Multi-View Deconvolution of Biomedical Images

Moritz Blume

Diploma Thesis in Computer Science

Technische Universität München
Department of Computer Science
Computer Aided Medical Procedures and Augmented Reality
in collaboration with
Siemens Corporate Research Inc., Princeton, NJ, USA

Moritz Blume

Multi-View Deconvolution of Biomedical Images

Diploma Thesis in Computer Science

Technische Universität München
Department of Computer Science

This thesis was conducted in collaboration with
Siemens Corporate Research Inc., Princeton, NJ, USA.

Author:	Moritz Blume
Director:	Prof. Dr. Nassir Navab
Supervisor:	Dipl. Inf. Darko Zikic
	Dipl. Inf. Wolfgang Wein
Date:	January 23th, 2007

Ich versichere, dass ich diese Diplomarbeit selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

Garching, den 23. Januar 2007

Moritz Blume

Abstract

Biomedical images are often degraded during image acquisition. Two very common degradations that occur in many biomedical imaging modalities are blur and noise. Corrupted images are more difficult to interpret by the physician or researcher and so the goal is to restore the original image.

Blurring can be mathematically modeled by convolution, and within this model the blur is characterized by the point-spread-function (PSF). Two kinds of blur can be distinguished: *shift-invariant* and *shift-variant* blur. Shift-invariant blur is constant over the whole image, and shift-variant blur is characterized by a PSF that is dependent on the spatial position in the image.

The process of restoring a blurred image is referred to as deconvolution in the literature. There are two major classes: *non-blind* deconvolution algorithms that assume that the PSF is known in advance and *blind* deconvolution algorithms that do not need any *a priori* knowledge regarding the PSF.

We propose a new blind deconvolution method. As an input, our method takes several observations of the same scene, each of them acquired from a different point of view, and their respective spatial transformations. The transformations are either directly known by dedicated image acquisition systems, or they can be retrieved by image registration or tracking. The output is the restored original image and PSF.

This multi-view setting occurs frequently in practice. However, most of the methods available in the literature do not specifically take advantage of the information provided by this setting.

We perform a quantitative analysis of our new method and compare it other deconvolution methods. There are two main contributions: Firstly, our algorithm is able to perform unparameterized blind deconvolution in the shift-variant case. To the best of our knowledge, known methods for blind shift-variant deconvolutions are dependent on an *a priori* blur model. Secondly, in the shift-invariant case, our algorithm allows for a better restoration at higher noise levels than comparable methods.

Acknowledgments

Many people contributed to this thesis. In the first place, I want to thank Prof. Nassir Navab who guided me to this exciting topic.

A very big thank you goes to Darko, my supervisor, who infected me with his interest on variational methods in optimization. He seemed to have an unlimited amount of time for discussions on mathematical concepts and reviewing the thesis. It was an extraordinary supervision.

Also many thanks to Wolfgang who gave me the possibility to spend almost four month of my thesis at Siemens Corporate Research in Princeton and stayed optimistic even though initial practical results were disappointing.

Thanks to Andreas Keil for various mathematical discussions in the early phase of this thesis. Thanks to Ben Glocker for gluing the title to the spine of the thesis. Thanks to Tobias Lasser for getting me back to mathematical reality when I thought that I know what I am doing.

And, of course, many thanks to Nuria...

Contents

Thesis Outline

Chapter 1: Introduction. In this thesis we deal with a multi-view setting: we suppose that several observations of the same object from different points of views are available for reconstruction. The multi-view setting is explained and motivated and a brief overview over the history of biomedical imaging is given.

Chapter 2: Mathematical Preliminaries. The basic mathematical concepts that are relevant for the rest of the thesis are presented.

Part I: Deconvolution - An Inverse Problem

Chapter 3: Classical Image Restoration. Classical image restoration is the process of restoring the original, non-degraded image based on a blurred and noisy observation and knowledge about the degradation. We review the most important methods in this field and discuss different regularization approaches.

Chapter 4: Blind Deconvolution. When no information about the degradation process is available, the task of restoring the original image from a blurred and noisy observation is denoted as blind deconvolution. A historically guided overview of literature on blind deconvolution is given.

Chapter 5: Classification of Deconvolution. The presented methods can be classified under several aspects. Among them are blind/non-blind, shift-variant/shift-invariant, two-step/iterative and frequency domain/spatial domain approaches.

Part II: A New Method for Multi-View Deconvolution

Chapter 6: Alternating Optimization for Multi-View Deconvolution. We present a mathematical modeling of the multi-view setting and use it for the definition of a cost functional that is subject to minimization. Due to a scale problem, the minimization is split into minimization of two separate cost functionals. Within this alternating optimization scheme, one of the cost functions is minimized by a numerical constrained optimization algorithm,

and the other is minimized by deriving the first-order optimality conditions using calculus of variations.

Chapter 7: Evaluation on Synthetic Data. We quantitatively evaluate our method for synthetically degraded data and compare it to another multi-view method.

Chapter 8: Application on Medical Data. Results for the reconstruction of a whole three-dimensional ultrasound dataset are shown.

1. Introduction

Images are measurements of real world quantities and thus - like all measurements - suffer from errors introduced during the acquisition process. Sources of errors are not rare. The most typical ones in biomedical imaging are blurring, noise and attenuation/occlusion. The process of recovering the original image based on one or more degraded observations is commonly known as *image restoration*. More specific denominations which refer to a special kind of degradation are *deconvolution* or *denoising*.

In this thesis, we will mainly deal with blur. Blur can complicate the detection of small features in any kind of images. There are many kinds of blur. In photography or microscopy, images are often blurred due to motion: the scene moves while the shutter is still open, which exposes one pixel of the camera's sensor to several passing spatial positions of the object of interest. Another very common kind of blur is due to defocus: when the plane of interest does not lie in the focal plane, it is blurred.

We consider a scenario where several observations of one object of interest are taken - each of them from different points of view. Figure (1.1) illustrates the general idea: the neuron cells are imaged two times from two different points of view. It can be clearly seen that all of them are degraded by blur, however, each of them differently. As described later, this scenario often occurs in biomedical applications.

We present a new algorithm that deals with this setting and is able to restore the original image from several shift-variantly blurred observations with only very little assumptions about the blurring characteristic.

There are two main contributions of our work:

1. Our algorithm allows for unparameterized *blind shift-variant* deconvolution. To the best of our knowledge, this has not yet been reported in the literature.

   Blind deconvolution algorithms only assume little *a priori* knowledge about the blurring process. The line between blind and non-blind deconvolution is not drawn strictly in the literature. Since many authors use a parameterized model of the blurring process and denote it as a blind method, we want to point out that we do not assume any blur model.

   Shift-variant methods allow for a blur characteristic that is changing with the spatial position in the image.

2. In the *shift-invariant* case, according to our experimental results, our algorithm is more robust against noise than other deconvolution methods.

In section 1.1, we briefly review some literature where the multi-view setting occurs. Deconvolution methods that can be used for this multi-view setting are discussed in section 1.2. Since our algorithm is not restricted to a specific imaging modality, section 1.3 gives an historical overview of biomedical imaging.

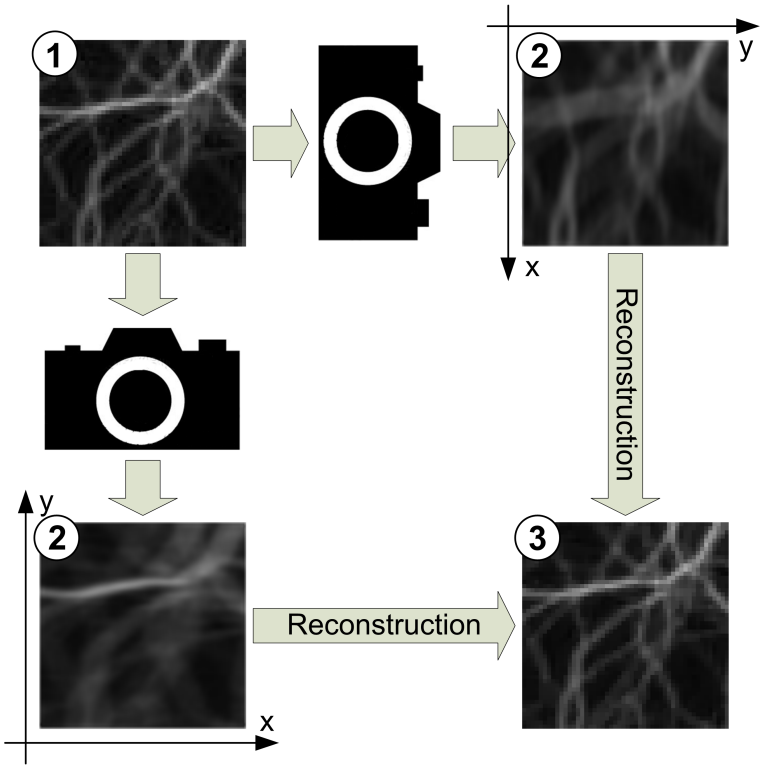


Figure 1.1.: Experimental setup: The object of interest (1) is imaged two times from different points of view. Each of the acquisitions leads to a blurred observation (2). Their coordinate systems are related by a spatial transformation - in this case a $90°$ rotation. The observed images together with the spatial transformation are the input for our new algorithm. The output is a restored image (3). As can be seen in the figure, the restoration is visually hardly distinguishable from its original counterpart.

1.1. Multi-View Imaging

The multi-view scenario frequently occurs in many three-dimensional biomedical applications.

In selective plane microscopy, a three-dimensional image of a transparent specimen is acquired with a conventional light microscope by optically sectioning it into numerous planes. Light of planes that are not in focus can be further reduced by just illuminating the in-focus plane. However, independent of the number of such planes, the axial resolution of conventional microscopes is much lower than the lateral resolution due to an anisotropic point-spread-function. In order to compensate this lack of information in the z-direction, a second, adequately rotated recording can be made. Several dedicated devices have been proposed in the literature. Shaw *et al.* were the first to develop a tilting stage and achieved an amelioration in the z direction by fusing two recordings from different angles. Kikuchi *et al.* constructed a Double-Axis Microscope, where two normal microscopes are orthogonally placed in

order to allow for a combined reconstruction [25]. Kozubek *et al.* describe specially dedicated tilting devices used in human genome research in order to acquire several views on the same object [28]. In [56], Swoger *et al.* present a Multiple Imaging Axis Microscope (MIAM) that is able to acquire more than two three-dimensional images. Within the same research group, Huisken *et al.* develope a Selective Plane Illumination Microscope (SPIM) [23]. They manually rotate the object of interest and then establish the transformation by methods of image registration.

Soler *et al.* work with three-dimensional ultrasound data [50, 51]. A one-dimensional linear array is embedded in a two-dimensional robotic platform which can linearly scan the tissue. Disposing of these slices and the corresponding spatial position, a three-dimensional image can be reconstructed. Similar to microscopy, the in-plane resolution is much better than the out-of-plane resolution. Soler *et al.* show that a second dataset can improve the resolution significantly.

1.2. Multi-View Restoration

In the following we briefly motivate the use of multiple views and then discuss the shortcomings of available multi-view restoration methods.

In image restoration, the literature distinguishes between two cases: the *non-blind* or classical image restoration problem, where the PSF is known in advance and the original image is subject to reconstruction, and the *blind* deconvolution problem, where neither the PSF nor the original image are known.

The vast majority of existing deconvolution methods are based on only one observation. Unfortunately, even for the non-blind case, deconvolution is an *ill-posed* problem and thus requires additional assumptions on the solution in order to render it *well-posed*. Rendering the problem well-posed is commonly known under the term *regularization*. A common way of regularization in deconvolution is to assume piecewise-smoothness of the original image.

The lack of information is even more severe in the blind case. There, regularization has to take into account both the original image and the PSF.

Many methods for regularization have been proposed, and we will discuss them in detail in chapter **??** and 4. Blind deconvolution works often well in low-noise scenarios. Unfortunately, some imaging modalities such as ultrasound suffer from a very high level of noise.

A natural way of regularization is to use multiple observations of the same scene. As we will show experimentally with our new method, this stabilizes the problem such that even for high noise levels a good restoration is accomplished. Furthermore, even *shift-variant* deconvolution can be accomplished.

For the multi-view setting, different methods were used for fusion of the two or more datasets.

Kikuchi *et al.* use synthetically generated projections and then apply a reconstruction method known fromComputed Tomography [26, 24]. Though this improves the quality of the reconstructed image with respect to the two input images, this is not a deconvolution method. Additionally, it assumes that the blur is *shift-invariant* and

depends on prior knowledge of its direction. So, this method is only applicable in specialized settings.

Swoger *et al.* use simple linear combination techniques like averaging [56].

Huisken *et al.* transform both images into the Fourier space and then combine them by choosing the maximum for each respective frequency [23]. Though the quality is generally improved, one has to note that also this is not a real deconvolution method. Furthermore, noise is significantly amplified.

In ultrasonography, Soler *et al.* assume a Gaussian shape of the PSF [50]. They justify their assumption based on measurements on a bubble phantom. The bubbles are assumed to be sphere shaped, and so their image gives rise to a method of estimation of the PSF. Since a theoretic model is chosen in advance, this approach is not generally applicable to other scenarios. In addition, assuming a Gaussian shape is only an approximation, and theoretical considerations suggest a different shape [38].

1.3. A Brief History of Biomedical Imaging

Biomedical imaging dates back to the end of the 17th century, when Robert Hooke published descriptions and illustrations of insects, sponges, bryozoans, foraminifera and bird feathers in his book *Microgphia*. More importantly, he was the first to state the term cell as a building block of plants. Hooke used a self-constructed compound microscope with a magnification factor of about thirty. Inspired by Hooke's work, Antony van Leeuwenhoek firstly observed bacteria, blood cells and other small biological speciem. Leeuwenhoek used a way simpler microscope with only one lens, however, able to magnify up to three hundred times [8].

Light microscopy is still one of the most used imaging modalities in todays biomedical imaging. For example, in order to study neural networks in mouse brains, neurons of trans genetic mice can be marked with the green fluorescence protein (GFP). This protein sends out fluorescence light when excited with laser light of adequate wavelength. Then, using confocal microscopy, a technique called optical sectioning allows for a three-dimensional point-by-point recording of the region of interest [40].

Also based on light, a recently developed technology is optical coherence tomography (OCT). However, unlike normal light microscopy, OCT takes advantage of light low-coherence interferometry and so is able to measure the distance of reflecting surfaces. Being a new technology, applications are seen e. g. in ophthalmology, where it allows for revealing the fine layered structure of the eye [22].

However, not only light is used for biomedical imaging. An older technology using a similar principle like OCT is ultrasonography. Instead of light waves, acoustic waves are sent into the tissue. Knowing the speed of sound, the received echo gives rise to the distance of the echoing object. Today, medical ultrasonography is widely used in almost all medical disciplines.

Only shortly after 1895, when Wilhelm Conrad Röntgen discovered a new kind of radiation, the X-rays (Nobel prize in physics, 1901), it was clear to physicians that medical surgery would benefit tremendously by this technique. In fact, the first

X-ray image made public (as shown in Figure 1.3) was the hand of the anatomist Albert von Kölliker - it was taken during Röntgen's initial lecture held at the Physical Medical Society in Würzburg [1].

Figure 1.2.: First X-ray image made public. It was the hand of the anatomist Albert von Kölliker and was taken during Röntgen's initial lecture at the Physical Medical Society in Würzburg

X-ray imaging layed the base for one of the most successful achievements in medical imaging: the computed tomography. Here, X-ray projections made around the object of interest give rise to enough information for image reconstruction. After some preliminary work in the sixties by Allan M. Cormack, Godfrey Hounsfield realized the first prototypes. The first recording of a real human being was taken in 1971. Both Cormack and Hounsfield received the Nobel Prize in medicine in 1979.

Also in the seventies, two other very important imaging modalities were brought to industrial production: Magnet Resonance Imaging (MRI) and Positron-Emission-Tomography (PET). Although PET and MR devices often look very similar to CT, the physical effects that are exploited are totally different: Magnet Resonance Imaging relies on the relaxation properties of spins of atomic nuclei when exposed to strong magnetic fields. As the relaxation time depends on the kind of tissue, MRI is especially useful to image soft tissue. PET is based on gamma radiation that is sent out by the annihilation of electrons and positrons. Since our surrounding matter is based on positive nuclei surrounded by negative electrons, if one can mark specific regions with positrons, annihilation takes place immediately. This is done by applying radionucleides (tracer) to the patient. By choosing the right tracer, one can mark the regions of interest. For example, in oncology, the most often used tracer is fluorodeoxyglucose. It is a glucose analog and taken up by all glucose consuming cells. Since cancer cells typically absorb glucose in large quantities, this radiotracer is perfectly suited in order to mark them.

2. Mathematical Preliminaries

2.1. Images, Gradients and Norms

Mathematically, can be modeled as functions

$$f : \Omega \to \mathbb{R} \tag{2.1}$$

that map from a spatial position to an intensity value. We define Ω to be a subset of a d-dimensional euclidean space: $\Omega = [0,1]^d \subset \mathbb{R}^d$. For convenience, we denote the functional space of all images by F. The variable d represents the dimension of the image, and in this thesis we consider the cases where $d = 2$ or $d = 3$.

The norm of a vectorial quantity $\mathbf{x} \in \mathbb{R}^n$ is

$$|\mathbf{x}| = \sqrt{\sum_{i=1}^{n} x_i^2} \ , \tag{2.2}$$

and similarly the norm of a continuous quantity $f \in \Omega \to \mathbb{R}$ is

$$\|f\| = \sqrt{\int_\Omega f^2(\mathbf{x}) \, d\mathbf{x}} \ . \tag{2.3}$$

We will often use the norm of an image in order to compare how similar it is to another one. For example, if we have two images u and z, the term $\|u - z\|$ gives rise to the similarity of these images. Since the square root is not convenient in differentiation, one usually prefers to use $\frac{1}{2}\|u - z\|^2$. This is called the *sum of squared differences*.

The gradient of a function f with respect to a vector $\mathbf{x} = (x_1, ..., x_n)^\top$ will be denoted by

$$\nabla_{\mathbf{x}} f = \left(\frac{\partial}{\partial x_1} f, ..., \frac{\partial}{\partial x_n} f \right)^\top \ . \tag{2.4}$$

For a one-dimensional signal, the gradient at a specific point is the slope of the function at that point. For a general function $f \in \Omega \to \mathbb{R}$, the gradient is a vector that is pointing towards steepest ascent of this function. This property is exploited by many optimization algorithms in order to find an extremal point.

We will often use the norm of the gradient $|\nabla_{\mathbf{x}} f|$. For biomedical images, the average gradient magnitude is usually relatively low. This is exploited by many image restoration algorithms for stabilization.

2.2. Convolution

Convolution is a mathematical concept that is frequently used in image processing for modeling blur. In fact, most of us might have made experiences with convolution: just remember the last time you took an image with your camera at night. Unless you used a tripod, it is pretty likely that the resulting photograph was blurred. Artists may like this effect, physicists and researchers in general do not...

The reason for blurring in this case is, that in a dark environment, the shutter of the camera has to open for a longer time in order to let enough light pass to the sensors. Depending on the degree of darkness, this may be up to seconds. If one does not keep still, or if the photographed object moves, then the same sensor element is exposed to light rays that come from the surrounding area of the initial ray. In the literature, this effect is referred to as *motion blur*.

Another kind of blur occurs when the object of interest is out of focus. The result is also a blurred image, and the literature speaks of *out-of-focus blur*. In conventional wide-field microscopy for example, not only the in-focus plane is displayed but also the out-of-focus planes above and below. The resulting observation thus is a sum of the intensities of all those regions which leads to a blurred image. This is why confocal microscopy, where only a certain point is illuminated and focused, becomes more and more popular in biomedical imaging. This increases the resolution dramatically and therefore allows for much higher magnification factors.

A way to describe convolution for a specific imaging device is to look at the image of a single point source: it is not a point anymore, but rather a point spread over the image plane. Accordingly, the image of this ideal point (which in practice does not exist) is called the point-spread-function (PSF) and commonly used in optics in order to characterize the way an imaging device convolves the original image.

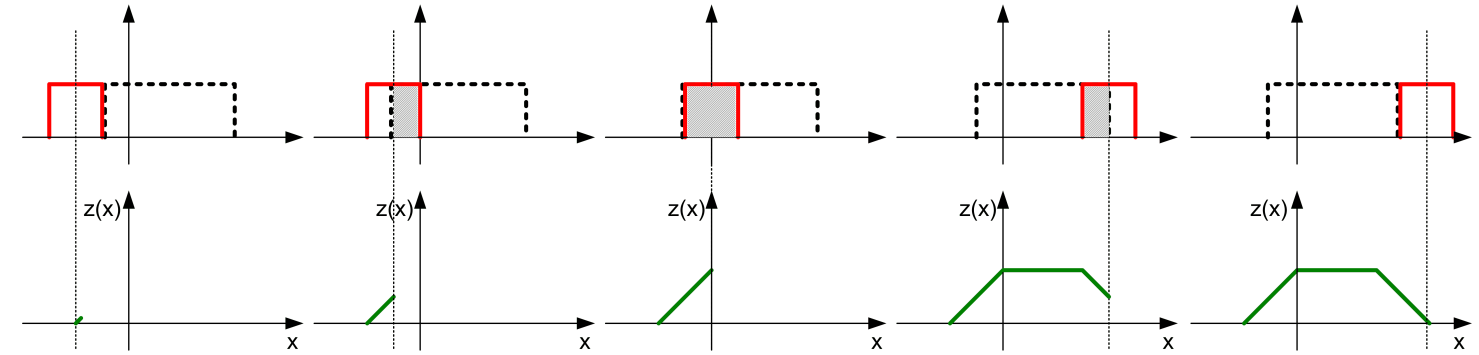
If the blurring is exactly equal for the whole recorded image, this is referred to as *shift-invariant* convolution. The other case, where the blurring depends on the spatial position in the image, is called *shift-variant*.

2.2.1. Shift-Invariant Convolution

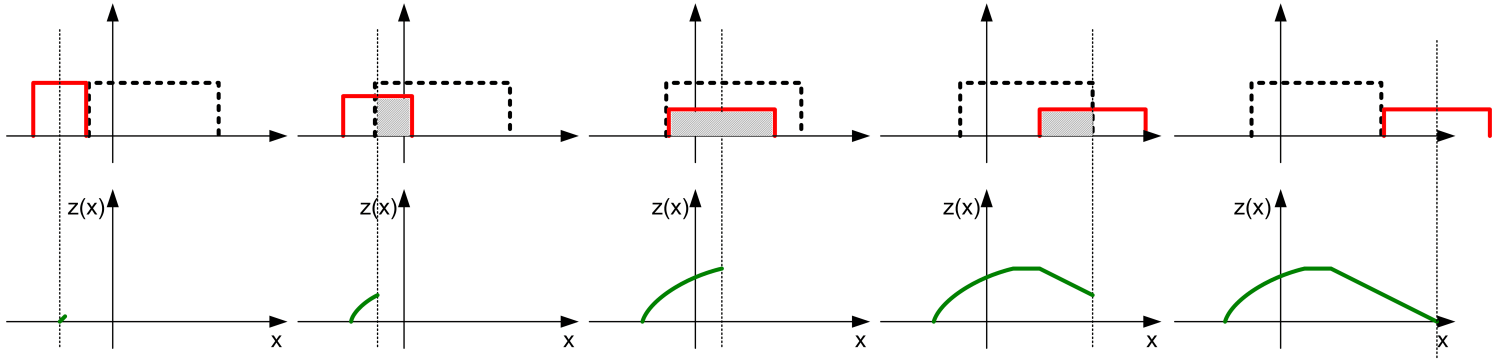
Mathematically, convolution can be described by the following equation:

$$z(\mathbf{x}) = (u * h)(\mathbf{x}) = \int_{\Omega} u(\mathbf{x} - \xi) \cdot h(\xi)\, d\xi \ . \tag{2.5}$$

The original image u is convolved with the *point-spread-function* (PSF) h (also known as the *filter* or the *kernel*) and the result is the convolved image z. Figure 2.1a illustrates this for a one-dimensional signal, and figure 2.2 shows how convolution acts on an image.

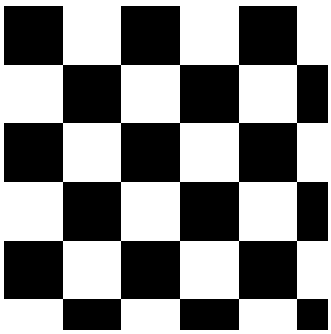


(a) Shift-invariant convolution

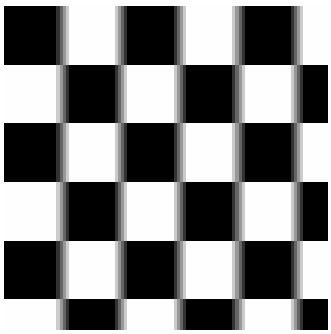


(b) Shift-variant convolution: The shape of h is changing with the spatial position.

Figure 2.1.: The original one-dimensional signal u (black dashed line) is convolved with the kernel h (red non-dashed line). The result is the convolved signal z. The sequences show how h is shifted over u. The value of z at a spatial position x is given by the overlapping area of u and h when h is shifted to that position x. This overlapping area corresponds to the integral in equation (2.5) and (2.7).



(a) Original image (99×99)



(b) Convolved image (same size as original image)

(c) Point-spread-function (9×9)

Figure 2.2.: Shift-Invariant Convolution

Typically, the PSF has values different from zero only for a very small area around $\xi = 0$. That means, that z at the position x is composed out of a linear combination of values of u around that position.

A very important theorem for deconvolution is the Convolution theorem

$$u * h = U \cdot H \quad , \tag{2.6}$$

where U is the Fourier transformed of u and H the Fourier transformed of h.

2.2.2. Shift-Variant Convolution

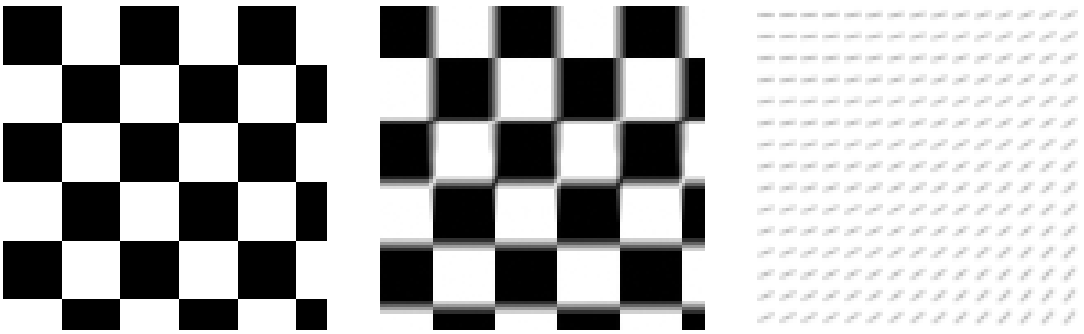
As mentioned, the above description only reflects a special case. A more general way is to let the PSF h depend on the spatial position. So, the original image u is convolved with a filter h that can be different for each spatial position in the image. The following equation reflects this idea:

$$z(\mathbf{x}) = u \star h = \int_{\Omega} u(\mathbf{x} - \xi) \cdot h(\mathbf{x}, \xi)\, d\xi \quad . \tag{2.7}$$

Figure 2.1b illustrates this for a one-dimensional signal, and figure 2.3 shows how a shift-variant rotating Gaussian distribution shape PSF acts on the checkerboard image.

2.3. Discretization

Throughout this thesis, discrete versions of two-dimensional images will be assumed to be of size $m \times n$. An image f will be represented by a vector $\mathbf{f}$ of size mn. Linearization is done in lexicographic order with respect to the function variables. The discrete PSF h will be represented by a vector $\mathbf{h}$ of size $mnop$, where $o \ll m$ and



(a) Original image (99×99)

(b) Convolved image (same size as original image)

(c) Shift-variant PSF: each pixel has its own 5×5 kernel. For the sake of visualization, only a small part is displayed.

Figure 2.3.: Shift-variant convolution

$p \ll n$ define the support of each filter. Unsurprisingly, this linearization is also performed in lexicographic order.

In discrete terms, the convolution - being a linear operation - can be expressed by a matrix-vector product $\mathbf{Hu}$ or $\mathbf{Uh}$. Since we assume that the discretized version of a PSF at a single image point is much smaller than the discretized image, $\mathbf{H}$ is an $mn \times mn$ sparse matrix. $\mathbf{U}$ is of size $mn \times mnop$ and is also highly sparse.

2.4. Optimization

It is beyond the scope of this thesis to give a well-founded introduction to optimization, as it can be found in [55, 39] or [49]. However, we will at least try to introduce to the principals that are necessary for the understanding of the rest of the thesis.

The basic goal of optimization is to find a vector $\mathbf{x}$ such that a cost function (also called the objective function) $C : \mathbb{R}^n \to \mathbb{R}$ has a minimal/maximal value $C(\mathbf{x})$ at $\mathbf{x}$. In the following, we will restrict our considerations without loss of generality to the problem of minimization of C, since maximization of C can be stated as the minimization of $\tilde{C}(\mathbf{x}) = -C(\mathbf{x})$.

It often happens that the cost function does not depend on a vector but on another function and its derivative:

$$\mathcal{C} : \mathbb{R} \times (\mathbb{R} \mapsto \mathbb{R}) \times (\mathbb{R} \mapsto \mathbb{R}) \mapsto \mathbb{R} \tag{2.8}$$

$$\mathcal{C}[x, f, f'] = \int_a^b F(x, f, f')\, dx \ . \tag{2.9}$$

In this case, $\mathcal{C}$ is called a *functional*. Then, similar to above, the goal is to find the function f that minimizes the functional $\mathcal{C}[x, f, f']$.

The space where solutions are to be found can be restricted to a certain domain. If this is the case, one speaks of constraints that are imposed on possible solutions $\mathbf{x}$. All $\mathbf{x}$ that fulfill the constraints define the *feasible region*.

2.4.1. Optimization of Functions

If the objective is a function of one variable, one can analyze its first and second derivative in order to get candidates for global extrema. Points where the first derivative is zero have a tangent with slope zero, and the second derivative gives rise to whether this is a local minimum or maximum or plateau. The first derivative is defined by

$$\frac{d}{dx}C(x) = \lim_{\alpha \to 0} \frac{C(x+\alpha) - C(x)}{\alpha} \quad . \tag{2.10}$$

In many practical cases, the objective function C is not known analytically. It has to be considered as a black box that returns its values for a given parameter vector if it is asked to do so. Many different kinds of numerical algorithms were developed in order to find solutions for such kind of problems. For a recent overview the interested reader is referred to [17].

Not having an analytical description of the cost function makes it more difficult - if not impossible - to find a global minimum, since it is not possible to evaluate C over the whole infinite parameter space.

Commonly used optimization methods can only guarantee that a solution is a global solution if C fulfills certain criteria. A popular class of objective functions are *convex* functions [7]. In this case, a local optimum is also a global optimum.

For cost functions whose parameter is a vector of several variables, the directional derivative gives rise to optima:

$$d_{\mathbf{x}}C(\mathbf{x};\boldsymbol{\nu}) = \lim_{\alpha \to 0} \frac{C(\mathbf{x}+\alpha\boldsymbol{\nu}) - C(\mathbf{x})}{\alpha} \quad ,|\boldsymbol{\nu}| = 1 \quad .$$

For $C \in \mathbb{R}^2 \mapsto \mathbb{R}$, the directional derivative is the slope of the function in direction $\boldsymbol{\nu}$. Similar to the one-dimensional case, a necessary condition for a minimum at a certain point is, that the directional derivative is zero at that point for all directions.

In the following, we will show a connection between the directional derivative and the gradient of a function. The main reason why we do this is to show an analogy between the directional derivative and the Gâteaux derivative in the next section. We first have to note that

$$d_{\mathbf{x}}C(\mathbf{x};\boldsymbol{\nu}) = \frac{d}{d\varepsilon}C(\mathbf{x}+\varepsilon\boldsymbol{\nu})\bigg|_{\varepsilon=0} \quad . \tag{2.11}$$

In words: The directional derivative at point $\mathbf{x}$ in direction $\boldsymbol{\nu}$ is equal to the scalar derivative at point $\epsilon = 0$ of the modified cost function $C(\mathbf{x} + \varepsilon\boldsymbol{\nu})$. The proof is very simple and shown in Appendix A.

Now, since we boiled the directional derivative down to a normal scalar derivative, we are able to apply all the rules we know from scalar derivation. With the help of the chain rule we get:

$$d_{\mathbf{x}}C(\mathbf{x};\boldsymbol{\nu}) = \frac{d}{d\varepsilon}C(\mathbf{x}+\varepsilon\boldsymbol{\nu})\Big|_{\varepsilon=0} \tag{2.12}$$

$$= \frac{\partial}{\partial\mathbf{x}}C\frac{d}{d\varepsilon}(\mathbf{x}+\varepsilon\boldsymbol{\nu})\Big|_{\varepsilon=0} \tag{2.13}$$

$$= (\nabla_{\mathbf{x}}C)^{\top}\boldsymbol{\nu} \tag{2.14}$$

2.4.2. Optimization of Functionals

We will now derive the first order optimality conditions for a functional

$$\mathcal{C}[x,f,f'] = \int_a^b F(x,f,f')dx \ . \tag{2.15}$$

The functional $\mathcal{C}$ depends on a scalar x, a function f and its first derivative f' (for the sake of an easier derivation we chose a scalar x rather than a vector).

As mentioned, the goal is to find a function $f_{\text{opt}} \in \mathbb{R} \mapsto \mathbb{R}$, such that the functional $\mathcal{C}$ has a minimal value at f_{opt}. This means that $\mathcal{C}[x, f_{\text{opt}}+\eta, f'_{\text{opt}}+\eta'] \geq \mathcal{C}[f_{\text{opt}}]$ for all perturbations $\eta \in \mathbb{R} \mapsto \mathbb{R}$. If we consider these perturbations η as a direction in the functional space, we arrive at a sort of directional derivative, called the Gâteaux derivative:

$$\delta_f\mathcal{C}[x,f,f';\eta] = \lim_{\alpha\to 0}\frac{\mathcal{C}[x,f+\alpha\eta,f'+\alpha\eta'] - \mathcal{C}[x,f,f']}{\alpha} \ . \tag{2.16}$$

Again, the goal is to find a tractable way in order to deal with this derivative. As for the directional derivative, we first note that (see Appendix A)

$$\delta_f\mathcal{C}[x,f,f';\eta] = \frac{d}{d\varepsilon}\mathcal{C}[x,f+\varepsilon\eta,f'+\varepsilon\eta'] \ . \tag{2.17}$$

Unfortunately, the rest of the derivation is not as easy as in the case of the directional derivative. In the 1750's, Leonhard Euler and Joseph-Louis Lagrange independently found a necessary condition that a solution has to fulfill [14]: the function f that minimizes $\mathcal{C}$ is found by solving the partial differential equation

$$\frac{\partial F}{\partial f} - \frac{d}{dx}\frac{\partial F}{\partial f'} = 0 \ . \tag{2.18}$$

The full derivation is given in Appendix B. It is very important to note that a solution to this equation is not necessarily a global optimum. Additional knowledge such that

the function is convex is necessary in order to make sure that the sought function f really optimizes the cost functional.

For the higher-dimensional case, where the objective function is

$$\mathcal{C} : \mathbb{R}^d \times (\mathbb{R}^d \mapsto \mathbb{R}) \times (\mathbb{R}^d \mapsto \mathbb{R}) \mapsto \mathbb{R} \quad (2.19)$$

$$\mathcal{C}\left[x_1, ..., x_n, f, \frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right] = \int_a^b F\left(x_1, ..., x_n, f, \frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right) dx_1 ... dx_n \ , \quad (2.20)$$

we get

$$\frac{\partial F}{\partial f} - \sum_{k=1}^{n} \frac{\partial}{\partial x_k} \frac{\partial F}{\frac{\partial f}{\partial x_k}} \ . \quad (2.21)$$

We will not give the derivation, since the general idea is the same as in the one-dimensional case. The interested reader is referred to [49].

Part I.

Deconvolution - An Inverse Problem

3. Classical Image Restoration

The goal of image restoration is to find the original, non-degraded image. This is an inverse problem and is typically modeled as follows (compare e. g. [16]):

$$z(\mathbf{x}) = \int_{\Omega} u(\mathbf{x} - \xi) \cdot h(\mathbf{x}, \xi)\, d\xi + \eta(\mathbf{x}) \ . \tag{3.1}$$

As before, z is the observed image, u is the original image which is convolved with a point-spread-function h. Additionally, usually noise is disturbing the observation which is modeled by the noise term η. If u, h and η are given and z is sought this would be a direct problem. However, the usual task is to restore the original image u when the rest of the variables is given or when only z is given. This is the inverse problem. A very important special case of (3.1) is the shift-invariant or *stationary* model:

$$z(\mathbf{x}) = \int_{\Omega} u(\mathbf{x} - \xi) \cdot h(\xi)\, d\xi + \eta(\mathbf{x}) \ . \tag{3.2}$$

Unfortunately, it is well known that the reverse solution of Fredholm integral equations of first kind like (3.1) or (3.2) is ill-posed in the Hadamard sense (see e. g. [34]). This is due to the fact that the inverse operator is not continuous and therefore small perturbations in the measured signal z can come from extremely large perturbations in the solution u. As mentioned, ill-posedness can be tackled by regularization. According to the historic time line, we will now first review some classical approaches, and then look at ways of rendering the restoration problem well-posed.

3.1. Inverse Problems

An inverse problem is one where the *effect* of a *system* is known, and, based on the known effect, its *cause* has to be concluded. This is called inverse problem since with respect to points in time, the effect of a system always comes after the cause. If one wants to conclude the effect from the cause, this is called a forward problem or a direct problem.

Let us look at a demonstrative example: A pirate ship that operates in the costal regions of the Maghreb shoots a canon in direction of an argosy. The cannonball almost hits the argosy. We suppose that the pirates are quite far away and windy weather is invoking huge waves, such that for the argosy crew it is not possible to see the pirates ship at any time. From the point of view of the pirates, the problem is to hit the argosy which they saw for some short moments. From the point of view

of the argosy crew, the problem is to determine from where the cannonball came in order to be able to shoot back properly.

The problem can be modeled as followed: A particle of mass m starts from a spatial position $\mathbf{x}$ with a velocity vector $\mathbf{v}$ and ends at a spatial position $\mathbf{x}'$ having a velocity $\mathbf{v}'$. The velocity vector encapsulates both direction and amount of the velocity at a certain point of time.

The pirates know both $\mathbf{x}$ and $\mathbf{x}'$ and have to calculate $\mathbf{v}$ in order to align the cannon properly. The cause of this system would be the initial position and the initial velocity vector. The effect is the final spatial position and the final velocity of the cannonball. Therefore, the pirates have to solve an inverse problem: they have to calculate the cause out of the intended/observed effect.

Having estimated $\mathbf{x}'$ properly, making some simplifications (e. g. that friction can be neglected) and knowing the cannon properties this is not a difficult task for the pirates. Now, fortunately for the argosy crew, the pirates did not well enough in order to estimate their position $\mathbf{x}'$, and so the cannonball makes an impact only close to the argosy.

Also the argosy crew has to solve an inverse problem: knowing the position of impact $\mathbf{x}'$, they have to solve for the initial position $\mathbf{x}$. If nothing else than the impact position is known, this is an *ill-posed* problem, since there exists an infinite number of solutions: for any initial position $\mathbf{x}$ one could find a velocity vector $\mathbf{v}$ such that the cannonball hits the target in $\mathbf{x}'$.

Inverse problems are often ill-posed. Hadamard [19] proposed the following three conditions for well-posed problems:

1. A solution exists
2. The solution is unique
3. The solution depends continuously on the data, in some reasonable topology.

Accordingly, ill-posed problems are characterized as those that are not well-posed. In order to be solvable by an numerical algorithm, they have to be restated in a certain manner. This is done by adding further information about the solution. In the literature, this is referred to as *regularization*.

Coming back to our example, the argosy crew could estimate the cannonball's velocity vector at the time of impact (e. g. by inspecting the water bubbles that it leaves behind). Then, at least the cannonball's trajectory could be determined. By further assuming that the cannonball was shot at approximately sea level height, the problem is rendered well-posed.

3.2. Classical Approaches

As mentioned, image restoration is an inverse problem. The system is the imaging device, the effect is the observed image and the cause is the original image. One knows the effect of the system and the goal is to find its cause. This is an ill-posed

problem and in the following sections we will see how early research on deconvolution deals with the issue.

3.2.1. Inverse Filter

The most naive approach for tackling deconvolution is to construct the inverse filter. This is easily done when looking at the Fourier transformation of equation (3.2) and disregarding the noise term. Together with the Convolution theorem we yield

$$Z(\omega) = U(\omega) \cdot H(\omega) \ , \tag{3.3}$$

where the Z, U and H are the Fourier transformations of z, u and h, and ω is a frequency value. Then, the original image can be reconstructed if the degradation kernel is known:

$$U(\omega) = \frac{Z(\omega)}{H(\omega)} \ . \tag{3.4}$$

There are two problems associated with this approach. Firstly, the inverse filter does not exist for all kernels [6]. In fact, the most important scenarios like motion-blur or out-of-focus blur do not have an inverse filter.

Secondly, noise is amplified quite heavily. This can be seen when the inverse filter is applied to the more realistic noisy scenario:

$$\begin{aligned} U(\omega) &= \frac{Z(\omega)}{H(\omega)} && (3.5)\\ &= \frac{U(\omega) \cdot H(\omega) + N(\omega)}{H(\omega)} && (3.6)\\ &= U(\omega) + \frac{N(\omega)}{H(\omega)} \ . && (3.7) \end{aligned}$$

Since the PSF is usually a low-pass filter, H becomes very small for high frequencies and accordingly its inverse is very large. Noise typically also has high magnitudes for high frequencies, and so it is amplified dramatically.

3.2.2. Wiener Deconvolution

The Wiener filter tackles the problem of high noise amplification. Let us suppose that we dispose of a restoration filter W. The error ε between the reconstructed image using the filter W and the original image is

$$\begin{aligned} \varepsilon(\omega) &= [Z(\omega) \cdot W(\omega)] - U(\omega) && (3.8)\\ &= [(U(\omega) \cdot H(\omega) + N(\omega)) \cdot W(\omega)] - U(\omega) \ . && (3.9) \end{aligned}$$

Now, the idea is to find the filter W that minimizes the *magnitude* of the *squared* error. Since W is a function, the magnitude is a functional, and in order to find a function that minimizes a functional, we have to make use of calculus of variations. However, at this point we do not want to enter into mathematical details and therefore just give the final solution according to [16]:

$$W(\omega) = \frac{H^*(\omega) \cdot U(\omega)}{|H(\omega)|^2 \cdot U(\omega) + N(\omega)} \ . \tag{3.10}$$

This construction relies on already knowing the frequency spectrum of u and thus is not directly applicable. In practice, Wiener Deconvolution works pretty well for a number of scenarios where H is supposed to be known in advance and the spectrum of U and N can be approximated properly.

3.2.3. Landweber Iteration

The Landweber iteration [32] (also known as Van Cittert's [11] or Bially [5] iteration) is based on the trivial identity

$$u = u + \alpha \left(z - h * u\right) \ . \tag{3.11}$$

Then, the following discretized iterative scheme is suggested:

$$\mathbf{u}_0 = \alpha \, \mathbf{z} \tag{3.12}$$

$$\mathbf{u}_{i+1} = \mathbf{u}_i + \alpha \left(\mathbf{z} - \mathbf{H}\mathbf{u}_i\right) \tag{3.13}$$

$$= \alpha \, \mathbf{z} + (\mathbf{I} - \mathbf{H})\mathbf{u}_i \ . \tag{3.14}$$

For any iterative scheme it is interesting to know if it does converge and if yes which is the limiting solution. It was found that

$$\mathbf{u}_\infty = \mathbf{H}^{-1}\mathbf{z} \ . \tag{3.15}$$

In other words, the limiting solution is found by applying the inverse filter to the blurred observation. However, an advantage of the Landweber iteration over the inverse filter solution is that the typical noise amplification can be avoided if the iteration is not continued until convergence. Another advantage is that the inverse filter itself does not have to be implemented.

Further eigenvalue analysis leads to a better understanding why an early cancellation of the iteration is often a good choice [32].

3.3. Regularization

As mentioned, the problem of deconvolution is ill-posed. In order to get a physically meaningful solution, one has to impose constraints on the set of possible solutions. Some of the classical approaches automatically incorporate some kind of regularization, e. g. the number of iterations in the Landweber approach. However, it was not before the sixties that mathematicians consciously addressed this issue and found general ways to render ill-posed problems well-posed.

3.3.1. Tikhonov-Phillips Regularization

A classical method for regularization was found in the sixties independently by Tikhonov and Phillips [41, 59]. Instead of seeking the perfect solution to (3.1), the idea is to construct a cost function that on the one hand encourages good reconstructions, and on the other hand penalizes solutions that are physically not meaningful. The cost function that is subject to minimization is

$$\|z - u \star h\|^2 \, d\mathbf{x} + \alpha \, J(u) \ . \tag{3.16}$$

Under the variety of possible solutions to (3.1), those u that minimize $J(u)$ are preferred. This renders the problem well-posed since now the problem is to solve a Fredholm integral equation of second kind and it is known that there exists a unique solution [34].

Typical choices for $J(u)$ are

$$J(u) = \int_\Omega u^2(\mathbf{x}) \, d\mathbf{x} \tag{3.17}$$

or

$$J(u) = \int_\Omega |\nabla u(\mathbf{x})|^2 \, d\mathbf{x} \ . \tag{3.18}$$

However, both of them are not adequate for image restoration since (3.17) produces spurious oscillations or ringing artifacts and (3.18) encourages smooth images and disregards edges.

3.3.2. Total-Variation Regularization

A better way of regularization for image restoration was found by Rudhin and Osher in their seminal paper on image noise removal algorithms [46]. In order to conserve edges while suppressing noise the so called total variation (TV) norm (a L_1 norm on the gradient) is used:

$$J_{\mathrm{TV}} = \int_\Omega |\nabla u(\mathbf{x})| \, d\mathbf{x} \ . \tag{3.19}$$

Unfortunately, as will be seen later, this makes the numerical solution much more difficult.

In variational frameworks, (3.19) often has to be derived with respect to u and so the problem arises that the absolute value is not differentiable at zero. In order to circumvent this problem, Vogel and Oman propose a slight modification of (3.19):

$$J_{\mathrm{TV}} = \int_{\Omega} \sqrt{|\nabla u(\mathbf{x})|^2 + \beta^2}\, d\mathbf{x} \quad . \tag{3.20}$$

For very small values of β this term approximates (3.19) but is still differentiable at zero.

4. Blind Deconvolution

Up to now, we only considered the problem of restoring the original image u by assuming that we know h. Though this is a very important problem, in practical applications the point-spread-function is often not directly available in advance. In astronomy for example, the PSF depends on atmospheric perturbations that are subject to change over time.

Even for cases where the PSF can be measured it is often better or more convenient if no knowledge of h has to be known in advance in order to apply the restoration:

- Measurements are error prone and image restoration is very sensitive to measurement errors.
- Measurements can be time consuming.
- Small parameter changes in the imaging system can change the PSF, which would make it necessary to repeatedly measure it again and again.

So, it is very appealing to be able to retrieve u from a convolved observation z without disposing of any knowledge about h. The scientific community recognized demand for such methods in the late sixties and aptly named this new class *blind deconvolution*.

Furthermore, theoretical predictions often depend on parameters which have to be determined (e. g. for motion blur or defocus blur). If it is known in advance that a theoretical model applies, blind deconvolution algorithms can incorporate this knowledge for restoration.

In fact, the question arises where to draw the line between blind and non-blind deconvolution. In the literature this line is not strictly drawn. Many algorithms denoted as "blind" make to some extent use of previous knowledge. Almost always it is assumed that both the image and the PSF are nonzero. Even algorithms where a theoretical model is assumed to be known are often denoted as "blind". In these cases, some authors prefer to use a more confined term such as *semi-blind deconvolution*.

Very good reviews over these approaches can be found in [6, 29] and [3]. Each of them uses a different classification in order to organize the available methods. However, many reasonable ways of classification are imaginable, and we feel that it is difficult to structure our introduction on the basis of a rigorous classification. To this end, we present the available literature ordered by a historic time line and only distinguish between deconvolution methods that restore shift-invariantly blurred images and deconvolution methods that restore shift-variantly blurred methods. However, a brief overview of possible classifications of blind deconvolution methods is given in the next chapter.

4.1. Shift-Invariant Blind Deconvolution

4.1.1. Frequency Domain Zeros

This class of approaches goes back to Slepian [48]. We look at the frequency domain representation

$$Z(\omega_1, \omega_2) = U(\omega_1, \omega_2) \cdot H(\omega_1, \omega_2) \tag{4.1}$$

of (3.2). The general idea is pretty simple: when $U(\omega_1, \omega_2)$ or $H(\omega_1, \omega_2)$ is zero, this implies that $Z(\omega_1, \omega_2)$ is also zero. If additionally a blur model for H is known which just depends on some few parameters, then these parameters can eventually be determined by looking at the zeros of Z.

This is the case for example for motion blur and out-of-focus blur. Horizontal motion blur can be modeled by

$$h(x, y) = \begin{cases} \frac{1}{2d} & -d \leq x \leq d \ \wedge \ y = 0, \\ 0 & \text{otherwise} \end{cases} . \tag{4.2}$$

The zeros of the Fourier transformed of h, namely H, are lying on lines perpendicular to the direction of motion blur and with constant distance of $1/d$.

For out-of-focus blur, the model usually used is

$$h(x, y) = \begin{cases} \frac{1}{R^2\pi} & \sqrt{x^2 + y^2} \leq R, \\ 0 & \text{otherwise} \end{cases} , \tag{4.3}$$

and the zeros occur on almost periodically appearing concentric circles around the origin.

By finding the parameter such that the zeros of H fit best with the zeros of Z, the point-spread-function can be determined and the problem reduces to the well studied classical image restoration problem where the filter is known.

The next step dates back to Sondi in 1972 [52]. Instead of looking at the Fourier transform of (3.2), he considers the cepstrum equation

$$C_z(p, q) = C_u(p, q) + C_h(p, q) \ . \tag{4.4}$$

The cepstrum is defined as the Fourier transform of the logarithm of the Fourier transform of a signal. The periodic or almost periodic patterns of motion blur and defocus blur models of H lead to large negative spikes in the ceptrum of h, namely C_h. Their pattern gives information for a determination of h.

However, Stockham and Cannon [53, 9] found that this kind of simple analysis only works for synthetically degraded data and is not robust when noisy real-life data is used. In order to make things more stable, they propose a subdivision of

the measured image into smaller, possibly overlapping rectangles. For each of those rectangles

$$Z_i(\omega_1, \omega_2) \approx U_i(\omega_1, \omega_2) \cdot H(\omega_1, \omega_2) \tag{4.5}$$

holds approximately. This is due to convolutional border effects. However, the approximation is accurate enough if the windows are large compared to the support of the point-spread-function. By averaging over the square of the magnitude of the Fourier transforms, a quite robust estimate for the power spectrum is received. This is mostly due to the fact that U_i varies for each of those windows while the flavor of H is retained in each of them. Now, looking at the ceptra of the power spectrum reveals much clearer spikes and therefore allows for a robust estimation of H and finally h.

4.1.2. Zero Sheet Separation

A fundamental contribution to the problem of blind deconvolution was made by Lane and Bates in 1987 [33]. They prove that blind deconvolution, while being theoretically impossible for one-dimensional signals, is achievable and even overdetermined for two or more dimensions. Their approach is not only very interesting due to its theoretic results, but also lead to several automatic algorithms that practically perform blind deconvolution. In the following, we will spend some space in order to illustrate the basic idea of an algorithm by Ghiglia and Romero [15], which, among these approaches, received the most attention in the literature.

Ignoring noise, the Fourier transformed of (3.2) is denoted by

$$Z(\omega_1, \omega_2) = U(\omega_1, \omega_2) \cdot H(\omega_1, \omega_2) \ . \tag{4.6}$$

Z is a two-dimensional polynomial that can be written as

$$Z(\omega_1, \omega_2) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} z(m, n)\, e^{-i2\pi\omega_1 m/M} e^{-i2\pi\omega_2 n/N} \tag{4.7}$$

By simply substituting $e^{-i2\pi\omega_1/M}$ by u and $e^{-i2\pi\omega_2/N}$ by v this shortens to

$$Z(u, v) = \sum_{n=0}^{N} \sum_{m=0}^{M} z(m, n)\, u^m v^n \ . \tag{4.8}$$

For one-dimensional polynomials, the fundamental theorem of algebra guarantees that a factorization is always possible. For two or more dimensions, a factorization is not generally possible. However, if it is possible, it is almost always unique. Thus, the two-dimensional blind deconvolution problem is equivalent to finding the factorization (the zero sheets) of a two-dimensional polynomial.

Unfortunately, finding the zero sheets of two-dimensional polynomials is not that easy. In fact, doing so is the most important step in the algorithm proposed by Ghiglia and Romero. Once this factorization is accomplished, U and H are known and simply applying the inverse Fourier transform leads to the image u and the PSF h.

Let us start by rewriting (4.8) as

$$Z(u,v) = \sum_{n=0}^{N-1} a_n(u) v^n \ , \tag{4.9}$$

where

$$a_n(u) = \sum_{m=0}^{M-1} z(m,n)\, u^m \ . \tag{4.10}$$

If we now fix the value of u, we dispose of a one-dimensional polynomial of degree N that only depends on the image data $z(m,n)$. Of course, this polynomial has N (complex) roots and any conventional root-finding algorithm can be used in order to find them. We fix the value of u to u_0, and the roots of (4.9) for this value are denoted by $v_1, v_2, ..., v_N$. Now, since we know that Z can be factored into U and H, the task is to assign the roots to either U or H. To this end, we define two sets S_1 and S_2. We start by arbitrarily putting v_1 into S_1. The goal is to classify the remaining v_i into either S_1 or S_2. However, we do not know yet if the roots assigned to S_1 belong to U or H. Once this part is accomplished, we will see later how U and H, and therefore also u and h can be derived.

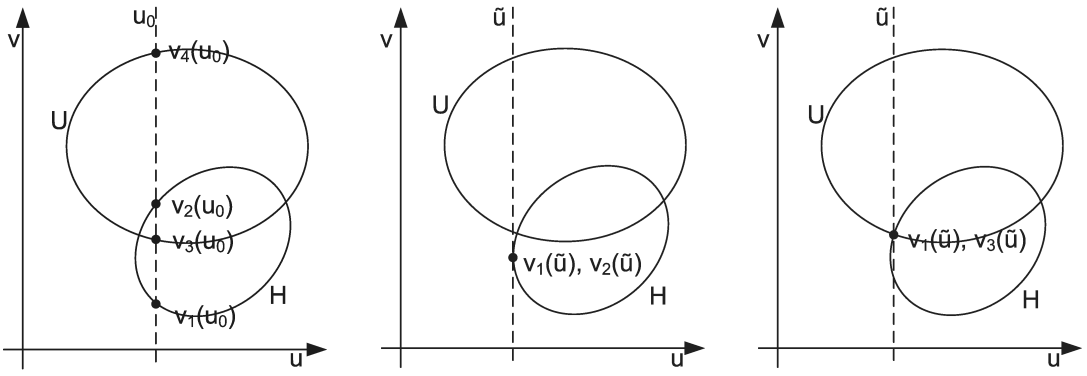
Figure 4.1 illustrates the classification procedure. For simplicity, the zero contours of U and H are supposed to be real values and not complex. We now want to know if v_i belongs to the same zero sheet as v_1 or not. Let us suppose that we dispose of a numerical procedure such that we can shift v_1 and v_i on a path until they come together at the point $u = \tilde{u}$ and $v = \tilde{v}$. This is not a trivial task and details are described in the referenced paper.

Knowing this point of coincidence, the following reflections will give the answer to the question whether v_i belongs to S_1 or S_2. We look at the case where both belong to the same zero sheet. Without loss of generality this shall be H, as displayed in Figure 4.1b. Then, the product rule gives

$$\frac{\partial}{\partial u} Z(\tilde{u},\tilde{v}) = \frac{\partial}{\partial u} U(\tilde{u},\tilde{v}) H(\tilde{u},\tilde{v}) + U(\tilde{u},\tilde{v}) \frac{\partial}{\partial u} H(\tilde{u},\tilde{v}) \tag{4.11}$$

and therefore $\frac{\partial}{\partial u} Z(\tilde{u},\tilde{v})$ is not zero in general. For the other case where v_1 and v_i belong to a different zero sheet (as illustrated in Figure 4.1c), $\frac{\partial}{\partial u} Z(\tilde{u},\tilde{v})$ is obviously always zero.

Ghiglia and Romero were not able to prove that $\frac{\partial}{\partial u} Z(\tilde{u},\tilde{v}) = 0$ implies that v_1 and v_i belong to different zero sheets under certain assumptions. In fact, without any



(a) Zeros of the one-dimensional polynomial $v(u_0)$.

(b) v_1 and v_2 belong to the same zero sheet.

(c) v_1 and v_3 belong to a different zero sheet

Figure 4.1.: The zeros of a the one-dimensional polynomial $v(u_0)$ are shifted until they merge to a single point. Then, it can be established whether they belong to the same polynomial or not.

assumptions, it is possible to construct two polynomials where this is not the case. However, Ghiglia and Romero argue that in the case of image deconvolution two such polynomials do not exist and hence this implication is true in general.

Now we dispose of the set S_1 that contains all $v_i(u)$ that belong to one of the two factors (either U or H) and S_2 that contains all $v_i(u)$ that belong to the other factor. Conceptually, this is the most difficult part. The remaining task is to use this information in order to find U and H. To this end, the same classification procedure is applied by exchanging the rules of u and v, which leads to $u_i(v)$ that belong to either a set T_1 or T_2. From the fundamental theorem of algebra we know that every one-dimensional polynomial can be expressed by its roots multiplied with a remaining arbitrary factor:

$$U(u,v) = A(u)\, Q_1(u,v) = A(u) \prod_{\{i|v_i \in S_1\}} (v - v_i(u)) \ . \tag{4.12}$$

Note that v_i depends on u. We suppose that by moving u beginning with $u = 1$ over the unit circle in the complex plane, the classification can be maintained with less effort than a complete reclassification for every u (e. g. by using sufficiently small steps). Again, this is not an easy procedure and the reader is referred to the original paper for further details.

If we supposed that the set T_1 is associated with U we can also write

$$U(u,v) = B(v)\, P_1(u,v) = B(v) \prod_{\{i|u_i \in T_1\}} (u - u_i(u)) \ . \tag{4.13}$$

Putting both equations together leads to

$$P_1(u,v) = \frac{B(v)}{A(u)} Q_1(u,v) \ . \tag{4.14}$$

$A(u)$ can be arbitrarily set to 1 for a certain u since we are only determining the polynomials up to a certain constant, and so B and finally U is sought. In practice a more robust numerical scheme that takes advantage of the fact that (4.14) must hold for every u it is applied to.
It remains to check if T_1 is really associated with U. This is done by evaluating (4.14) for different values of u and comparing the results for B. If B is equal up to an arbitrary scale factor, then the assumption is true. If not, the same steps have to be applied based on the assumption that T_1 is associated with H.
The same procedure is applied in order to seek H.

Discussion. Although this method gives insight into the theoretical concept of blind deconvolution, it has several drawbacks that limit practical applicability. Firstly, the method is very sensitive to noise. Secondly, it has a high complexity which makes it difficult to deal with growing datasets. Finally, the algorithm is prone to numerical instability for large datasets.

4.1.3. Autoregressive Moving Average

In the Autoregressive Moving Average (ARMA) approaches, the original image is modeled as a autoregressive process

$$u(x,y) = \sum_{l,m \in R_a/(0,0)} a(l,m) u(x-l, y-m) + \nu(x,y) \ . \tag{4.15}$$

a is the set of autoregressive parameters and ν is an error term. R_a is the support of the signal a and determines the causality of the random process. Most often a so-called nonsymmetric half-plane (NSHP) causality is used in order to model the process. Experience shows that for this causality, only three parameters (a(0, 1), a(1, 0) and a(1, 1)) are usually sufficient. This autoregressive model is valid in situations where u is smooth and homogeneous, such as in photography. Equation (4.15) can be rewritten in matrix vector notation by representing the two-dimensional functions as vectors:

$$\mathbf{u} = \mathbf{A}\mathbf{u} + \boldsymbol{\nu} \ . \tag{4.16}$$

The degraded and observed image z is defined as usual:

$$z(x,y) = \sum_{l,m \in R_h} h(l,m) u(x-l, y-m) + \eta(x,y) \ . \tag{4.17}$$

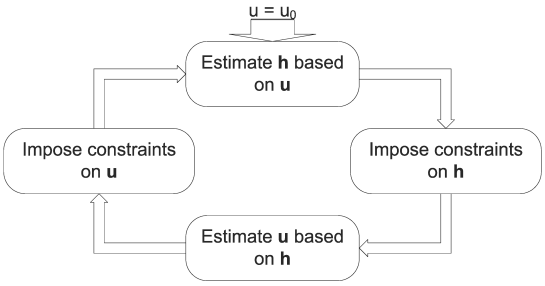


Figure 4.2.: Principle of Iterative Blind Deconvolution

Very similar to above, this can be expressed in terms of matrix-vector notation:

$$\mathbf{z} = \mathbf{H}\mathbf{u} + \eta \ . \tag{4.18}$$

Both (4.16) and (4.18) can be combined to

$$\mathbf{z} = \mathbf{H}(\mathbf{I} - \mathbf{A})^{-1}\nu + \eta \ , \tag{4.19}$$

where $\mathbf{I}$ is the identity matrix. The task is now to find the parameters of both A and H. Knowing h, any image restoration technique can be applied in order to restore the original image.

In order to overcome practical problems like instability and non-unique solutions often an parametric model of the PSF is chosen. Methods in order to finally determine the parameters divide into Maximum-Likelihood approaches (like e. g. [57, 31]) and Generalized Cross-Validation approaches. A more detailed overview of ARMA methods is beyond the scope of this thesis. For this overview, and also for more literature references, the interested reader is referred to the excellent overview paper of Kundur and Hatzinakos [29].

4.1.4. Iterative Blind Deconvolution

The basis for a new class of deconvolution algorithms was laid by Ayers and Dainty in 1988 [2]. Figure (4.1.4) illustrates the general idea. One alternates between an estimation of the kernel h based on the supposedly known image u and an estimation of the image u based on the supposedly known PSF h. In order to make this iteration successful, *a priori* constraints are imposed on u and h, e. g. that both of them are nonnegative everywhere. Ayers and Dainty impose those constraints in the spatial domain. After the estimation of u, they simply set all negative values to zero. The same is done for h after its estimation.
The estimation itself is done by transformation into the Fourier space and then constructing the inverse filter. Of course, all the known problems of inverse filtering (e. g. noise amplification, problem of dealing with values close to zero) apply to their

approach. For details on how Ayers and Dainty deal with these problems the reader is referred to their original paper.

Davey *et al.* [12] improve this somehow ingenuous approach in order to be able to deal with noise and complex-valued images by introducing a Wiener-type filter. They found that limiting the support of the PSF (another constraint on h) leads to an accelerated and more stable convergence of the iteration.

Miura and Baba successfully applied a variant of this method for the deconvolution of astronomical speckle interferograms [37]. Instead of just one convolved observation, they dispose of N equally blurred observations. For each observation, one deconvolution cycle is applied. This leads to N restorations (though not good ones, since only one cycle was passed). Now, the new input for the next cycle is calculated by averaging over these N images. Further improvements were made by Tsumaraya *et al.* [61].

4.1.5. Simulated Annealing

Simulated Annealing [27] is an optimization procedure that has been successfully applied to many scientific fields. The general idea is that one wants to find a parameter vector $\mathbf{P}$ that minimizes a given cost function $C(\mathbf{P})$. In Simulated Annealing a sequence of parameters $\mathbf{P}_0, ..., \mathbf{P}_k, ...\mathbf{P}_n$ is evaluated. The parameter vector $\mathbf{P}_{k+1}$ is found as follows: $\mathbf{P}_k$ is perturbed by a small random variation resulting in $\tilde{\mathbf{P}}$. If $\Delta Q = Q(\tilde{\mathbf{P}}) - Q(\mathbf{P_k})$ is below zero then $\mathbf{P_{k+1}}$ is set equal to $\tilde{\mathbf{P}}$. Otherwise, the parameter is still accepted with probability $e^{-\Delta Q/T}$. T is called the temperature and is initially set to a high value. Then, it decreases with the number of iterations (this simulates the process of metal annealing and gives the method its name).
Simulated Annealing is a global optimization procedure, which means that it can find a global minimum usually without getting stuck in local minima (however, one can never be sure if this is the case). This is due to the temperature T which also allows for parameter updates that increase the cost function.

McCallum [36] used this optimization scheme for the problem of blind deconvolution. To this end, he defines the cost functional

$$Q(u, h) = \|u * h - z\|^2 \quad . \tag{4.20}$$

Both f and h can be discretized and encoded into one parameter vector. Then, after an optimization with simulated annealing, decoding of this parameter vector reveals the sought f and g.

Though this algorithm delivers reasonable results even for noisy data, its computational complexity impedes its application on large data sets.

4.1.6. Nonnegativity and Support Constraints Recursive Inverse Filtering

Kundur and Hatzinakos presented a novel blind restoration technique in 1998 [30]. The only *a priori* knowledge they request before their method can be applied is that the original image u is nonnegative, that the object of interest is surrounded by a

uniform background of a known intensity L_B (e. g. this could be simply a black background) and that its support is known (the support is defined as the smallest rectangle that surrounds the object of interest). They construct a cost function $J(h_{\text{inv}})$ which is subject to minimization. The parameter h_{inv} that minimizes the cost function is an inverse filter that can then directly be used in order to restore the original image. Accordingly, they name their new approach "Nonnegativity and Support Constraints Recursive Inverse Filtering" (NAS-RIF).

The cost function J is constructed in such way that the violation of the required constraints will be punished. In order to do so, the current estimated reconstruction $\hat{u} = z * h_{\text{inv}}$ is projected into the space of images that do not violate any of the constraints. This projection is denoted by $\hat{u}_{\text{NL}}$ and the cost function subject to minimization is

$$J(h_{\text{inv}}) = \|\hat{u}_{\text{NL}} - \hat{u}\|^2 \ . \tag{4.21}$$

The exact definition of the projection is

$$\hat{u}_{\text{NL}}(x,y) = \begin{cases} \hat{u}(x,y) & \text{if } \hat{u}(x,y) \geq 0 \text{ and } (x,y) \in D_{\text{sup}} \\ 0 & \text{if } \hat{u}(x,y) < 0 \text{ and } (x,y) \in D_{\text{sup}} \\ L_B & \text{if } \in \overline{D_{\text{sup}}} \end{cases} , \tag{4.22}$$

where spatial positions within the set D_{sup} belong to the region of support and those within $\overline{D_{\text{sup}}}$ are not part of it.

Kundur and Hatzinakos minimize this cost function with the method of conjugate gradients. They also prove convexity of the cost functional J which implies that convergence is guaranteed. For cases in which the support is not known, they also provide a novel algorithm that determines this support automatically. Finally, they compare their results with IDB blind restoration techniques. Simulation results of the NAS-RIF method have demonstrated the algorithm's improved speed of convergence over the existing techniques, its superior convergence properties, and its lower computational complexity.

4.1.7. Minimization of a Cost-Functional

In 1996, You and Kaveh [67] pick up an idea that originally was used in the simulated annealing approach described previously. The idea is to formulate a cost function that is subject to minimization:

$$Q(u,h) = \|u * h - z\|^2 \ . \tag{4.23}$$

This cost functional makes sure that the reconstructed image u convolved with the reconstructed kernel h is similar to the observed image z.

However, remembering that deconvolution is an ill-posed problem and even more the blind case, constraints have to be imposed in order to make this optimization successful. Since natural images usually consist of low frequency components, it seems reasonable to encourage such solutions by punishing high frequency components in both u and h. To this end, (4.23) is modified to

$$\begin{aligned} Q(u,h) =& \|w_1 \cdot (u * h - z)\|^2 \\ &+ \|w_2 \cdot (a * u)\|^2 \\ &+ \|w_3 \cdot (b * h)\|^2 \ , \end{aligned} \tag{4.24}$$

where a and b are high-pass filters and w_1, w_2 and w_3 are spatially varying weighing parameters. The weighing allows for balancing the influence of the different regularization terms.

Unfortunately, a direct minimization of (4.24) by making use of numerical minimization methods like steepest descent or conjugate gradients has not been successful. This is due to a scale problem whose identification is another contribution of You and Kaveh in this paper. In order to circumvent this problem, similar to the IDB method of blind deconvolution, a numerical scheme that alternates between the restoration of u and h is used.

A different way for stabilizing the minimization of (4.23) is to use total variation. As described in chapter 3, total variation punishes noise while conserving sharp edges. In [62], Vogel and Oman show that it is sufficient to add a total variation term to the cost function in order to blindly restore both u and h:

$$Q(u,h) = \|u * h - z\|^2 + \alpha \int_\Omega |\nabla_{\mathbf{x}} u(\mathbf{x})| d\mathbf{x} \ . \tag{4.25}$$

Additionally, they derive a fixed-point iterative scheme and solve the resulting system of linear equations with a novel preconditioned conjugate gradient method.

Total variation can also be applied to both the image u and the kernel h [10]:

$$Q(u,h) = \|u * h - z\|^2 + \alpha \int_\Omega |\nabla_{\mathbf{x}} u(\mathbf{x})| d\mathbf{x} + \beta \int_\Omega |\nabla_{\mathbf{x}} h(\mathbf{x})| d\mathbf{x} \ . \tag{4.26}$$

This makes it even possible to restore a shift-variantly blurred image [68], though only with the limitation of having an *a priori* blur model at hand.

4.1.8. Multi-Channel Deconvolution

As described above, Lane and Bates showed in 1987 that blind deconvolution from a single image is principally possible. Their considerations hold for the case where the convolved signal is completely observable. This scenario applies for some applications, e. g. in astronomy, where the object of interest is embedded in a uniform background. However, in many applications it is not possible to receive a full observation of the convolved signal since the field of view is cut at the edges. Even

though many algorithms work with this partial data, no theoretic results were established until 1999.

In 1999, Harikumar and Bresler extend theoretic results established for the 1-D case by Gurelli and Nikias [18] to the 2-D partial data signal recovery [21, 20]. They show that even for partially observable data it is principally possible to get a perfect reconstruction if there are at least three or more differently blurred observations of the same scene available. The setting then can be denoted as

$$z_i = u * h_i \ , \tag{4.27}$$

where z_i is the i-th observation, u the original image and h_i the i-th kernel.
Harikumar's and Bresler's proves are based on some assumptions on the original image and the kernels. For example, the h_i have to be co-prime, which means that the only common factor of their Fourier transforms is a scalar constant. That describes the disparity of the channels that is necessary for the reconstruction. However, these are quite mild assumptions and happen to apply in most of the considered scenarios. Based on these considerations, they extend the EVAM algorithm proposed by Gurelli and Nikias and show results on synthetically degraded data.

However, though the algorithm developed by Harikumar and Bresler has proven to be successful for low-noise data in some practical applications [45], it has problems with higher noise levels. In order to deal with this case, Šroubek and Flusser state a cost function which includes necessary conditions derived by Harikumar and Bresler but also punishes noise [42]:

$$\frac{1}{2}\sum_{i=1}^{N} \|u * h_i - z_i\|^2 + \alpha\, Q(u) + \beta\, R(h_1, ..., h_N) \tag{4.28}$$

$Q(u)$ is a term that stabilizes u (a good choice is total variation) and

$$R(h_1, ..., h_N) = \sum_{1 \le i < j \le N} \|z_i * h_j - z_j * h_i\|^2 \tag{4.29}$$

imposes a condition from Harikumar and Bresler on the kernels. Intuitively it should be clear that R has to be zero for the perfect solution: since $z_i = u * h_i$ it follows that by using the z-transform that $U = Z_i/H_i$ and therefore $Z_i/H_i = Z_j/H_j$. Cross multiplication and back transformation then yields R.

The final cost function that Šroubek and Flusser minimize then denotes as

$$\frac{1}{2}\sum_{i=1}^{N} \|u * h_i - z_i\|^2 + \alpha \int_\Omega |\nabla u(\mathbf{x})| + \beta \sum_{1 \le i < j \le N} \|z_i * h_j - z_j * h_i\|^2 \ . \tag{4.30}$$

Optimization theory provides several ways for seeking a minimum. Šroubek and Flusser derive first order optimality conditions by making use of calculus of variations.

4.2. Shift-Variant Blind Deconvolution

As mentioned, for many real-life scenarios the shift-variant model is more adequate. However, since for real applications a spatially variant PSF is often difficult to determine, the shift-variant models mostly lead to the blind deconvolution problem.

For the case of blind deconvolution, the lack of information is getting tremendous. For each measured pixel, the original pixel intensity value as well as a respective PSF have to be estimated.

As noted, many theoretical results about shift-invariant blind deconvolution rely on some considerations made in the frequency space. The convolution theorem is essential in order to get the connection between the spatial domain and the frequency domain:

$$u * h = \mathcal{F}(u) \cdot \mathcal{F}(h) = U \cdot H \quad . \tag{4.31}$$

Unfortunately, for the shift-variant convolution this theorem does not hold. This implies that many of the theoretic results of blind shift-invariant deconvolution are not relevant.

Perhaps it is due to this fact, that literature for blind shift-variant deconvolution is rare. To the best of our knowledge, currently there are no theoretical results that state that shift-variant blind deconvolution is principally possible. The proposed algorithms, as we will further see, seem to confirm this result: even for a noiseless scenario heavy regularization techniques which provide a great deal of a priori information are needed.

4.2.1. Image Motion Degradation and Restoration

In 1972, Sawchuk studied the restoration of images that are spatially blurred due to motion [47]. Motion blur in general is shift-variant unless the plane of interest is moving parallel to the imaging plane without rotating. Accordingly, parts of the object of interest that are close to the imaging plane suffer from more motion blur than those that are farer away. If rotation is in the game, even parallel plane motion is shift-variant.

Sawchuk states that at the time of writing his paper a shift-variant restoration is computationally too complex even if the blur was known exactly. To this end, he shows that by applying a coordinate transformation it is possible to spatially transform the observed image such that only a shift-invariant restoration is necessary. After this shift-invariant restoration the image can be back transformed. He describes several typical special cases for shift-variant motion blur and applies his general transformation method on these models.

4.2.2. Restoration of Blurred Images by Decision-Directed Filtering

Tekalp *et al.* describe a decision-directed filtering algorithm [58]. They assume that a number of known shift-invariant point-spread-functions is available in advance

and that the shift-variant blur can be described by a matching of each pixel to one of the shift-invarinat PSFs. After accomplishing this model-based segmentation into regions of distinct blurs, a region-by-region reconstruction of the original image is performed.

4.2.3. Identification of a Class of Image Blur

Three years later, Özkan *et al.* make use of the local Fourier transform and so reconstruct the image part by part [69]. For each rectangle, the already known method of analyzing the zero crossings of the Fourier transformation is applied in order to estimate a shift-invariant PSF. This is possible by a piecewise shift-invariant approximation of an *a priori* blur model. Methods in order to estimate the real shift-variant model from the approximated piecewise invariant model are presented. Finally, Özkan *et al.* show some results for shift-variant motion blur.

It is important to note that this method only works for a slowly spatially varying image blur.

4.2.4. Block-Adaptive Image Identification and Restoration

In the same year, Lagendijk and Biemond propose a method that similarly to the previous one assumes that the PSF is only slowly spatially varying such that a piecewise reconstruction is justified. With difference to Özkan's method, no *a priori* blur model is needed. Instead, an expectation maximization algorithm working with an ARMA model is applied.

The main problem that arises with this approach is that due to the subdivision in several blocks, only very few data that can be used for reconstruction is available. The larger the blocks are, the more data is on-hand and the more stable is the method - however, at the cost of a worse approximation to the unknown shift-variant PSF.

4.2.5. Identification and Restoration of Motion Blurs in Sequential Images

In 1992, Trussel and Fogel investigate on a more general method for blind motion blur deconvolution than the one given by Sawchuck [60]. No *a priori* model for the blur is required and so, on the one hand, much theoretical preliminary work is spared. On the other hand, this method is also applicable in cases where a theoretical model is not available.

Of course, all these advantages do not come for nothing: Trussel and Fogel assume that a sequence of images is available. As before, these images are subdivided into smaller blocks. Two images of a sequence suffice in order to robustly determine the direction and magnitude of the movement for each of the blocks. This is done by a previously published method of Fogel. Then, the velocity vectors of each block give rise to the direction of blurring by assuming that the blurring is shift-invariant for each of the blocks.

Once the PSF for each block is established, a modified version of the Landweber method (denoted as "Sectioned Landweber Iteration" by Trussel and Fogel) is used

for reconstruction.

4.2.6. POCS-Based Restoration of Blurred Images

Computational complexity of the restoration, even if the point-spread-function is known in advance, is a big challenge. To this end, Özkan *et al.* propose a new algorithm based on a modification of the method of projection onto convex sets (POCS) in order to make it applicable to the shift-variant case [70]. Different from many previous methods, this allows for a computationally efficient reconstruction for any type of spatially varying PSF, if it is known in advance. Thus, literally speaking, each pixel can have its own shift-invariant PSF.

The remaining problem however is still the estimation of the PSF, and Özkan *et al.* rely on their previously described paper that assumes that a blur model is known in advance. Its parameters then are estimated by dividing the image into small subregions and estimating shift-invariant PSFs for each subregion. Finally, the shift-variant PSF is sought according to the model based on the subregions' PSFs.

Özkan *et al.* show experimental results for parameterized out-of-focus blur and motion blur. They use synthetically degraded data in order to compare the mean-squared-error (MSE) of the reconstruction to the MSE of the original image. Reconstructions for the actually known blur function are compared to the restoration with an estimated blur function (using their previously published method).

Since, according to their results, the restoration with the known blur function is far better compared to the estimated blur, it can be concluded that a remaining problem is the robust estimation of shift-variant blur.

4.2.7. A Regularization Approach to Blind Restoration of Images

As described for the blind and shift-invariant case, it is possible to construct a cost function which is then subject to minimization. You and Kaveh show that the same principle can be applied to shift-variant deconvolution [66]. As in the shift-invariant case, the image is regularized by assuming piecewise smoothness, e. g. by making use of total variation.

However, some additional assumptions have to be made. The tremendous lack of information in order to restore the shift-variant PSF has to be filled by regularization. To this end, You and Kaveh propose to also use piecewise smoothness assumptions. However, if we recall that in the two-dimensional case, the PSF is a function of four variables (two in order to determine the spatial position, and two for the two-dimensional shape), the question arises with respect to which of those variables piecewise smoothness has to be assumed and what the meaning is. Concerning the transitions, piecewise smoothness greatly reduces the space of feasible solutions, since its incorporation into the cost function implies that quickly spatially varying functions are discouraged. Concerning the remaining variables, piecewise smoothness has the same effect as for the image. There is a huge amount of practical situations where this also makes sens (e. g. for motion blur and out-of-focus blur).

All these assumptions still do not suffice for a good restoration, and so You and Kaveh additionally assume that a theoretic blur model is available. They present a result for horizontal motion blur applied to a synthetic image.

This method is extended and theoretically more thoroughly studied in [67] and [68].

5. Classification of Deconvolution

We already mentioned that there exist many reasonable ways of classification. We did not want to structure our review based on a rigorous classification since an unbalanced structure would have been the result. Nevertheless a knowledge about some possible classes is important for a deeper understanding of deconvolution.

5.1. Blind vs. Non-Blind

The classical image restoration problem, where one wants to estimate the original image based on an observation and the PSF is usually referred to as the non-blind deconvolution problem. As already discussed, applications for this problem are rare and a more interesting question is to retrieve both the original image and the PSF from one or more observations.

5.2. Shift-Variant vs. Shift-Invariant

Shift-Invariant deconvolution comprises deconvolution methods that try to deconvolve images based on a shift-invariant model. Accordingly, shift-invariant deconvolution assumes a shift-invariant model, where the PSF is varying over the spatial coordinate space and is not constant. The two cases were already thoroughly discussed.

5.3. Two Step vs. Iterative

This classification only holds for blind restoration methods. Two step procedures are characterized by a strict separation of the estimation of h and u. First, h is sought. Then, any of the well studied classical image restoration methods can be applied. Typical representatives of two step procedures are the frequency domain methods.

Most of the studied methods are iterative and the retrieval of u and h can not be separated. Experience has shown that iterative methods are more stable and usually show a better performance for noisy data.

5.4. Frequency Domain vs. Spatial Domain

Some of the reviewed methods work in the frequency domain, while others stay in the spatial domain. There are also hybrid methods that make use of both domains, as for example the Iterative Blind Deconvolution algorithm, where the estimation of

the image or the PSF takes place in the frequency domain and the constraints are imposed in the spatial domain.

5.5. Regularization

Whether blind or non-blind, image deconvolution is an ill-posed problem. That means, without adding information many solutions exist. Adding enough information renders the problem well-posed. This addition of information is referred to as regularization.

All of the presented methods use some kind of regularization (even if the authors do not explicitly mention the term "regularization"). In classical image restoration regularization focuses on the image (since the PSF is assumed to be already known). In blind image restoration, not only the image but also the PSF have to be sought, and so both of them form part of the parameter space that has to be regularized.

In the following, we will propose classes for the regularization of the image and the PSF. Sometimes an algorithm applies to only one of these classes, but more often several methods of regularization are chosen and accordingly it can be categorized in more than one class.

5.5.1. Regularization of the Image

- **Early Abortion of Iteration.** We already noted that aborting e. g. the Landweber iteration before convergence is a form of regularization.
- **Smoothness.** Often smooth images are encouraged as a solution by adding a dedicated term to a cost function.
- **Total variation** proved to be better than smoothing since it preserves sharp edges.

5.5.2. Regularization of the Point-Spread-Function

- **Theoretic Blur Model.** By assuming a known theoretic blur model, the space of feasible solutions greatly reduces and the only remaining task is to find the model's parameters. Apart from motion blur or out-of-focus blur it is often difficult to find the right model. In many cases such a model even does not exist.
- **Multiple Observations.** If possible, a more natural way of adding information is to use multiple differently blurred observations of the same scene. There are two possible subclassifications:
  - Each of the observed images is blurred differently. This setting applies e. g. in astronomy, where the imaging characteristics change over time and thus differently blurred observations are easily obtained by just leaving a certain amount of time in between the acquisitions.

  - Each of the observed images is blurred equally, but the observations are spatially transformed (e. g. rotated, translated) one to each other. This setting is the one we aim at in this thesis. Further details will be given in the next chapters.

- **Degree of Shift-Variance.** For a shift-variant PSF, a possible assumption is that the PSF varies only slowly in the spatial domain. In discrete terms, where one PSF is assigned to each pixel, this means that PSFs of the close neighborhood of a pixel will be very similar one to each other.

Part II.

A New Method for Multi-View Deconvolution

6. Alternating Optimization for Multi-View Blind Shift-Variant Deconvolution

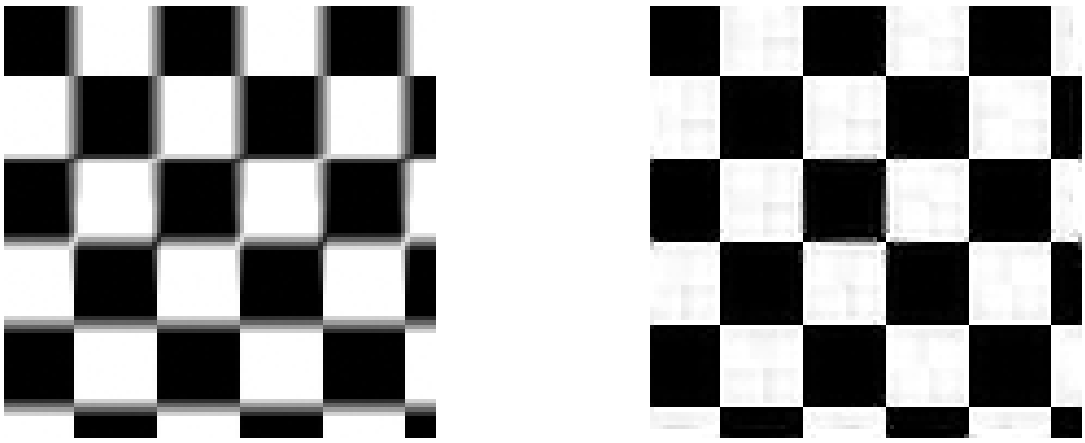
As seen, for many practical applications it is necessary to assume a shift-variant model. The PSF has to be either theoretically determined, measured, or a blind deconvolution algorithm has to be used. Until the current day, all of these approaches comprise disadvantages (depending on the scenario). On the one hand, a theoretical determination or a measurement of the PSF involves a detailed knowledge about the physics of the imaging process and can be difficult, time-consuming or not even possible. On the other hand, the shift-variant blind deconvolution has - to the best of our knowledge - so far only been addressed by making *a priori* assumptions about the shape of the PSF, for example that the PSF is a Gaussian or an instance of motion- or out-of-focus blur.

Consequently, for the shift-variant model, up to now there exists no general method that accomplishes deconvolution without having to adapt it to the particular imaging process.

To this end, we propose a new method which overcomes these drawbacks by introducing a general framework for blind deconvolution. The main idea is to use multiple observations of the same scene. By using multiple images, our method accounts for the severe lack of information present in the standard blind deconvolution problem. We experimentally show that few observations can yield a high-quality restoration of both the original image and the shift-variant PSF. Since no parametrization or estimation of the PSF is needed, the method is easily applicable in many different settings.

The proposed framework incorporates the information about the transformation between the single observations. Here any arbitrary transformation, rigid as well as non-rigid, can be used. This is important in biomedical settings where non-rigid deformations of the object can occur between the single scanning procedures. Furthermore, the method automatically uses the maximum number of available measurements for a certain image point. This way, even outside the overlapping area of all images, the maximally available information is automatically used. This also allows us to include information about damaged parts of an image in a straightforward manner.

An example for the general scenario to which our method is applicable is shown in Figure 1.1. The object of interest is rotated and translated between the observations. The observations together with the transformations are used as input for our method. This scenario is already present or can easily be constructed for numerous applications in biomedical imaging modalities such as confocal microscopy, two-photon microscopy, optical coherence tomography and ultrasound. Transformations between the single observations can be acquired by image registration [23, 35] or



(a) One of the observations (b) Reconstruction by our method

Figure 6.1.: Simple example for shift-variant blind deconvolution. Please note that the observation depends on the spatial position because of the spatially-variant PSF.

tracking. A more comfortable and less error-prone way is to use dedicated devices which have in-built capabilities of providing the transformation (as done by Swoger *et al.* [56]).

Our method contributes to a large class of multi-view applications by improving the process of image restoration and thus allowing one to acquire images of higher quality by using the existing hardware setup.

6.1. Model

We use the following model of image formation

$$z_i = (u \circ \varphi_i) \star h + \eta \ , \tag{6.1}$$

where $z_i \in F$ is the actual measurement, $u \in F$ is the original image, $\varphi_i : \Omega \to \Omega$ is the spatial transformation between measurement and original image, h is the point-spread-function and $\eta \in F$ is the noise.

In presence of noise, we cannot assume that $z_i = u(\varphi_i) \star h$ holds exactly, so we formulate the cost function which is subject to be minimized as

$$\mathcal{D}[u, h] = \frac{1}{2} \sum_{i=1}^{P} \int_{\Omega_i} e_i(\mathbf{x})^2 \, d\mathbf{x} \ , \tag{6.2}$$

where we suppose to dispose of P measurements. $e_i(\mathbf{x})$ is an error term that measures the difference of the convolved real image to the measurements at position $\mathbf{x}$

$$e_i(\mathbf{x}) = (u \circ \varphi_i) \star h - z_i \tag{6.3}$$

$$= \int_{\Omega} u(\varphi_i(\xi)) h(\mathbf{x}, \xi - \mathbf{x}) \, d\xi - z_i(\mathbf{x}) \ . \tag{6.4}$$

The set Ω_i denotes the overlapping domain of the original image u and the image transformed by φ_i.

The problem of minimizing the functional from Equation (6.2) is ill-posed. In order to render the problem well-posed we introduce regularizing terms $\mathcal{U}$ and $\mathcal{H}$. While $\mathcal{U}$ operates on the image u, the term $\mathcal{H}$ imposes restrictions on the point-spread-function h.

So we define the regularized functional $\mathcal{J} : F \times F \to \mathbb{R}$ to be minimized which models the problem as

$$\mathcal{J}[u,h] = \mathcal{D}[u,h] + \alpha\mathcal{U}[u] + \beta\mathcal{H}[h] \ . \tag{6.5}$$

Here, the positive real coefficients α and β determine the influence of the respective regularization term.

Additionally, since we assume to deal with an intensity PSF, we want to restrict h to zero and positive values, that is, $h(\mathbf{x},\mathbf{s}) \geq 0$. We also assume that the imaging system neither emits nor absorbs energy and so the PSF has to be normed to one at every spatial position $\mathbf{x}$ in the image , that is $\int_\Omega h(\mathbf{x},\mathbf{s})\, d\mathbf{s} = 1$. The final goal then is to find functions u and h that minimize $\mathcal{J}$ subject to constraints, that is, compute

$$\underset{u,h}{\operatorname{argmin}} \ J[u,h] \tag{6.6}$$

subject to

$$h(\mathbf{x},\mathbf{s}) \geq 0 \qquad \forall \mathbf{x},\mathbf{s} \in \Omega \tag{6.7}$$

$$\text{and} \int_\Omega h(\mathbf{x},\mathbf{s})\, d\mathbf{s} = 1 \qquad \forall \mathbf{x} \in \Omega \ . \tag{6.8}$$

Regularization Term $\mathcal{U}$. Total variation [46] has proven to be a very successful method for image regularization since it discourages noise while edges are preserved

$$\mathcal{U}[u] = \int_\Omega |\nabla_{\mathbf{x}} u(\mathbf{x})|\, d\mathbf{x} \ . \tag{6.9}$$

This regularization term penalizes too large gradients in the image to be reconstructed.

Regularization Term $\mathcal{H}$. As mentioned, the lack of information due to the shift-variant PSF is tremendous. In order to account for this lack of information and the resulting ill-posedness of the problem, we introduce a regularization term $\mathcal{H}$ which operates on the PSF. Coming back to the idea that a shift-variant PSF can be considered as a set of shift-invariant PSFs for each pixel $\mathbf{x}$, the regularization term $\mathcal{H}$ penalizes the difference between the shift-invariant PSFs at neighboring pixels. It is important to note that this regularization *will not* impose smoothness with respect to the second parameter. This means that not the single shift-invariant PSFs but the transitions between them are required to be smooth.

In summary, the regularization term $\mathcal{H}$ penalizes large changes of the PSF between neighboring points by measuring the norm of the gradient with respect to the first parameter of h

$$\mathcal{H}[h] = \iint_{\Omega} |\nabla_{\mathbf{x}} h(\mathbf{x}, \mathbf{s})|^2 \, d\mathbf{s} \, d\mathbf{x} \ . \tag{6.10}$$

6.2. Algorithm

The main difficulty of blind deconvolution in general is that we neither have u nor h. For this type of problem, You and Kaveh identify a scale problem [67] and so justify an algorithmic scheme that is alternating between the calculation of h and u. For the calculation of h we assume that u is given, and for the calculation of u we assume h to be given. These two steps are executed alternately until convergence

$$\text{(i)} \qquad h \leftarrow \operatorname*{argmin}_{h} \mathcal{J}[u, h] \text{ subject to } (6.7) \text{ and } (6.8) \tag{6.11}$$

$$\text{(ii)} \qquad u \leftarrow \operatorname*{argmin}_{u} \mathcal{J}[u, h] \ . \tag{6.12}$$

What remains now is to define the optimization methods for steps (6.11) and (6.12). This will be done in the next two sections.

6.2.1. Solving for the Shift-Variant PSF

Assuming that u is a fixed constant, (6.11) can be expressed as

$$h \leftarrow \operatorname*{argmin}_{h} \ \mathcal{D}[u, h] + \beta \mathcal{H}[h] \text{ s. t. } (6.7) \text{ and } (6.8) \ . \tag{6.13}$$

The regularization term $\mathcal{H}$ from (6.10) can be rewritten as

$$\iint_{\Omega} \sum_{i} \left[\frac{\partial}{\partial x_i} h(\mathbf{x}, \mathbf{s}) \right]^2 \, d\mathbf{s} \, d\mathbf{x} \ . \tag{6.14}$$

After discretization, the minimization problem can be expressed by the following constrained least-squares problem

$$\mathbf{h} \leftarrow \operatorname*{argmin}_{\mathbf{h}} \left\| \begin{bmatrix} \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_N \\ \mathbf{G}_{x_1} \\ \mathbf{G}_{x_2} \end{bmatrix} \mathbf{h} - \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_N \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right\|^2 \tag{6.15}$$

subject to

$$h_i \geq 0 \quad \forall i \in N \tag{6.16}$$

$$\text{and } \ \mathbf{N}\mathbf{h} = \mathbf{1} \ . \tag{6.17}$$

As mentioned earlier, $\mathbf{U_i}$ is a discrete representation of the convolution operator. The transformation φ_i is already encoded in $\mathbf{U_i}$. Note that each row has at most $o \cdot p$ non-zero entries and thus $\mathbf{U_i}$ is highly sparse. $\mathbf{G}_{x_i}$ is a sparse matrix that represents the partial differential operator $\frac{\partial}{\partial x_i}$.

6.2.2. Solving for the Image

If we assume that h is a fixed constant, (6.12) can be expressed as

$$u \leftarrow \underset{u}{\operatorname{argmin}} \ \mathcal{D}[u, h] + \alpha \mathcal{U}[u, \nabla_{\mathbf{x}} u] \ . \tag{6.18}$$

A necessary condition for a minimum of a functional is that the Gâteaux derivative vanishes for all perturbations

$$\delta_f \mathcal{F}[f; \phi] = \frac{\mathcal{F}[f + \varepsilon\phi] - \mathcal{F}[f]}{\varepsilon} = 0 \ . \tag{6.19}$$

By making use of the linearity of (6.18) we get

$$\delta_u \mathcal{D}[u, h; \phi] = -\alpha \ \delta_u \mathcal{U}[u, \nabla_{\mathbf{x}} u; \phi] \ . \tag{6.20}$$

The derivation of the left side is similar to the one shown by You and Kaveh [68]. The difference is that in our case we have a sum over several error terms and a transformation of u by φ_i. Adapting this derivation (as shown in the appendix), the left side turns out to be

$$\int_\Omega \left[\sum_i \int_{\Omega_i} e_i(\xi + \varphi_i^{-1}(\mathbf{x})) h(\xi + \varphi_i^{-1}(\mathbf{x}), \xi) \, d\xi \right] \phi(\mathbf{x}) \, d\mathbf{x} \ . \tag{6.21}$$

A derivation of the right side is shown by Rudin *et al.* in [46]. The problem of the total variation being not differentiable everywhere is circumvented by using an approximation of the Euclidean norm by $|\cdot|_\epsilon = \sqrt{\sum_i f(\mathbf{x_i})^2 + \epsilon^2}$ (compare [62]). Then, (6.9) becomes

$$\mathcal{U}_\beta[u] = \int_\Omega \sqrt{|\nabla_{\mathbf{x}} u(\mathbf{x})|^2 + \beta^2} \, d\mathbf{x} \ , \tag{6.22}$$

and its Gâteaux derivative is

$$\int_\Omega \nabla_{\mathbf{x}}^\top \frac{1}{\sqrt{u_{x_1}^2 + u_{x_2}^2 + \beta^2}} \nabla_{\mathbf{x}} u \cdot \phi(\xi) \, d\xi \ . \tag{6.23}$$

Putting both sides together and taking into account that the resulting equation must hold for any arbitrary perturbation ϕ we finally yield the following partial differential equation (PDE)

$$\begin{aligned} \sum_i \int_{\Omega_i} e_i(\xi + \varphi_i^{-1}(\mathbf{x})) h(\xi + \varphi_i^{-1}(\mathbf{x}), \xi) \, d\xi \\ = \ -\alpha \nabla_{\mathbf{x}}^\top \frac{1}{\sqrt{u_{x_1}^2 + u_{x_2}^2 + \beta^2}} \nabla_{\mathbf{x}} u \ . \end{aligned} \tag{6.24}$$

The image which is subject to reconstruction is found by solving this PDE.

Discretization. As already mentioned, in discrete terms, we represent functions by vectors. Integrals and partial derivatives (by using finite differences) can then be easily expressed as matrix vector products. The discretization of (6.24) then is

$$\sum_i \mathbf{T}_i^\top \mathbf{H}^\dagger \mathbf{Z}_i \left(\mathbf{H}\mathbf{T}_i\mathbf{u} - \mathbf{z}_i\right) = -\alpha\, \mathbf{G}_\mathbf{x}^\top \mathbf{D}(\mathbf{u}) \mathbf{G}_\mathbf{x}\mathbf{u} \ . \tag{6.25}$$

On the left side, the term $\mathbf{H}\mathbf{T}_i\mathbf{u} - \mathbf{z}_i$ approximates e_i. $\mathbf{T}_i$ discretizes the transformation φ_i. $\mathbf{H}$ encodes the shift-variant convolution with h. $\mathbf{Z}_i$ is a diagonal matrix with entries one or zero and so takes care of the fact that the error term is only evaluated on the set Ω_i. $\mathbf{H}^\dagger$ denotes the outer integral. Instead of first transforming both e_i and h and then integrating, the discretization first integrates and then applies the inverse transformation φ_i^{-1} (discretized by $\mathbf{T}_i^\top$) on the result.

On the right side of (6.25), the matrix vector product $\mathbf{G}_\mathbf{x}\mathbf{u}$ is the finite difference approximation of $\nabla_\mathbf{x} u$. $\mathbf{D}(\mathbf{u})$ is a square matrix that represents the fraction.

Note that since there are only slight differences between the outer integral of (6.25) and the convolution integral in e_i (the first parameter of h depends on the integration variable ξ, the sign of the integration variable ξ is the same in both e and h) it is not surprising that the matrix $\mathbf{H}^\dagger$ is related to $\mathbf{H}$. In fact, if we assume that the PSF h varies only slowly with the spatial position, $\mathbf{H}^\dagger$ can be approximated by $\mathbf{H}^\top$.

Numerical Solution Scheme. Having discretized the problem in equation (6.25), the last step is to present the actual numerical scheme for solving the problem. The dependency of $\mathbf{D}$ on $\mathbf{u}$ makes the problem nonlinear. In order to deal with this problem, we formulate a fixed point equation $\mathbf{u} = \rho(\mathbf{u})$ as

$$\begin{aligned}\mathbf{u} = &\left\{ \left[\sum_i \mathbf{T}_i^\top \mathbf{H}^\dagger \mathbf{Z}_i \mathbf{H} \mathbf{T}_i \right] + \alpha \mathbf{G}_\mathbf{x}^\top \mathbf{D}(\mathbf{u}) \mathbf{G}_\mathbf{x} \right\}^{-1} \\ &+ \sum_i \mathbf{T}_i^\top \mathbf{H}^\dagger \mathbf{Z}_i \mathbf{z}_i \ .\end{aligned} \tag{6.26}$$

Here, the inverse is merely a notation for a method solving a linear equation. In practice we use the method of conjugate gradients. This fixed point scheme directly leads to the following iteration

$$\mathbf{u}^{(i+1)} = \rho(\mathbf{u}^{(i)}) \ . \tag{6.27}$$

7. Evaluation on Synthetic Data

In order to perform a quantitative evaluation we use a simulation of the degradation process with a known ground truth. Since we want to simulate a realistic scenario, we use real images to represent the actual object to be scanned. Using synthetic data has several advantages over the real-life scenario:

- The original data is known and so the restored data can be compared directly.
- One is sure that the assumed model (convolution and noise) is given, and no other disturbing effects are present.
- Many sources of errors, e. g. misregistration, are avoided.

There are several possible criteria for evaluation. Some are related with the result itself, like the quality of reconstruction. Others are giving rise about the algorithm, for example its computational complexity or memory complexity.

In this evaluation, the attention is drawn to the quality of restoration. To be more precise, we measure the quality of reconstruction by computing the similarity of the reconstructed image to the original image. However, this is not yet precise enough. In fact, it is very difficult to measure similarity from the point of view of the human visual system objectively. It is beyond the scope of this thesis to give an overview of the respective literature, and as a starting point for the interested reader we recommend [13, 63] and [64]. In the end, the reader has to form his personal opinion about the restoration quality by visual inspection and therefore we always present the reconstructed results.

In the existing literature, various measures are used. Among them are the mean squared error (MSE) [70], the signal-to-noise ratio improvement [67], the percentage mean squared error (PMSE) [30, 42] and the percentage root mean squared error (PRMSE) - just to name a few.

We decided to use the PMSE, since it is independent of the image size and sampling values:

$$\mathrm{PMSE}_{\mathbf{v}} = 100 \frac{|\mathbf{v} - \mathbf{v}^{\mathrm{orig}}|^2}{|\mathbf{v}^{\mathrm{orig}}|^2} \quad . \tag{7.1}$$

This similarity measure can be used to compare the similarity of the reconstructed image to the original image as well as the similarity of the reconstructed PSF to the original PSF. In case of the image similarity, it is important to note that our algorithm strongly relies on overlapping areas. So, for parts where no information from other observations is available, the reconstruction is worse. To take this into account, $\mathbf{v}$ contains only those parts of the image for which corresponding elements in all the other observations are available.

Among many factors that might influence the restoration quality, we focus on the following ones:

1. Level of noise
2. Spatial transformation
3. Number of available observations
4. Point-spread-function
5. Image subject to restoration

There is an unlimited number of combinations of these factors. For instance, it is possible that the number of observations does not influence the general performance in the noise-free case, but for noisy images it has a big influence. Or, a certain type of PSF might work well in combination with some spatial transformation, and it does not for others. Especially in the shift-variant case it can be expected that translations have a big effect on the restoration, whereas in the shift-invariant case only rotations or other deformations will have influence. Under these aspects it is very difficult to give a good picture about the connection of these factors. Even worse, there may be more factors we did not think of.

For the simulation, we perform the following steps for in order to generate the degraded observations:

- Rotation and translation
- Convolution/Blurring
- Addition of white Gaussian noise

We state the noise level by the signal to noise ratio SNR $= P_{\mathbf{z}_i}/P_{\text{Noise}}$, where the signal power $P_{\mathbf{z}_i}$ is the power of the measurements $\mathbf{z}_i$ and P_{Noise} is the noise power. With the noise power defined as the variance of the noise σ^2, and using the logarithmic decibel scale we yield

$$\text{SNR(dB)} = 10 \log \frac{\sum_{i=1}^{P} \mathbf{z}_i^\top \mathbf{z}_i}{P \cdot m \cdot n \cdot \sigma^2} \quad . \tag{7.2}$$

7.1. Two-Dimensional Shift-Invariant Blind Restoration

Level of noise. We applied our method to a photograph of a cameraman that is often used in the literature. The photograph is blurred by a PSF that consists of a horizontal strip of one pixel height and seven pixel width. Two observations are used, the second one is rotated by ninety degrees. Different levels of noise were applied. Figure (7.1) shows the results of the reconstruction and Figure (7.1) shows the evaluation of the PMSEs.

Spatial transformation. In order to find out about the relation between the reconstruction quality and the spatial transformation, in this experiment we rotate the second observation only about $45°$ and compare it to the previously about $90°$ rotated example. Figure (7.3) and (7.1) show that although visually the two results are hardly distinguishable, the reconstruction for $90°$ worked better. Additionally, the convergence in the $90°$ case is faster.

Number of Available Observations As shown by Figure (7.5) and Figure (7.1), using four observations instead of only two significantly improves the reconstruction result.

Point-Spread-Function. As mentioned, our algorithm takes advantage of additional information that is available in several observations. However, whether rotated observations really provide more information depends on the PSF. Therefore, we will expect that for example it is not possible to restore a blurring by a disc-shaped PSF. Figure (7.7) and Figure (7.1) confirm this assumption.

Image subject to restoration. Figures (7.9) and (7.1) show the evaluation for other images. As expected, the restoration quality and the speed of convergence depend on the provided image. However, the question remains which image characteristics exactly are responsible for that and how they influence the restoration.

7.2. Two-Dimensional Shift-Variant Blind Restoration

Also for the shift-variant case we can suppose that the restoration results depend on the level of noise, the spatial transformation, the number of available observations, the point-spread-function and the image subject to restoration. Due to time constraints, we are not able to study the restoration results in the same detail as we did for the shift-invariant case and focus on the in our opinion most interesting aspects: the dependency of the reconstruction quality on the noise level and the image subject to reconstruction.

Level of Noise In our opinion it is quite obvious that the interesting question is not whether noise affects the restoration results, but up to which noise level reasonable reconstructions are still possible. Figure (7.11) and Figure (7.2) show that even up to a SNR of 10 dB a better result is obtained. We also see that for higher levels of noise it is better to stop the algorithm earlier. This might be due to the fact that there, the deconvolution part of the method converges earlier. Then, only denoising is performed at each iteration, and since the image is already denoised after few iterations, any further denoising worsens the result.

Image Subject to Reconstruction Especially for the shift-variant case, we expect that the image itself plays an important role. The main reason is that the image

has to provide information that can be used for estimation of the PSF at every spatial position. Large uniformly colored areas for example can not provide any information about the PSF. In the shift-invariant case, this was not a problem since other parts of the image can still provide enough information.

This expectation is confirmed by Figure 7.13 and Figure 7.2, which show restorations of a random pattern. The random pattern provides structure at any spatial position, and the shift-variant setting greatly benefits from this additional information.

7.3. Comparison to Other Blind Deconvolution Methods

It is very difficult to compare our algorithm to other methods. The vast majority of available methods make use of one observation, and therefore dispose of fewer data than we do in our setting. It is intuitively clear that this makes them much more prone to noise and our review of the more recent literature confirms this feeling.

7.3.1. Shift-Invariant Blind Methods

Chan and Wong work with noisy real-life data, but do not make a quantitative evaluation for noisy synthetic data [10]. Kundur and Hatzinakos mostly apply the NAS-RIF algorithm on noise-free data [30]. The lowest SNR that they evaluate is 40 dB. At this level, their algorithm is instable and if the iteration with the best result is manually chosen, the PMSE is approximately 1. Both You and Kaveh [68] and Bar *et al.* [4] use a parameterized PSF. The results shown compare the similarity of the PSF, but not the similarity of the reconstructed image to the original image. Additionally, the lowest noise level investigated in their work is 30 dB.

There are only very few methods we are aware of that deal with our multi-view scenario. Among many of them no quantitative comparison between the reconstructed image and the original image is made [26, 24, 54, 56, 23, 50].

The only multi-view publications we are aware of that provide a quantitative evaluation of the reconstruction quality of the image are by Harikumar and Bresler [21, 20] and Šroubek and Flusser [45, 42, 43, 44]. The MC-AM algorithm of Šroubek and Flusser can be seen as an extension of the EVAM algorithm by Harikumar and Bresler, and so we will only compare our method to the MC-AM algorithm.

Note that a comparison of our method to the MC-AM has to be taken with a pinch of salt. On the one hand, the MC-AM can be applied to our setting, but it is more general: it assumes that each of the observations is blurred by a different PSF (refer to chapter 4 for an overview of their method). On the other hand, we already saw that the performance of the reconstruction depends on various factors like the content of the image and the PSF.

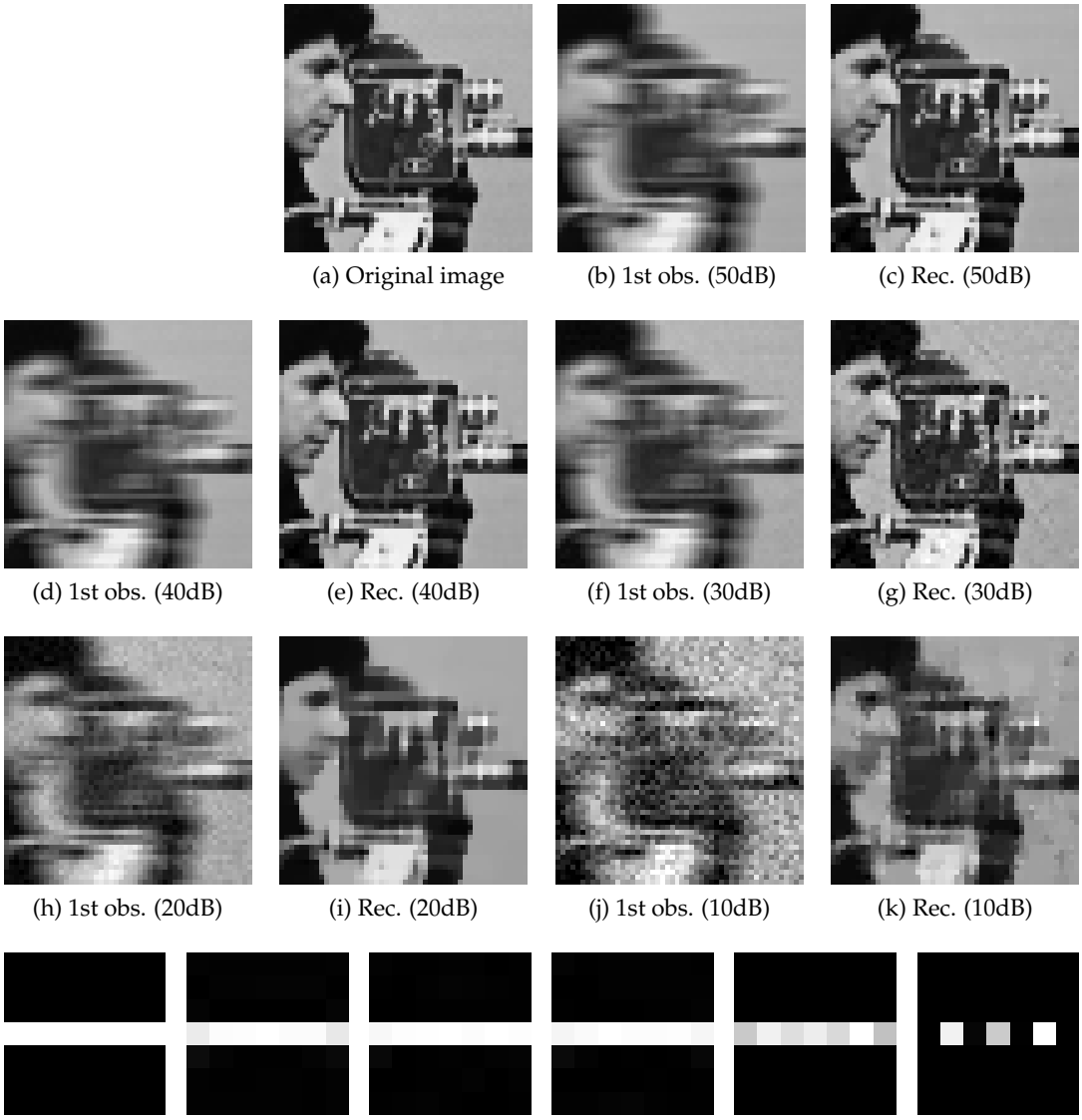
Since the setting for our method is more restrictive in the shift-invariant case, it can be expected that the additional knowledge we incorporate leads to better results of our algorithm for high noise levels. Figure (7.3.1) shows some of Šroubek's and

Flusser's restorations and seems to confirm our expectation. This is further backed by Table (7.3).

7.3.2. Blind Shift-Variant Methods

Comparison with shift-variant blind methods is even more difficult. There is very few literature that is showing results. We did not see any publication that is dealing with a multi-view setting in the shift-variant case. All the papers we are aware of use parameterization in order to face the tremendous lack of information present in this case.

Nevertheless we want to show at least one other example of blind shift-variant restoration. You and Kaveh assume horizontal motion blur [68]. They degrade a synthetic image of three discs with a horizontal motion blur that is stronger at the vertical middle of the image than at the vertical borders



(a) Original image (b) 1st obs. (50dB) (c) Rec. (50dB)

(d) 1st obs. (40dB) (e) Rec. (40dB) (f) 1st obs. (30dB) (g) Rec. (30dB)

(h) 1st obs. (20dB) (i) Rec. (20dB) (j) 1st obs. (10dB) (k) Rec. (10dB)

(l) Original PSF (m) Rec. (50dB) (n) Rec. (40dB) (o) Rec. (30dB) (p) Rec. (20dB) (q) Rec. (10dB)

Figure 7.1.: Results of our blind deconvolution method for the cameraman with different levels of noise. Two observations were used for reconstruction, only the first one is shown in this Figure. The second observation is rotated by 90 degrees and is distorted with the same blur and SNR as the first one.

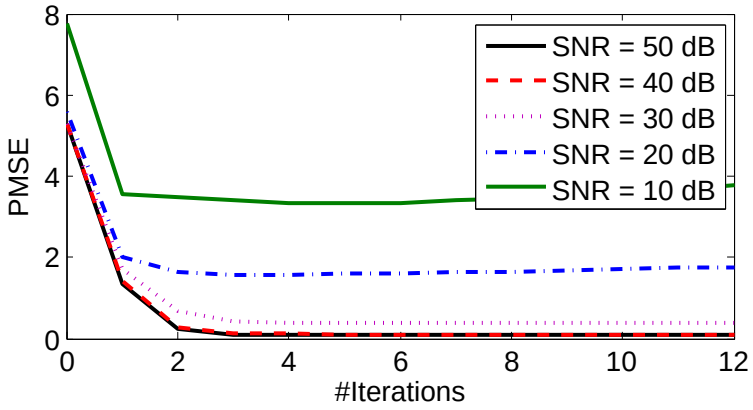
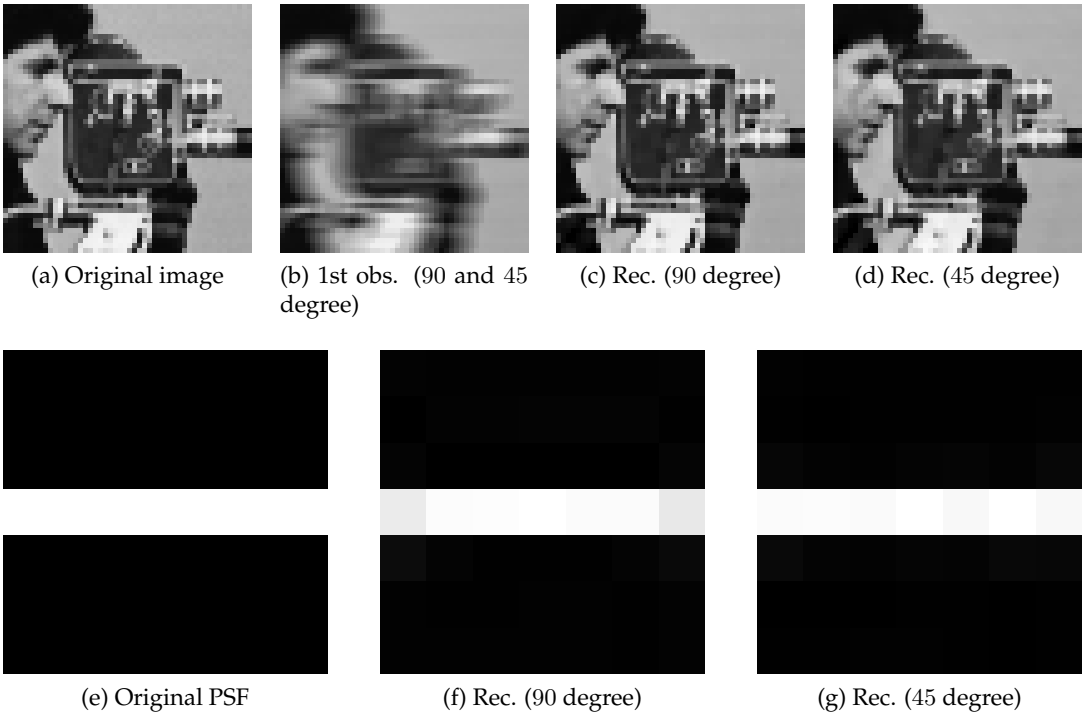


Figure 7.2.: Evaluation of the percentage mean squared error for the images shown in Figure (7.1). The 0-th iteration is the initial guess for the image.



(a) Original image (b) 1st obs. (90 and 45 degree) (c) Rec. (90 degree) (d) Rec. (45 degree)

(e) Original PSF (f) Rec. (90 degree) (g) Rec. (45 degree)

Figure 7.3.: Results of our blind deconvolution method for the cameraman at 50 dB noise level for two different angles. For each case, two observations were used for reconstruction, only the first one is shown in this Figure.

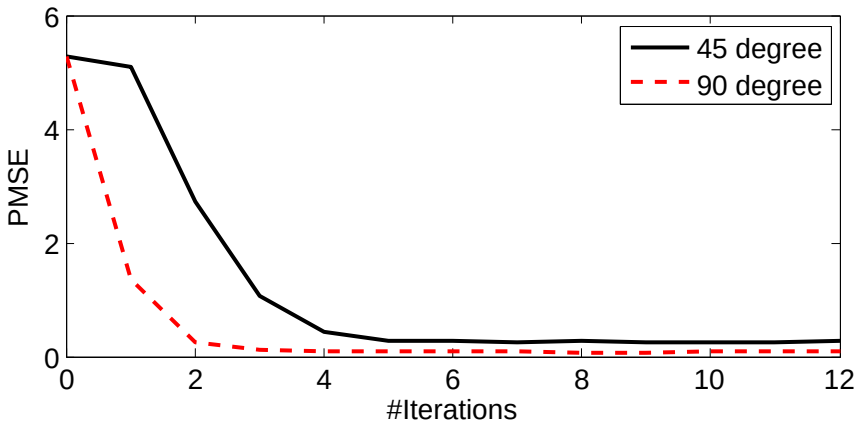
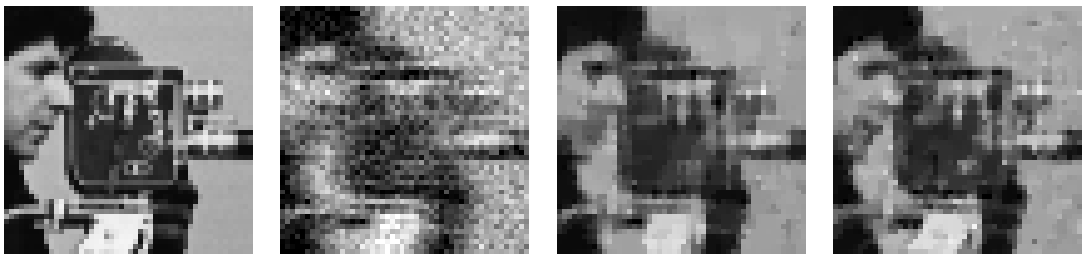
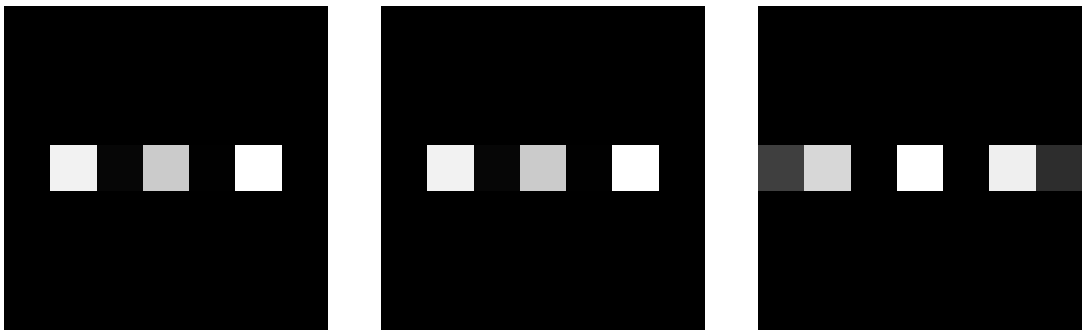


Figure 7.4.: Evaluation of the percentage mean squared error for the images shown in Figure (7.3). The 0-th iteration is the initial guess for the image.



(a) Original image (b) 1st obs. (two and four obs.) (c) Rec. (two obs.) (d) Rec. (four obs.)



(e) Original PSF (f) Rec. (two obs.) (g) Rec. (four obs.)

Figure 7.5.: Results of our blind deconvolution method for the cameraman at 10 dB noise level. In the first case, two observations (at 0 and 90 degree) are used. In the second case, four observations (at 0, 45, 90 and 135 degree) are available for restoration.

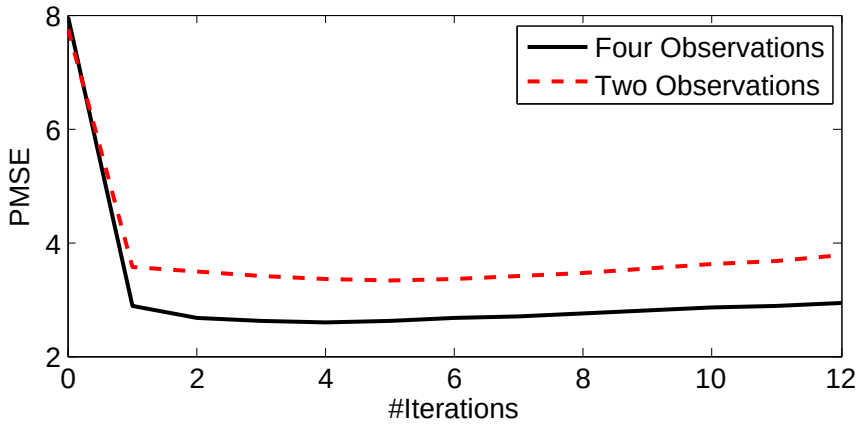


Figure 7.6.: Evaluation of the percentage mean squared error for the images shown in Figure (7.5). The 0-th iteration is the initial guess for the image.

(a) Original image (b) 1st obs. (PSF a) (c) Rec. (PSF a)

(d) 1st obs. (PSF b) (e) Rec. (PSF b) (f) 1st obs. (PSF c) (g) Rec. (PSF c)

(h) Orig. PSF a (i) Rec. PSF a (j) Orig. PSF b (k) Rec. PSF b (l) Orig. PSF c (m) Rec. PSF c

Figure 7.7.: Results of our blind deconvolution method for the cameraman with different point-spread-functions. In each case, three observations at angles of 0, 45 and 135 degrees are used for reconstruction.

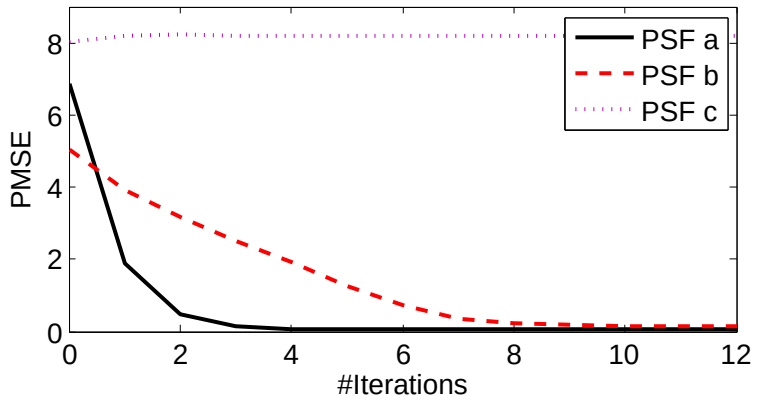


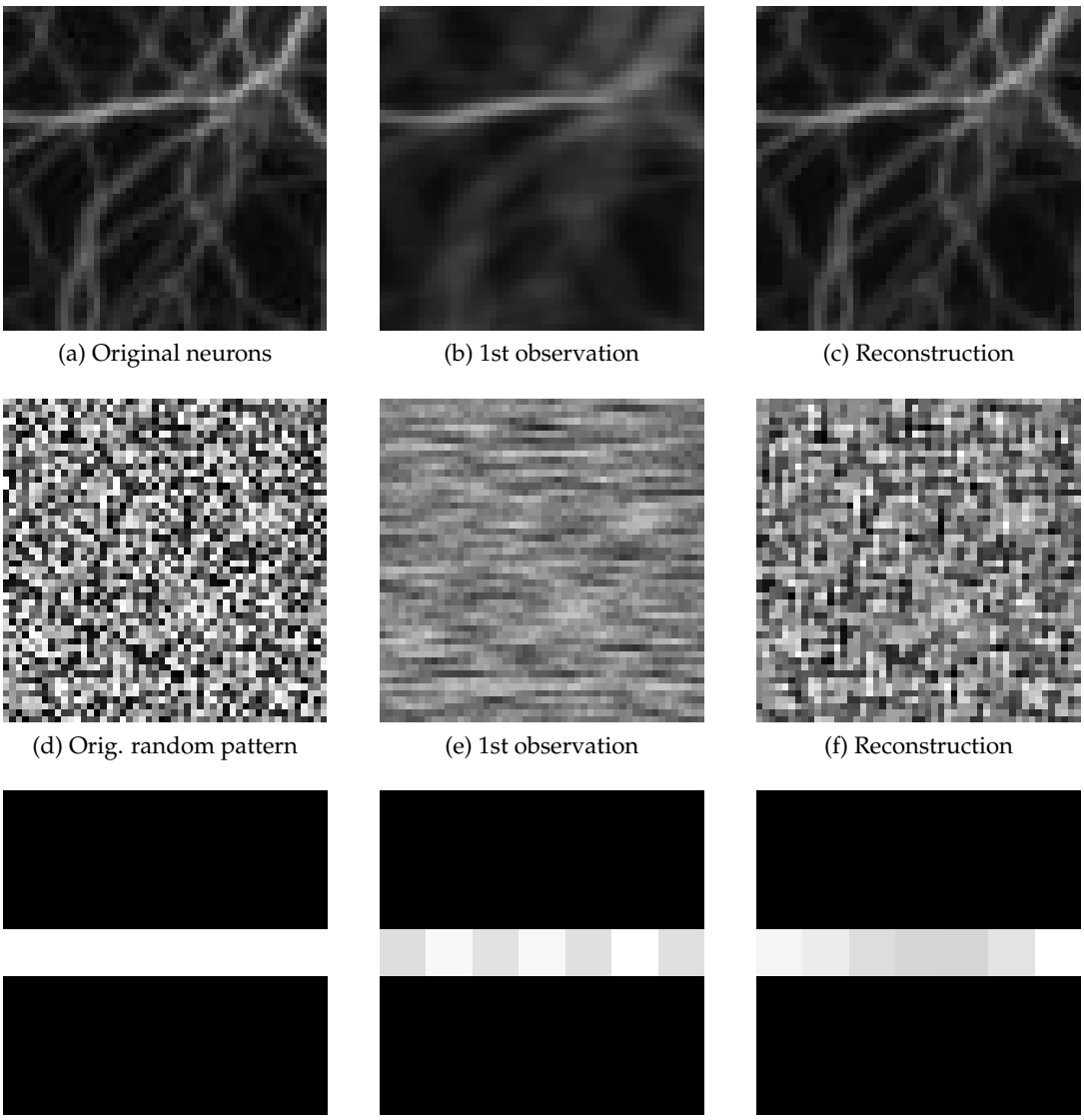
Figure 7.8.: Evaluation of the percentage mean squared error for the images shown in Figure (7.7). The 0-th iteration is the initial guess for the image.

Description	PMSE
Cameraman, 50 dB, two observations (90 deg), horizontal PSF	0.091
Cameraman, 40 dB, two observations (90 deg), horizontal PSF	0.096
Cameraman, 30 dB, two observations (90 deg), horizontal PSF	0.38
Cameraman, 20 dB, two observations (90 deg), horizontal PSF	1.8
Cameraman, 10 dB, two observations (90 deg), horizontal PSF	3.8
Cameraman, 50 dB, two observations (45 deg), horizontal PSF	0.26
Cameraman, 10 dB, four observations (45 deg, 90 deg, 135 deg), horizontal PSF	0.047
Cameraman, 50 dB, three observations (45 deg, 90 deg, 135 deg), PSF a)	2.9
Cameraman, 50 dB, three observations (45 deg, 90 deg, 135 deg), PSF b)	0.14
Cameraman, 50 dB, three observations (45 deg, 90 deg, 135 deg), PSF c)	8.19
Neurons, 50 dB, two observations (90 deg), horizontal PSF	0.41
Random pattern, 50 dB, two observations, (90 deg), horizontal PSF	6.85

Table 7.1.: Overview of final PMSE for the blind shift-invariant case.

Description	PMSE
Random Pattern, 50 dB	3.8079
Random Pattern, 40 dB	3.8258
Random Pattern, 30 dB	4.0666
Random Pattern, 20 dB	4.6372
Random Pattern, 10 dB	14.2224

Table 7.2.: Overview of final PMSE for the blind shift-variant case.



(a) Original neurons (b) 1st observation (c) Reconstruction

(d) Orig. random pattern (e) 1st observation (f) Reconstruction

(g) Orig. PSF (h) Rec. PSF neurons (i) Rec. PSF random pattern

Figure 7.9.: Results of our blind deconvolution method for an image of hippocampal rat neurons and a random pattern. In each case, two observations at angles of 0 and 90 degrees are used for reconstruction. The noise level is 50 dB.

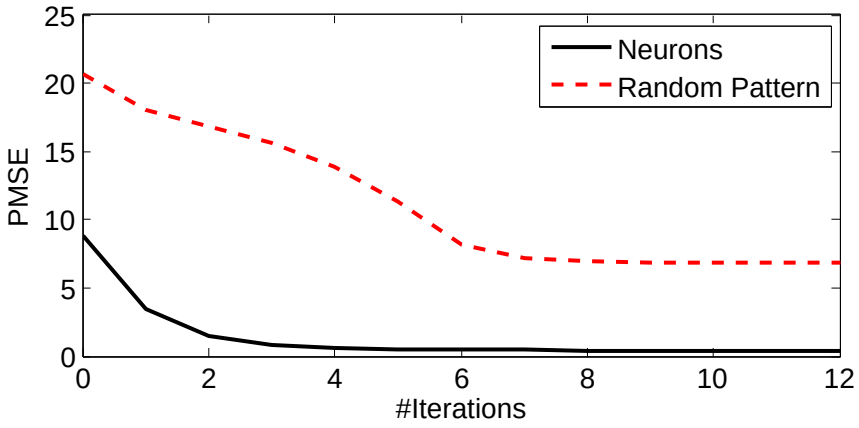
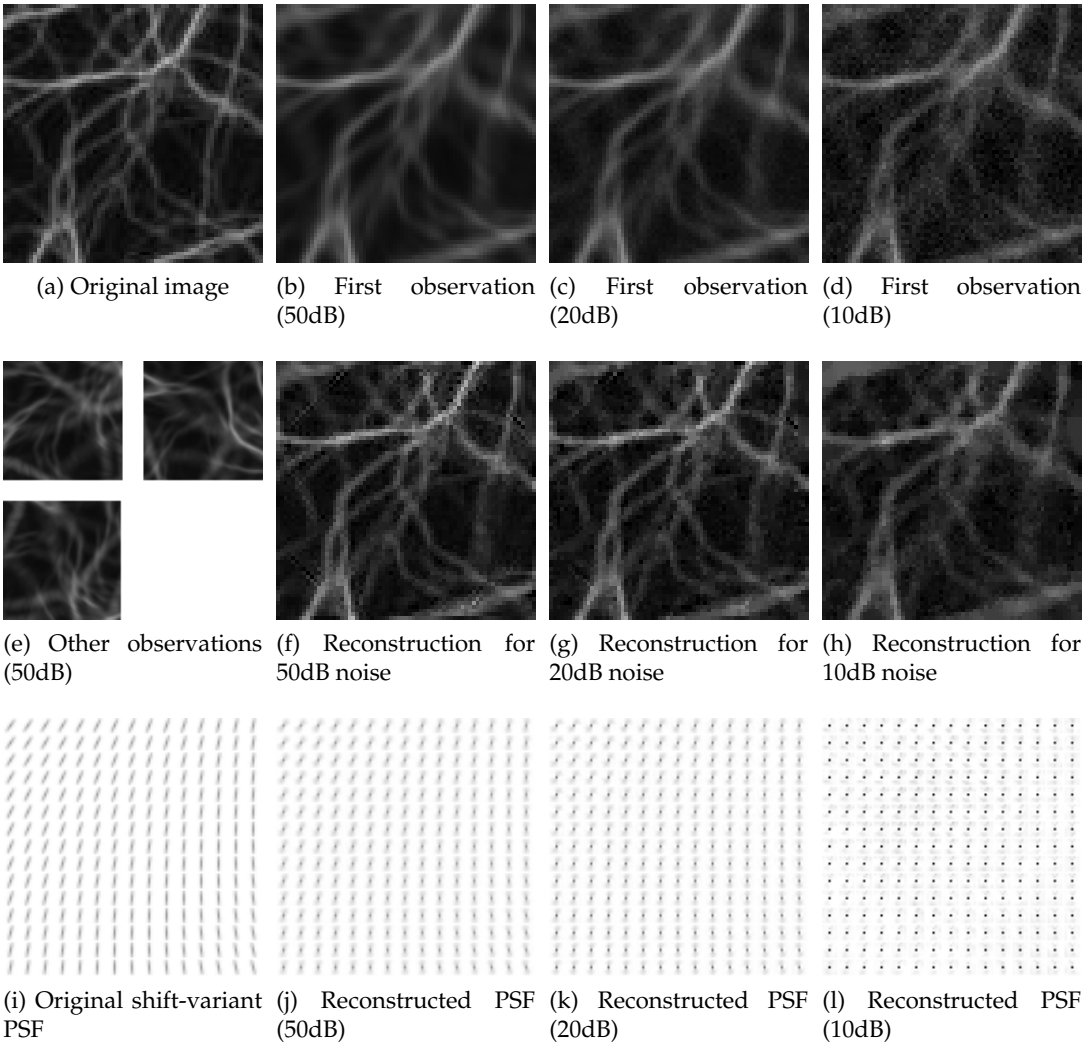


Figure 7.10.: Evaluation of the percentage mean squared error for the images shown in Figure (7.9). The 0-th iteration is the initial guess for the image.

SNR	PSME (EVAM)	PMSE (MC-AM)	PMSE (Our Method)
50 dB	2.31	2.29	0.091
40 dB	6.90	4.04	0.096
30 dB	20.92	7.03	0.38
20 dB	n/a	12.93	1.8
10 dB	n/a	21.86	3.8

Table 7.3.: Comparison of our method to the EVAM algorithm by Harikumar and Bresler [21, 20] and the MC-AM by Šroubek and Flusser [42]. Note that the comparison has to be taken with care, since from a visual inspection we suppose that the initial PMSE used in the EVAM and MC-AM examples is much higher. Additionally, our method was applied to a different setting (see text for details).



(a) Original image (b) First observation (50dB) (c) First observation (20dB) (d) First observation (10dB)

(e) Other observations (50dB) (f) Reconstruction for 50dB noise (g) Reconstruction for 20dB noise (h) Reconstruction for 10dB noise

(i) Original shift-variant PSF (j) Reconstructed PSF (50dB) (k) Reconstructed PSF (20dB) (l) Reconstructed PSF (10dB)

Figure 7.11.: Results of our blind-deconvolution method on an image of hippocampal rat neurons. We yield high-quality reconstructions of the images and the PSF. **(a)** The original image. **(b)(c)(d)** First observations for different noise levels (50dB, 20dB and 10dB respectively). Shift-variant filtering with a synthetic ground truth PSF is applied to the first observations. **(e)** Shows the other three observations. Here, the shift-variant PSF is applied after rotation and translation and subsequently uniform noise of the respective level is added. **(f)(g)(h)** Reconstruction computed by our method from the observations for the respective noise levels. **(i)** Ground truth shift-variant PSF used for generating the observations. The shift-variant PSF is composed of a non-isotropic Gaussian distribution for every image point, with strongly varying orientations. The single Gaussians are discretized by a 7×7 filter. Due to space limitation, only a certain area of the PSF is displayed. Note that in this small area the orientation of the Gaussians already changes by approximately 45 degrees. **(j)(k)(l)** Reconstructions of the PSF for the respective noise levels.

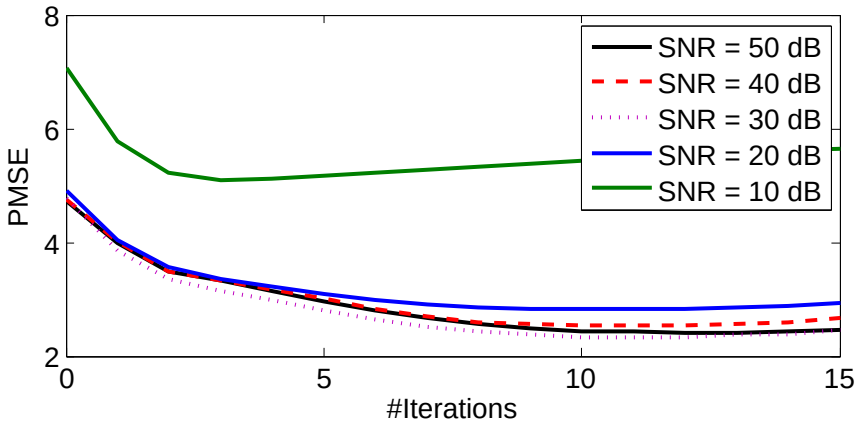
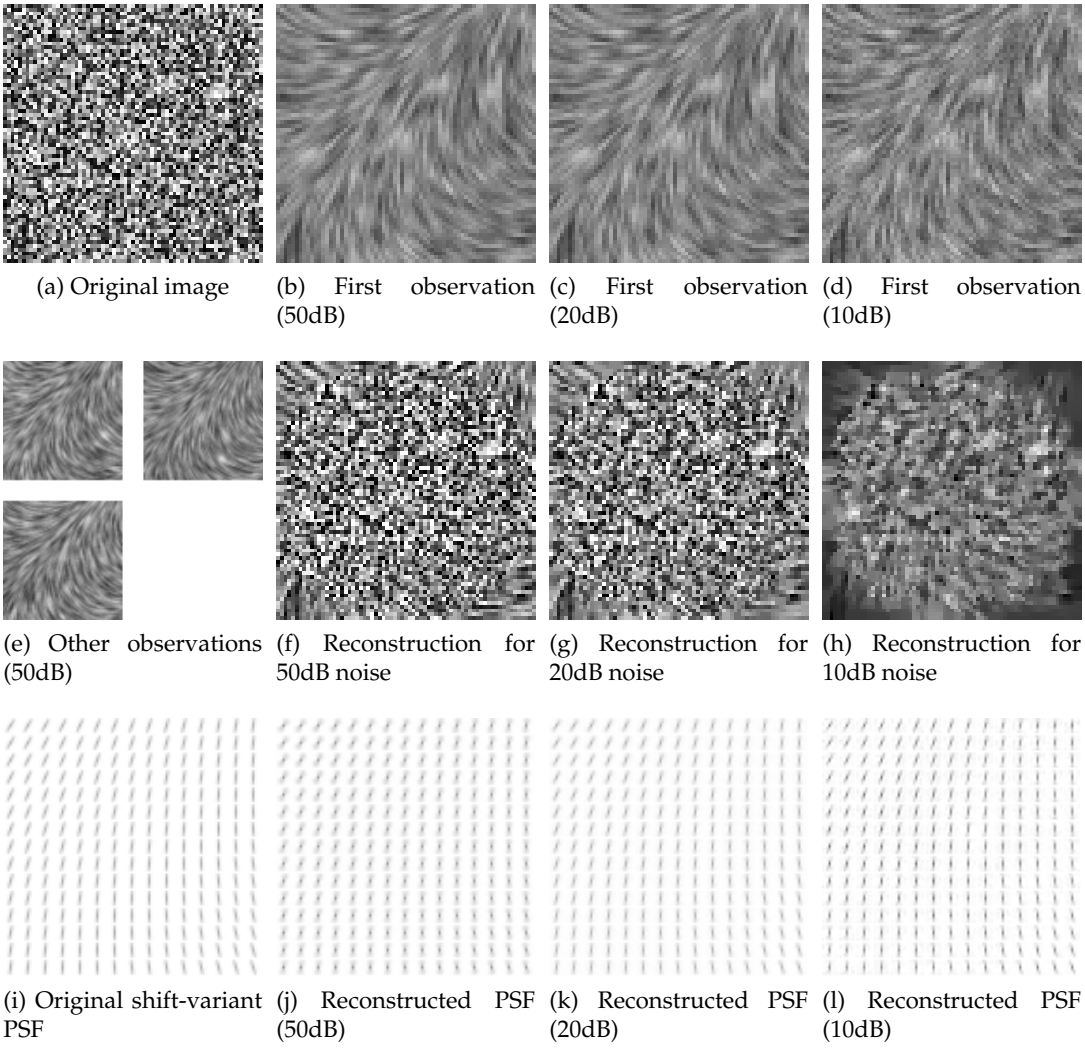


Figure 7.12.: Evaluation of the percentage mean squared error for the images shown in Figure (7.11). The 0-th iteration is the initial guess for the image.



(a) Original image (b) First observation (50dB) (c) First observation (20dB) (d) First observation (10dB)

(e) Other observations (50dB) (f) Reconstruction for 50dB noise (g) Reconstruction for 20dB noise (h) Reconstruction for 10dB noise

(i) Original shift-variant PSF (j) Reconstructed PSF (50dB) (k) Reconstructed PSF (20dB) (l) Reconstructed PSF (10dB)

Figure 7.13.: Results of our blind-deconvolution method for a random pattern. The process of image formation and the PSF is the same as in the previous neuron example.

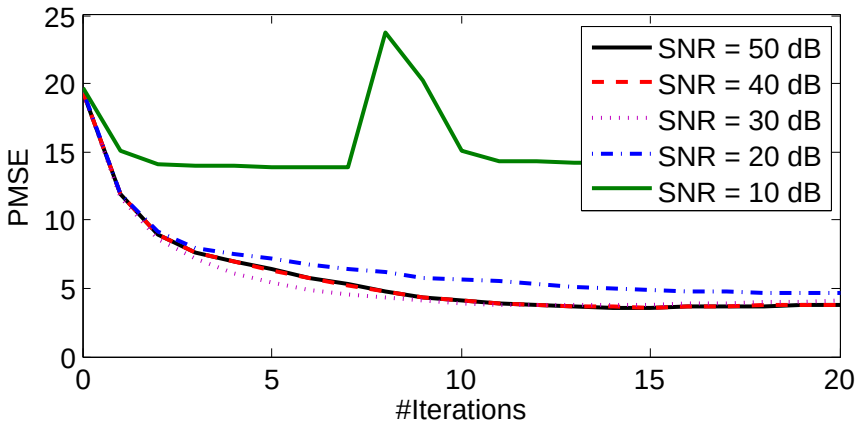
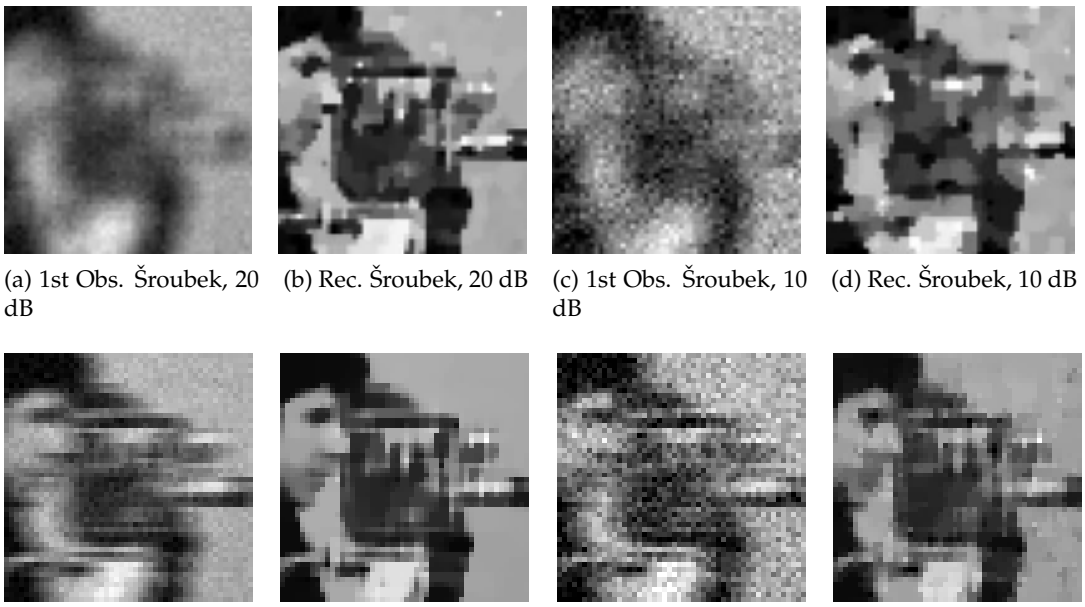


Figure 7.14.: Evaluation of the percentage mean squared error for the images shown in Figure (7.2). The 0-th iteration is the initial guess for the image.



(a) 1st Obs. Šroubek, 20 dB (b) Rec. Šroubek, 20 dB (c) 1st Obs. Šroubek, 10 dB (d) Rec. Šroubek, 10 dB

(e) 1st Obs. Our method, 20 dB (f) Rec. Our method, 20 dB (g) 1st Obs. our method, 10 dB (h) Rec. Our method, 10 dB

Figure 7.15.: Comparison of our method to the MC-AM by Šroubek and Flusser [42]. For the MC-AM algorithm, three observations are used in each case. Our method is applied on two observations, the second one rotated by $90°$. The percentage mean squared error between the original image and the our reconstructions is significantly better (compare Table (7.3)). However, it can be clearly seen that the degradation in **(a)** and **(c)** is heavier than in **(b)** and **(d)**.

8. Application on Medical Data

We apply our method on two different ultrasonic datasets. The first one is acquired by ourselves with a Siemens SONOLINE Antares © three-dimensional ultrasound machine, shown in Figure 8.1a. The second one is taken from Soler *et al.* [51].

8.1. Abdominal Phantom

The Siemens SONOLINE Antares © is a three-dimensional ultrasound machine that acquires a fan of two-dimensional planes, as shown in Figure 8.2a. Depending on the parameters, the acquisition takes a few seconds. Knowing the geometry of the setup, a three-dimensional image can be reconstructed.

From visual inspection of the reconstructed volumes, the in-plane resolution is significantly better than the out-of-plane resolution. Our assumption is that the reason for this lack of resolution is blurring by an anisotropic PSF that is extended in the direction perpendicular to the imaging planes.

In order to apply our deconvolution method, we acquire a second volume rotated by an angle of $90\,\mathrm{deg}$ about the axial direction. So, the poor out-of-plane resolution of the first volume can be compensated by the better in-plane resolution of the second volume.

For our experiments, we use an interventional 3D abdominal phantom by CIRS (see Figure 8.1b). The key advantage of a phantom over in-vivo data is to avoid problems that arise due to motion. Common kinds of motion that occur in human tissue are respiratory or cardiac motion. As we stated initially, our algorithm is able to deal with any kind of transformation, including non-rigid transformations. However, due to time constraints we were not able to implement a deformable registration and so we have to restrict on rigid transformations as they are given with the phantom.

We use a shift-invariant version of our algorithm in order to block-wise reconstruct the ultrasound volume. Though it would be better to use the shift-variant version, our implementation is not yet able to deal with such a huge amount of data. Figures 8.3 and 8.4 show the reconstruction of selected slices.

8.2. Breast Tissue

Soler *et al.* use a two-dimensional robotic platform, in which a linear transducer array is embedded. A series of slices is obtained by moving the transducer in one direction. In the same manner, another series of slices can be obtained at a different angle. Figure 8.5 shows reconstruction results for breast tissue of both Soler's method and our method.

(a) Siemens SONOLINE Antares ©

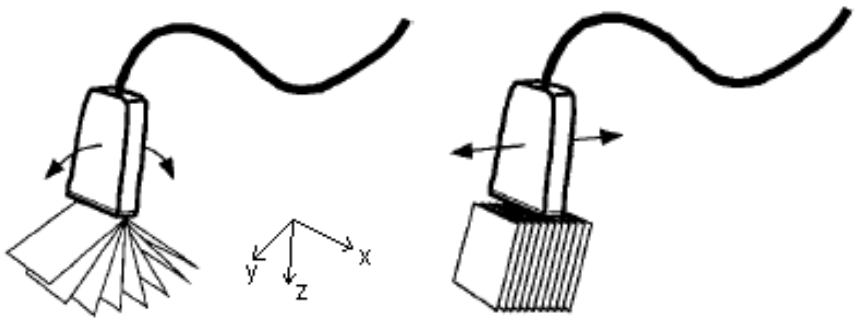
(b) Phantom

Figure 8.1.: The interventional 3D abdominal phantom by CIRS is scanned with the Siemens SONOLINE Antares ©

8.3. Discussion

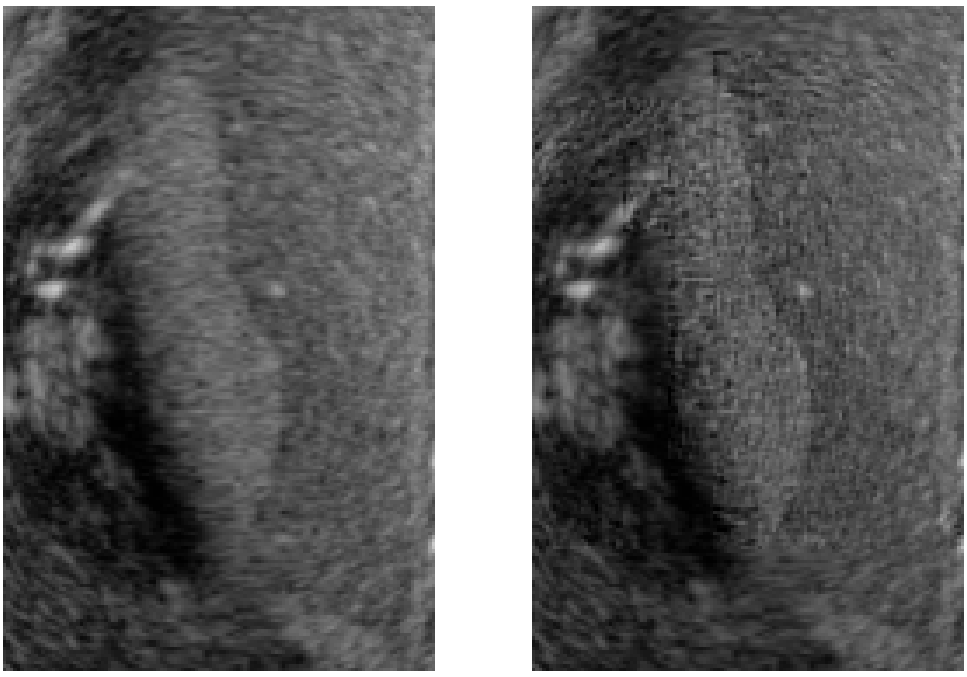
The results show clearly that the reconstructed images are significantly sharper than the original input image, and edges are more defined and better locatable. Whether this location is really correct can not be proved within the scope of this thesis due to time constraints.

With respect to the comparison of our results to the ones of Soler *et al.* , at this moment we are not able to give an evaluation since we just recently got aware of this publication.



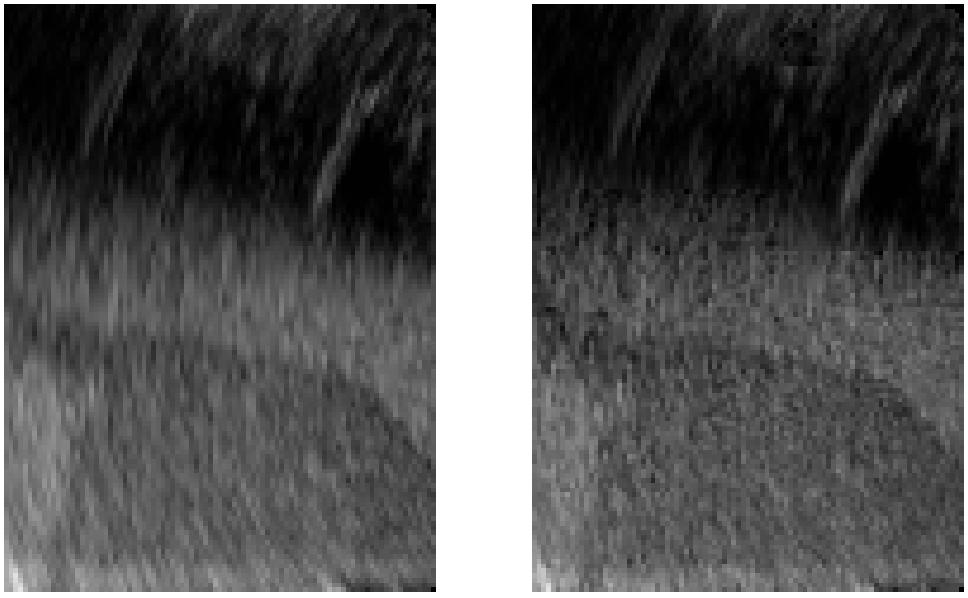
(a) Tilted scanning (b) Linear scanning

Figure 8.2.: Three-dimensional images can be gathered by different mechanical scanning approaches. In each case, two-dimensional planes are acquired and then combined to a volume.



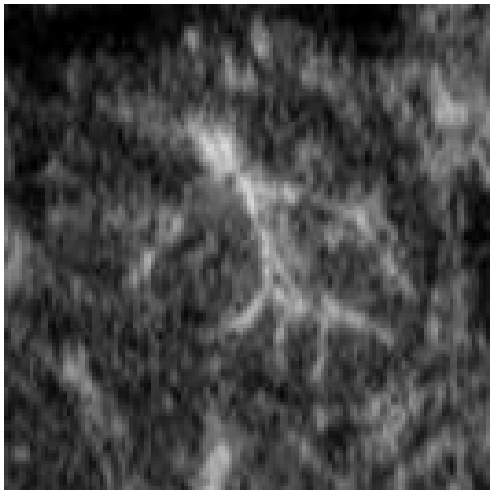
(a) First observation (b) Reconstruction

Figure 8.3.: Reconstruction of a selected x-y-slice of three-dimensional acquisition of abdominal phantom.

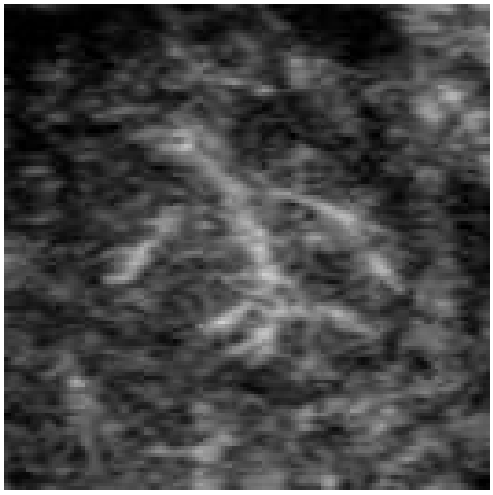


(a) First observation (b) Reconstruction

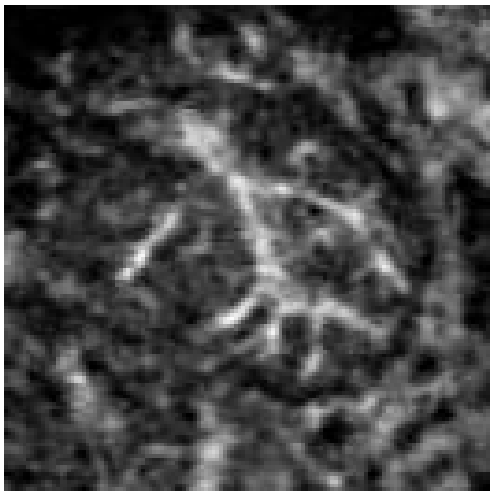
Figure 8.4.: Reconstruction of a selected y-z-slice of three-dimensional acquisition of abdominal phantom.



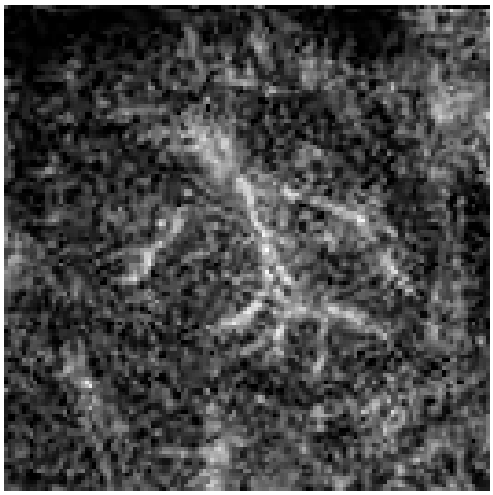
(a) First observation



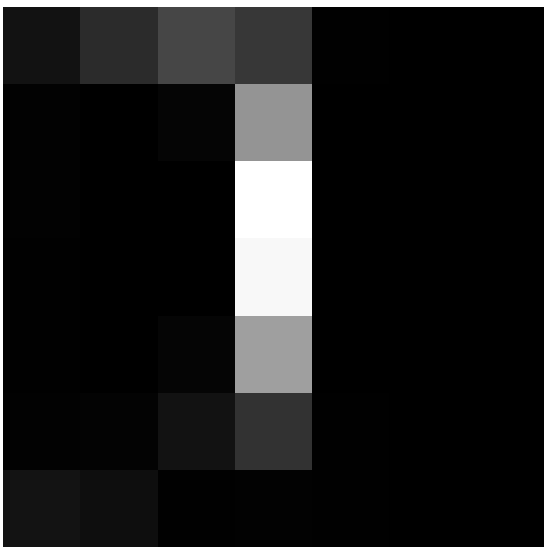
(b) Second observation



(c) Multi-View Reconstruction (Soler)



(d) Multi-View Reconstruction (our algorithm)



(e) Reconstructed PSF (our algorithm)

Figure 8.5.: Breast tissue acquired by a two-dimensional robotic platform constructed by Soler *et al.* . Two slices for two different scans transformed by an angle of $90\,\mathrm{deg}$ are used for reconstruction.

Conclusion and Future Work

Deconvolution is characterized by a tremendous lack of information. A very promising way to fill this lack is to use multiple differently blurred observations. In this thesis, we experimentally show that this even allows for blind shift-variant unparameterized deconvolution under the assumption that the transitions of the shift-variant point-spread-function are smooth.

Currently, our algorithm is implemented in Matlab. Matlab is very well suited for rapid prototyping, but it is difficult to be used for large datasets due to memory and performance limitations. So, in order to reconstruct a whole three-dimensional dataset we have to go back to the shift-invariant version of our method and reconstruct the volume block-wise.

In order to achieve a reconstruction of a whole three-dimensional volume without subdividing it in smaller blocks, our next step is a memory efficient implementation in a programming language like C++.

As mentioned, our method is not restricted to a certain imaging modality. We are currently starting collaborations in order to apply it to confocal microscopy.

Part III.

Appendix

A. From the Directional/Gâteaux Derivative to the Scalar Derivative

In chapter 2 we used a relationship between the directional derivative and a scalar derivative of a slightly modified cost function. We stated that

$$d_{\mathbf{x}}C(\mathbf{x};\nu) = \frac{d}{d\varepsilon}C(\mathbf{x}+\varepsilon\nu)\bigg|_{\varepsilon=0} \quad . \tag{A.1}$$

From the definitions of the scalar and directional derivative this is easy to see [65]:

$$d_{\mathbf{x}}C(\mathbf{x};\nu) = \frac{d}{d\varepsilon}C(\mathbf{x}+\varepsilon\nu)\bigg|_{\varepsilon=0} \tag{A.2}$$

$$\Leftrightarrow \lim_{\alpha\to 0}\frac{C(\mathbf{x}+\alpha\nu)-C(\mathbf{x})}{\alpha} = \lim_{\alpha\to 0}\frac{C(\mathbf{x}+(\varepsilon+\alpha)\nu)-C(\mathbf{x})}{\alpha}\bigg|_{\epsilon=0} \quad . \tag{A.3}$$

The same idea applies to the Gâteaux derivative [49]:

$$\delta_f\mathcal{C}[f;\eta] = \frac{d}{d\varepsilon}\mathcal{C}[f+\varepsilon\eta]\bigg|_{\varepsilon=0} \tag{A.4}$$

$$\Leftrightarrow \lim_{\alpha\to 0}\frac{\mathcal{C}[\mathbf{x}+\alpha\eta]-\mathcal{C}[\mathbf{x}]}{\alpha} = \lim_{\alpha\to 0}\frac{\mathcal{C}[\mathbf{x}+(\varepsilon+\alpha]\eta)-\mathcal{C}[\mathbf{x}]}{\alpha}\bigg|_{\epsilon=0} \quad . \tag{A.5}$$

B. Derivation of the Euler-Lagrange Equation

The Euler-Lagrange equations give a necessary but not sufficient condition for local optima of the functional

$$\mathcal{C}[x, f, f'] = \int_a^b F(x, f, f')\, dx \quad . \tag{B.1}$$

We would like to seek functions where the Gâteaux derivative of $\mathcal{C}$ is zero:

$$\delta_f \mathcal{C}[x, f, f'; \eta] \stackrel{!}{=} 0 \quad . \tag{B.2}$$

The following remarks are taken from [49]. From equation (2.17) we know that the Gâteaux derivative can be expressed as a scalar derivative of a slightly modified cost function. By making further use of the chain rule and linearity of differentiation, (B.2) can be rewritten as

$$0 \stackrel{!}{=} \frac{d}{d\varepsilon} \mathcal{C}[x, f + \varepsilon\eta, f' + \varepsilon\eta'] \Big|_{\varepsilon=0} \tag{B.3}$$

$$\Leftrightarrow 0 \stackrel{!}{=} \int_a^b \frac{\partial F}{\partial f}\eta + \frac{\partial F}{\partial f'}\eta'\, dx \tag{B.4}$$

$$\Leftrightarrow 0 \stackrel{!}{=} \int_a^b \frac{\partial F}{\partial f}\eta\, dx + \int_a^b \frac{\partial F}{\partial f'}\eta'\, dx \quad . \tag{B.5}$$

Now, the right term can be integrated by parts:

$$\int_a^b \frac{\partial F}{\partial f'}\eta'\, dx = \left[\frac{\partial F}{\partial f'}\eta\right]_a^b - \int_a^b \frac{d}{dx}\frac{\partial F}{\partial f'}\eta\, dx \tag{B.6}$$

Putting this back to (B.5) we yield

$$0 \stackrel{!}{=} \int_a^b \frac{\partial F}{\partial f}\eta\, dx + \left[\frac{\partial F}{\partial f'} \cdot \eta\right]_a^b - \int_a^b \frac{d}{dx}\frac{\partial F}{\partial f'}\eta\, dx \quad . \tag{B.7}$$

Since $\eta(a) = 0$ and $\eta(b) = 0$ this simplifies to

$$0 \stackrel{!}{=} \int_a^b \eta \left[\frac{\partial F}{\partial f} - \frac{d}{dx}\frac{\partial F}{\partial f'}\right] dx \quad . \tag{B.8}$$

Because η is an arbitrary perturbation we can now make use of the fundamental theorem of calculus of variations [49] and finally arrive at the Euler-Lagrange equation:

$$\frac{\partial F}{\partial f} - \frac{d}{dx}\frac{\partial F}{\partial f'} \stackrel{!}{=} 0 \ . \tag{B.9}$$

C. Gâteaux Derivative of $\mathcal{D}$

According to (2.17) The Gâteaux derivative can be expressed by a scalar derivative

$$\delta_u \mathcal{D}[u, h; \phi] = \left. \frac{\partial}{\partial \varepsilon} \mathcal{D}[u + \varepsilon\phi, h] \right|_{\varepsilon=0} \quad . \tag{C.1}$$

Then, the term (C.1) can be transformed as follows.

$$\sum_i \int_{\Omega_i} e_i(\xi) \left. \frac{\partial}{\partial \varepsilon} e(\xi) \right|_{\varepsilon=0} d\xi \tag{C.2}$$

$$= \sum_i \int_{\Omega_i} e_i(\xi) \int_{\Omega} h(\xi, \mathbf{s}) \phi(\varphi_i(\xi - \mathbf{s}))\, d\mathbf{s}\, d\xi \tag{C.3}$$

$$= \int_{\Omega} \sum_i \int_{\Omega_i} e_i(\xi) h(\xi, \mathbf{s}) \phi(\varphi_i(\xi - \mathbf{s}))\, d\xi\, d\mathbf{s} \tag{C.4}$$

Now, substituting $\varphi_i(\xi - \mathbf{s})$ by $\mathbf{x}$, the Gâteaux derivative is

$$\int_{\Omega} \left[\sum_i \int_{\Omega_i} e_i(\xi) h(\xi, \xi - \varphi_i^{-1}(\mathbf{x}))\, d\xi \right] \phi(\mathbf{x})\, d\mathbf{x} \quad . \tag{C.5}$$

Finally, yet another substitution leads to

$$\int_{\Omega} \left[\sum_i \int_{\Omega_i} e_i(\xi + \varphi_i^{-1}(\mathbf{x})) h(\xi + \varphi_i^{-1}(\mathbf{x}), \xi)\, d\xi \right] \phi(\mathbf{x})\, d\mathbf{x} \quad .$$

Bibliography

[1] Assmus. Early history of x rays. Technical report, Stanford Linear Accelerator Center, 1995.

[2] G. R. Ayers and J. C. Dainty. Interative blind deconvolution method and its applications. *Opt. Lett.*, 13:547–, 1988.

[3] M.R. Banham and A.K. Katsaggelos. Digital image restoration. *Signal Processing Magazine, IEEE*, 14(2):24–41, March 1997.

[4] L. Bar, N. Sochen, and N. Kiryati. Semi-blind image restoration via mumford-shah regularization. *Image Processing, IEEE Transactions on*, 15(2):483–493, Feb. 2006.

[5] H. Bially. Iterative behandlung linearer funktionalgleichungen. *Arch. Ration. Mech. Anal.*, 4:166–176, 1959.

[6] J. Biemond, R.L. Lagendijk, and R.M. Mersereau. Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883, May 1990.

[7] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[8] Encyclopaedia Britannica. *Encyclopaedia Britannica 2007 (Encyclopaedia)*. Encyclopaedia Britannica (UK) Ltd, 2007.

[9] M. Cannon. Blind deconvolution of spatially invariant image blurs with phase. *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on*, 24(1):58–63, February 1976.

[10] T.F. Chan and Chiu-Kwong Wong. Total variation blind deconvolution. *Image Processing, IEEE Transactions on*, 7(3):370–375, March 1998.

[11] P. H. Van Cittert. Zum einfluss der spaltbreit auf die intensitätsverteilung in spektrallinien ii. *Z. Physik*, 69:298–308, 1931.

[12] B. L. K. Davey, R. G. Lane, and R. H. T. Bates. Blind deconvolution of noisy complex-valued image. *Optics Communications*, 69(5, 6):353–356, 1989.

[13] A. M. Eskicioglu and P. S. Fisher. Image quality measures and their performance. *IEEE Transactions on Communications*, 43(12):2959–2965, December 1995.

[14] Leonhard Euler. *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive Solutio problematis isoperimetrici latissimo sensu accepti*. Lausannæ, Genevæ, apud Marcum-Michaelem Bousquet & socios, 1744.

[15] D. C. Ghiglia and L. A. Romero. Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America A*, 11:107 –, 1994.

[16] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, 2 edition, January 2002.

[17] Nick Gould, Dominique Orban, and Philippe Toint. Numerical methods for large-scale nonlinear optimization. *Acta Numerica*, 14:299–361, 2005.

[18] M.I. Gurelli and C.L. Nikias. Evam: an eigenvector-based algorithm for multichannel blind deconvolution of input colored signals. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 43(1):134–149, Jan. 1995.

[19] MI Hadamard. On problems in partial derivatives, and their physical significance. Technical report, Bull. Princeton Univ, 1902.

[20] G. Harikumar and Y. Bresler. Exact image deconvolution from multiple fir blurs. *Image Processing, IEEE Transactions on*, 8(6):846–862, June 1999.

[21] G. Harikumar and Y. Bresler. Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms. *Image Processing, IEEE Transactions on*, 8(2):202–219, Feb. 1999.

[22] D Huang, EA Swanson, CP Lin, JS Schuman, WG Stinson, W Chang, MR Hee, T Flotte, K Gregory, CA Puliafito, and al. et. Optical coherence tomography. *Science*, 254(5035):1178–1181, November 1991.

[23] Jan Huisken, Jim Swoger, Filippo Del Bene, Joachim Wittbrodt, and Ernst H. K. Stelzer. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305(5686):1007 – 1009, August 2004.

[24] Susumu Kikuchi, Kazuo Sonobe, and Nagaaki Ohyama. Three-dimensional microscopic computed tomography based on generalized radontransform for optical imaging systems. *Optics Communications*, 123:725–733, Feburary 1996.

[25] Susumu Kikuchi, Kazuo Sonobe, Daizo Shinohara, Nagaaki Ohyama andShinro Mashiko, and Yasushi Hiraoka. A double-axis microscope and its three-dimensional image position adjustment based on an optical marker method. *Optics Communications*, 129:237–244, August 1996.

[26] Susumu Kikuchi, Kazuo Sonobe, Liliek S. Sidharta, and Nagaaki Ohyama. Three-dimensional computed tomography for optical microscopes. *Optics Communications*, 107:432–444, May 1994.

[27] S. Kirkpatrick, Jr. Gelatt, C. D., and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[28] M. Kozubek, P. Matula, H. Eipel, and M. Hausmann. Automated multi-view 3d image acquisition in human genome research. In *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*, pages 91–98, 19-21 June 2002.

[29] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, May 1996.

[30] D. Kundur and D. Hatzinakos. A novel blind deconvolution scheme for image restoration usingrecursive filtering. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 46(2):375–390, February 1998.

[31] R. L. Lagendijk, A. M. Tekalp, and J. Biemond. Maximum likelihood image and blur identification - a unifying approach. *Optical Engineering*, Vol. 29:422–435, May 1990.

[32] L. Landweber. An iteration formula for fredholm integral equations of the first kind. *Am. J. Math. 1951.*, 73:615–624, 1951.

[33] R. G. Lane and R. H. T. Bates. Automatic multidimensional deconvolution. *Journal of the Optical Society of America A*, 4:180–188, 1987.

[34] M M Lavrentev. *Ill Posed Problems of Mathematical Physics*. American Mathematical Society, 1986.

[35] Petr Matula, Michal Kozubek, and Pavel Matula. Applications of image registration in human genome research. In *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*, Lecture Notes in Computer Science, pages 376–384, 2004.

[36] B. C. McCallum. Blind deconvolution by simulated annealing. *Optics Communications*, 75(2):101–105, February 1990.

[37] Noriaki Miura and Naoshi Baba. Extended-object reconstruction with sequential use of the iterative blind deconvolution method. *Optics Communications*, 89(5-6):375–379, May 1992.

[38] J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee. Modeling ultrasound imaging as a linear, shift-variant system. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 53(3):549–563, March 2006.

[39] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization (Springer Series in Operations Research)*. Springer, Berlin, 2000.

[40] J. Pawley. *Handbook of Biological Confocal Microscopy*. Springer, Berlin, 2006.

[41] David L. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. ACM*, 9(1):84–97, 1962.

[42] F. Šroubek and J. Flusser. Multichannel blind iterative image restoration. *Image Processing, IEEE Transactions on*, 12(9):1094–1106, Sept. 2003.

[43] F. Šroubek and J. Flusser. Multichannel blind deconvolution of spatially misaligned images. *Image Processing, IEEE Transactions on*, 14(7):874–883, July 2005.

[44] F. Šroubek and J. Flusser. Resolution enhancement via probabilistic deconvolution of multiple degraded images. *Pattern Recognition Letters*, 27(4):287–293, March 2006.

[45] F. Šroubek, J. Flusser, T. Suk, and S. Simberova. Multichannel blind deconvolution of the short-exposure astronomical images. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 49–52vol.3, 3-7 Sept. 2000.

[46] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[47] A.A. Sawchuk. Space-variant image motion degradation and restoration. *Proceedings of the IEEE*, 60(7):854–861, July 1972.

[48] D. Slepian. Restoration of photographs blurred by image motion. *Bell Syst. Tech. J.*, 46:2353–2362, Dec 1967.

[49] Donald R. Smith. *Variational Methods in Optimization*. Dover Publications, 1998.

[50] P. Soler, G. Delso, N. Villain, I. Bloch, and E. D. Angelini. Wavelet based spatial compounding for 3d ultrasound breast imaging.

[51] Pau Soler, Gaspar Delso, Nicolas Villain, Elsa Angelini, and Isabelle Bloch. Superresolution spatial compounding techniques with application to 3d breast ultrasound imaging. In Stanislav Emelianov and William F. Walker, editors, *Medical Imaging 2006: Ultrasonic Imaging and Signal Processing*, volume 6147 of *Proceedings of SPIE*, 2006.

[52] M. M. Sondhi. Image restoration: The removal of spatially invariant degradations. *Proceedings of the IEEE*, 60(7):842–853, July 1972.

[53] Jr. Stockham, T.G., T.M. Cannon, and R.B. Ingebretsen. Blind deconvolution through digital signal processing. *Proceedings of the IEEE*, 63(4):678–692, April 1975.

[54] Kurt Sätzler and Roland Eils. Resolution improvement by 3-d reconstructions from tilted views in axial tomography and confocal theta microscopy. *Bioimaging*, 5:171–182, 1997.

[55] Rangarajan K. Sundaram. *A First Course in Optimization Theory*. Cambridge University Press, 1996.

[56] J. Swoger, J. Huisken, and E. H. K. Stelzer. Multiple imaging axis microscopy improves resolution for thick-sample applications. *Optics Letters*, 28:1654–1656, 2003.

[57] A. Tekalp, H. Kaufman, and J. Woods. Identification of image and blur parameters for the restoration of noncausal blurs. *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on*, 34(4):963–972, August 1986.

[58] A. M Tekalp, H. Kaufman, and J. W. Woods. Model-based segmentation and space-variant restoration of blurred images by decision-directed filtering. *Signal Process.*, 15(3):259–269, 1988.

[59] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* John Wiley, 1977.

[60] H.J. Trussell and S. Fogel. Identification and restoration of spatially variant motion blurs in sequential images. *IEEE Transactions on Image Processing*, 1(1):123–126, January 1992.

[61] F. Tsumuraya, N. Miura, and N. Baba. Iterative blind deconvolution method using lucy's algorithm. *Astronomy and Astrophysics*, 282(2):699–708, Feb 1994.

[62] C.R. Vogel and M.E. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *Image Processing, IEEE Transactions on*, 7(6):813–824, June 1998.

[63] Zhou Wang and A. C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, March 2002.

[64] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[65] David V. Widder. *Advanced Calculus: Second Edition (Prentice-Hall Mathematics Series)*. Dover Publications, 1989.

[66] Ilu-Li You and M. Kaveh. A regularization approach to blind restoration of images degraded by shift-variant blurs. *IEEE*, pages 2607 – 2610, 1995.

[67] Yu-Li You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *Image Processing, IEEE Transactions on*, 5(3):416–428, March 1996.

[68] Yu-Li You and M. Kaveh. Blind image restoration by anisotropic regularization. *Image Processing, IEEE Transactions on*, 8(3):396–407, March 1999.

[69] Mehmet K. Özkan, A. Murat Tekaip, and M. Ibrahim Sezant. Identification of a class of space-variant image blur. In Mehmet R. Civanlar, Sanjit K. Mitra, and Robert J. Moorhead II, editors, *Image Processing Algorithms and Techniques II*, volume 1452 of *Proceedings of SPIE*, pages 146–156, June 1991.

[70] M.K. Özkan, A.M. Tekalp, and M.I. Sezan. POCS-based restoration of space-varying blurred images. *IEEE Transactions on Image Processing*, 3(4):450–454, July 1994.