson-Lucy algorithm, which makes it a less attractive stopping criterion.
Chapter 5

Applications of Fluorescence Image Restoration

This chapter discusses several applications of image restoration in fluorescence microscopy. The first section tries to provide insight into why constrained image restoration algorithms perform better than linear algorithms. We measure the performance of these algorithms as function of the background estimate used by these algorithms. The performance measured outside the microscope's bandwidth is used to measure the "superresolution" capabilities of non-linear constrained restoration algorithms.

In the next section we show how the performance of image processing algorithms can be improved by reducing the noise influence on the restoration. A procedure, which we call prefiltering, reduces this noisy sensitivity after the image has been acquired. It uses a local smoothing filter to suppress those parts of the image spectrum where the noise dominates the signal. We show, when using a Gaussian filter as smoothing filter that the extra blurring of this filtering can be compensated for by convolving the point spread function with the same Gaussian filter.

We continue with a study in which the influence of image restoration prior to quantitative image analysis is investigated. It shows that the accuracy of integrated intensity measurements performed on a spherical object in the neighborhood of another object is increased considerably.
5.1 Background Estimation

The performance of iterative image restoration algorithms is strongly dependent on the accurate determination of several important parameters. In the previous chapter we have investigated methods to determine the regularization parameter. We have shown the dependency of the algorithms on the first estimate and discussed criteria for stopping these iterative algorithms. Furthermore, van der Voort (van der Voort & Strasters, 1995) has shown the dependency of the performance of the ICTM algorithm on the accuracy with which the point spread function is estimated.

5.1.1 Introduction

In this section we investigate the influence of the estimation of the background on the performance of several image restoration algorithms. As explained in Chapter 2 we model the image formation in a fluorescence microscope as a convolution of an object with the microscope's point spread function. We add a (constant) background and distort the image by noise. Using this image formation model the image restoration algorithms as described in Chapter 3 attempt to recover the object given a priori knowledge about the type of noise, the point spread function and the background.

The removal of the background from the acquired image is incorporated differently in the Tikhonov-Miller, ICTM, Carrington, Richardson-Lucy and the Conchello algorithms. In the Tikhonov-Miller, ICTM and Carrington algorithms the assumption of additive Gaussian noise removes the dependency of the noise from the object and background. Therefore the background can be removed by subtracting it from the acquired image (see Chapter 3 for a rigorous mathematical treatment). The Richardson-Lucy and Conchello algorithms model the noise as Poisson noise. This prevents the removal of the background by subtraction. Instead the background is incorporated in the conditional expectation of the translated Poisson process used to model the object intensities in the image (see section 3.2 in (Snyder & Miller, 1991) and Appendix A of this thesis). The EM algorithm uses the relative weights of the intensity of the object and of the background to estimate the object intensity from the intensities found in the acquired image.

In both approaches an overestimation of the background will lead to large errors in the estimation of the object intensity. In the case of the ICTM and Carrington algorithms the subtraction of an overestimated background reduces low object intensities to negative values which are clipped to zero after the first iteration. This leads to an incorrect estimation of the object's shape reducing the performance of the algorithm. Similarly an
overestimation of the background in the EM case will lead to an underestimation of the object’s intensities producing a poor performance.

On the other hand an underestimation of the background yields undesired properties as well. Not only will it lead to an incomplete removal of the background but low object intensities will - in the case of the ICTM and Carrington algorithms - not be set to near zero values in the first estimate. Therefore, the number of points in the object being clipped in each iteration is considerably reduced. Since the clipping operation of these algorithms is the only non-linear operation that distinguishes these algorithms from the linear Tikhonov-Miller filter, a reduction of the number of pixels being clipped will decrease the performance benefits of these algorithms.

In the extreme case where no points are being set to zero, both the ICTM and the Carrington algorithms reduce to an iterative, conjugate-gradient algorithm producing a solution that is identical to the one produced by the linear Tikhonov-Miller restoration filter. In other words, without clipping the ICTM and Carrington algorithms produce a solution equal to the linear solution, thus without restoring information beyond the bandwidth of the point spread function. In that case the ICTM and Carrington algorithms will not produce the sometimes claimed “super-resolution” results.

![Image of image, gradient magnitude mask, inverse mask, boundary region, object & background region]

**Figure 5.1** Generation of object, boundary and background region grey-weighted masks using the gradient magnitude of the image.

### 5.1.2 Experiment

In this section we present results from a simulation experiment in which the performance of several image restoration algorithms is measured as function of the estimated background. We not only measured the overall performance but also measured the performance in specific regions in both the frequency and spatial domain. We used the footprint of the optical transfer function (OTF) (see Chapter 2) as a mask to measure the performance inside and outside the bandwidth of the point spread function. This can give
some insight to what extent these non-linear algorithms restore information beyond the microscope’s resolution.

In the spatial domain we measured the performance in three regions in the image: inside the object, at its boundary, and in the background of the image. To avoid aliasing effects, we used bandlimited masks to select these regions. The mask for the boundary region has been generated by normalizing the gradient magnitude of the acquired image (see Figure 5.1). The gradient of the image has been computed by convolving the image with Gaussian derivatives. We used a sigma of 0.9 pixels for the derivatives. We segmented the inverse of the boundary mask into two masks: an object mask and a background mask (see Figure 5.1).

![Image](image.png)

**Figure 5.2** Mean-square-error performance of the ICTM, Tikhonov-Miller, Richardson-Lucy, RL-Conchello, clipped Tikhonov-Miller, Carrington and unclipped ICTM algorithms together with the performance of the unrestored data as a function of the estimated background value.

We have tested the linear Tikhonov-Miller filter, the ICTM algorithm, the Richardson-Lucy algorithm and the Conchello algorithm. We also included the ICTM algorithm with the clipping after each iteration disabled to show that this variant of the ICTM algorithm is nothing but a conjugate gradient algorithm to find the linear Tikhonov-Miller result. The ICTM algorithm performs better than linear filters by clipping after each iteration. One could question whether the result obtained this way is different from clipping the result of the linear Tikhonov-Miller filter. We have therefore included the result of this procedure as well (referred to as clipped Tikhonov-Miller).
Figure 5.3 Enlargement of Figure 5.2 around the true background.

All images have been generated with the same constant background intensity of 16.0 ADU. However, we have varied the estimate of the background intensity, which is an input to the restoration algorithms, from 0.0 to 32.0 ADU. We have used a simulated confocal point spread function with a NA of 1.3, a refractive index of 1.515, an excitation wavelength of 488 nm, an emission wavelength of 520 nm, and a pinhole diameter of 300 nm. The images, sampled at twice the Nyquist rate, are 64x64x32 pixels large which corresponds to an image size of 1.55x1.55x2.73 μm. We used spheres with an intensity of 200.0 ADU and a diameter of 500 nm as objects. In this experiment we used a signal-to-noise ratio of 1.0 which corresponds to a conversion factor of 0.22 ADU/photon.

5.1.3 Results

Figure 5.2 and Figure 5.3 clearly show that the performance of the tested non-linear restoration algorithms is strongly dependent on the estimation of the background. A (severe) underestimation of the background yields a performance of these algorithms that is not significantly better than the linear Tikhonov-Miller filter. An overestimation has an even more dramatic influence on the performance. The performance drops quite significantly for relatively small overestimations (< 25 %), drops below the performance of the linear filter for an overestimation of 25-50%, and even drops below the performance of the unrestored (acquired) image for an overestimation larger than 50%.

A lower performance corresponds with a higher mean-square-error, which is plotted in Figure 5.4.
Figure 5.4 Mean-square-error performance of the ICTM, Tikhonov-Miller, Richardson-Lucy, RL-Conchello, and clipped Tikhonov-Miller algorithms measured inside the bandwidth of the OTF.

Figure 5.5 Mean-square-error performance of the ICTM, Tikhonov-Miller, Richardson-Lucy, RL-Conchello, and clipped Tikhonov-Miller algorithms measured outside the bandwidth of the OTF.

The figures also show that the performance of the unclipped ICTM is identical to the linear Tikhonov-Miller filter, and that the Carrington algorithm yields a performance characteristic very similar to the ICTM algorithm. We have therefore not included the results of the unclipped ICTM algorithm and the Carrington algorithm in further analysis.
The performance inside the bandwidth of the OTF, as shown in Figure 5.4, shows a similar characteristic as found for the overall performance. Roughly ninety percent of the total mean-square-error is measured in this region. This is simply explained by the fact that most of the object energy is found in this region. The remaining ten percent, found outside the bandwidth of the OTF shows a slightly different characteristic (see Figure 5.5). The mean-square-error of the linear Tikhonov-Miller is here lower than all the non-linear algorithms except for the RL-Conchello algorithm. The mean-square-error of the Tikhonov-Miller filter is simply the energy of the object in this region of the Fourier domain. As expected, the non-linear algorithms add frequency components outside the bandwidth of the OTF to improve the performance inside the bandwidth. These restored components however do not always improve the performance outside the bandwidth as well.

The performances in the spatial domain on the object, edge and background are shown in Figure 5.6, Figure 5.7 and Figure 5.8. We conclude from these figures that the largest gain in performance is found in the background, the region that contains the most pixels and the lowest pixel intensities.
Figure 5.7 Mean-square-error performance of the ICTM, Tikhonov-Miller, Richardson-Lucy, RL-Conchello, and clipped Tikhonov-Miller algorithms measured around the edges of the object.

Figure 5.8 Mean-square-error performance of the ICTM, Tikhonov-Miller, Richardson-Lucy, RL-Conchello, and clipped Tikhonov-Miller algorithms measured in the background.

The ICTM algorithm produces the best performance in the object and edge region for an overestimation of 12.5 - 25%, when some of the low intensity edges of the object are being clipped introducing (correlated) high frequency components (Figure 5.5). These
figures also clearly show that only the performance of the clipped Tikhonov-Miller algorithm in the background region is influenced by the background estimation. Figure 5.9 shows the square-error at every frequency in the center $\omega_0\omega_0$ slice between the original object and the restoration results of the ICTM, RL-Conchello and clipped TM algorithms for three values of the estimated background (12.0, 16.0 and 20.0). It clearly shows that errors due to noise (stochastic errors) are replaced by “deterministic” errors at low frequencies.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image}
\caption{The mean-square-error in the center $\omega_0\omega_0$ slice of the ICTM (top), RL-Conchello (middle) and clipped TM (bottom) algorithms for an estimated background of 12.0 (left), 16.0 (center) and 20.0 (right). The grey scaling is constant over all nine images.}
\end{figure}
5.1.4 Methods for background estimation

The experiments presented in the previous paragraph show the strong influence of the background estimation on the performance of non-linear restoration algorithms. In this section we discuss how the background can be estimated in real images.

Given the diversity of the samples being imaged with fluorescence (confocal) microscopy, the characteristics of the acquired images are similarly diverse. Therefore it is impossible to construct a single background estimation algorithm suitable for such a wide range of images. Not only do the samples vary from sparse (like our images of spheres) to very dense, the background can be different as well.

In general, the background can be characterized as being of low intensity and of a low spatial frequency. In images of sparse objects the majority of the pixels are background pixels, and a histogram-based algorithm can be used to estimate the background. Figure 5.10 shows the histogram of one of the simulated confocal images used in the experiments as described in section 5.1.2. Assuming that the distribution of the dominant noise source in the image is unbiased and unimodal (which holds for Poisson noise), we can estimate the background using the position of the maximum of the histogram.

![Histogram of an simulated confocal image](image)

**Figure 5.10** Histogram of an simulated confocal image as used in the experiment described in section 5.1.2.

In images of dense objects or in images with a non-constant background, the histogram-based approach will not be very accurate (still the maximum of the lowest intensity peak might give a reasonable estimate). In these cases an approach based on mathematical morphology or fitting of a polynomial might work.

---

7 This can be implemented by fitting a parabola through the maximum values in the histogram.
In the first approach an opening operation can be used to estimate the shape of the background (Verbeek et al., 1988). The second approach fits a low order polynomial through the pixels (Young, 1970). The accuracy and the bias of the fit will be improved if the fit is done through background pixels only. Therefore the image needs to be segmented (coarsely) in object and background pixels. One way of doing this is to use a criterion based on the noise variance. Given a first (histogram-based) estimate of the background intensity a pixel can be labeled as object pixel if its intensity is more than \( n \) (for example 2 or 3) times the standard deviation of the noise. The noise variance in the image can be estimated using the technique described in section 4.3.

**Figure 5.11** The mean-square-error between the acquired image and the blurred restoration result as function of the background.

We propose however an alternative method for estimating the background using the dependency of the non-linear restoration algorithms on the background estimation. As illustrated by Figure 5.4, the performance of the non-linear algorithms inside the bandwidth of the OTF is (strongly) dependent on the background. Therefore a measure that uses this part of the frequency domain can be used to measure the performance of a non-linear algorithm as function of the background estimation. By optimizing this measure as function of background, the optimal background can be determined.

We propose to use the mean-square-error between the acquired image and the restoration result blurred by the microscope’s OTF with the added background,

\[
\sum (g - (H \hat{f} + b))^2
\]  
(4.29)
to measure the performance as function of the background. The values of this measure as discussed in the experiments in the previous section are shown in Figure 5.11 for the ICTM, clipped-TM, RL-Conchello and Richardson-Lucy algorithms.

Using the proposed measure we can find the optimal background value by increasing the background value until the mean-square-error increases significantly. For a constant background, the optimal background value will be found in the interval between zero and to the mean intensity of the acquired image.

A disadvantage of the proposed method for background estimation is that it requires a few restoration results obtained with, for example, the ICTM algorithm. This could lead to a high computational complexity, resulting in an unacceptably long processing time. This could be solved by using the clipped Tikhonov-Miller filter instead. As can be observed from Figure 5.2 the mean-square-error of the clipped Tikhonov-Miller filter is also minimal for the correct value of the background.

5.2 Prefiltering: Reducing the Noise Sensitivity

The performance of the restoration results is strongly dependent on the amount of noise in the image. The Richardson-Lucy algorithm is an unregularized maximum likelihood estimator for Poisson distorted data. It can therefore produce results that are adapted to the noise in the image. In previous experiments (van Kempen et al., 1997 and Chapter 4 of this thesis) we observed that the performance of the Richardson-Lucy is strongly dependent on the signal-to-noise ratio of the acquired image. Furthermore, results obtained at low signal to noise ratios showed a strong adaptation to the noise realization in the acquired image.

On the other hand, regularized algorithms like the ICTM algorithm or the Carrington algorithm balance their fit to the data with a regularization term making them less sensitive to the noise in the image. The balance between the data and the regularization term is highly dependent on the signal-to-noise ratio in the acquired image. These regularized algorithms may therefore produce solutions that are smoother than desired.

The obvious way to improve the performance of the image restoration algorithms is to increase the signal-to-noise ratio of the image. Given that the noise induced by photon counting is the dominant source of noise in a microscope fluorescence image, this can be achieved by collection of more photons per pixel. This however is not always possible. Limits on the total acquisition time, saturation of the fluorescence molecules, and bleaching can limit the signal-to-noise ratio of the acquired image.
We have therefore investigated a method that reduces this noisy sensitivity after the image has been acquired. This method, which we call prefiltering, suppresses those parts of the image spectrum that do not contain any signal information (or where the noise contribution is much larger than the signal contribution). These frequencies will prevent signal recovery and only amplify noise in the final result.

We have suppressed these (high frequency) parts of the spectrum by convolving the acquired image with a Gaussian filter. Although the Gaussian filter will mainly suppress high frequencies, lower (object) frequencies are also effected. Being a linear filter we can compensate for the extra blurring of the Gaussian filter by convolving the PSF with the same Gaussian filter, as shown in Figure 5.12.

![Diagram](image)

**Figure 5.12** The incorporation of the proposed Gaussian prefiltering in an image restoration procedure.

The above defined Gaussian prefiltering can also be described analytically by incorporating it in the functionals that are optimized by the various restoration algorithms. For the ICTM algorithm and Carrington’s algorithm the Tikhonov functional with prefiltering yields

\[
\Phi(\hat{f}) = \|G(\mathbf{H}\hat{f} - g)\|^2 + \lambda \|C\hat{f}\|^2
\]  

(4.30)

with G the Gaussian prefilter matrix. We refer to chapter 3 for a derivation of the functionals stated in this section. The incorporation of prefiltering in the functional derived from the assumption of Poisson noise distorted data (used by the Richardson-Lucy algorithm) yields

\[
L(f) = -\sum GHf + (Gm)^T \ln(GHf + b)
\]  

(4.31)
Chapter 5  Applications of Image Restoration

Using these functionals the prefiltering can easily be incorporated in the algorithms based on these functionals.

The Gaussian prefiltering is a linear convolution filter to suppress the noise in the recorded image. Noise suppression by local smoothing however can also be obtained using non-linear filters like the median filter (see Figure 5.13). It is therefore interesting to compare the performance of Gaussian prefiltering with a median filter.

![Figure 5.13 Power spectrum of the original object, the recorded image, and the spectra after smoothing the recorded image by a median and a Gaussian filter.](image)

We have performed a simulation experiment to test the influence of the proposed prefiltering methods on the performance of the Richardson-Lucy and ICTM algorithm to restore spheres convolved with a confocal PSF and distorted by Poisson noise. The performance of the algorithms is measured with the mean-square-error and I-divergence measure as a function of the signal-to-noise ratio. We have generated spheres with a radius of 1.0 μm, an object intensity of 200.0 ADU and a background of 40.0 ADU. For the prefiltering we have used a Gaussian with a sigma of 1 pixel in all dimensions and a median filter with a size of three pixels in all three dimensions.
Figure 5.14 show the I-divergence as well as the mean-square-error performance of the Richardson-Lucy and ICTM algorithms as function of the signal-to-noise ratio. The graphs clearly show the strong dependence of the performance of these algorithms on the signal-to-noise ratio. Furthermore, it shows the improvement of the Gaussian prefiltering on both algorithms. The proposed median prefiltering however does not, in general, improve the results. The center x-y and x-z slices of the restored images of a confocal image with a signal-to-noise ratio of 16.0 are shown in Figure 5.15.

![Graphs showing I-divergence and mean-square-error performance](image)

**Figure 5.14** The I-Divergence (top) and mean-square-error (bottom) performance of the Richardson-Lucy algorithm (left) and the ICTM algorithm (right) with and without median or Gaussian prefiltering.
Figure 5.15 The center x-y and x-z slices of the results of Richardson-Lucy (top) and the ICTM algorithm (bottom) are shown without pre-filtering (left), and with median (middle) or Gaussian pre-filtering (right). The images are 128 x 128 x 32 pixels in size, the radius of the sphere is 1.0 μm and the SNR 16.0. The images are displayed with eight-bit grey scale resolution, without stretching the intensities.

We performed a second experiment to test the influence of the proposed prefiltering on the performance of the Richardson-Lucy and ICTM algorithm to restore two dimensional textured disks convolved by an in-focus wide-field incoherent PSF and distorted by Poisson noise. The performance of the algorithms is measured with the mean-square-error and l-divergence measures as a function of the signal-to-noise ratio.

We have generated disks with a radius of 1.0 μm, an object intensity of 200.0 and a background of 20 (Figure 5.16). We have added texture to the disk by modulating the disk in both dimensions with a sine. We have used sine of with a period equal to 1/7th of the image size, and a modulation depth of 20% of the image intensity.
The images are 256x256 pixels in size, the signal-to-noise ratio ranges from 16 to 1024 (12 dB - 30 dB). We have restored the textured disks using the two restoration methods with and without (using three methods of) prefiltering: Gaussian prefiltering (with a sigma of 2.0 pixels), median prefiltering (with a filter size of 5 pixels), and a combined median-Gaussian filtering (median size of three pixels and a sigma of 1.4 for the Gaussian filter). Figure 5.9 shows the mean-square-error and 1-divergence of the restoration results as function of the signal-to-noise ratio.

![Image](image.png)

**Figure 5.16** The top row shows the textured disk *(left)* and its image *(right)*, the bottom row shows the same images but with their intensities stretched to make the texture more pronounced.

The restoration results are shown in Figure 5.17 for an image with a signal-to-noise ratio of 16.0. In this figure we have zoomed in on the texture, by showing only the intensities of the texture. Cross-sections of the centerline of the images shown Figure 5.17 together with that of the textured disk are shown in Figure 5.18.
Figure 5.17 The results of Richardson-Lucy (top) and the ICTM algorithm (bottom) are shown without prefiltering and with median, Gaussian, and combined median-Gaussian prefiltering.
Figure 5.18 Cross sections of the center line of the textured disk together with the result of the Richardson-Lucy algorithm (left) and the ICTM algorithm (right) with and without Gaussian, median or combined median-Gaussian prefiltering. The grey line indicates intensities of the original object.
5.3 Improvement of Quantitative Measurements

In this section we present results of a feasibility study to determine the applicability of confocal microscopy to the analysis of mitotic cells in tumor tissue. This study has been a collaboration with the section Quantitative Pathology of the VU hospital (Amsterdam).

Proliferation markers are an important tool in the prognosis of breast cancer (ten Kate et al., 1993; Belien et al., 1997). The number of mitotic cells per unit area in a representative histological section is an important proliferation marker. To improve reproducibility and to reduce the time associated with mitoses counting, automated counting of mitoses in histological sections using image processing techniques has been developed (Belien et al., 1997). Histological sections stained with Feulgen are acquired with a camera attached to a wide-field fluorescence microscope. The stoichiometric
staining with Feulgen allows the fluorescence intensity within each cell to be related to the amount of DNA in the cell. The increased amount of DNA in a cell discriminates the mitotic cells from the non-mitotic cells.

This procedure, however, does not allow the stoichiometric data to be related to morphological characteristics of the three-dimensional cell, since this information is lost by the sectioning of the cell. One possible way to combine both characteristics is to acquire a three dimensional image of the cell using a fluorescence confocal microscope.

We performed a feasibility study to investigate the applicability of using a fluorescence confocal microscope in the analysis of mitotic cells (van Dis, 1996). One of the problems we encountered was a very poor axial resolution in the images of human liver cells. This resulted in images in which the intensities of neighboring cells merged into each other, making segmentation of these cells and a reliable measurement of their DNA content impossible (see Figure 5.20). We have therefore tried to improve the axial resolution of these images by means of image restoration. Furthermore, we have performed a simulation experiment to gain more insight in the influence of image restoration on quantitative measurements.

5.3.1 Determination of the point spread function
The image restoration algorithms investigated in this thesis require a priori knowledge of the point spread function of the microscope used to acquire the images. This a priori knowledge can be obtained from a theoretical description of the image formation. We have devoted Chapter 2 to the modeling of the image formation of a fluorescence microscope based on the assumption of diffraction limited optics.

However, such a theoretical description of the PSF will only produce reliable estimates of the microscope’s PSF if the assumptions made are satisfied by the conditions under which the images are acquired (van der Voort & Strasters, 1995).

This is, however, not the case in the images under investigation here (see Figure 5.20). The images are blurred to a larger extent than can be expected from diffraction limited point spread function. Apparently factors like the imperfect setup of the microscope (suboptimal pinhole size), specific sample preparation (mismatches in refractive indices) or other aberrations contribute to the decreased (axial) resolution.
We therefore had to measure the PSF of the microscope. In principle the PSF can be measured by acquiring an image of a point object (or of an object significantly smaller than the resolution of the microscope). Images of such small fluorescence beads (diameter < 0.1 µm) are however too noisy for a reliable estimation of the PSF.

We have therefore used the method suggested by (van der Voort & Strasters, 1995). Assuming that the image formation can be modeled as a convolution of the object with the point spread function distorted by noise, the PSF can be restored if the object is known. In section 4.2 we describe a method to generate bandlimited spheres. Using this method we can restore the PSF from an image of a fluorescent bead with a known diameter using the ICTM or the Richardson-Lucy algorithm. Although this method does not require images of very small beads, the beads cannot be too large as this will hamper the estimation of the high frequencies of the PSF. The signal-to-noise ratio of the image from which the PSF is restored can be improved by averaging several images of (various) fluorescence beads.

We have acquired ten images of volume-stained latex beads with a diameter of 0.28 µm (Molecular Probes, Oregon, USA). These images have been aligned with subpixel accuracy before averaging. We have aligned the images by measuring the center of mass of the beads and shifting them to the center of the image. The center of mass of the beads has been measured by taking only the pixels with an intensity two sigma above the average background intensity into account. The mean and standard deviation of the background intensity has been estimated in regions at the border of the image (van Dis, 1996).
Figure 5.21 Center regions of the center x-y (top) and x-z (bottom) slices of the measured PSF. The x-y region is 4.2 \mu m \times 3.5 \mu m in size, the x-z region 8.3 \mu m \times 12.8 \mu m.

shows the center x-y and x-z planes of the PSF that has been measured using the described procedure. shows the restoration result of the ICTM algorithm using the measured PSF of the image of human liver cells as shown in Figure 5.20. Clearly the restored image shows
an improved resolution in the axial direction, resulting in an improved separability of the cells.

\[ x-y \text{ slices} \quad x-z \text{ slices} \]
Figure 5.22 The same five x-y and x-z slices of the restoration result of the ICTM algorithm of the image displayed in Figure 5.20.

5.3.2 Improving volume measurements by image restoration

The results presented in the previous paragraph show the improvement in (axial) resolution obtained by image restoration, which results in an improved separability of the individual cells. In this section we investigate how image restoration influences measurements performed on one object in the vicinity of another object.

Although the image restoration algorithms investigated in this thesis minimize the mean-square-error or the I-divergence, these measures are not conclusive when the goal of image restoration is to improve the quantitative analysis of microscopic images. Therefore, we have also examined the performance of these restoration procedures with an analysis-based performance measure.

The total amount of fluorescence inside a closed object is in many applications a useful measurement. It could for example represent the total amount of DNA inside a cell nucleus. Image blur however does not permit this measurement when objects lie close together. In this situation image restoration algorithms can be applied prior to the measurement to reduce this blur. A useful performance measure is how the intensities are being reshuffled without mixing intensities that come from different objects. We have implemented this performance measure by assigning the intensity of a pixel to the object with the shortest distance from the pixel to the object's surface.

In this section we test the ability of the Richardson-Lucy and the ICTM algorithm to improve the quantitative measurement of the total amount of fluorescence inside one sphere in the vicinity of a second sphere as a function of the distance between the spheres. We have generated one sphere with a radius of 0.4 μm and an intensity of 200.0 ADU and a second sphere with a radius of 0.6 μm and an intensity of 50.0 ADU (with a background
set to 20.0 ADU). A fixed lateral distance of 0.05 μm is used between the two spheres and we varied the axial distance from 0.0 μm to 2.0 μm. We choose to vary the axial distance as most of the blurring occurs in this direction. The total amount of fluorescence inside the spheres is measured before convolving it with the confocal PSF to obtain the ground truth. This is then compared to the measurements on the simulated images with and without applying image restoration prior to the measurement. We measured the total amount of fluorescence inside a sphere by assigning the intensity of a pixel to the sphere with the shortest distance from the pixel to the sphere's surface. Figure 5.23 depicts an overview of the experiment.

![Diagram](image)

**Figure 5.23** Overview of the simulation experiment performed to test the influence of image restoration on the volume measurement of a sphere in the vicinity of a second sphere.

In the simulation experiment we used a sampling distance of 46.0 nm in both the lateral and axial direction. The images have a size of 64 x 64 x 128 pixels. A signal-to-noise ratio of 16.0 was used. Figure 5.24 shows the error of the total intensity measured before and after restoration relative to the total intensity measured before convolving the spheres with the confocal PSF. Two spheres (with an axial distance of 0.2 μm) and their confocal image are shown in Figure 5.25 together with the restoration results of the filtered Richardson-Lucy and the ICTM algorithms. We used the GCV and CLS algorithms to determine the regularization parameter of the ICTM algorithm.
Figure 5.24 The relative error of the total intensity of one sphere as a function of axial distance to a second sphere. The intensity is measured before and after restoration with the filtered Richardson-Lucy or the ICTM algorithm (using GCV and CLS to determine $\lambda$).

Figure 5.25 Two spheres. From the left to the right, we show the center $x$-$z$ slice of the object, image and of the restoration results of the prefiltered R-L and ICTM algorithm. The axial distance between the two spheres is 0.2 $\mu$m.
The most computational intensive part of the image restoration algorithms used in this thesis is the convolution with the point spread function. This convolution can be implemented efficiently using the fast Fourier transform. The fast Fourier transform is an example of the class of separable image processing algorithms. In this chapter we show that a straightforward implementation of separable image processing algorithms on modern workstations will give the worst possible performance regarding data cache utilization on large images. Modern workstations are equipped with fast cache memory to enable the CPU to access the relatively slow main memory without noticeable delay. However, two typical cache characteristics, limited associativeness and power of two based memory address mapping on cache lines, severely hamper the performance of separable image processing algorithms. We present three methods based on transposing the image to improve the data cache usage for both write-through and write-back caches. Experiments with a 3x3 uniform filter and the fast Fourier transform performed on a range of Sun workstations show that the proposed methods improve the performance considerably.
6.1 Prologue

Iterative image restoration algorithms are computationally expensive. Both the large number of iterations (typically ranging from ten up to a thousand) as well as the computational complexity of a single iteration contribute to this. The complexity of a single iteration of the investigated algorithms is dominated by a small number of convolutions with the microscope's point spread function. A discrete convolution can be computed as a neighborhood operation or by multiplying the Fourier transforms of the two images that need to be convolved. The complexity of the local neighborhood implementation is

\[ O\left( \prod_{i=1}^{D} m_i \right) \]  \hspace{1cm} (6.1)

with \( m_i \) the size of the neighborhood and \( D \) the dimensionality of the image. The complexity using the fast Fourier transform is

\[ O\left( \sum_{i=1}^{D} n_i \log n_i \right) \]  \hspace{1cm} (6.2)

with \( n_i \) the size of the image in each dimension. Given the spatial extent of the PSF (see chapter 2), the convolution can not be computed efficiently with a local neighborhood operation. Instead it can be computed more efficiently using the fast Fourier transform.

The fast Fourier transform therefore contributes significantly to the computational complexity of the iterative image restoration algorithms. Table 6.1 lists the relative amount of time that the ICTM, Carrington, RL-Conchello, and Richardson-Lucy algorithms spent on computing the fast Fourier transform during the experiments described in section 4.3.

<table>
<thead>
<tr>
<th></th>
<th>ICTM</th>
<th>Carrington</th>
<th>Richardson-Lucy</th>
<th>RL-Conchello</th>
</tr>
</thead>
<tbody>
<tr>
<td>relative time (%)</td>
<td>41.8</td>
<td>45.5</td>
<td>52.1</td>
<td>40.1</td>
</tr>
</tbody>
</table>

Table 6.1 The relative amount of time that the ICTM, Carrington, RL-Conchello, and Richardson-Lucy algorithms spent on computing the fast Fourier transform.

From Table 6.1 it is clear that the execution time of these image restoration algorithms is largely dependent on the time spent on computing the fast Fourier transform. A speedup of the Fourier transform is thus an efficient way to speed up the overall execution time of these algorithms.
The three dimensional Fourier transform,

$$F(\omega_x, \omega_y, \omega_z) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy dz \, f(x,y,z) e^{-i(\omega_x x + \omega_y y + \omega_z z)}$$  \hspace{1cm} (6.3)$$

can be separated in three consecutive one-dimensional Fourier transforms,

$$F(\omega_x, \omega_y, \omega_z) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} dz e^{-i\omega_z z} \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dy e^{-i\omega_y y} \frac{1}{(2\pi)^1} \int_{-\infty}^{\infty} dx f(x,y,z) e^{-i\omega_x x}$$  \hspace{1cm} (6.4)$$

This reduces the complexity of the Fourier transform from $O(\Pi_{i=1}^{3} f_i)$ to $O(\Sigma_{i=1}^{3} f_i)$ with $f$, the complexity per pixel of the one dimensional Fourier transform.

The Fourier transform is an example of a class of image processing algorithms whose multi-dimensional operation can be separated in consecutive one-dimensional operations. Other examples of separable image processing algorithms are the Gaussian filter and the rectangular uniform filter. In this chapter we discuss how separable filters, like the fast Fourier transform, can be implemented efficiently on modern workstations.
Optimal cache usage for separable image processing algorithms on general purpose workstations

Klamer Schutte and Geert M.P. van Kempen

(Signal Processing, vol. 59, no. 1, pp. 113-122, May 1997)

6.2 Introduction

For years, the efficiency of image processing algorithms has been expressed in terms of the number of operations. Until the early nineties, the speed of standard computer hardware for image processing tasks was dominated by the CPU speed and thus limited by the number of operations.

In the past decade, the CPU speed of workstations has increased much faster in comparison to memory speed (access time). This has resulted in a memory bandwidth bottleneck. Currently, modern workstations are equipped with a fast cache memory to circumvent this bottleneck. Ideally, a cache enables the CPU to access the main memory without noticeable delay.

Two cache characteristics of modern CPU designs however cause the complete class of separable image processing algorithms to give the worst possible performance regarding data cache usage on large images. These characteristics are the limited number of locations at which a memory address can be stored in cache and a power of two based address arithmetic used for mapping memory addresses to data cache addresses.

Separable image processing algorithms are algorithms of which their $k$-dimensional ($k$-D) operation on an image can be replaced by $k$ consecutive one-dimensional operations. In general this reduces the computational complexity per pixel from $O\left(\prod_{i=1}^{k} f_i\right)$ to $O\left(\sum_{i=1}^{k} f_i\right)$ with $f_i$ the size of the environment considered in each dimension (Groen et al., 1988). Examples of separable algorithms are the fast Fourier and Hartley transform, convolution with Gaussian and boxshaped kernels and morphological operations with appropriate structuring elements.

Often such algorithms are executed by first processing the input image in the row direction, followed by processing the output of the first step in the column direction. This
is followed in the 3D case by a processing step in the slice direction. In this paper we will refer to this method of processing as the "classical" method. We will present methods to improve the data cache utilization which are based on transposing the image after it has been processed in one of its dimensions. This transposition is optimized for both write-through as well as write-back cache designs. The presented methods can be used for processing images of arbitrary dimension. (Wada, 1994) developed a system to circumvent the same problem of cache trashing. However, the solution he proposes needs a different image layout, which is hard to implement in current software systems. Both (Wittenbrink & Somani, 1993) and (Dykes et al., 1994) propose a tiling mechanism. These approaches however are not aimed at separable algorithms but rather on local neighborhood algorithms. (Alperen et al., 1994) have developed a Uniform Memory Hierarchy model. Based on this model, they derive a "communication efficient" way to perform a fast Fourier transform by using matrix transposition. (Lam et al., 1991), (Agarwal et al., 1994), and (Ling, 1993) propose a blocking approach for separable algorithms. The difference with the presented approach is that these approaches ignore the effects, which occur at image sizes of the power of two, which are commonly used in image processing.

The next section describes the cache organization for modern workstations and explains the problems that occur when using such machines for image processing. In section 6.4, we present our methods to circumvent these problems. Experiments are described in section 6.5 and we conclude in section 6.6.

6.3 Cache Organization

This section gives a brief explanation of caches found on general-purpose workstations. More detailed information can be found in textbooks on computer architecture such as (Hennessy & Patterson, 1990). We will look into the performance of workstations with first (L1) level and optionally second (L2) level caches. A second level cache is located between the first level cache and the main memory. The first level cache often is divided into separate parts for instructions and data. In this paper, we only consider the data cache.

Copies of main memory areas (referred to as blocks) are stored in a cache. Each copied block is stored in a cache line. A cache line typically has the size of 32 bytes. At the time of writing, the size of first level caches ranges between 4k and 64k bytes; second level caches have sizes between 256k and 1M (Dobberpuhl, 1992; Sun Microsystems, 1992; Moore, 1993; Heinrich, 1994; Sun Microsystems, 1995). Cache architectures differ in the number of locations at a memory block can be stored in a cache. Distinction is made
between *direct mapped*, *N-way set associative*, and *fully associative* caches. In a direct mapped cache, each block can be stored in only one cache line. The mapping is usually the block-frame address (\( X L \) in Figure 6.1) modulo the number of cache lines (\( L \) in Figure 6.1). In current cache designs, the cache mapping uses the lower memory address bits for addressing within a cache line (\( y \) in Figure 6.1) and the medium address bits for addressing the appropriate cache line (\( L \) in Figure 6.1), making the range of \( L \) and of \( y \) powers of two. If a block can be placed in a restricted set of places in the cache, the cache is said to be *set associative*. A set is a group of two or more cache lines. A block is first mapped onto a set, and then the block can be placed anywhere within the set. If there are \( N \) cache lines in a set, the cache is called *N-way set associative*. In the case of a fully associative cache, each memory address can be mapped to every cache line. However, fully associative caches are usually not found in modern workstations due to speed constraints and design complexities. Typically, such machines are equipped with direct mapped or 2,4-way set associative first level cache (Dobberpuhl, 1992; Sun Microsystems, 1992; Moore, 1993; Heinrich, 1994; Sun Microsystems, 1995), or with a first level cache in conjunction with a direct mapped second level cache (Sun Microsystems, 1992; Heinrich, 1994; Sun Microsystems, 1995).

\[
\begin{array}{c}
\text{cache} \\
\begin{array}{cccc}
X & y \\
0 & a & b & c & d \\
1 & & & & \\
2 & & & & \\
3 & & & & \\
\end{array}
\end{array}
\begin{array}{c}
\text{memory} \\
\begin{array}{cccc}
Ly & 0 & 1 & 2 & 3 \\
0 & a & b & c & d \\
1 & & a & b & c & d \\
2 & & & a & b & c & d \\
3 & & & & a & b & c & d \\
\end{array}
\end{array}
\]

*Figure 6.1* The mapping of blocks on a direct mapped cache. The cache has four lines (0-3) of 4 entries (a-d). For the part of the memory shown, \( X \in A, \ldots, F, L \in 0, \ldots, 3 \) and \( y \in a, \ldots, d \). Block addresses A0a and B0a map to the same cache line.

Data caches can operate in two modes for writing data to memory: write-through and write-back. In write-through mode, the data will be written to both the cache and to lower-level memory (main memory, or L2 cache). In write-back mode, the data will be written to cache only. The data is written to lower-level memory only when it is replaced. In write-back cache designs, writing of data occurs at the speed of the cache memory, and multiple writes within a cache line require only one write to lower-level memory. The
current generation of workstations is equipped with write-back caches. In the case that a
L1 and a L2 cache are both present, the L1 cache often works in write-through mode.
Cache trashing occurs when a cache line is discarded from the cache before it can be used
again. This can result in a situation where most time is spent on filling the cache. Many
image processing applications use images with a size of $2^n$ in every dimension. In the
situation that $2^n > R(L_y)$, with $R(L_y)$ the range of $L_y$, several pixels in the same
column/slice of the image will be mapped to the same cache line. ($R(L_y)$ equals sixteen in
Figure 6.1) Therefore, when we process an image in these directions, we use the same
cache line multiple times for different addresses. This will cause the previous contents of
the cache line to be discarded. As the previous contents are part of the same column/slice
and a cache line spans more then one pixel, this means that we discard pixels in following
column/slices in the cache before using them. This results in worst case cache utilization,
as for every pixel a complete cache line must be fetched.

6.4 Optimization of Cache Utilization

The problem described in the previous section can be circumvented by performing all
processing in the row direction. By doing so, we make optimal use of the content of cache
lines: by accessing a pixel, its neighboring pixels in the row direction will be placed in the
cache as well, as in general the sizes of cache lines are larger than the number of bytes
used to represent a pixel value. Furthermore, we access those blocks that will be mapped
on consecutive cache lines. This results in more efficient cache usage as all cache lines
are used.

![Figure 6.2](image)

*Figure 6.2* If the images are transposed when written, reading only happens in the
row direction, resulting in optimal performance for write-through cache systems.

6.4.1 2D images

The situation in which we perform all the processing steps in the row direction can easily
be obtained by transposing the result of each processing step (see Figure 6.2) As
explained in section 6.3, write-through cache speeds up memory reads only. Therefore, the solution for this type of cache is to read in row order, and to write the results transposed to memory. This is the optimal solution as all the read operations are performed optimally. In the remainder of this paper, we will refer to this method as the "transpose" method.

**Figure 6.3** Stripe processing: a limited number of lines is processed into a temporary buffer. The lines in this buffer are transposed while being copied to the output image.

In the case of a write-back cache, we can also improve the speed up of memory write operations by consecutively writing those pixels to the output image that are mapped on the same cache line. This can be done if the transposition of the image is performed stripe wise (see Figure 6.3) First, we process image lines scanwise to a temporary buffer, followed by a writing the buffer columnwise to the output image. As the pixels in the same column of the buffer are mapped on the same cache line in the output image, we perform the write operations optimally. By adjusting the number of row lines that will be copied into the small buffer, we optimize the transpose operation for write-back caches. Improvement is made if the length of each line in the buffer is not a power of two (see section 6.3) This is obtained by applying an offset to the start of each new line in the buffer (see Figure 6.4). By choosing the appropriate offset or skew, one can prevent a situation where within the same column more then one line of the buffer is mapped onto the same cache line, thus ensuring optimal performance while reading from this buffer. We will call this approach the "stripe" method in the remainder of this paper.
A disadvantage of both the transpose and stripe method is that both methods require extra memory for their operation, whereas the classical method can in many cases be performed "in place". Both need an extra buffer at the size of the image being processed, and the stripe method needs extra memory for its stripe buffer as well. This could be a problem when processing images that are large compared to the total available main memory. A solution to this problem is to use the "double stripe" method. This means that the image is transposed to a buffer, processed, and transposed back to memory, with both transpose operations using the "stripe" method described above (see Figure 6.5). We apply this method only in the column direction, and the classical method in the row direction, in which no trashing will occur. For separable image processing this can be performed in place, so that no separate input and output image are needed.

6.4.2 Optimal stripe size

In this section we will derive the optimal sizes for the stripe $T$. The upper bound of $T$ is given by relation that the skew per line times $T$ should not exceed the cache size. This upper bound for $T$ is not optimal for a direct mapped, write-back cache. If we use a fixed stripe buffer and process the complete image from top to bottom, the cache lines used for
the stripe buffer are overwritten $M$ times with pixels of the input image, with $M$ the factor that the image is larger than the cache. This means that we get $MTC/I$ cache write faults because the dirty cache lines of the stripe buffer are written back to main memory, with $C$ the number of bytes in one image line, and $I$ the cache line length.

![Diagram of cache and stripe buffer](image)

**Figure 6.6** Cache conflicts when copying the stripe buffer to the output image. The mapping from the stripe buffer and the stripe in the output image to the cache are shown. The black area is the conflict area.

A similar process takes place when copying the stripe buffer into the output image, as in Figure 6.6. Here we have $T^2/I$ cache faults for every stripe. Since there are $R/T$ stripes to be processed we get $R/T/I$ cache faults for a complete image, with $R$ the number of image lines of the image. Combining the cache faults due to overlap in the filling and the copy stage of the stripe buffer we get $(MC+R)I/I$ cache write faults. Without these cache faults a maximum stripe size should be optimal due to minimization of loop overhead. If they are taken in account the optimum is dependent on the actual compiler and instruction set used.

### 6.4.3 Extension to higher dimensional images

The transpose rules given in the section on 2D images can easily be extended to the 3D case. Again, read and write operations in the column and particularly in the slice direction should be avoided as much as possible by performing all the processing in the row direction. If we process all the rows in the $xz$ plane, store the result in the column direction of the $xy$ plane, and perform this for all the $xz$ planes, we process the complete image in the $x$ direction, and also rotate the image such that $(x,y,z) \rightarrow (z,x,y)$. If we repeat this operation again we process in the row direction which now corresponds to the original $z$ direction. This will rotate $(z,x,y) \rightarrow (y,z,x)$. If we process again we rotate $(y,z,x) \rightarrow (x,y,z)$, and have processed the original $x$, $y$ and $z$ directions (see Figure 6.7). This solution is optimal for write through caches, as all read operations are performed in the row direction. In the case that a the size of a single $xy$ plane is smaller than that of the
cache, this algorithm is also optimal for write back caches. In the case of larger 3D images, we should use the stripe method.

Figure 6.7 Rotation in 3D. After processing every dimension we rotate the image while writing the image data.

Having extended the 2D case to 3D, it is easy to derive a solution for the general kD case. As in the 3D case, we process all lines in dimension $x_i$ in the $x_i x_j$ plane, where $x_i$ is the highest dimension, and write the result in the $x_i x_j$ plane in the $x_j$ direction. If we process all the $x_i x_j$ planes we get the rotation $(x_{i,p}, x_{i,q}, \ldots, x_{i,p}, x_{i}) \rightarrow (x_{i,p}, x_{i,q}, x_{i,p}, \ldots, x_{i,p})$. After this, we process the $x_{i,k,k}$ plane in the $x_i$ direction, which gives the rotation $(x_{i,k}, x_{i,p}, x_{i,q}, x_{i,p}, x_{i,p}) \rightarrow (x_{i,k}, x_{i,p}, x_{i,q}, x_{i,p}, x_{i,k})$. If we repeat this for all $k$ dimensions we complete the processing. Similar to the 3D case, in the case of write-back caches, we have to execute the rotations in a striped fashion if a plane is too big to fit in the cache.

6.5 Experiments

We have performed experiments using the latest four generations of Sun workstations: The Sparc IPX, SuperSparc I, SuperSparc II and UltraSparc I CPUs with L1 or both L1 and L2 caches. (Table 6.2). For the experiments, we have chosen to use a 3x3 uniform filter and the fast Fourier transform (FFT) algorithm as benchmarking filter operations. The 3x3 uniform filter is a filter with one of the lowest computational complexities (Groen et al., 1988). The FFT, on the other hand, is a separable filter with a relatively
high computational complexity. Therefore, this set of filters gives a good indication of the performance gain for the proposed cache methods.

<table>
<thead>
<tr>
<th>machine</th>
<th>CPU</th>
<th>frequency</th>
<th>L1 cache</th>
<th>L2 cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun Ultra 140</td>
<td>UltraSparc I</td>
<td>143 MHz</td>
<td>16k,1,wt,32</td>
<td>512k,1,wb,64</td>
</tr>
<tr>
<td>Sun Sparc 20</td>
<td>SuperSparc II</td>
<td>60 MHz</td>
<td>16k,4,wt,64</td>
<td>1M,1,wb,128</td>
</tr>
<tr>
<td>Sun Sparc 10</td>
<td>SuperSparc I</td>
<td>40 MHz</td>
<td>16k,4,wb,64</td>
<td>none</td>
</tr>
<tr>
<td>Sun Sparc IPX</td>
<td>MB86903</td>
<td>40 MHz</td>
<td>64k,1,wt,32</td>
<td>none</td>
</tr>
</tbody>
</table>

Table 6.2 The fields "L1 cache" and "L2 cache" indicate the cache size (in bytes), the associativity (1-, 2- or 4-way), the cache mode: write-through (wt) or write-back (wb), and the size of a cache line (in bytes).

All experiments are performed on machines running a UNIX operating system (SunOS 4.1.3, Solaris 2.4 and Solaris 2.5) and all test software is written in C and compiled with gcc 2.7.2 (optimization level 2). The execution times of the various methods have been determined by measuring the user time to eliminate the influence of other processes. In the experiments we have taken special care to avoid a situation where the corresponding pixels of the input and output image map to positions in the cache which have a distance smaller than one cache line. This would result in a considerable amount of cache trashes for the classical method. We used a skew of 32 bytes and a buffer size of 32 image lines in all the experiments.

6.5.1 Influence of computer architecture
In the first experiment, we test the performance of four different cache designs on images with sizes ranging between four and eight megabytes. These sizes were used to guarantee that none of the images would fit completely in one of the caches. We timed the execution of the 3x3 uniform filter on byte and float (4 bytes) images and the FFT on complex (8 bytes) images. The uniform filter tests were performed on both power-of-two sized images and on images with a size of a prime number that is at least one cache line size larger than the power of two sized image.

We have normalized the execution times of the classical, transpose, stripe and double stripe method by dividing these times by the execution time of filtering the image in the row direction only (performed two times for the 2D images, and three times for the 3D images). The relative times obtained this way give an indication of the decrease in performance when the images are processed in directions suboptimal in terms of cache usage. The results are shown in Table 6.3, Table 6.4 and Table 6.5.
<table>
<thead>
<tr>
<th></th>
<th>Ultra 140</th>
<th>Sparc 20</th>
<th>Sparc 10</th>
<th>Sparc IPX</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>2048</td>
<td>2083</td>
<td>2048</td>
<td>2083</td>
</tr>
<tr>
<td>classical</td>
<td>4.11</td>
<td>2.29</td>
<td>3.96</td>
<td>2.02</td>
</tr>
<tr>
<td>transpose</td>
<td>4.24</td>
<td>2.30</td>
<td>3.49</td>
<td>1.75</td>
</tr>
<tr>
<td>stripe</td>
<td>1.32</td>
<td>1.32</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>double stripe</td>
<td>1.29</td>
<td>1.31</td>
<td>1.35</td>
<td>1.36</td>
</tr>
<tr>
<td>scanwise (µs)</td>
<td>0.24</td>
<td>0.24</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>size</td>
<td>128</td>
<td>149</td>
<td>128</td>
<td>149</td>
</tr>
<tr>
<td>classical</td>
<td>3.99</td>
<td>3.29</td>
<td>2.60</td>
<td>1.98</td>
</tr>
<tr>
<td>transpose</td>
<td>1.26</td>
<td>1.26</td>
<td>1.37</td>
<td>1.35</td>
</tr>
<tr>
<td>stripe</td>
<td>1.45</td>
<td>1.46</td>
<td>1.38</td>
<td>1.37</td>
</tr>
<tr>
<td>double stripe</td>
<td>1.39</td>
<td>1.39</td>
<td>1.43</td>
<td>1.43</td>
</tr>
<tr>
<td>scanwise (µs)</td>
<td>0.37</td>
<td>0.36</td>
<td>0.86</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 6.3 Uniform 3x3 filtering of byte images. The normalized processing times of the classical, transpose, stripe and double stripe method are listed as well as the absolute processing time in the row direction ("scanwise"). The field "size" specify the size (in pixels) of each dimension of the 2D or 3D images.

The results show that for filtering 2D images both stripe methods give the best performance for write-back caches (the first three machines in Table 6.2). Furthermore, it shows that for write-through caches, the transpose method gives the best performance. On 3D byte images, the transpose method is the fastest for most machines. This can be explained by the fact that a single plane of the image is still so small that it fits entirely in cache. Therefore we do not have to copy a part of the image plane to a temporary buffer to prevent cache trashing in the column direction, as is done by the stripe method. The results for 3D float images show that the stripe methods are the most effective method on power of two sized images for machines with write-back caches. The transpose and stripe methods are significantly faster than the classical method, as these methods do not access the image in the slice direction.
Chapter 6  Optimal Cache Usage for Separable Filters

<table>
<thead>
<tr>
<th></th>
<th>Ultra 140</th>
<th>Sparc 20</th>
<th>Sparc 10</th>
<th>Sparc IPX</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>1024</td>
<td>1033</td>
<td>1024</td>
<td>1033</td>
</tr>
<tr>
<td>classical</td>
<td>3.01</td>
<td>2.09</td>
<td>2.73</td>
<td>1.80</td>
</tr>
<tr>
<td>transpose</td>
<td>2.30</td>
<td>2.11</td>
<td>2.43</td>
<td>1.71</td>
</tr>
<tr>
<td>stripe</td>
<td>1.32</td>
<td>1.29</td>
<td>1.32</td>
<td>1.33</td>
</tr>
<tr>
<td>double stripe</td>
<td>1.31</td>
<td>1.29</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>scanwise (μs)</td>
<td>0.40</td>
<td>0.40</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>size</td>
<td>128</td>
<td>137</td>
<td>128</td>
<td>137</td>
</tr>
<tr>
<td>classical</td>
<td>2.78</td>
<td>2.39</td>
<td>2.26</td>
<td>1.58</td>
</tr>
<tr>
<td>transpose</td>
<td>1.34</td>
<td>1.25</td>
<td>1.40</td>
<td>1.25</td>
</tr>
<tr>
<td>stripe</td>
<td>1.22</td>
<td>1.24</td>
<td>1.27</td>
<td>1.22</td>
</tr>
<tr>
<td>double stripe</td>
<td>1.34</td>
<td>1.33</td>
<td>1.26</td>
<td>1.37</td>
</tr>
<tr>
<td>scanwise (μs)</td>
<td>0.61</td>
<td>0.61</td>
<td>1.33</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Table 6.4 Uniform3x3 filtering of float images. The normalized processing times of the classical, transpose, stripe and double stripe method are listed as well as the absolute processing time in the row direction ("scanwise").

<table>
<thead>
<tr>
<th></th>
<th>Ultra 140</th>
<th>Sparc 20</th>
<th>Sparc 10</th>
<th>Sparc IPX</th>
</tr>
</thead>
<tbody>
<tr>
<td>classical</td>
<td>1.70</td>
<td>1.37</td>
<td>1.12</td>
<td>1.04</td>
</tr>
<tr>
<td>transpose</td>
<td>1.60</td>
<td>1.29</td>
<td>1.17</td>
<td>1.00</td>
</tr>
<tr>
<td>stripe</td>
<td>1.13</td>
<td>1.11</td>
<td>1.05</td>
<td>1.09</td>
</tr>
<tr>
<td>double stripe</td>
<td>1.18</td>
<td>1.11</td>
<td>1.05</td>
<td>1.09</td>
</tr>
<tr>
<td>scanwise (μs)</td>
<td>1.55</td>
<td>3.86</td>
<td>6.36</td>
<td>18.49</td>
</tr>
</tbody>
</table>

Table 6.5 Fast Fourier transform of complex images. The normalized processing times of the classical, transpose, stripe and double stripe method are listed as well as the absolute processing time in the row direction ("scanwise"). The 2D complex images have a size of 1024x1024 pixels.

6.5.2 Influence of image size

The second experiment tests the performance of the three cache methods as function of the image size. We measured the absolute processing times per pixel of classical, transpose and both stripe methods for a 3x3 uniform filter on byte images as function of the image size. Figure 6.8 shows that for machines with write-back caches, the time per pixel of the stripe method remains constant for large images, both for prime sized as well as for power-of-two sized images. The times of the classical and transpose methods on the machine with only an L1 write-back cache (Sun Sparc 10) show a sudden increase at size
256. From this point, the image sizes are larger than the cache size, causing cache trashing when processing the image in the column direction. A similar effect can be seen twice on the graphs of the Sun Sparc 20 and the UltraSparc at image sizes of 256 and 1024. At the first point, the image size is larger than the L1 cache, and at 1024, the image is also larger than the L2 cache.

![Graphs showing processing time per pixel](image)

*Figure 6.8* Processing time per pixel of a 3x3 uniform filter on 2D byte images as function of the row size of a square image. The processing times of the scanwise, classical, transpose, stripe and double stripe method are listed.

### 6.6 Conclusions

The presented experiments validate the theory that on modern workstations separable image processing algorithms execute with poor cache usage when applied to large images. This is caused by two typical characteristics of cache designs: the limited number of locations where a memory address can be stored in cache and the power of two based address arithmetic used for mapping memory addresses to cache addresses. The latter characteristic causes processing of large power of two sized images, commonly found in image processing, to give the worst performance regarding cache usage.

We have presented the transpose, stripe method and double stripe method, to improve the cache usage for both write-through and write-back caches. The methods are based on transposing the image between consecutive processing steps. We have optimized the transpose operation for write-back cache by writing the result of each image line to a
skewed temporary buffer. Alternatively, the double stripe method can be used if no temporary image should be used.

We have performed experiments with a 3x3 uniform filter and with the fast Fourier transform (performed on respectively byte, float and complex images) on both 2D and 3D images. The results show a considerable increase in performance when using the stripe method on machines equipped with write-back caches. Specifically, on images with power of two sizes, we can achieve a speed up of a factor of three for both byte and float images on these machines.

The optimal method to use based on the theory and experiments presented in this paper are:

- If you use a machine with write-through cache, use the transpose method.
- If the image is large compared to the cache size, use the stripe method or the double stripe method.
- Use the classical method otherwise.

Acknowledgment

We are grateful to Prof. Dr. I.T. Young, Dr. Ir. P.P. Jonker and Ir. J.G.E. Olk for their thorough reading of the draft version of this article. This work was partially supported by the Foundation for Computer Science in the Netherlands (SION) as part of the "Programme Model Controlled Image Sequence Processing" (K.S.).
Chapter 7

Conclusions

This thesis is devoted to the application of non-linear image restoration in fluorescence microscopy. We have investigated the iterative constrained Tikhonov-Miller (ICTM) algorithm, the Carrington algorithm, the Richardson-Lucy algorithm and the RL-Conchello algorithm. We have studied the influence of various parameters, like the regularization parameter and background estimation, on the performance of these algorithms. We showed that for a proper choice of the parameters, the restoration algorithms strongly reduce the distortions imposed by a fluorescence microscope on the image. This will result in a significant improvement of the accuracy of quantitative analysis of these images, when image restoration is applied prior to the measurements. Finally we have proposed an implementation for the complete class of separable filters, which includes the fast Fourier transform, which optimizes the usage of the computer’s data cache. We showed that this can lead to speed-ups of up to a factor of three.
Chapter 7  Conclusions

7.1 Image Formation

The three-dimensional image formation in a fluorescence microscope has been modeled as a convolution of the sample with the point spread function of the microscope on a background and distorted by Poisson noise. We derived a model for the point spread function of a general fluorescence microscope having both a finite-sized illumination and detection pinhole. The point spread function of this microscope can be expressed in terms of a diffraction limited point spread function and its pinholes. The point spread functions of a confocal fluorescence microscope and a wide-field fluorescence microscope are derived as special cases in which, in the confocal case, the illumination pinhole is reduced to a point size and, in the wide-field case, the illumination pinhole is infinitely large.

The assumption of diffraction limited optics does not take aberrations into account. For example, Sheppard (Sheppard, 1997) has studied the influence on the point spread function due to aberrations in high aperture optics. Aberrations induced by mismatches in refractive index (Hell et al., 1993; Sheppard & Torok, 1997) have also been studied. These aberrations, however, will make the point spread function space-variant. The blurring of the image by a space-variant point spread function cannot be modeled with a convolution. Although most restoration algorithms can incorporate blur by a shift-variant point spread function, the computational complexity will in such a case increase enormously making the image restoration very slow.

7.2 Image Restoration Algorithms

In contrast with linear image restoration, the non-linear restoration algorithms discussed in this thesis optimize their functionals iteratively while imposing a non-negativity constraint.

Based on an image formation model (Chapter 2), the various algorithms have been derived. The ICTM algorithm and Carrington algorithm minimize the Tikhonov functional, which assumes that additive Gaussian noise is the dominant noise source. The ICTM algorithm imposes a non-negativity constraint by clipping after each iteration. The Carrington algorithm does this by constraining the Tikhonov functional using the Kuhn-Tucker conditions.

The Richardson-Lucy algorithm employs the expectation-maximization (EM) algorithm to optimize the likelihood functional of Poisson-distorted data. The RL-Conchello algorithm incorporates Tikhonov regularization in the Richardson-Lucy algorithm. Results presented in this thesis show that these algorithms are all capable of strongly reducing the blur imposed by the image formation. For well chosen values of their parameters such as the regularization parameter and background level the algorithms
show a similar performance. All non-linear methods perform better than the linear Tikhonov-Miller filter.

In most of our experiments the ICTM and Carrington algorithm required a significantly fewer number of iterations compared to the EM based algorithms, resulting in faster execution times for a restoration (van Kempen et al., 1997). The implementations we used for these two algorithms however do require more memory.

We could have used the ICTM formalism (clipping after each conjugate gradient descent iteration) to minimize a functional based on the assumption that Poisson noise is the dominant noise source. Results presented in this thesis however show that the type of noise used in the functional will not strongly influence the performance of restoration algorithms. Therefore an ICTM-like algorithm based on the assumption of Poisson noise, will probably not lead to a significant improvement in performance over the ICTM algorithm when applied to Poisson distorted data.

7.3 Restoration Parameters

The performance of non-linear iterative image restoration algorithms is dependent on various parameters. In this thesis we have investigated the impact of the signal-to-noise ratio, regularization parameter, background estimation, choice of first estimate, and the stop criterion. We found that the signal-to-noise ratio, regularization parameter, and background estimation strongly influence the performance of these algorithms.

Regularization Parameter

We have compared various methods for determining the regularization parameter which has been developed by Galatsanos (Galatsanos & Katsaggelos, 1992) for linear Tikhonov restoration of images distorted with additive Gaussian noise. We have tested the ability of these algorithms to determine the regularization parameter of non-linear restoration algorithms for the restoration of images distorted with Poisson noise. Our results show that the method of Generalized Cross Validation (GCV) provides an efficient way for producing a value of the regularization parameter which is close to its optimal value. We defined the optimal value to be that value that produced the minimum mean-square-error between the restoration result and the original object. We determined this optimal value using the Golden search algorithm (Press et al., 1992), a numerical minimization algorithm. Over a large range of the signal-to-noise ratio, the GCV method produced practically the same value for the regularization parameter as was obtained with this search algorithm. (This “optimal” procedure can only be performed on synthetic data since it requires the non-blurred data as well). This indicates that the assumption of
additive Gaussian noise and the use of the linear Tikhonov restoration filter by the GCV method does not influence its performance even when used on images distorted with Poisson noise at a low signal-to-noise ratio.

**Background Estimation**

We model the deterministic part of the image formation as a convolution of the original object with the point spread function on a background. Since we require that restoration algorithms restore the original object from the acquired image, both the blurring and the background need to be removed from the image. The restoration algorithms discussed in this thesis are non-linear since they constrain their result to non-negative values. This constraint will only be effective when the intensities in the restoration result have values near zero. We have shown that the effectiveness of this constraint is strongly influenced by the background estimation which is an input parameter in all restoration algorithms. We showed that an underestimation of the background will make the constraint ineffective which results in a performance of these non-linear algorithms which does not differ much from the performance obtained by linear restoration filters. An overestimation of the background however is even more dramatic since it results in a clipping of object intensities. We showed that this will dramatically deteriorate the performance of these non-linear restoration algorithms.

**First Estimates**

We have proposed a method for estimating the background based on measuring the performance of the restoration algorithm as a function of the estimated background. We showed that the mean-square-error between the acquired image and the restoration result, blurred with the point spread function and with the added background, will strongly increase when the background is being overestimated.

We have investigated the influence of various choices of a first estimate as used by the iterative image restoration algorithms investigated in this thesis. We found that most algorithms converged to the same solution for the tested choices. Only the Richardson-Lucy algorithm failed to converge to a steady solution but showed a decrease in performance for a large number of iterations. This can be explained by the unregularized nature of this algorithm making it sensitive to the noise realization present in the acquired image.

**Stopping Criterion**

We did not test different criteria for stopping the iterative restoration algorithms but discussed the effects of our choice, stopping when the change of the functional drops
below a predetermined threshold. This criterion has disadvantages, especially when used on a large variety of images. We observed that the absolute value of the threshold depends of the type of images being restored (sparse/complicated structures). Furthermore the signal-to-noise ratio and the amount of regularization imposed on the restoration result do influence this stopping criterion as well. Still our choice for the stopping criterion is a practical one and, as the results in this thesis show, will yield good restoration results.

A more generally applicable and objective criterion for stopping the iteration of the various restoration algorithm is not yet known, although the method proposed by Perry (Perry & Reeves, 1994) based on the GCV algorithm might prove to be a good alternative.

In this thesis we have investigated the influence of several parameters on the performance of the restoration algorithms by measuring the performance as a function of these parameters. We have, for each parameter, separately investigated methods to determine their optimal value. When this procedure is regarded as a multi-parameter optimization problem, one will realize that this procedure will not be optimal since it ignores the cross correlation between the parameters, i.e. it assumes the parameters to be independent. It is very unlikely that the regularization parameter, stop criterion, and background estimation have a completely independent influence on the restoration performance. Although theoretically the optimal values of these parameters are then found by using a multi-parameter optimization algorithm, such an algorithm will be too complex to be usable.

### 7.4 Application of Image Restoration

We have presented two applications of image restoration in fluorescence microscopy: a reduction of the noise sensitivity of the image restoration algorithms by prefiltering, and a feasibility study of the application of confocal imaging to the analysis of mitosis.

**Prefiltering**

We define prefiltering as being a smoothing operation performed on the acquired image prior to the image restoration to reduce the influence of the noise on the restoration result. The suppression of noise in the acquired image will make the restoration result produced by the Richardson-Lucy algorithm less sensitive to the noise realization in the acquired image and will relax the influence of regularization in the restoration result for the regularized algorithms. We propose to implement this smoothing with a Gaussian convolution filter. This filtering will add an extra blur to the image but can be compensated for by blurring the point spread function with the same filter. We have also tested the effectiveness of the median filter, a non-linear smoothing filter, as a prefilter.
Chapter 7  Conclusions

We found however that it does not produce results as good as those obtained with a Gaussian prefilter. Particularly, the mean-square-error performance of the median prefilter is much worse than that of the Gaussian prefilter. This can partially be attributed to the fact that a median filter will shift edges of a curved object inwards (van Vliet, 1993) and therefore will not conserve the total intensity in the image. The mean-square-error criterion penalizes the median filter for this, whereas this property of the median filter does not have a large impact on the visual performance of the restoration result. The filtering of the acquired image will, however, make the remaining Poisson noise correlated, violating one of the assumptions used in the derivation of the various restoration algorithms. In our experiments we did not find any suggestions that this negatively influenced the performance of the algorithms.

Improvement of the analysis of mitotic cells

The second application of image restoration presented in this thesis investigates the improvement of quantitative analysis of fluorescence images when image restoration is applied prior to the measurements. We presented a feasibility study of the application of confocal microscopy for the analysis of mitotic cells in tumor tissue. We found that the blur in the recorded images made a direct analysis of the cells impossible. We have therefore applied image restoration on the images. We measured the point spread function of the microscope used in this experiment. The images restored with the ICTM algorithm showed a much-increased axial resolution as expressed by a clear separation of nearby cells. In a simulation experiment, performed in parallel with these investigations, we studied the influence of blur on the measurement of the total intensity in one spherical object in the neighborhood of another object. We showed that measurements performed on restored data yield a greatly reduced systematic error. In this simulation study we have only studied the influence of image restoration on the measurement. We have not included the influence of segmentation. Instead we used the segmentation results from the original objects in all measurements. When segmentation had been taken into account the improvement on the measurement result would probably have increased since the acquired image is significantly more difficult to segment than the restored result.

7.5 Cache-optimized Separable Filters

The most computationally complex part of the restoration algorithm is the convolution with the point spread function. Although the convolution can be computed efficiently using the fast Fourier transform, we found that the algorithms spent 40 to 50% of their total execution time on computing the fast Fourier transform.
The fast Fourier transform is an example of the class of separable image operations. We have show that a straightforward implementation of these filters will lead to a worst case use of the data cache of modern workstations. This is caused by the limited associativity of the cache and the power-of-two based map of memory locations to cache locations. We have proposed procedures, based on transposing the data, which will result in an optimal use of the cache, both for write-through design of the cache as well as a write-back cache design. In particular, the proposed double buffering procedure can be implemented with minimal alterations to existing software algorithms. Measurements performed on the various proposed procedures show that speed-ups can be obtained of a factor of three on modern workstations.
Appendix A

The Conditional Expectation of a Translated Poisson Process

In this appendix we will follow the derivation of the conditional expectation of a translated Poisson process as found in (Snyder & Miller, 1991) and show that

\[ E[F|m] = \left( \frac{H^T f}{Hf + b} \right)^T m \]  \hspace{1cm} (A.1)

Let \( \{F(A): A \subseteq X\} \) be a Poisson process on the complete data space \( X \) with an integrable intensity function \( f(E[F] = f(x)) \). Points of this input point-process are translated to the incomplete output space \( Y \) to form the output point-process \( m(E[m] = \mu(y)) \) where each point is independently translated according to the transition density \( h(x | y) \) (Snyder & Miller, 1991).

If in addition points of an independent (background) Poisson process with intensity \( b(y) \) are superimposed in the output space with the points translated from the input space, then \( m \) is a Poisson process with intensity

\[ \mu(y) = \int dx h(y | x) f(x) + b(y) \]  \hspace{1cm} (A.2)

The proof is given by (Snyder & Miller, 1991).

Let \( \{Y_1, Y_2, ..., Y_k, ...\} \) be disjoint sets in \( Y \) such that \( Y = \bigcup_{k=1}^{\infty} Y_k \). Define \( m(Y_k) \) to be the number of points in \( Y_k \) and let \( m = \{m(Y_1), m(Y_2), ..., m(Y_k), ...\} \) denote histogram data derived from points in the output space. Furthermore let \( m(A; Y_k) \) be the number of points that are translated from \( A \). Then

\[ F(A) = \sum_{k=1}^{\infty} m(A; Y_k) \]  \hspace{1cm} (A.3)

from which it follows that

129
\[
E[F(A) | m] = E[F(A) | m(Y_1, m(Y_2), ..., m(Y_k), ...)
\]
\[
= \sum_{k=1}^{\infty} E[m(A; Y_k) | m(Y_1, m(Y_2), ..., m(Y_k), ...)]
\]
\[
\text{(A.4)}
\]

For use in (A.4) we have (Snyder & Miller, 1991),
\[
E[m(A; Y_k) | m(Y_1, m(Y_2), ..., m(Y_k), ...)
\]
\[
= E[m(A; Y_k) | m(Y_k)]
\]
\[
= \frac{m(A; Y_k)}{m(A; Y_k) + m(A^c; Y_k) + \int_{Y_k} dy b(y)}
\]
\[
\text{(A.5)}
\]

where
\[
m(A; Y_k) = E[m(A; Y_k)] = \int_{Y_k} dy \int_A dx h(y|x)f(x)
\]
\[
\text{(A.6)}
\]
is the average number of points in \(Y_k\) that were translated from \(A\) and \(m(A^c; Y_k)\) is the average number of points in \(Y_k\) that were translated from the complement \(A^c = X - A\) of \(A\). Substituting (A.5) in (A.4) yields
\[
E[F(A) | m(Y_1, m(Y_2), ..., m(Y_k), ...)
\]
\[
= \sum_{k=1}^{\infty} \left[ \frac{m(A; Y_k)}{m(A; Y_k) + m(A^c; Y_k) + \int_{Y_k} dy b(y)} \right] m(Y_k)
\]
\[
\text{(A.7)}
\]

When the size of the subsets of \(Y\) in \(\{Y_1, Y_2, ..., Y_k, ...\}\) tends toward zero and \(A\) equals \(X\), (A.7) becomes
\[
E[F(X) | m] = \int_Y \left[ \frac{\int_X h(y|x)f(x)dx}{\int_X h(y|x)f(x)dx + b(y)} \right] m(dy)
\]
\[
\text{(A.8)}
\]
or written in the matrix notation of chapter 3
\[
E[F|m] = \left( \frac{H^T f}{Hf + b} \right)^T m
\]
\[
\text{(A.9)}
\]

Q. E. D.
References


References


134


References


Summary

This thesis presents iterative, non-linear image restoration techniques for application on three-dimensional fluorescence microscope images. The goal of this research is to gain a better understanding of the behavior of non-linear image restoration algorithms and to develop novel methods to improve their performance in such a way that measurements can be performed more accurately on three dimensional fluorescence images.

The formation and acquisition of a three-dimensional image by means of (confocal) fluorescence microscopy blurs the image and disturbs it with noise. These distortions hide fine details in the image hampering both the visual and the quantitative analysis of the image. The purpose of image restoration is to invert this and to suppress the noise restoring the fine details in the image, which results in an improved analysis of the image.

Chapter 1  Image Restoration in Fluorescence Microscopy
In the first chapter we introduce the principles of fluorescence microscopy and discuss the properties of the three-dimensional image formation in a fluorescence microscope. In the second part of this chapter we introduce the principles of image restoration. We give an overview of various restoration techniques used in fluorescence microscopy and discuss the influence of regularization and the estimation of the background on the performance of non-linear image restoration algorithms.

Chapter 2  Image formation in Fluorescence Microscopy
This chapter describes the image formation in a fluorescence microscope based on the wave description and the quantum nature of light. Using the wave description of light, the finite resolution of an image obtained with a microscope is derived. We discuss the conditions under which a fluorescence microscope can be modeled as a linear shift invariant system. Diffraction theory is used to model the field of incident light near focus. Using this model, we derive the image formation in a general fluorescence microscope, having both a finite sized illumination and detection aperture. The image formation in a confocal fluorescence microscope and in a wide-field fluorescence microscope are derived as special cases. We discuss sampling theory to formulate the conditions for an error free conversion of an analog image into a digital representation. Using the quantum nature of light we describe the noise properties of a light detection system. Both intrinsic
and extrinsic noise sources are treated, as well as the photon-limited characteristics of scientific-grade light detectors.

Using both descriptions of light we model the image acquired by a fluorescence microscope as the original image blurred by a translation-invariant point spread function and distorted by noise.

Chapter 3  Methods for Image Restoration

In this chapter several methods for image restoration are discussed. The Wiener filter is the linear filter that minimizes the mean square error between the original image and its restored estimate, assuming that the image is distorted by additive Gaussian noise. The Tikhonov-Miller filter is the linear filter found when minimizing the Tikhonov functional. This functional is the squared difference between the acquired image and a blurred estimate of the original object regularized by a Tikhonov energy bound. Both the Wiener filter and the Tikhonov-Miller filter are linear operations on the recorded image. Therefore they cannot restrict the domain in which the solution is to be found nor can they restore information at frequencies that are set to zero by the image formation process. These restrictions are tackled by algorithms discussed in the second part of this chapter. Both the iterative constrained Tikhonov-Miller algorithm and the Carrington algorithm iteratively minimize the Tikhonov functional. They differ however in the way the non-negativity constraint is incorporated. We conclude this chapter with a discussion of the Richardson-Lucy algorithm. This iterative algorithm finds the maximum likelihood solution using the EM algorithm when the acquired image is distorted by Poisson noise.

Chapter 4  Iterative Image Restoration Algorithms

This chapter deals with various aspects that play a role when testing and comparing iterative image restoration algorithms. We start by defining two performance measures, the mean-square-error and I-divergence, that we use for measuring and comparing the performance of image restoration algorithms. Many of the tests we present have been performed on simulated images. We discuss the properties of objects generated using an analytical description of their Fourier transform. Images are created distorting the objects with Poisson noise at a predetermined signal-to-noise ratio. We continue with a comparison of different methods to determine the regularization parameter for the Tikhonov functional. These methods use different criteria to balance the fit of the restored data to the measured image with the imposed regularization on the restored data. The last two sections deal with the iterative character of the discussed algorithms. In the first section we compare different choices for a first estimate. In the second section stop criteria for iterative optimizations are discussed.
Chapter 5  Application of Image Restoration in Fluorescence Microscopy
This chapter discusses several applications of image restoration in fluorescence microscopy. The first section tries to give some insight as to why constrained image restoration algorithms perform better than linear algorithms. We measure the performance of these algorithms as function of the background estimate used by these algorithms. The performance measured outside the microscope’s bandwidth is used to measure the “superresolution” capabilities of non-linear constrained restoration algorithms. In the next section we show how the performance of image processing algorithms can be improved by reducing the noise influence on the restoration. We continue with a study in which the influence of image restoration prior to quantitative image analysis is investigated. It shows that the accuracy of integrated intensity measurements performed on a spherical object in the neighborhood of another object is increased considerably. The final section of this chapter presents the results of the application of image restoration on confocal fluorescence images of the microscopic network structure of gel-like food samples.

Chapter 6  Optimal Cache Usage for Separable Filters
The most computationally intensive part of the image restoration algorithms used in this thesis is the convolution with the point spread function. This convolution can be implemented efficiently using the fast Fourier transform. The fast Fourier transform is an example of the class of separable image processing algorithms. In this chapter we show that a straightforward implementation of separable image processing algorithms on modern workstations will give the worst possible performance regarding data cache utilization on large images. Modern workstations are equipped with fast cache memory to enable the CPU to access the relatively slow main memory without noticeable delay. However, two typical cache characteristics, limited associativeness and power of two based memory address mapping on cache lines, severely hamper the performance of separable image processing algorithms. We present three methods based on transposing the image to improve the data cache usage for both write-through and write-back caches. Experiments with a 3x3 uniform filter and the fast Fourier transform performed on a range of Sun workstations show that the proposed methods considerably improve the performance.
Samenvatting

Dit proefschrift behandelt de toepassing van beeldrestauratietechnieken op drie-dimensionale microscopische fluorescentiebeelden. Het doel van dit onderzoek is om meer kennis te verwerven over het gedrag van niet-lineaire beeldrestauratietrategies en om nieuwe methoden te ontwikkelen die gericht zijn op het verbeteren van de prestaties van deze algoritmen om zodoende nauwkeurigere metingen te kunnen verrichten op drie-dimensionale fluorescentiebeelden.

De beeldvorming en de registratie van een drie-dimensionaal beeld met behulp van een (confocale) fluorescentiemicroscoop vervaagt het beeld en verstoort het met ruis. Deze verstoringen verbergen de fijne details in het beeld. Dit vermoedelijk zowel de visuele als kwantitatieve analyse van het beeld. Het doel van beeldrestauratie is om deze vervagingen ongedaan te maken en de ruis te onderdrukken zodat fijne details in het beeld weer detecteerbaar zijn om zodoende tot een verbeterde analyse van het beeld te komen.

Hoofdstuk 1 Beeldrestauratie in Fluorescentiemicroscopie

In dit hoofdstuk introduceren wij beeldrestauratie en haar toepassing in drie-dimensionale fluorescentiemicroscopie. In het eerste deel van dit hoofdstuk beschrijven we de principes van fluorescentiemicroscopie en bediscussieren we de eigenschappen van de drie-dimensionale beeldvorming in een fluorescentiemicroscoop. We bespreken de beeldvorming in zowel een conventionele microscoop als in een confocale microscoop.

Ondanks de significante verbetering in laterale en axiale resolutie van een confocale microscoop ten opzichte van een conventionele microscoop, verstoort de beperkte axiale resolutie van een confocale microscoop het opgenomen beeld in aanzienlijke mate. Deze verstoringen zullen de betrouwbaarheid van zowel visuele als kwantitatieve metingen aan deze beelden verminderen.

In het tweede deel van dit hoofdstuk worden de principes van beeldrestauratie behandeld. Beeldrestauratiemethoden zijn erop gericht om de vervagingen in het beeld veroorzaakt door de beeldopname ongedaan te maken. Door de aanwezigheid van ruis in het beeld zullen alleen geavanceerde beeldrestauratietrategies goede resultaten produceren. Met de snelle toename van zowel de rekenkracht als de geheugencapaciteit van moderne computers kunnen steeds betere, maar tegelijkertijd ook rekenintensievere algoritmen worden toegepast voor de restauratie van fluorescentie beelden. Een overzicht van deze
modern beeldrestauratiemethoden wordt gegeven. In het laatste deel van dit hoofdstuk wordt ingegaan op de invloed van regularisatie en de schatting van de achtergrondintensiteit op de prestaties van niet-lineaire beeldrestauratiealgoritmen.

Hoofdstuk 2 Beeldvorming in Fluorescentiemicroscopie

Dit hoofdstuk beschrijft de beeldvorming in een fluorescentiemicrocoop, gebruikmakend van zowel het golf- als quantumkarakter van licht. Gebruikmakend van de golfbeschrijving van licht wordt de eindige resolutie afgeleid van een, met een microcoop, opgenomen beeld. Wij bediscussiëren de condities waaronder een lichtmicrocoop beschreven kan worden als een lineair translatie-invariant-systeem. Diffractietheorie wordt gebruikt om het veld van het invallend licht nabij focus te modelleren. Gebruikmakend van dit model leiden wij de beeldvorming af van een algemene fluorescentiemicrocoop. Deze microcoop heeft zowel een eindige belichtingsapertuur als een eindige detectieapertuur. De beeldvorming van een confocale microcoop en van een conventionele microcoop worden afgeleid als specifieke gevallen van deze algemene microcoop.

Wij bediscussiëren bemonsteringstheorie om de condities voor een foutloze conversie van een analoog beeld naar een digitale representatie te formuleren.

De quantumbeschrijving van licht worden gebruikt om de ruiseigenschappen van een lichtdetectiesysteem te modeleren. Zowel intrinsieke als externe ruisbronnen worden behandeld, alsnog de fotongelimiteerde ruiskarakteristieken van hoogwaardige lichtdetectoren.

In de laatste paragraaf worden beide beschrijvingen van licht gecombineerd om zodoende te komen tot een model voor de totale beeldvorming in een microcoop.

Hoofdstuk 3 Methoden voor Beeldrestauratie

Hoofdstuk 4 Iteratieve Beeldrestauratiealgoritmen

Dit hoofdstuk behandelt verscheidene aspecten die een rol spelen bij het testen en vergelijken van iteratieve beeldrestauratiealgoritmen. Wij beginnen met twee prestatie maten, de gemiddelde-kwadratische-fout en de I-divergentie, die we gebruiken om de prestaties van restauratiealgoritmen te meten en te vergelijken. De meeste van de testen die we presenteren, zijn uitgevoerd op gesimuleerde beelden. Wij bediscussiëren de eigenschappen van gesimuleerde objecten die gegenereerd zijn gebruikmakend van een analytische beschrijving van hun Fourier getransformeerde. Beelden worden gecreëerd door Poisson ruis toe te voegen aan de objecten met een vantevoren ingestelde signaalruis-verhouding. Wij vervolgen met een vergelijking van verschillende methoden om de regularisatieparameter van de Tikhonov functionaal te bepalen. Deze methoden gebruiken verschillende criteria om de fit van de restauratie schatting aan het opgenomen beeld te balanceren met de opgelegde regularisatie van de schatting. De laatste twee paragrafen behandelen het iteratieve karakter van de bediscussiëerde algoritmen. In de eerste paragraaf vergelijken we verschillende keuzes voor de beginschatting en in de tweede sectie worden stopcriteria behandeld.

Hoofdstuk 5 Beeldrestauratie toegepast in Fluorescentiemicroscopie

Dit hoofdstuk beschrijft verscheidene toepassingen van beeldrestauratie in fluorescentiemicroscopie. In de eerste paragraaf wordt geprobeerd inzicht te verkrijgen in de vraag waarom niet-lineaire beeldrestauratiealgoritmen beter presteren dan lineaire algoritmen. Wij hebben de prestaties van deze algoritmen gemeten als functie van de geschatte achtergrondintensiteit die als input dient voor deze algoritmen. De prestaties, gemeten buiten de bandlimiet van de microscoop, worden gebruikt om de “superresolutie” capaciteiten te meten van de niet-lineaire beeldrestauratiealgoritmen. In de volgende paragraaf laten we zien hoe de prestaties van iteratieve algoritmen door middel van “prefiltering” verbeterd kunnen worden. Deze methode vermindert de ruisafhankelijkheid van de restauratiealgoritmen. Wij besluiten met een studie waarin de
invloed van beeldrestauratie op een kwantitatieve analyse wordt onderzocht. Het toont aan dat de meting van de totale intensiteit van een bolvormig object in de nabijheid van een ander object sterk verbeterd wordt door, voorafgaand aan de meting, het beeld te restaureren.

Hoofdstuk 6 Optimaal Gebruik van Cachegeheugens voor Separabele Filters
Het meest rekenintensieve onderdeel van beeldrestauratiealgoritmen, zoals gebruikt in dit proefschrift, is de convolutie van een beeld met de puntspreidingsfunctie van de microscoop. Deze convolutie kan efficiënt worden geïmplementeerd door middel van de fast Fourier transformatie. De fast Fourier transformatie is een voorbeeld van de klasse van separabele beeldverwerkingsoperaties. Separabele beeldverwerkingsoperaties zijn meerdimensionale operaties die op te splitsen zijn in meerdere opeenvolgende eendimensionale operaties waardoor de rekencomplexiteit van deze operaties sterk verminderd wordt. Wij tonen aan dat een recht toe-recht aan implementatie van separabele beeldverwerkingsoperaties op moderne workstations leidt tot de slechtstmogelijke prestaties met betrekking tot het gebruik van het data-cache-geheugen. Moderne workstations zijn uitgerust met snelle cachegeheugens om het de centrale processor (CPU) in een computer mogelijk te maken op volle snelheid het relatief trage hoofdgeheugen te benaderen zonder merkbare vertragingen. Echter, twee karakteristieken van cachegeheugens, de beperkte associativiteit en de tweemachtgebaseerde afbeelding van hoofdgeheugenadressen op cachelijnen, limiteren in hoge mate de prestaties van separabele beeldverwerkingsoperaties. Wij presenteren drie methoden, gebaseerd op het transponeren van het beeld, om het gebruik van het cachegeheugen te optimaliseren voor zowel “write-through” als “write-back” cachegeheugens. Experimenten met een 3 bij 3 uniform filter en de fast Fourier transformatie, uitgevoerd op een serie van Sun workstations, laten zien dat de voorgestelde methoden de prestaties in hoge mate verbeteren.
Dankwoord

Ik heb het doen van promotieresarch wel eens vergeleken met een bezoek aan Yosemite Valley. Bij aanvang van je onderzoek sta je op de Valley floor en kijk je tegen meer dan duizend meter hoge, steile, rotsvanden van kennis aan. Je realiseert je dat je eerst daar tegenop zult moeten klimmen om een bijdrage aan de wetenschap te kunnen leveren. Ik heb mijn *Four Year Trail* niet alleen gelopen. Daarom wil ik graag een aantal mensen bedanken voor de ondernomen hike.

Ted, wat ik vooral in je waardeer, is je rotsvaste vertrouwen in een goede afloop. Ook twijfel je als AIO soms danig aan een succesvolle promotie, jij blijft het licht aan het einde van de tunnel zien. Dat gaf mij vertrouwen en motivatie om verder te gaan.

Lucas, je bent na een jaar mijn begeleider geworden, op een moment dat de weg nog lang was. Door tijdelijk ook aan CGH te werken, heb ik mijn restauratie draad weer op kunnen pakken en onder jouw bezielende begeleiding tot een pracht einde kunnen brengen.

Peter, als PH-AIO zat je het verste weg, maar toch ook zeer dichtbij. Daar droeg niet alleen ons beide, zeer verwante, promotie onderwerpen aan bij, maar vooral de vele duizende e-mails die wij hebben uitgewisseld. Uit onze samenwerking ik heb veel steun en inspiratie kunnen putten.

Ik wil graag heel PH bedanken voor een fantastische tijd die ik als student en AIO in de vakgroep heb mogen doorbrengen, jullie weten als geen ander, gezellige professionaliteit af te wisselen met professionele gezelligheid!

Ik vrees dat ik ook de buitengewone gezelligheid van alle PH AIOs zal gaan missen. Met smart zal ik terugdenken aan de TP Kafee borrels en de vele daarop volgende diners, de BBQ’s, de dagelijkse Dilbert, de na-de-lunch-koffie, de lange, vaak nutteloze “mailwars” op de PH-AIO maillist, de PH-Radio (on the net & in the air), de altijd gezellige (ASCI) conferenties, en nog veel meer dat ik alweer vergeten ben.

Mijn F136 kamergenootjes, Karel, Klamer, Erik, Michael, Peter & Judith wil ik graag bedanken voor de vaak broodnodige niet-wetenschappelijk conversaties. Helaas is er wel veel veranderd: jullie drinken nu niet eens meer koffie!

Mike, de ontwikkeling van 3D-Pro was een leerzame en boeiende onderneming, en dat komt voor een groot deel door onze samenwerking, waarvoor dank.
Samenvatting

De QM-AIOs, Frank, Hans, Stephanie en Erik bedank ik niet alleen voor de vaak lange, maar toch boeiende QM-AIO sessies, maar vooral voor de gezelligheid tijdens de vele workshops en conferenties die wij als QMers bevoorrecht waren te mogen bijwonen.

Gaarne wil ik ook Johan Haverkamp en Han Blonk bedanken voor het me in staat stellen om tijdens mijn post-docschap bij Unilever mijn promotie te kunnen afronden.

Mijn ouders wil ik bedanken voor hun immer en onvoorwaardelijke steun en liefde. Finally, Lise, min Elskede, you have shown me that there is much more out there!
Curriculum Vitae &

Bibliography

Geert M.P. van Kempen was born in Haarlem, the Netherlands, on November 20th 1968. He obtained his Gymnasium diploma from the Petrus Canisius College in Alkmaar. In 1987 he went to Delft to study Applied Physics at the Delft University of Technology where he received his Master’s degree (Ingenieur’s diploma) in 1993. For his Masters thesis he studied the application of image restoration on the removal of 1/f-noise as present in images obtained with a scanning tunneling microscope (STM). This research was a collaboration between the Applied Crystallography Group of Prof. dr. ir. F. Tuinstra and the Pattern Recognition Group of Prof. dr. I.T. Young.

In the autumn of 1992 he spent three months as a trainee at the Gesellschaft für Schwerionenforschung (GSI) in Darmstadt, Germany, where he worked on the analysis of the decay of $^{26}$Mg.

In 1993 he started as a Ph.D. student in the Pattern Recognition Group where he worked on the application of iterative image restoration techniques on three-dimensional microscopic fluorescence images.

During this period he played an important role in the development of 3D-Pro, a multidimensional scientific image processing library.

Since January 1998 he is employed as a post-doc by Unilever Research in Vlaardingen on a joint project with the Pattern Recognition Group to analyze the microscopic network structure of water structuring systems by means of three-dimensional microscopy.
Bibliography

Journal articles


Conference proceedings


