Master's Thesis

Superresolution for a Hyperspectral Line-Scan Imaging System

Kurt Weingartmann, BSc

Institute for Computer Graphics and Vision
Graz University of Technology

Thesis supervisor: Univ.-Prof. Dipl.-Ing. Dr. techn. Horst Bischof
Thesis advisor: Univ.-Ass. Dipl.-Ing. Dr. techn. Thomas Pock
Graz, Austria, December 2013

Abstract

Hyperspectral imaging systems are well-suited to identify materials based on their material properties which are commonly used in several industrial separation processes. In contrast to RGB cameras, hyperspectral cameras has a higher spectral resolution and capture intensity information over a broader spectral waveband. This advantage means a deterioration in spatial resolution of a limited sensor size. The aim of this master thesis is to analyse the HELIOS camera based hyperspectral line scan imaging system and to develop a superresolution approach which improves the spatial and temporal image quality of the final high resolution images. For this, we assume that the acquisition width of a pixel of the line scan camera is lager than the distance covered by the conveyor belt per one camera acquisition step. This additional information about the overlapping area of two temporal adjacent pixels is incorporated into two maximum a posteriori approaches. They are solved by an efficient primal dual algorithm which offers a global solution for the two convex problems. The detailed evaluation of the performances with synthetic and real data shows that the developed algorithms are superior to the bicubic interpolation approach in several scenarios and metrics and underlines the practical usefulness of the proposed approaches.

Kurzfassung

Bei industriellen Matrialtrennungsprozessen kommen unter anderem bildgebende hyperspektral Systeme zum Einsatz, da sich diese zum Identifizieren von Materialien anhand ihrer Oberflächeneigenschafen eignen. Im Gegensatz zu herkömmlichen RGB-Kamerasystemen wird mit Hyperspektralkameras ein größerer elektromagnetischer Wellenbereich höher aufgelöst aufgenommen. Da optische Systeme durch ihre Sensorgröße beschränkt sind, führt die höhere spektrale Auflösung zur Verringerung der örtlichen Auflösung. Das Ziel dieser Masterarbeit ist es, den Aufnahmeprozess eines HELIOS Hyperspektral-Zeilen-Kamera-Systemes zu analysieren und einen Superresolution-Ansatz zu entwickeln, welcher die örtliche und zeitliche Auflösung und somit die Qualität des aufgenommenen Bildes erhöht. Zu diesem Zweck nehmen wir an, dass die örtliche Aufnahmebreite der Kamera auf dem Förderband für einen Aufnahmevorgang größer ist, als das Förderband in einem Aufnahmevorgang weiterbewegt. Dadurch enthalten zeitlich aneinandergrenzende Pixel teilweise Informationen vom selben Teilbereich. Diese zusätzliche Information wird verwendet, um zwei maximum a posteriori Ansätze zu entwicklen. Diese werden dann mithilfe des Primal-Dual-Algorithmuses gelöst, welcher eine globale Lösung für die beiden konvexen Problemstellungen liefert. Die detaillierte Auswertung der Performanz der Ansätze mit synthetischen und echten Datensätzen sowie verschiedenen Metriken zeigt, dass die entwickelten Superresolution-Methoden den bikubischen Interpolationsansatz übertreffen, und unterstreichen den praktischen Nutzen der vorgestellten Ansätze.

Acknowledgements

First of all, I have to thank my life partner who encouraged and supported me all the time. I would also like to thank my both sons, who let me write my work at home even though playing with dad is more fun. Another thank you goes to to Prof. Horst Bischof for supervising my master thesis. And last but not least, thanks to Dr. Thomas Pock for technical support and reviewing this thesis.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kentlich gemacht habe.

Graz, am..... .....

(Unterschrift)

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

....

date (signature)

Contents

1 Introduction **12**

1.1 Motivation and Problem Description 12
1.2 Organisation of the Master's Thesis 13
1.3 Spectral Imaging 14
1.3.1 Spectroscopy 14
1.3.2 Remote Imaging 14
1.3.3 Hyperspectral Imaging 15
1.3.4 Spectral Data Representation 15
1.4 A Hyperspectral Line-Scan Imaging System 16
1.5 High Resolution Images 18
1.5.1 Why High Resolution Images? 18
1.5.2 Increasing Resolution by Hardware Modification 18
1.5.3 Increasing Resolution by Image Processing Methods . . . 20
1.6 Superresolution 22
1.6.1 Superresolution, a Higher Image Restoration Problem . . 22
1.6.2 General Superresolution Model 24

2 Related Work **26**

2.1 Frequency Domain Approach 27
2.2 Nonuniform Interpolation Approach 28
2.3 Projection onto Convex Sets (POCS) 29
2.4 Regularized Superresolution Approaches 31
2.4.1 Constrained Least Square 31
2.4.2 Statistical Approaches 32
2.5 Iterative Backprojection Approach 34
2.6 Superresolution for Hyperspectral Images 34

2.6.1 Image Fusion 35
2.6.2 Spectral Unmixing 35

3 HELIOS Superresolution Approach 38
3.1 Preliminaries 38
3.1.1 Definitions 38
3.1.2 Total Variation (TV) 39
3.2 HELIOS Image Acquisition Process 40
3.3 Assumptions for the Superresolution Model 42
3.4 Discrete Setting 44
3.4.1 Discretisation of a continuous Image 44
3.4.2 Discrete Images $f^{(d)}$ and $f^{(u)}$ 45
3.5 HELIOS Superesolution Model 45
3.6 Discrete Convex Optimisation Problem 46
3.6.1 Downsampling Operator 47
3.6.2 Blurring Operator 49

4 Algorithms 51
4.1 Preliminaries 51
4.1.1 Definitions 51
4.1.2 Duality 53
4.1.3 Conjugate Gradient Method 54
4.1.4 General Primal-Dual Algorithm 54
4.2 ROF Superresolution 57
4.2.1 From Primal to Primal-Dual 57
4.2.2 ROF Proximal Operators 58
4.2.3 Iterative ROF Superresolution Primal-Dual Algorithm . . 59
4.2.4 Step Size 59
4.3 TVL_1 Superresolution Primal-Dual 61
4.3.1 From Primal to Primal-Dual 61
4.3.2 TVL_1 Proximal Operators 61
4.3.3 Iterative L_1 Superresolution Primal-Dual Algorithm . . . 62
4.3.4 Step Size 62

5 Experimental Results **64**
5.1 Environment 64
5.2 Synthetic Data Evaluation 65
5.2.1 Evaluation Metrics 66
5.2.2 Hyperspectral Metrics 66
5.2.3 Performance with varying Shifting Factor 67
5.2.4 Performance with varying Sampling Rate 68
5.3 Real Data Evaluation 69
5.3.1 Camera System Setup 69
5.3.2 Calibration Pattern 72
5.3.3 Determining Shifting Factor s 73
5.3.4 Determining λ 76
5.3.5 Determining the Blur Kernel 76
5.3.6 Real Data Evaluation Results 78
5.4 Summary 82

6 Conclusion and Outlook **83**
6.1 Conclusion 83
6.2 Outlook 84

List of Figures

1.1 The Electromagnetic Spectrum 15
1.2 Data Cube 16
1.3 Data Cube: Spectral Data from a Specific Location 17
1.4 Data Cube: Spatial Data from a Specific Spectral Band 17
1.5 HELIOS Camera 19
1.6 HELIOS Camera-Based Material Separation System 20
1.7 Image Acquistion System 21
1.8 Superresolution Model 23
1.9 General Superresolution Model 24

2.1 Projection onto Convex Sets 30
2.2 Endmember-based TV vs. Quantum TV 37

3.1 HELIOS Image 41
3.2 Line-Scan Camera Acquisition Process 43
3.3 Downsampling 48

4.1 Proximal Operator 52
4.2 Legendre-Fenchel Transform 53
4.3 Duality 54

5.1 Synthetic Reference Image 65
5.2 Shifting Factor Evaluation 67
5.3 Sampling Rate Evaluation 68
5.4 Material Signature 73
5.5 Calibration Grid 74
5.6 Captured Grid 75

5.7 Shifting Factor Estimation 77
5.8 Point Spread Function Estimation 78
5.9 Real Data Evaluation 80
5.10 Real Data Evaluation 81

List of Tables

5.1 Noise Free Synthetic Data Results (Pompoms) 70
5.2 Noise Free Synthetic Data Results (Beads) 71
5.3 Shifting Factors 76
5.4 CLS Blur Kernel 79
5.5 Gaussian Blur Kernel 79

Chapter 1

Introduction

Contents

1.1	Motivation and Problem Description	12
1.2	Organisation of the Master's Thesis	13
1.3	Spectral Imaging	14
1.4	A Hyperspectral Line-Scan Imaging System	16
1.5	High Resolution Images	18
1.6	Superresolution	22

1.1 Motivation and Problem Description

With the beginning of the space age in the 1950s and 1960s and the possibility to save information digitally, the idea of observing the earth from space was born in order to obtain information about the earth's renewable and nonrenewable resources. For this, the spectral imaging was initially developed which gather information from the electromagnetic energy field arising from the earth and the occurring spectral and spatial variations in this field. This allows to drawing conclusions about the earth's surface and the materials in this observed field. [24] This technology was refined over the years, and over the last years spectral imaging has become more important in industrial application. The reason for this trend is, that spectral imaging has the ability to analyse the surface materials of different objects, which is a central problem in various and especially industrial fields. For example in the food industry, spectral imaging can

be used to distinguish between ripe and bad fruits because the electromagnetic energy received from ripe and bad fruits differs. A difference in the reflected electromagnetic energy can also be observed in the recycling industry to distinguish between certain types of plastic, which allows to separate them although they may look the same. The centrepiece of all spectral imaging systems is a sensor which captures detailed spectral information about the scene in addition to spatial information. Thus, conclusions can be drawn regarding the materials. Because of limitations in optics and sensor size, spectral imaging systems have generally a relatively low spatial resolution compared to classical optical imaging systems. This is a major drawback because especially for separation processes in industrial application, a high spatial resolution is needed to achieve an accurate result. In order to find a possibility to overcome the low resolution drawback, the purpose of this master thesis is to analyse the image acquisition process of the HELIOS[1] camera based hyperspectral line-scan imaging system, which is used for industrial application, and to develop a superresolution algorithm, which increases the spatial resolution in order to get higher resolved hyperspectral images with better quality.

1.2 Organisation of the Master's Thesis

The following section 1.3 provides an overview of spectral imaging, including the explanation of the terms spectroscopy (section 1.3.1), remote imaging (section 1.3.2), hyperspectral imaging (section 1.3.3) and the definition of the typical spectral imaging data representation called data cube (section 1.3.4). Section 1.4 is concerned with the hyperspectral line-scan imaging system HELIOS and its image acquisition process. Section 1.5 explains why high resolution images are desirable for many imaging applications. The last section in this chapter gives an introduction to the widely used superresolution approach and presents the general superresolution model which forms the basic model for our HELIOS superresolution approach discussed in chapter 3

Chapter 2 gives an overview of different superresolution and hyperspectral superresolution approaches over the last few decades.

In chapter 3 the idea of variational superresolution for a specific hyperspectral line-scan imaging system is introduced and the pros and cons of diverse

[1] HELIOS is a trademark of the company EVK.

variations from the approach are discussed.

In chapter 4, the general primal dual algorithm is introduced which is an efficient method to solve a class of problems including variational superresolution approaches. In section 4.2 and 4.3, the concrete implementation and the convergence of our proposed algorithms are discussed.

Chapter 5 contains all the information about our experimental results, including the description of our test environment in section 5.1 and information about the used metrics, which are discussed in detail in section 5.2.1.

Finally, chapter 6 draw some conclusions about the results and discuss the possibilities for future work.

1.3 Spectral Imaging

Spectral Imaging in general can be seen as a fusion of spectroscopy and remote imaging technologies. Through the combination of these complex sensors, so-called imaging spectrometers are developed which are able to collect multispectral and spatial information.

1.3.1 Spectroscopy

Historically, spectroscopy has always been concerned with the dispersion of (visible) light into its component colors through e.g. a prism. Nowadays, the term describes the study of the interaction from electromagnetic radiation with matter and the variance in energy with wavelength [27]. In figure 1.1, the electromagnetic spectrum with its varying wavelength is depicted. In terms of optic remote sensing such as earth remote sensing, the task of spectroscopy is to measure the reflected sunlight from materials which are on the earth's ground. For this, a so-called spectrometer disperse reflects light in many narrow wavebands and measures the intensity of incident energy for each waveband.

1.3.2 Remote Imaging

Remote imaging generally describes the process of collecting visual information about an object without contacting it. Since the objective of this method is to gather as much visual information as possible, the spatial resolution is traditionally. As a consequence of sensor size limitations, remote imaging systems use

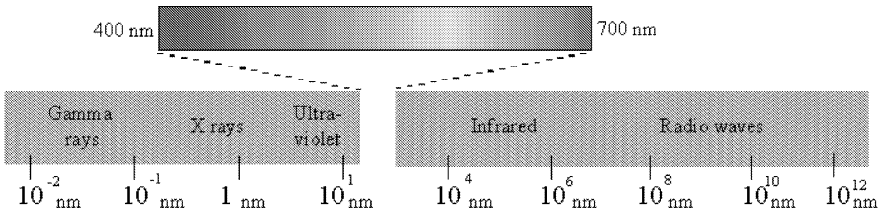


Figure 1.1: The electromagnetic spectrum (taken from [39])

one or only a few relatively broad spectral wavebands [26]. A typical example of remote imaging systems can be found in the waste industry in which different materials are separated. For this task, remote image systems uses as usual RGB cameras. They acquire the information about the visual light spectrum in three spectral bands (red, green and blue) about an area to distinguish the materials by colour. But there are a many applications where different materials have similar or the same color. In these cases, RGB camera-based systems fails. For these applications, (hyper)spectral imaging systems, like HELIOS camera-based systems (see 1.4), are an alternative to solve this kind of problem because they combine the multispectral approach from the spectroscopy with the spatial information from remote imaging technology.

1.3.3 Hyperspectral Imaging

As described in the previous sections, spectral imaging systems are able to handle many spectral wavebands. A particular class of such systems is based on so-called hyperspectral sensors. This sensor mostly divides a certain waveband into hundreds of continuous narrow bands. According to [39], it is not the number of wavebands but rather it is the continuous and narrowness of the measured wavelengths which defines a sensor as hyperspectral. For more information, we refer to [39, 27]

1.3.4 Spectral Data Representation

Typically, data from spectral imaging system have three (two spatial and one spectral) dimensions. For this, a data structure called data cube is a common instrument to visualise such data. Each dimension is an axis of the data cube and the intensity is represented by colour. In figure 1.2, a data cube is shown.

According to Manolakis et al. in [26], if we extract all voxel values from the

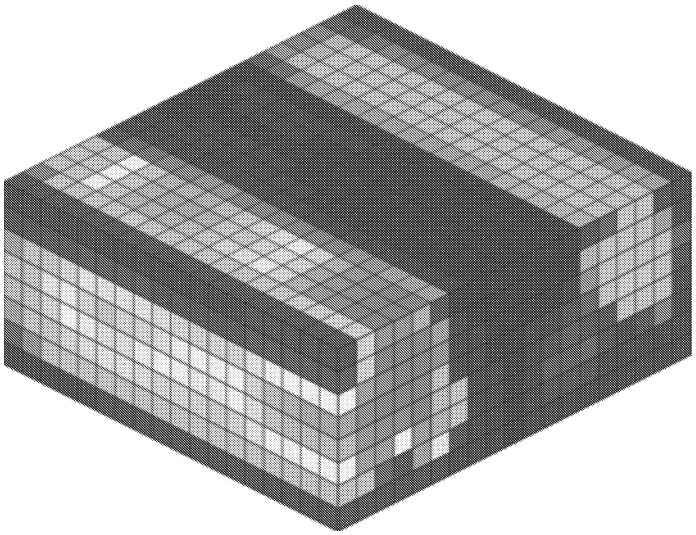
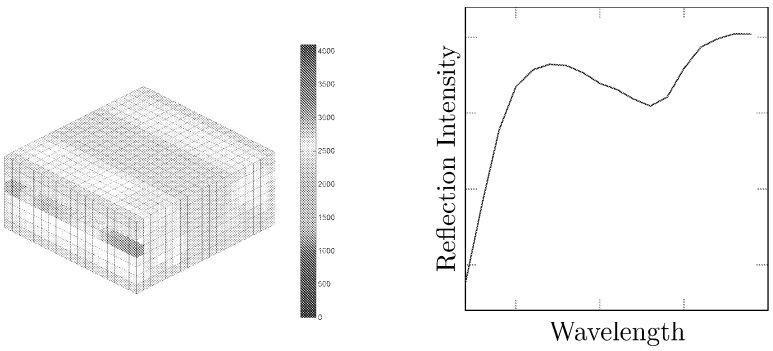


Figure 1.2: Spectral data represented by a data cube

same spatial location and plot them as a function of wavelength, we receive the average spectrum from the materials in the corresponding area. Figure 1.3 shows this relationship. Contrary, if we pick all values in the same waveband, we obtain a grayscale image which represents the spatial distribution of the reflection intensity of the observed scene in the chosen waveband. This is shown in figure 1.4.

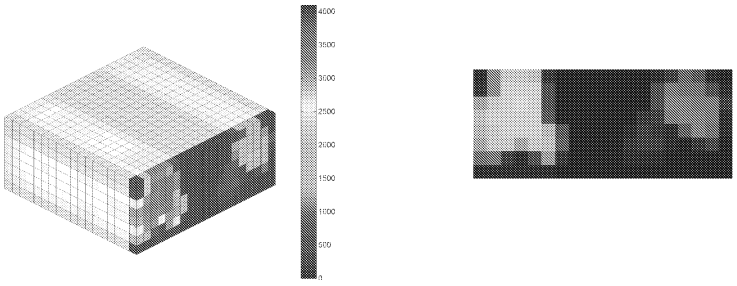
1.4 A HELIOS Camera-based Hyperspectral Line-Scan Imaging System

This section introduces the HELIOS camera-based hyperspectral line-scan imaging system, which is a trademark of EVK. The HELIOS camera-based systems are commonly used for separating objects based on their surface material properties. For this, such a system uses a line-scan technique where only one line is captured in each acquisition step. Therefore, the data collected by the HELIOS camera has a spatial, spectral and temporal dimension in which the spatial dimension is represented through the captured row and a temporal dimension denotes the point of time. Figure 1.5 shows the layout and the functional prin-



(a) **(b)**

Figure 1.3: Figure 1.3a shows the spectral data selection from a specific spatial location and figure 1.3b shows the selected data plotted as a function of wavelength.



(a) **(b)**

Figure 1.4: Figure 1.4a shows the spatial data selection from a specific spectral band and figure 1.4b presents the selected data as a grayscale image.

ciple of a HELIOS camera. This camera system is usually used in combination with a conveyor belt which transports the materials and a radiation source illuminating the scene. Figure 1.6 shows a typical HELIOS camera-based material separation system. The used radiation source are typically halogen headlights because the emitted energy has a relatively constant intensity throughout the electromagnetic spectrum. This energy interacts with the surfaces of the materials on the conveyor belt. Caused by this interaction, parts of the energy will be absorbed and reflected. Thus, the electromagnetic spectrum from the reflected energy differs from the emitted energy. Finally, the camera sensor converts the received energy into digital measurements.

1.5 High Resolution Images

In this section, we discuss why high resolution is necessary for imaging systems and which possibilities exist to archive them.

1.5.1 Why High Resolution Images?

In many imaging applications, high resolution images are required especially in medical, military and satellite imaging systems. For example, in medical images systems like magnetic resonance tomography (MRT), the exact position and size of a tumor can only be found if the spatial resolution of the MRT image is sufficiently high. There are several ways to enhance the spatial resolution of imaging systems which are discussed in the following.

1.5.2 Increasing Resolution by Decreasing the Pixel Size on Sensor or Chip

The most direct approach is to increase the number of pixels on the sensor by decreasing the pixel size. As a result, the available light per pixel decreases. This circumstance leads to quality loss because noise generally increases if light decreases. So it is obvious that there are limitations to increase the amount of pixels on the sensor. The other possibility is to increase the capacity of the sensor by increasing the chip size. But this approach has problems with the speed-up of the charge transfer rate. Hence, there are again limitations for

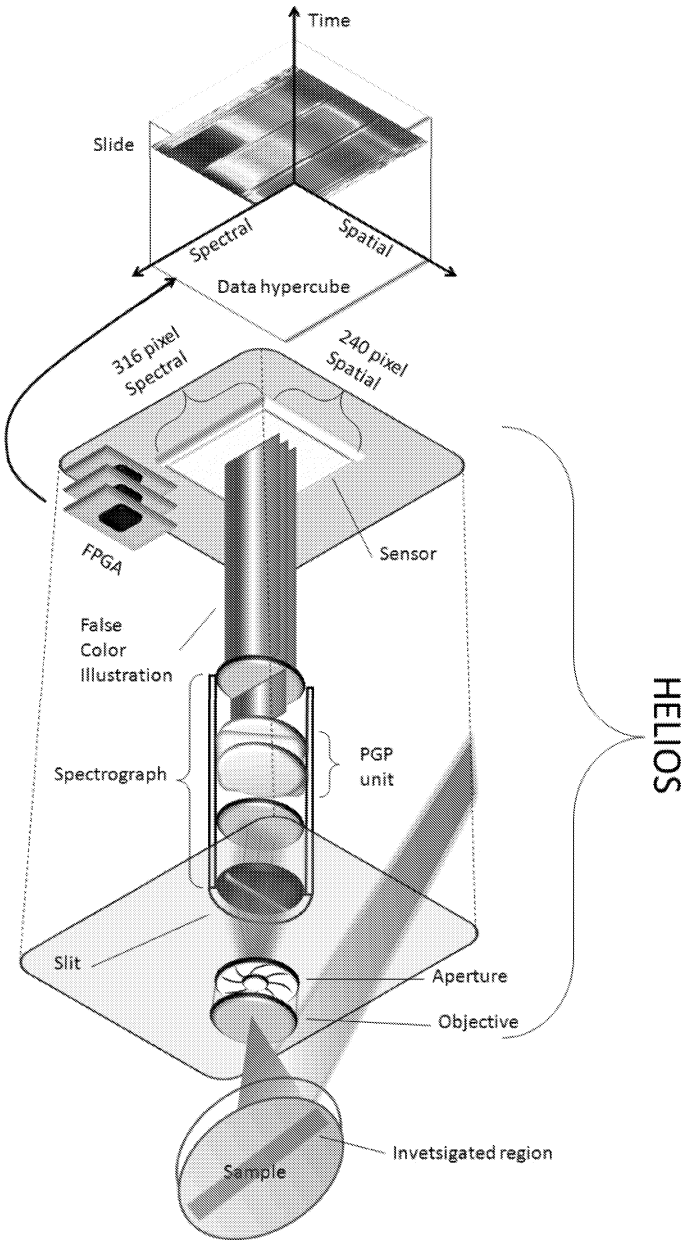


Figure 1.5: Layout of the hyperspectral line-scan imaging camera called HELIOS [29]

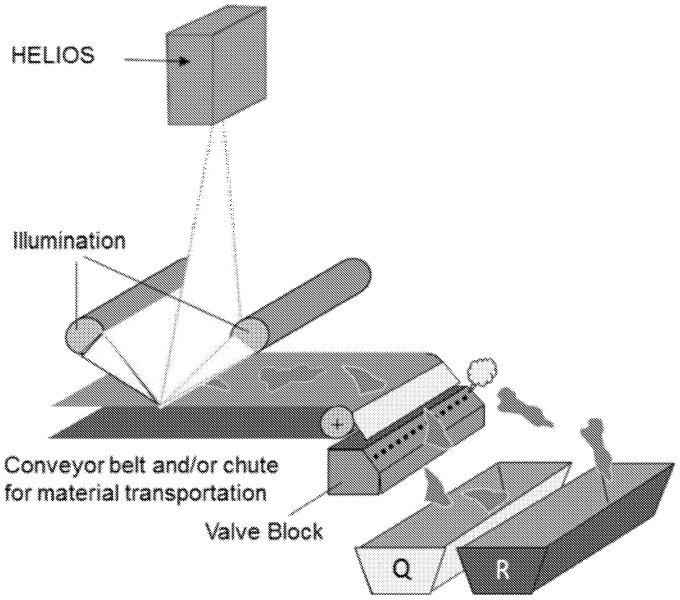


Figure 1.6: Layout of a HELIOS camera-based material separation system [25]

increasing the resolution. Furthermore, highly accurate optics and sensors imply high costs, costs always being an important factor for industrial applications.

1.5.3 Increasing Resolution by Image Processing Methods

Another possibility to increase the pixel density of an image is to use image processing methods which enhance the resolution by applying mathematical methods. Generally, one can distinguish between two different approaches called interpolation and superresolution. These two approaches are briefly discussed in the following subsections.

1.5.3.1 Superresolution

A widely used approach to overcome the limitations in optics and sensor size is to create a high resolution image from several low resolution images. This approach is called "Superresolution", "High Resolution Reconstruction" or "Resolution Enhancement" and has been well studied in the last years. The basic idea for this approach is that multiple slightly different images captured from the same

scene contain more information than one image alone. For this approach, the information from the multiple images are combined to form a higher resolution image. It is obvious that more images contain more information, and so, the quality of the enhanced resolution image depends on the number of available images. A further aspect of the high resolution reconstructed image quality is the quality of the low resolution images. In the digital image acquisition process, there typically are several factors which result in quality loss. These factors are optical distortion, aliasing, motion blur and noise. Figure 1.7 schematically shows the image acquisition process. Methods in image processing which are

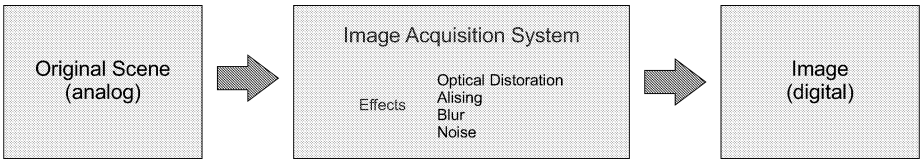


Figure 1.7: Image acquisition system

attempts to solve this problem are called "Image Restoration Methods". These methods try to overcome the problems which occur during the digital image acquisition process. In contrast to superresolution methods, image restoration methods only try to restore a blurred, aliased and noisy image and do not change the size of the image. From this point of view, superresolution methods can be seen as a higher-level image restoration problem because they solve the image restoration problem and increase the spatial resolution.

1.5.3.2 Interpolation

The other approach which increases the size of the image is called "Image Interpolation" and in contrast to the superresolution approach, it obtains the information from a single image. This implies that image interpolation methods have less information about the scene, so they cannot be as good as superresolution methods. As a result, they cannot restore high frequencies lost during the image acquisition process. Therefore, they cannot be referred as superresolution or high resolution reconstruction methods, and it is obviously that such methods are not as powerful as superresolution methods.

1.6 Superresolution

In this section, we focus on the superresolution approach and the underlying model, which forms the basement of the following chapters.

1.6.1 Superresolution, a Higher Image Restoration Problem

As stated in section 1.5.3.1, superresolution combines the two image processing techniques image resolution enhancement and image restoration to generate a discrete high resolution image $f^{(u)}$ with size $M^{(u)} \times N^{(u)}$ from K discrete low resolution images $f_i^{(d)}$ where $i = 1..K$ and $K \geq 2$ with size $M^{(d)} \times N^{(d)}$ where $M^{(u)} > M^{(d)}$ and $N^{(u)} > N^{(d)}$.

1.6.1.1 Image Resolution Enhancement

Image resolution enhancement methods try, as the name already suggests, to enhance the resolution of a image. For this, they need additional information because increasing the pixel density without further information results in a dramatic loss of quality. Thus, the basis assumption in image resolution enhancement or rather superresolution methods is the availability of two or more low resolution images $f_i^{(d)}$ from the same scene which contains slightly different information. This means that the images have subpixels shifts among themselves because if the images are shifted in exact integer steps, there is no additional information about the scene in the overlapping area, and so they are useless. The low resolution images $f_i^{(d)}$ are then registrated to each other or the movement between the images are calculated. Afterwards, the collected information is used to generate a high resolution image $f(u)$. In figure 1.8, this basic superresolution approach is shown. Obviously, the quality of the reconstructed high resolution image $f^{(u)}$ depends on the number and the quality of the low resolution images $f_i^{(d)}$. Thus, superresolution methods additionally use image restoration methods to improve the quality of $f^{(u)}$.

1.6.1.2 Image Restoration

By reducing or removing effects caused by an image acquisition process (e.g. blurring, optical disortation or noise) from the degraded image, restoration

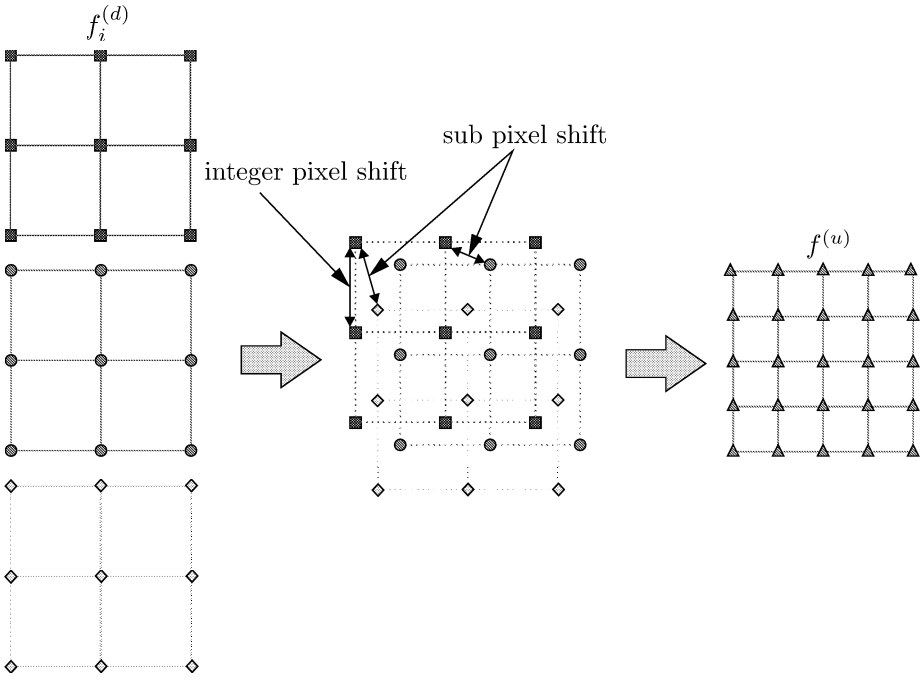


Figure 1.8: This figure shows the fundamental resolution enhancement or rather superresolution approach [30]. The pictures on the left are multiple low resolution images $f_i^{(d)}$ from the same scene with sub pixel shifts. These images are registered to each others in the registration step. After that, a superresolution image $f^{(u)}$ from the scene is created which has a higher pixel density than the multiple low resolution images.

methods increase the quality. In this context, superresolution methods can be seen as a higher level image restoration problem because they restore the image and simultaneously increase the resolution of the image by boosting the pixel density.

1.6.2 General Superresolution Model

As described in [30, 28], the first step for a general superresolution is to analyse the general image formulation process to obtain a model which relates the high resolution image $f^{(u)}$ to the low resolution images $f_i^{(d)}$. The high resolution image $f^{(u)}$ is an ideal image (without any degradation) which is sampled from the continuous scene. This image $f^{(u)}$ is the origin of the image formulation process which leads to multiple low resolution images $f_i^{(d)}$. A particular low resolution image $f_i^{(d)}$ originates from a high resolution image $f^{(u)}$ by passing the steps warping, blurring, downsampling and at least by adding noise. So, as described before, this process is called image formulation process and can be denoted as the forward problem. The inversion of this process is an ill-posed problem and is called superresolution. In figure 1.9, the image formation process and the general superresolution model are schematically shown.

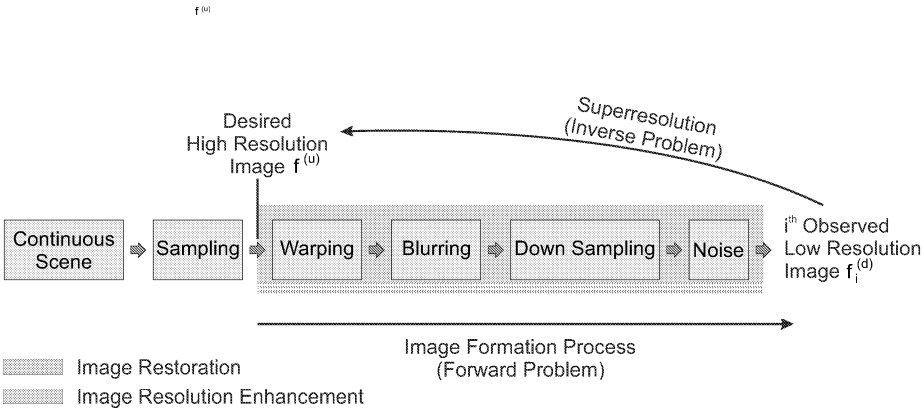


Figure 1.9: This figure shows the general superresolution model which is the inversion of the image formation process. The red and the blue shaded areas show the assignment to the Image Restoration or Image Resolution Enhancement Methods. [30]

1.6.2.1 Mathematical Formulation

As stated by Mizel et al. in [28] and Park in [30], the general superresolution model is commonly described by a series of linear transformation. With using matrix-vector notation, the model can be written in the following form

$$f_i^d = D_i B_i W_i f^u + \eta_i \tag{1.1}$$

where D, B and W represent the operators for downsampling, blurring and warping and η_i describes the noise. Since D_i, B_i and W_i are linear operators, they can be combined to one linear operator A_i so that the equation (1.1) becomes

$$f_i^d = A_i f^u + \eta_i \tag{1.2}$$

where $A_i = D_i B_i W_i$.

Chapter 2

Related Work

Contents

2.1 Frequency Domain Approach 27

2.2 Nonuniform Interpolation Approach 28

2.3 Projection onto Convex Sets (POCS) 29

2.4 Regularized Superresolution Approaches 31

2.5 Iterative Backprojection Approach 34

2.6 Superresolution for Hyperspectral Images 34

As we discussed in chapter 1, one of the major tasks in computer vision is creating a higher resolution image from two or more low resolution images which contain slightly different information about the same scene. This task is called superresolution and is a very active area because it is possible to overcome limitations in optics and sensors by combining the information about the scene from the low resolution images. In several industrial, medical, military and science applications a high resolution is necessary to locate objects or generally to obtain information of the scene, so for this applications, superresolution may be a possibility to improve the performance. In the following, we will give a general outline of superresolution approaches and their extensions based on the general superresolution model in section 1.6 proposed in recent years.

2.1 Frequency Domain Approach

The first superresolution approach dates back to Huang and Tsay [19], presented a method which reconstructs a high resolution image from several low resolution images. For this, they use a frequency domain approach which is also described by Yang et al. [46] and Park et al. [30]. The basic principles of this method are

- the shifting property of the continuous fourier transform (CFT) and
- the aliasing property between the CFT and discrete fourier transform (DFT)

By defining a continuous image as $g(k,l)$, shifts between several continuous images can be described by

$$g_i(k,l) = g(k + \Delta_{i1}, l + \Delta_{i2}) \tag{2.1}$$

where $i = 1, 2 \dots K$ and Δ_{i1} and Δ_{i2} denote the global shift in horizontal and vertical direction which are assumed to be known. If $\mathcal{G}(u,v)$ is the CFT from the continuous image $g(k,l)$, the shifted CFT images can be calculated by the CFT shifting property, which leads to

$$\mathcal{G}_i(u,v) = e^{j2\pi(\Delta_{i1}u + \Delta_{i2}v)}\mathcal{G}(u,v) \tag{2.2}$$

For the low resolution images, impulse sample periods ζ_1 and ζ_2 are defined. Hence, the observed low resolution images $f_i^{(d)}[m,n]$ are related to the continuous images $g_i(k,l)$ by

$$f_i^{(d)}[m,n] = g_i(\zeta_1 m + \Delta_{i1}, \zeta_2 n + \Delta_{i2}) \tag{2.3}$$

where $m = 0, 1 \dots M-1$ and $n = 0, 1 \dots N-1$. By exploiting the aliasing relationship between the CTF and the DFT, the DFT of a low resolution image which is defined by $\mathcal{F}_i^{(d)}[x,y]$ can be represented in the following way

$$\mathcal{F}_i^{(d)}[x,y] = \frac{1}{\zeta_1\zeta_2}\sum_{j_1=-\infty}^{\infty}\sum_{j_2=-\infty}^{\infty}\mathcal{G}_i\left(\frac{2\pi}{\zeta_1}\left(\frac{x}{M} - j_1\right), \frac{2\pi}{\zeta_2}\left(\frac{y}{N} - j_2\right)\right) \tag{2.4}$$

By assuming that $\mathcal{G}(u,v)$ is band-limited, which means that $|\mathcal{G}(u,v)| = 0$ for $|u| \geq \frac{M\pi}{\zeta_1}$,$|v| \geq \frac{N\pi}{\zeta_2}$, the linear equation system

$$\mathcal{F}^{(d)} = A\mathcal{G} \tag{2.5}$$

can be determined by incorporating (2.2) into (2.4) and expressing it in matrix form. For archiving a simpler formulation, the subscripts are dropped out in the above formula. $\mathcal{F}^{(d)}$ denotes a column vector with K entries where the f^{th} entry relates to the DFT coefficient $\mathcal{F}_i^{(d)}[x, y]$. A is a matrix of size $K \times MN$ and contains the relationship between the continuous fourier transform of the continuous image and the discrete fourier transform of the low resolution images, and $\mathcal{G}$ is a column vector with $MN \times 1$ and represents the coefficients of CFT of the continuous image which will be obtained. The equation system in (2.5) is solved with respect to $\mathcal{G}$, and subsequently the high resolution image $f^{(u)}$ is reconstructed by using the inverse discrete fourier transform.

The model behind the described approach is not able to handle blurring or noise. Furthermore, only global translations which are known a priori are modeled. To overcome this problem several extensions of the initial frequency domain approach are presented. Kim et al. [22] took account of the blurring and noise problem and formulate an approach which can handle low resolution images with the same noise and blurring. Later, Kim et al. [23] extended this method to handle different blurrings for each low resolution image and using the Tikhonov regularisation to provide a stable method. Tom et al. [42] presented a frequency domain approach which uses a maximum likelihood framework to formulate the image registration and blurring problem and solves the problem by concurrently using the expectation maximization algorithm.

The frequency domain approach was the first superresolution method in history and it can be computed efficiently and parallel. Nevertheless, because of difficulties with incorporating prior information into the model and the restriction to a global translation model, most later works about resolution enhancement methods are spatial domain approaches. [30, 46]

2.2 Nonuniform Interpolation Approach

The nonuniform interpolation approach is a simple and intuitive non-iterative forward approach which consists of three steps [46, 30]. The fist step is motion estimation from several low resolution images. After that, a non-uniform interpolation step calculates a high resolution image, and finally, an image restoration step is responsible for noise and blurring reduction. These approaches can be implemented computationally very efficiently, which is a major advantage be-

cause of real-time application. Nevertheless, through step-by-step processing, an optimal solution cannot be granted because of ignoring errors from previous stages.

2.3 Projection onto Convex Sets (POCS)

The projection onto convex sets method is an iterative approach for which prior knowledge is used to create constraints where each of the constraints restrict the solution to be an element of a closed convex set C_i. If the sets have a non-empty intersection, the solution is in the intersection set C_s. As stated in [30], the paper by Stark et al. [40] was the first where the projection onto convex sets method was applied for superresolution reconstruction. In this case, as described in [4], POCS finds a superresolution image $\hat{f^{(u)}} \in C_s$, which is in the intersection of M closed convex sets. [4], [38]

$$C_s = \bigcap_{i=1}^{M} C_i \tag{2.6}$$

C is a subset of the Hilbert space $\mathbb{H}$ (for the definition of $\mathbb{H}$, we refer to [38]) and is called convex if for any two elements $f_1, f_2 \in C$ the subset C contains the element $f = \theta f_1 + (1-\theta) f_2$ where $\theta \in [0,1]$. C is closed if the limit element of any sequence of elements in C is also in C. For each subset C_i, a projection operator $P_i : \mathbb{H} \to C_i$ is defined, so that

$$\|f - P_i f\| = \min\|f - h\| \ \forall h,\ h \in C_i \tag{2.7}$$

This means that the nearest element to f, which is in C_i, is $P_i f$. With the projection operators P_i, a sequence of images $\{{f^{(u)}}^n\}$ can by generated by

$${f^{(u)}}^{n+1} = P_M P_{M-1} \dots P_2 P_1 {f^{(u)}}^n \tag{2.8}$$

This sequence $\{{f^{(u)}}^n\}$ converges to the image $\hat{f^{(u)}}$, which is shown in figure 2.1. An image ${f^{(u)}}^0$ is projected by the projector P_1 onto the nearest point on C_1 and subsequently it is projected onto C_2 by P_2. Thereby the first iteration is completed. By iterating this process, the image converges to the solution set C_s and at the end, it reaches the red point which represents the image $\hat{f^{(u)}}$, which is on the convex set C_s.

Based on the assumption that the motion information is exact, Teklap et al. [41] extend the POCS method from Stark et al. [40] to include observation

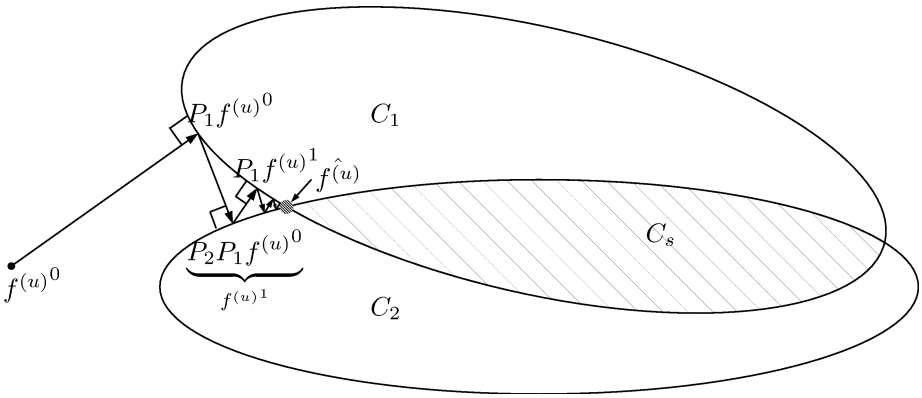


Figure 2.1: This figure shows the projection onto convex sets method. An arbitrary image ${f^{(u)}}^0$ is projected by the projection operator P_1 to the nearest image in the closed convex set C_1. Subsequently P_2 projects the image $P_1 {f^{(u)}}^0$ onto the set C_2, and by repeating these steps, the projected image converges to $\hat{f^{(u)}}$, which is the set C_s. [4], [38]

noise. For this, they define closed convex constraints to restrict $f^{(u)}$ to the sets $C_{i;j}$. Index i refers to the low resolution image $f_i^{(d)}$ and the index j addresses the pixel of the image $f_i^{(d)}$ which is in vector notation. Thus, for each pixel j of each low resolution image $f_i^{(d)}$ a closed convex constraint exists which restricts $f^{(u)}$ to the set $C_{i;j}$.

$$C_{i;j} = \{f^{(u)} : |r_{i;j}^{(u)}| \leq \delta_0\} \tag{2.9}$$

where

$$r_{i;j}^{(u)} = f_{i;j}^{(d)} - [A_i f^{(u)}]_j \tag{2.10}$$

and $A_i = W_i D_i B_i$ and $\delta_0 = \alpha\sigma_n$. σ_n denote the standard deviation of the noise, and $\alpha \geq 0$ is a suitable constant which defines a statistical bound. δ_0 can be considered as confidence in the observation. The projection $P_{i;j}$ onto the convex set $C_{i;j}$ in [41, 43] is defined by

$$P_{i;j} = \begin{cases} f^{(u)} + \frac{a_{i;j}(r_{i;j}^{(u)} - \delta_0)}{\|a_{i;j}\|^2} & \text{if } r_{i;j}^{(u)} > \delta_0 \\ f^{(u)} & \text{if } -\delta_0 < r_{i;j}^{(u)} < \delta_0 \\ f^{(u)} + \frac{a_{i;j}(r_{i;j}^{(u)} + \delta_0)}{\|a_{i;j}\|^2} a_{i;j} & \text{if } r_{i;j}^{(u)} < -\delta_0 \end{cases} \tag{2.11}$$

where $a_{i;j}$ is the column vector which contains the j^{th} row of the matrix A_i.

A further extension to this is described by Patti et al. in [31] where for the first time a POCS approach is able to handle motion blur and arbitrary sampling lattice for video resolution enhancement. Eren et al. proposed in [11] a robust object-based method which extends the previous approach in the case of multiple moving objects in a scene. They assume that inaccurate motion vectors appear at moving object boundaries and occlusion areas. By creating a validity map, they ensure that only pixels with accurate motion information are projected by the associated projector. Furthermore, they create a segmentation map to distinguish between different moving objects. So, they are able to create motion models for each object, which leads to more accurate motion information, which also improves the quality of the image. A different approach is proposed by Elad et al. [10]. They analyse the ML, MAP and POCS approach and present a general hybrid approach which combines the advantages from the analysed methods. Because of combining these methods, incorporating prior knowledge is simple and an optimal solution is found, which in contrast cannot be ensured by the POCS approach.

The POCS approach is simple and powerful but the optimality of the solution cannot be guaranteed and the methods are slow convergence. Thus, high computational power is needed.

2.4 Regularized Superresolution Approaches

Finding a high resolution image for our superresolution model in section 1.6.2 typically results in a large set of possible solutions. In order to restrict and stabilize the solution set, regularized superresolution approaches have been developed. In the following, the maximum a posteriori (MAP), the maximum likelihood (ML) and the constrained least square (CLS) approach are described. It should be noted that the ML and CLS are equal to the MAP approach under certain conditions. For more information, we refer to Park et al. in [46].

2.4.1 Constrained Least Square

The constrained least square approach make use of deterministic prior information of the high resolution image to stabilise the solution. If the matrices A_i are

estimated, the CLS Problem can be represented for example by [2, 30]

$$\hat{f^{(u)}} = \min \left\{ \sum_{i=1}^{K} \|A_i f^{(u)} - f_i^{(d)}\|^2 + \lambda \|H f^{(u)}\|^2 \right\} \tag{2.12}$$

where $\hat{f^{(u)}}$ is the estimated optimal high resolution image. With the a priori knowledge that images normally tend to be smooth, we can for example set $H = \nabla$, therefore the regularisation term ($\|\nabla f^{(u)}\|^2$) represents the smoothness of the solution and provides a suitable counterweight to the data term ($\sum_{i=1}^{K} \|A_i f^{(u)} - f_i^{(d)}\|^2$), which represents the fidelity to the data. The weighting between those terms is controlled by the regularisation parameter λ. This basic approach is adapted from Hong et al. [18], who presented an iterative regularized algorithm for video sequences with an assumed spatially invariant points spread function where λ is calculated for each iteration. Hardie et al. [17] proposed an approach for Infrared Imaging System where knowledge of the optical system and detector array are incorporated into the observation model. Based on this, a constrained least squares problem is formulated which is subsequently solved by a gradient descent and a conjugate gradient procedure.

2.4.2 Statistical Approaches

Statistical approaches like ML or MAP generally treat the enhanced and low resolution images as stochastic variables. Using this assumption, the Bayesian framework can be used where the probability of $f^{(u)}$ given $f_1^{(d)}, \ldots, f_K^{(d)}$ is

$$Pr(f^{(u)}|f_1^{(d)}, \ldots f_K^{(d)}) = \frac{Pr(f^{(u)}) Pr(f_1^{(d)}, \ldots, f_K^{(d)}|f^{(u)})}{Pr(f_1^{(d)}, \ldots, f_K^{(d)})} \tag{2.13}$$

where $Pr(f^{(u)}|f_1^{(d)}, \ldots f_K^{(d)})$ is the a posterori, $Pr(f^{(u)})$ is the a priori, $Pr(f_1^{(d)}, \ldots, f_K^{(d)}|f^{(u)})$ is the likelihood and $Pr(f_1^{(d)}, \ldots, f_K^{(d)})$ is the normalisation term. For the equation (2.13) exist two common ways to find an optimal solution $\hat{f^{(u)}}$ by maximising either the likelihood or the a posteriori term.

2.4.2.1 Maximum A Posteriori (MAP)

The MAP estimator maximises the a posteriori term with respect to $f^{(u)}$ to find an optimal solution $\hat{f^{(u)}}$.

$$\hat{f^{(u)}} = \arg\max_{f^{(u)}} \left\{ Pr(f^{(u)}|f_1^{(d)}, \ldots, f_K^{(d)}) \right\} \tag{2.14}$$

Using the logarithmic function and Bayes conditional probability theorem, equation (2.14) leads to

$$\hat{f^{(u)}} = \arg\max_{f^{(u)}} \left\{ \ln Pr(f_1^{(d)}, \ldots, f_K^{(d)} | f^{(u)}) + \ln Pr(f^{(u)}) \right\} \tag{2.15}$$

Under the assumption that noise is an independent identically distributed zero mean Gaussian distribution, the likelihood term can be expressed by the following equation.

$$Pr(f_1^{(d)}, \ldots f_K^{(d)} | f^{(u)}) = e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{K} \| A_i f^{(u)} - f_i^{(d)} \|^2} \tag{2.16}$$

If we use Gibbs density function [13] [37] for the a priori term

$$Pr(f^{(u)}) = \frac{1}{Z} e^{-\frac{1}{\lambda} U(f^{(u)})} \tag{2.17}$$

where U is called energy function and Z and λ are constants, the MAP estimation can be represent in the following form

$$\hat{f^{(u)}} = \arg\min_{f^{(u)}} \frac{1}{2\sigma^2} \sum_{i=1}^{K} \| A_i f^{(u)} - f_i^{(d)} \|^2 + \frac{1}{\lambda} U(f^{(u)}) \tag{2.18}$$

where λ is a regularisation parameter between the data term and the regulation term $U(f^{(u)})$.

2.4.2.2 Maximum Likelihood (ML)

The ML maximises, as the name suggest, the likelihood term in equation (2.14). This leads to an optimal solution $\hat{f^{(u)}}$ under the assumption that the prior term is uniform, hence, it has no effect for the maximisation. The optimal solution using ML estimator, under the assumption that noise is an independent identically distributed zero mean Gaussian distribution, is given by

$$\hat{f^{(u)}} = \arg\min_{f^{(u)}} \frac{1}{2\sigma^2} \sum_{i=1}^{K} \| A_i f^{(u)} - f_i^{(d)} \|^2 \tag{2.19}$$

Tom et al. [42] use an ML superresolution approach to address the registration and blurring problem in the frequency domain and solve it by an EM algorithm. Elad et al. [10]. formulate an ML (MAP), POCS hyprid approach, which combines the POCS approach with the statistical approaches ML and MAP.

The ML approach is simpler than MAP, but it is particularly sensitive concerning noise and errors from the registration and blurring estimation. Furthermore, it is not possible to incorporate prior knowledge into the problem, which makes the MAP approach a more robust alternative.

2.5 Iterative Backprojection Approach

Another approach is presented by Irani and Peleg in [20]. They formulate an iterative back projection method to reconstruct superresolution images. For this, the back projection method used in computed tomography image reconstruction serves as a model. The main idea is to calculate the difference between i simulated low resolution images $f_i^{(d)}$ and i observed low images ${f_i^{(d)}}^n$. This error is multiplied with a back projection kernel b^{BP} and subsequently back projected to estimate a high resolution image ${f^{(u)}}^{n+1}$. These steps are iterated, which leads to the following update schema

$${f^{(u)}}^{n+1}(x) = {f^{(u)}}^n(x) + \sum_{y \in \bigcup_i Y_{i,x}} (f_i^{(d)}(y) - {f_i^{(d)}}^n(y)) \frac{(b_{xy}^{BP})^2}{\alpha \sum_{y' \in \bigcup_i Y_{i,x}} b_{xy}^{BP}} \quad (2.20)$$

where $Y_{i,x}$ describes the set $\{y \in g_i | y \text{ is influenced by } x\}$ and α is a normalisation factor and $b_{xy}^{BP} = b^{BP}(x - z_y)$. z_y denotes the centre of the receptive field of y in ${f^{(u)}}^n$. This update schema minimises the error between the simulated and observed low resolution image, but an optimal solution cannot be guaranteed. Another issue is the choice of b^{BP} which, as described in [20], can be chosen arbitrarily. However, the choice influences the solution if more possible solutions exist. Hence, b^{BP} may be utilised as an additional constraint so that the reached solution has a particular property (e.g. smooth). In contrast to the POCS or regulised approaches, it is hard to integrate prior knowledge into the iterative back projection method, but it is an intuitive and easy to understood method.

2.6 Superresolution for Hyperspectral Images

Superresolution methods are typically applied to grayscale images, nevertheless several methods exist which are able to handle colour images. Next, we will show how superresolution techniques can be used for hyperspectral imaging to overcome the relatively low spatial resolution of these images.

2.6.1 Image Fusion

In many papers, as described by Gu et al. [15], the images from a hyperspectral camera are combined with images from an additional sensor. [33, 14, 34]. The advantage of both sensors are merged to obtain high resolution hyperspectral images. These approaches will not further be discussed here because this would go beyond the scope of this thesis and it is not planned to equip the HELIOS camera-based hyperspectral imaging systems with an additional sensor.

2.6.2 Spectral Unmixing

Another way to overcome the ill-posedness of the superresolution inverse problem of hyperspectral image is to incorporate prior knowledge about the spectral signature of the material in the scene. For this approach, it is assumed that all materials which are in the scene can be determined by their a priori known spectral signature. These a priori known materials are denoted in literature as endmember. As hyperspectral images generally have a low spatial resolution, it is assumed that a mixing of the spectral signature in adjacent spatial pixels occur. This is typically represented by the linear mixing model

$$f = \sum_{i=1}^{N} a_i m_i + e = Ma + e \tag{2.21}$$

where f is a vector which contains the pixels of the mixed spectrum, M is a matrix which contains the information about the endmember, a are the mixing coefficients and e denotes the error vector. By adding the non-negative (2.22) and sum-to-one constraint (2.23)

$$a_i \geq 0 \tag{2.22}$$

$$\sum_{i=1}^{N} a_i = 1 \tag{2.23}$$

to the linear mixing model (2.21), physically impossible solutions are excluded. The linear mixing model is typically solved by minimizing the constrained least squares model:

$$\min_{a} \|Ma - f\|_2^2 \text{ subject to } a_i \geq 0;\ \sum_{i=1}^{N} a_i = 1 \tag{2.24}$$

The solution for this minimisation problem is vector a, which contains the information about the mixing of the endmember. This information is then used to enhance the resolution of hyperspectral images.

Gu et al. [15] propose an approach in which the information of the unmixing step is combined with the superresolution mapping. In this case, the unmixing step is first calculated, and subsequently, the authors train a neuronal network with the information of the mixing and the spatial correlation tendency of landscapes. Therefore, the network uses spatial and spectral information to increase the resolution of hyperspectral images. In another paper Guo et al. [16] propose two methods called endmember-based TV model and quantum TV model. For both methods, first the L1 unmixing step is calculated, which makes the pixels "pure". This is done by

$$\min_a \|a\|_1 + \frac{\lambda}{2}\|Ma - f\|_2^2 \text{ subject to } a_i \geq 0 \tag{2.25}$$

where $\|a\|_1$ is used instead of the sum-to-one constraint. After the unmixing step Guo et al. use the TV energy superresolution model for both methods. The endmember-based TV model is then

$$\min_u \; E = \sum_{j=1}^{d} \int_\Omega \|\nabla u_j\|_2 dx + \frac{\alpha}{2} \sum_{j=1}^{d} (DBu_j - f_j)^2 dx + \frac{\beta}{2} \int_\Omega \prod_{i \in I(x)} \|u - m_i\|_2^2 \tag{2.26}$$

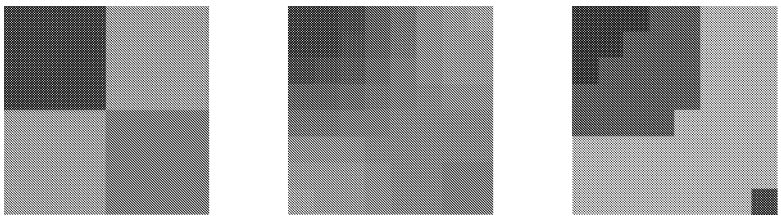
where the term

$$\frac{\beta}{2} \int_\Omega \prod_{i \in I(x)} \|u - m_i\|_2^2$$

penalises pixel which cannot be assigned to an endmember.
The Quantum TV Model is an iterative update schema which first zooms the image using the TV Model, and afterwards, each pixel is rounded up to the nearest endmember signal.

$$v^{n+1} = \arg\min_u \sum_{j=1}^{d} \int_\Omega \|\nabla u_j\|_2 dx + \frac{\alpha}{2} \sum_{j=1}^{d} (DBu_j - f_j)^2 dx \tag{2.27}$$

$$u^{n+1}(x) = \arg \min_{\{m_i(x): i \in I(x)\}} \|m_i(x) - v^{n+1}(x)\|_2 \tag{2.28}$$



(a) (b) (c)

Figure 2.2: Shows the original image (a) and the results from the endmember-based TV (b) and quantum TV (c) models on a synthetic 177-band 2x2 image with 4x zoom (images are taken from [16])

.

Chapter 3

HELIOS Superresolution Approach

Contents

3.1	**Preliminaries**	**38**
3.2	**HELIOS Image Acquisition Process**	**40**
3.3	**Assumptions for the Superresolution Model**	**42**
3.4	**Discrete Setting**	**44**
3.5	**HELIOS Superesolution Model**	**45**
3.6	**Discrete Convex Optimisation Problem**	**46**

This chapter presents a detailed insight into the HELIOS superresolution approach in order to achieve a high resolution reconstruction from some low resolution hyperspectral line scan images. The discussed approach uses a variational superresolution model which is extended for hyperspectral images.

3.1 Preliminaries

3.1.1 Definitions

Definition Hadamard product: As defined in [21], the Hadamard product $A \circ B$ of two matrices $A, B \in \mathbb{R}^{M \times N}$ where $A = [a_{i,j}]$ and $B = [b_{i,j}]$ where

$1 \leq i \leq M$ and $1 \leq j \leq N$ is given by

$$A \circ B = [a_{i,j} \cdot b_{i,j}] \tag{3.1}$$

Definition Gradient: The gradient of the function $f : \mathbb{R}^N \to \mathbb{R}$ is defined as

$$\text{grad}\ (f) = \nabla f = \begin{pmatrix} \frac{\delta f}{\delta x_1} \\ \vdots \\ \frac{\delta f}{\delta x_n} \end{pmatrix} \tag{3.2}$$

Definition Divergence: The nabla operator is also used to define the divergence operator, which is the negative adjoint of ∇.

$$\text{div} = -\nabla^* \tag{3.3}$$

3.1.2 Total Variation (TV)

The total variation in image reconstruction was first used by Rudin et al. [36] and is typically used as a regularizer for MAP approaches.

Definition Total Variation: The total variation in a continuous setting from a function $u \in L^1(\Omega)$ defined by

$$J(u) = \sup\left\{ -\int_\Omega u(x)\text{div}\xi(x)\text{d}x : \xi \in C^1(\Omega;\mathbb{R}^2), |\xi(x)| \leq 1\ \forall x \in \Omega \right\} \tag{3.4}$$

In the case that $J(u) < +\infty$, which means that the function has Bounded Variation, and if $u \in C^1$, the total variation is

$$J(u) = \int_\Omega |\nabla u| dx \tag{3.5}$$

3.1.2.1 Discretisation of the Total Variation

For the discrete total variation, we use the same formulation as described in [44, 5, 6]. Therefore, an image is represented by a regular grid with the size $M \times N$. The indexes i, j describe the position on the grid and h_1 and h_2 define the height and width of a pixel. So we can define our grid as

$$(ih_1, jh_2) : 1 \leq i \leq M, 1 \leq j \leq N \tag{3.6}$$

We also define a vector space $X = \mathbb{R}^{MN}$ and the discrete gradient operator $\nabla : X \to Y$ where $Y = X \times X$ and

$$(\nabla u)_{i,j} = \begin{pmatrix} (\nabla u)^x_{i,j} \\ (\nabla u)^y_{i,j} \end{pmatrix}$$

where

$$(\nabla u)^x_{i,j} = \begin{cases} \frac{u_{i+1,j}-u_{i,j}}{h_1} & \text{if } i < M \\ 0 & \text{if } i = M \end{cases}$$

$$(\nabla u)^y_{i,j} = \begin{cases} \frac{u_{i,j+1}-u_{i,j}}{h_2} & \text{if } j < N \\ 0 & \text{if } j = N \end{cases}$$

In this setting, the discrete total variation is defined by

$$J(u) = \|(\nabla u)_{i,j}\|_{2,1} \tag{3.7}$$

where

$$\|q\|_{2,1} = \sum_{0\leq i,j\leq M,N} \sqrt{((q)^x_{i,j})^2 + ((q)^y_{i,j})^2}$$

Additionally, we define the discret diverence operator div: $Y \to X$ as follows:

$$(\text{div } p)_{i,j} = (\text{div } p^x)^x_{i,j} + (\text{div } p^y)^y_{i,j}$$

where

$$((\text{div } p^x)^x_{i,j} = \begin{cases} 0 & \text{if } i = 0 \\ \frac{p_{i,j}-p_{i-1,j}}{h_1} & \text{if } 0 < i < M \\ \frac{-p_{i-1,j}}{h_1} & \text{if } i = M \end{cases}$$

$$(\text{div } p^y)^y_{i,j} = \begin{cases} 0 & \text{if } j = 0 \\ \frac{p_{i,j}-p_{i,j-1}}{h_2} & \text{if } 0 < j < N \\ \frac{-p_{i,j-1}}{h_2} & \text{if } j = N \end{cases}$$

3.2 HELIOS Image Acquisition Process

In this section, we discuss in detail the image acquisition process from a typical HELIOS camera-based material separation system and precisely define the images which are delivered by such a system. HELIOS camera systems use a line-scan technique in combination with a hyperspectral sensor. Thus, in one acquisition step l, a matrix $f_l^{(d)}$ of size $M \times S$ is captured where M describes the number of pixels and S denotes the number of spectral wavebands. In the following, the matrix $f_l^{(d)}$ is also denoted as line because in the spatial dimension, only a line is captured. Each acquisition step from the camera requires a certain amount of time which is described by t. This is commonly the inverse of the camera frequency. The camera is directed at the conveyor belt which transports material with speed v. By concatenating of N successive captured

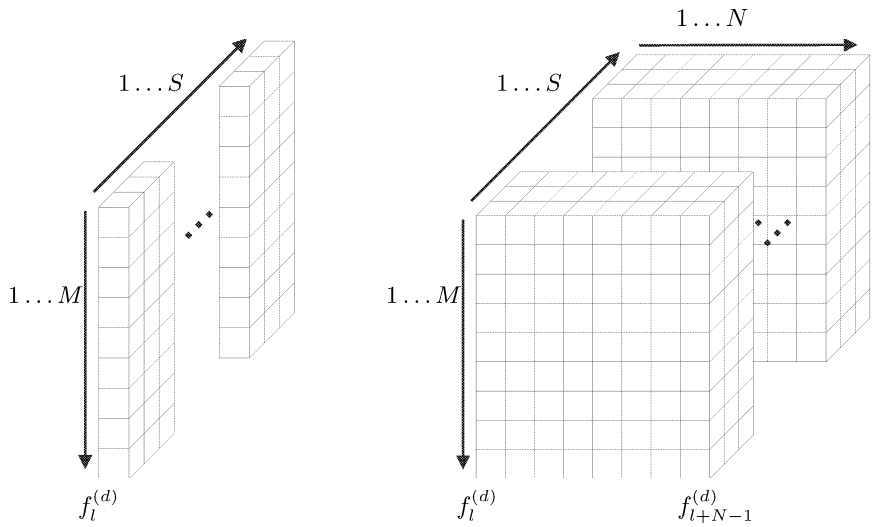


Figure 3.1: The left figure shows the matrix $f_l^{(d)}$ with size $M \times S$ which is captured from one HELIOS image acquisition step at time l. The right figure shows the image $f^{(d)}$ with size $M \times S \times N$ from the HELIOS camera system where N successive lines $f_l^{(d)} \dots f_{l+N-1}^{(d)}$ are concatenated.

lines, an image $f^{(d)}$ over time with size $M \times S \times N$ is generated. This image $f^{(d)}$ contains three dimensional data and is typically represented as a data cube described in section 1.3.4. The HELIOS camera system can generally have three settings where conveyor belt speed v is either smaller, equal or greater than the horizontal spatial coverage (line width) w of the line scan camera on the conveyor belt divided by its acquisition process time t. These different settings are shown in the columns of figure 3.2, which outlines the image acquisition process from a line scan camera where the spectral dimension was omitted in order to simplify the figure. The grey arrows show the moving directions of the conveyor belt, and the red squares denote the captured pixels in one line acquisition step. By viewing the first four entries from a column, we can see that different parts of the green triangle are captured by the line-scan camera over time. Depending on the horizontal spatial coverage (line width) w of the line scan camera on the conveyor belt, its acquisition process time t and the conveyor belt speed v, the HELIOS image $f^{(d)}$ contain different information. The correlation between these factors are shown in the fifth row of figure 3.2. The last row in the figure shows the theoretical image if the HELIOS Image $f^{(d)}$ is correctly aligned. In the case where $v < \frac{w}{t}$, which means that the conveyor belt moves the object horizontally less than the horizontal spatial coverage of the camera in one acquisition step. We can see that an overlap between temporal adjacent pixel occurs. The opposite is shown in the case $v > \frac{w}{t}$, whereby the object is horizontally moved by the conveyor belt, which is more than the horizontal spatial coverage of the camera in one acquisition step. Thus, gaps occur because not all parts of the object are captured by the camera.

3.3 Assumptions for the Superresolution Model

As we know from section 1.6, the basic premise for superresolution is that two or more sub-pixel shifted images from the same scene are available. In the case of a material separation system based on the HELIOS camera system, shown in figure 1.6, we have analysed in section 3.2 that three possibilities for the image acquisition process exist. But only if the moving distance of the conveyor belt during one camera acquisition step is lower than the horizontal spatial coverage (line width) of the HELIOS line scan camera, superresolution is possible. In this case, an overlap between adjacent captured lines occurs if the lines are

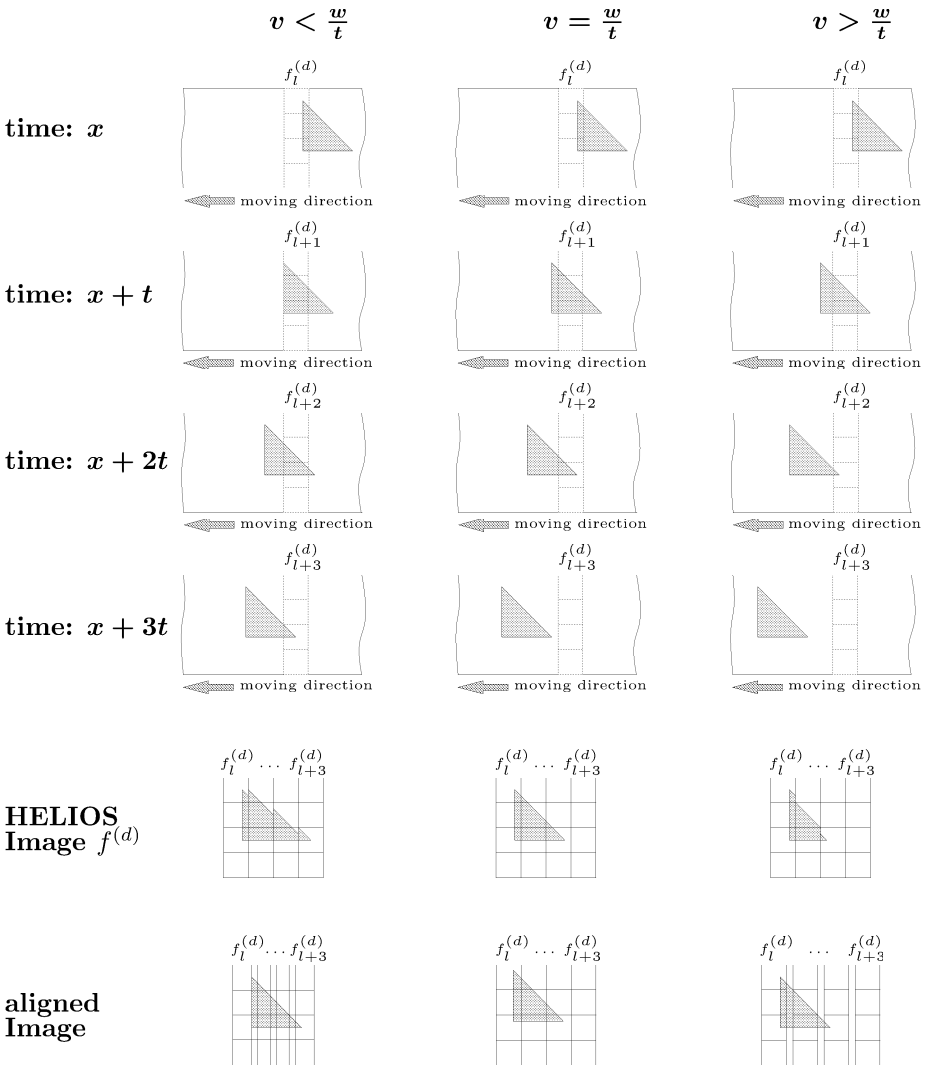


Figure 3.2: The figure is structured as a table and shows the image acquisition process from three different HELIOS camera settings. The columns denote these settings where the conveyour belt speed v is either smaller, equal or greater than the horizontal spatial coverage (line width) w of the line scan divided by the acquisition process time t. The first four rows show the image acquisition process of four successive acquisition steps for each of the HELIOS settings. Row number five shows the whole captured HELIOS Image from the four steps and the last row shows the theoretical image if we bring the captured lines to the correct position.

aligned, which means that the information of this overlapping area is contained in each of the overlapping lines. Thus, superresolution is possible because we have more information about the same area. So, for the following, we assume this scenario.
A further assumption is that the resolution of the spectral dimension of the HELIOS input images is high enough, so there is no need to increase the spectral resolution of the image. Thus, the input image and the superresolution image have the spectral resolution.

3.4 Discrete Setting

3.4.1 Discretisation of a continuous Image

As discussed in section 1.6.2, an image captured by a camera can be seen as a discretisation from a continuous image. In our case, we use three dimensional discrete images, thus, we have to model the discretisation step from a three-dimensional continuous image $g : \Omega \to \mathbb{R}$ to a three-dimensional discrete image f, which is given on a three-dimensional grid of size $M \times S \times N$. For this, we extend the thoughts about the discretisation step b Unger et al. [44], who describe the two-dimensional discretisation with squared pixel areas to model a general three-dimensional discretisation schema.
We first define a voxel of the discrete image $f_{j,k,l}$ at position $(h_1 j, h_2 k, s h_3 l)$ with height h_1, depth h_2, width h_3 and shifting s where $h_1, h_2, h_3 > 0$ and $s \geq 0$. In the case of $h_1 = h_2 = h_3$, the voxel is a cube, in the other case it is a cuboid. For simplifying the following definition, we define a vector $h = [h_1, h_2, h_3]^T$, which contains the information about the height, depth and width of a voxel. s is defined as a shifting factor for the temporal dimension, which means if $s = 0$ there is no shifting, hence voxel $f_{j,k,l}$ and voxel $f_{j,k,l+1}$ are at the same position. In the case of $0 < s < 1$, there is an overlap between the voxels $f_{j,k,l}$ and $f_{j,k,l+1}$, and if $s > 1$, horizontal gaps between those voxels exist. As you can see in equation 3.8, the voxel value $f_{j,k,l}$ is calculated from the continuous image g by taking the mean in the voxel region $\Delta^{h,s}_{j,k,l}$.

$$f_{j,k,l} = \frac{1}{h_1 h_2 h_3} \int_{\Delta^{h,s}_{j,k,l}} g(x) dx \tag{3.8}$$

where the voxel region itself is defined by

$$\Delta_{j,k,l}^{h,s} = (h_1 j, h_2 k, s h_3 l) + ([-\frac{h_1}{2}, +\frac{h_1}{2}) \times [-\frac{h_2}{2}, +\frac{h_2}{2}) \times [-\frac{h_3}{2}, +\frac{h_3}{2})) \quad (3.9)$$

.

3.4.2 Discrete Images $f^{(d)}$ and $f^{(u)}$

For our problem, we define the discrete images $f^{(d)}$ and $f^{(u)}$, which are both samplings from the same continuous image g, but each has a different voxel size vector h and shifting factor s. The image $f^{(d)}$ describes the input image which is captured from the HELIOS camera system and the image $f^{(u)}$ is the calculated high resolution image from our approach. To distinguish between the parameter for the low- or high resolution image, we again use the superscript letters $^{(d)}$ and $^{(u)}$. In order to make superresolution possible for our approach, we assume that image $f^{(d)}$ with size $h^{(d)}$ has an overlap between two adjacent pixels in the temporal dimension. Thus, we define a shifting factor $s^{(d)} : 0 < s^{(d)} < 1$. The resulting superresolution image $f^{(u)}$ has no overlap or gap, so $s^{(u)} = 1$. A further assumption is that the spectral resolution of the image $f^{(u)}$ and $f^{(d)}$ is equal. With the described assumptions, we define the relationship between the sampling of the high resolution image $f^{(u)}$ and low resolution $f^{(d)}$ by a vector $\zeta = [\zeta_1, 1, \zeta_3]^T$ where $\zeta_1, \zeta_3 \geq 1$. Furthermore, we define that $h^{(d)} = \zeta \circ h^{(u)}$ where $\circ$ denotes the Hadamard product, which is defined in 3.1. This means that

$$\begin{pmatrix} h_1^{(d)} \\ h_2^{(d)} \\ h_3^{(d)} \end{pmatrix} = \begin{pmatrix} \zeta_1 h_1^{(u)} \\ h_2^{(u)} \\ \zeta_3 h_3^{(u)} \end{pmatrix}$$

We can clearly see that ζ contains the scaling information between the superresolution image $f^{(u)}$ and the low resolution $f^{(d)}$. This information is used by the downsampling operator D discussed in section 3.6.1.

3.5 HELIOS Superesolution Model

As we have defined our discrete images $f^{(d)}$ and $f^{(u)}$, we develop a HELIOS superresolution model based on the image formation process from the HELIOS camera system. We know from the general superresolution model that the starting point for the image formation process is the high resolution image $f^{(u)}$.

In the general superresolution model, a wrapping operator W is applied to this image in order to align the sub-pixel shifts between several images. In our case, the high resolution image $f^{(u)}$ is calculated from the low resolution image $f^{(d)}$, so there is only one input image. Thus we can omit the wrapping operator W. The next step takes into account that by each image acquisition, negative effects like image distortion, blurring etc. occurs, for this a blurring operator B is applied to the high resolution image $f^{(u)}$. Afterwards, the downsampling operator D is applied to the blurred high resolution image $f^{(u)}$, which describes the relationship between $f^{(u)}$ and the low resolution image $f^{(d)}$. Finally, we assume that each image is degraded by noise, for this, a noise term η is added to the HELIOS superresolution model which can mathematically be described by the following equation

$$f^{(d)} = DBf^{(u)} + \eta \tag{3.10}$$

Based on on equation 3.10, we can define our discrete convex optimisation problem, which is described in the next section.

3.6 Discrete Convex Optimisation Problem

For our convex optimisation problem, we choose a MAP approach to find a solution for the superresolution model described in equation 3.10. As mentioned in section 2.4.2.1, MAP approaches finding an optimal solution and allows incorporating prior knowledge into the problem. This is done by modeling a regularisation term which restricts the solution in a specific way. In our case, we use the discrete Total Variation, which is described in detail in section 3.1 for regularisation. As we defined in 3.4, the amount of spectral bands from $f^{(u)}$ and $f^{(d)}$ are equal, and so we handle each spectral band separately. We define an index j where $1 \leq j \leq S$, which addresses the spectral dimension, which means that $f_j^{(u)}$ and $f_j^{(d)}$ contains all information about the j^{th} spectral band of the high and low resolution image. Furthermore, we transform $f_j^{(u)}$ and $f_j^{(d)}$ to vectors by stacking the columns, so $f_j^{(u)}$ has then size $M^{(u)}N^{(u)} \times 1$ and $f_j^{(d)}$ has size $M^{(d)}N^{(d)} \times 1$. Based on these thoughts, we can formulate our discrete convex optimisation problem as

$$\min_{f_j^{(u)}} = \left\{ \lambda \| DBf_j^{(u)} - f_j^{(d)} \|_{p_1} + \| \nabla f_j^{(u)} \|_{p_2} \right\}; \forall j = 1 \ldots S \tag{3.11}$$

where $\|.\|_p$ defines the p-norm. The operators D, B in equation (3.11) are linear and describe the downsampling and blurring operators. The convex minimisation problem from above can be solved by setting $p_1, p_2 = 1$ which leads to the TVL^1 superresolution method, or by setting $p_1 = 2$ and $p_2 = 1$ to the ROF superresolution method.

3.6.1 Downsampling Operator

The purpose of the downsampling operator D is to model the downsampling process from a discrete high resolution image $f^{(u)}$ to a discrete low resolution image $f^{(d)}$. In section 3.4.2, we precisely discussed how we can obtain these discrete images from a continuous image g.
We have defined in section 3.4.2 the relationship between the voxel sizes of image $f^{(u)}$ and $f^{(d)}$, which is

$$\begin{pmatrix} h_1^{(d)} \\ h_2^{(d)} \\ h_3^{(d)} \end{pmatrix} = \begin{pmatrix} \zeta_1 h_1^{(u)} \\ h_2^{(u)} \\ \zeta_3 h_3^{(u)} \end{pmatrix}$$

and the shifting factors, which are

$$s^{(u)} = 1$$

$$0 \leq s^{(d)} \leq 1$$

As we assumed in section 3.3, the number of spectral bands for the superresolution image $f^{(u)}$ is the same as for the input image $f^{(d)}$, hence, we only have to model the downsampling operator D for the spatial and temporal dimension. For a certain spectral band i, we can calculate the voxel value $f_{j,i,l}^{(d)}$ from the downsampled image $f^{(d)}$ with the following equation, which can be found in its two-dimensional form for squared pixel sizes in the paper [44] by Unger et al.

$$f_{j,i,l}^{(d)} = \frac{1}{\mu(\Delta_{j,i,l}^{\zeta h^{(u)}, s^{(d)}})} \sum_{0 \leq x,z \leq M^{(u)} N^{(u)}} \mu(\Delta_{x,i,z}^{h^{(u)}, s^{(u)}} \cap \Delta_{j,i,l}^{\zeta h^{(u)}, s^{(d)}}) f_{x,i,z}^{(u)} \qquad (3.12)$$

In equation (3.12), $\Delta_{k,l}^{\zeta h^{(u)}, s^{(d)}}$, $\Delta_{k,l}^{h^{(u)}, s^{(u)}}$ denotes the voxel regions which are defined in section 3.4.1, and $\mu(\Delta)$ is the Lebesgue measure, which denotes the volume of a voxel region. As you can see, the downsampling calculation for

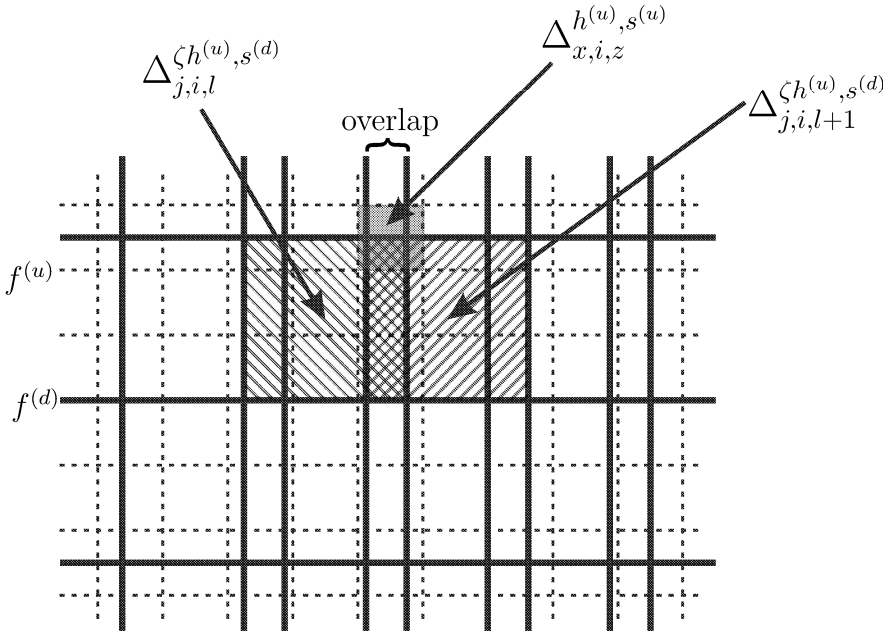


Figure 3.3: The figure shows the assumption about the downsampling process from a discrete high resolution image $f^{(u)}$ to a discrete low resolution image $f^{(d)}$ where the pixels in one dimension of the low resolution image contain partly the same information

$f_{j,i,l}^{(d)}$ is a weighted area integral over the region of $\Delta_{j,i,l}^{\zeta h^{(u)},s^{(d)}}$. For a better understanding, figure 3.3 graphically shows our assumptions concerning the samplings. This calculation is stored in a Matrix D with size $M^{(d)}N^{(d)} \times M^{(u)}N^{(u)}$, which represent our downsampling operator.
Unlike "normal" downsamplings (e.g. [44, 28]), sometimes the same subregion from a superresolution image voxel is included two ore more times in the calculation of lower sampled voxel. That is the case in the case if voxel region $\Delta_{x,i,z}^{h^{(u)},s^{(u)}}$ intersects with the overlap between $\Delta_{j,i,l}^{\zeta h^{(u)},s^{(d)}}$ and $\Delta_{j,i,l+1}^{\zeta h^{(u)},s^{(d)}}$. This is also shown in figure 3.3. This multiple used subregion can be calculated more accurately through the inverse downsampling process (upsampling process) because there are two ore more low sampled voxels which contains information about this subregion.

3.6.2 Blurring Operator

A crucial factor for superresolution is finding a blurring operator B, which contains all blurring effects caused by imperfect optics and sensors. For this, a blur kernel b is typically either calculated by optimisation methods (e.g. constrained least square (CLS)) or estimated as a Gaussian function. As described in previous sections each spectral band is calculated separately, hence, blurring kernel b is two-dimensional and furthermore it is assumed that all spectral have the same blurring. Once a blur kernel b is found, the blurring operator B can be constructed as a matrix with size $M^{(u)}N^{(u)} \times M^{(u)}N^{(u)}$ based on it. In the following subsections, we describe in short the blur kernel estimation via Gaussian function and a estimation via non-blind constrained least square optimisation.

3.6.2.1 Blur Kernel via Gaussian Blurring

If no knowledge about blurring is available, an appropriate method is to assume the blurring as a Gaussian function [44, 28, 16]. For example, a two-dimensional Gaussian blur kernel can be calculated through

$$b(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2}\right)} \tag{3.13}$$

where σ_x and σ_y describe the variances in x and y direction and μ_x and μ_y are the means in their respective directions.

3.6.2.2 Blur Kernel via Constrained Least Square

Another way to find the Blur kernel is the calculation via optimisation. In this case, we need a calibration pattern and a blurred image recorded by the imaging system to derive a non-blind optimisation problem. The relationship between a spectral band of the high resolution image and a spectral band of the low resolution image is

$$f_j^{(d)} = DBf_j^{(u)} \tag{3.14}$$

as discussed in section 3.6. To calculate the two-dimensional blurring kernel $b(x, y)$ with size $M^{(b)} \times N^{(b)}$, we rewrite $Bf_j^{(u)}$ so that

$$Bf_j^{(u)} = Ub \tag{3.15}$$

where U is a matrix of size $M^{(u)}N^{(u)} \times M^{(b)}$ and b is a vector of size $M^{(b)}N^{(b)} \times 1$, which represents the blur kernel $b(x, y)$, hence equation (3.14) is

$$f_j^{(d)} = DUb \tag{3.16}$$

This equation can be solved with the constrained least square optimisation which is

$$\min_b = \|DUb - f_j^{(d)}\|_2 \text{ subject to } b_i \geq 0 \text{ and } \sum_i b_i = 1 \tag{3.17}$$

The blur kernel b can be determined by the iterative gradient descent update schema

$$b^{n+1} = b^n + \alpha(U^T D^T DUb - U^T D^T f_j^{(d)}) \tag{3.18}$$

$$b_i^{n+1} = \frac{\max(b_i^{n+1}, 0)}{\sum_i b_i^{n+1}} \tag{3.19}$$

where α is a constant and controls the step size for the gradient descent step, and equation (3.19) is for fulfilling the $\sum_i b_i = 1$ and $b_i \geq 0$ constraints. A sightly different version of this is used by Yuan et al. in [47].

Chapter 4

Algorithms

Contents

4.1 Preliminaries **51**

4.2 ROF Superresolution **57**

4.3 TVL_1 Superresolution Primal-Dual **61**

In this chapter, we will discuss in detail the algorithms which are used to solve the convex optimisation problem from the previous chapter. For this, we first introduce some mathematical concepts which will be used from the discussed algorithms.

4.1 Preliminaries

4.1.1 Definitions

Definition Indicator function: The indicator function $\delta_\Sigma(x)$ for the set Σ is defined by

$$\delta_\Sigma(x) = \begin{cases} 0 & \text{if } x \in \Sigma \\ \infty & \text{else} \end{cases} \tag{4.1}$$

Definition Subdifferential: The subdifferential $\partial f(x)$ in a point $x \in \mathbb{R}^n$ of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is given by the set

$$\partial f(x) = \{g \in \mathbb{R}^n : f(x) + \langle g, y - x\rangle \leq f(y), \forall y \in \mathbb{R}^n\} \tag{4.2}$$

which means that the subdifferential is a generalisation of the gradient to non-differential functions. In figure 4.2, the subdifferential of a function and its representation in the dual space is shown.

Definition Subgradient: An element g is called subgradient in a point $x \in \mathbb{R}^n$ if $g \in \partial f(x)$ where $\partial f(x)$ is the subdifferential defined in (4.2).

Definition Proximal operator: The resolvent operator or proximal operator for a convex function f and $\tau > 0$ is defined as

$$x = (I + \tau \partial f)^{-1}(y) = \arg\min_x \left\{ \frac{\|x - y\|^2}{2\tau} + f(x) \right\} \tag{4.3}$$

which means that x is the value where function $f(y)$ reaches its minimum value. In figure 4.1, an example is shown where the proximal operator is calculated for two functions $F(x)$ and $G(x)$ and the domain of x is restricted to $[-0.5, 0.5]$.

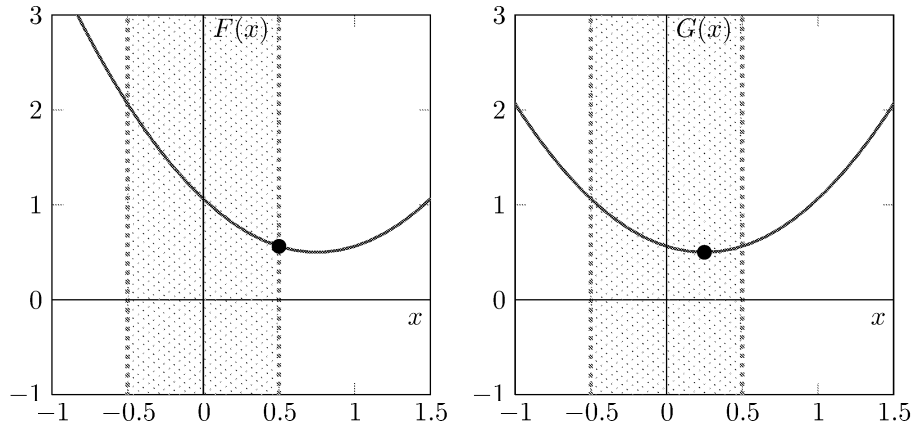


Figure 4.1: For the functions $F(x) = (x - 0.25)^2 + 0.5$ (left) and $G(x) = (x - 0.75)^2 + 0.5$ (right) where $x \in [-0.5, 0.5]$ (hatched area), the proximal operator finds the minimum values (black dots) with in the domain of x.

Definition Legendre-Fenchel transform: The convex conjugate f^* of a function $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is defined through the Legendre-Fenchel transform [35]

$$f^*(p) = \sup_{x \in dom f} \{\langle x, p \rangle - f(x)\} \tag{4.4}$$

which is the maximum gap between the linear function $\langle x, p \rangle$ and $f(x)$ in the case that f is differentiable $p = f'(x)$. Figure 4.2 shows an example of the Legendre-Fenchel transform with a continuous but not differentiable function.

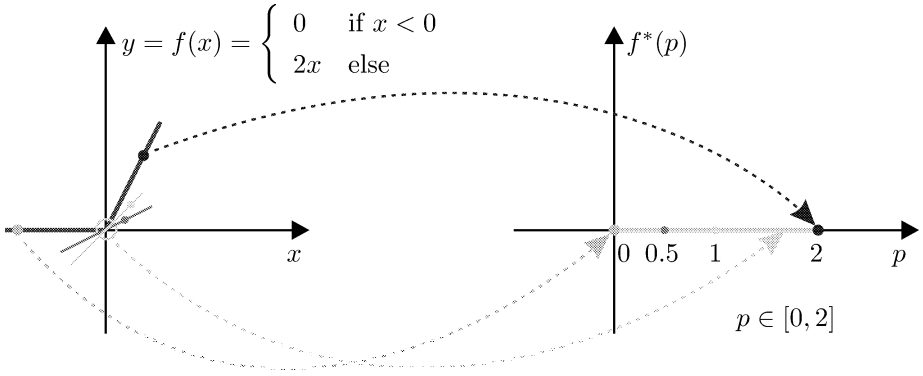


Figure 4.2: The figure shows the Legendre-Fenchel transform for a continuous function (blue) which is not differentiable at $x = 0$. For this point, the subdifferential is the closed interval $[0, 2]$ because it contains all possible tangents in that point (this is schematically shown with the red and green tangent). The subddifferential at point $x = 0$ can be seen as a line (yellow) in the dual space. For any point $x < 0$, the slope $p = 0$ (cyan) and for any point $x > 0$ $p = 2$ (dark blue).

4.1.2 Duality

The duality principle means that we are able to bring a function, problem or concept into another related formulation and vice versa [1]. Depending on the viewpoint, we talk about the primal or dual formulation. The Legendre-Fenchel Transformation, which is defined in 4.1.1, is such a transformation. It changes the representation of continuous but not necessary differentiable function, which is usually defined through points so that every point on a function is defined by $(x, f(x))$, to a representation so that the function is defined through its slope and the convex conjugate $(p, f^*(p))$. In short: $(x, f(x)) \Leftrightarrow (p, f^*(p))$. This is represented graphically in figure 4.3. For more information about the duality and the Legendre-Fenchel transformation, we refer to [35].

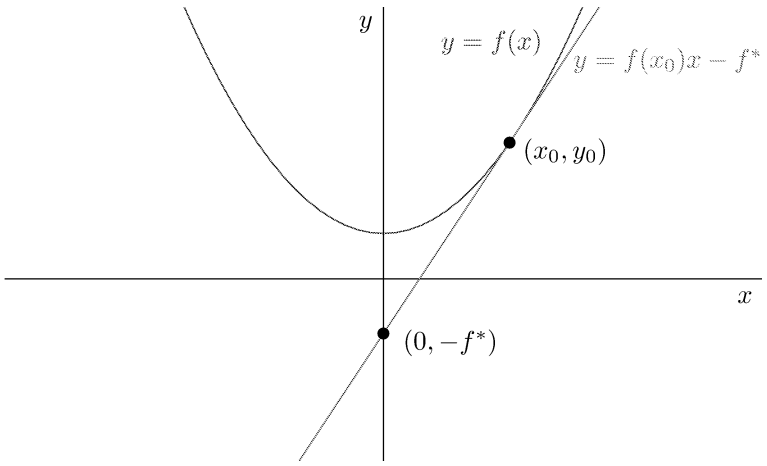


Figure 4.3: The figure shows the relationship between a function (blue), its tangent (red) on point (x_0, y_0) and the convex conjugate f^* for this point. f^* is the intersection between the tangent of the point (x_0, y_0) and the negative y-axis.

4.1.3 Conjugate Gradient Method

The conjugate gradient method is an old and one of the best known methods to solve a linear equation system with the form $Ax = b$ to x. We use this method to accelerate the calculation of the proximity of $G(x)$ from the ROF method in section 4.2. As stated in the document by Barett et. al. [3], the conjugate gradient method belongs to a class called non-stationary iterative methods and is an effective way to solve symmetric positive definite systems. In contrast to stationary methods, non-stationary methods use information for computation which changes at every iteration. In algorithm 1, the conjugate gradient method is described, which can also found in [3]. In the document by Barett et. al. [3], further information about the conjugate gradient and many more iterative methods for solving linear systems can be found.

4.1.4 General Primal-Dual Algorithm

In the following, we will explain the general primal-dual algorithm which can be used to solve a particular class of optimisation problems in an efficient way. One of the major advantages is the possibility to parallelise this algorithm, which

Algorithm 1 conjugate gradient algorithm

$x^0 =$ initial guess
$r^0 = b - Ax^0$
$maxiter =$ max. number of iterations
$tolerance =$ tolerance of method
for $i = 1$ to $maxiter$ **do**
  $\rho_{i-1} = {r^{(i-1)}}^T r^{(i-1)}$
  if $i = 1$ **then**
    $p^1 = r^0$
  else
    $\beta_{i-1} = \frac{\rho_{i-1}}{\rho_{i-2}}$
    $p^i = r^{i-1} + \beta_{i-1} p^{i-1}$
  end if
  $q^i = Ap^i$
  $\alpha_i = \frac{\rho_{i-1}}{{p^i}^T q^i}$
  $x^i = x^{i-1} + \alpha_i p^i$
  $r^i = r^{i-1} - \alpha_i q^i$
  convergence check: break if $convergence < tolerance$
end for

has become more important in the last years.

4.1.4.1 General Primal-Dual Formulation

First, we define two vector spaces X, Y, which are finite dimensional vector spaces. Furthermore, we define a linear operator $K : X \to Y$, which has an operator norm

$$\|K\| = \sup\{\|Kx\| : x \in X \text{ with } \|x\| < 1\} \tag{4.5}$$

For these definitions, a general class of optimisation problems can be represented in the following form [7, 1]:

$$\min_{x \in X} F(Kx) + G(x) \tag{4.6}$$

F and G are proper, convex, lower-semicontinuous functions. If we replace the term $F(Kx)$ with its convex conjugate which is defined in 4.1.1 where $\langle ., . \rangle$ denotes the standard inner product, the problem can be turned into a saddle-point problem

$$\min_{x \in X} \max_{y \in Y} \{\langle Kx, y \rangle - F^*(y) + G(x)\} \tag{4.7}$$

which is a primal-dual formulation of the primal problem in (4.6).

4.1.4.2 General Primal-Dual Algorithm

The saddle-point problem in equation (4.7) can be implemented through following algorithm, which is discussed in the paper [1], [44], [7], [32] [6], [9].

Algorithm 2 General Primal Dual Algorithm

Initialisation:

$\tau, \sigma > 0;\ \theta \in [0, 1];\ x^0, y^0 \in X \times Y$

Iterations:

$x^{n+1} = (I + \tau\delta G)^{-1}(x^n - \tau K^* y^n)$

$y^{n+1} = (I + \sigma\delta F)^{-1}(y^n + \sigma K(x^{n+1} + \theta(x^{n+1} - x^n))$

4.1.4.3 Convergence

There are two common ways to ensure the convergence of the general primal dual algorithm, which are described in the following sections.

Primal Dual with estimated σ and τ

As stated in [7], the general primal-dual algorithm converges under the condition that $\theta = 1$ and τ and σ are chosen so that the inequality

$$\tau\sigma L^2 < 1 \tag{4.8}$$

holds. In the above formula $L = ||K||$, which means that L is the operator norm of operator K. For the proof and further reading we refer to [7] and [5].

Preconditioned Primal Dual

In some cases where the operator norm of K is hard to compute, we can use an alternative method to find τ and σ. This method is described by Pock et. al. in [32]. The constants σ and τ are replaced with diagonal matrices Σ and T where the diagonal consists of $(\sigma_1, \ldots, \sigma_n)$ and $(\tau_1, \ldots, \tau_n)$ respectively. These elements are called diagonal preconditioners and calculated through the following formulas

$$\tau_j = \frac{1}{\sum_{i=1}^{m} |K_{i,j}|^{2-\alpha}}, \ \sigma_j = \frac{1}{\sum_{j=1}^{n} |K_{i,j}|^{\alpha}} \tag{4.9}$$

where $\alpha \in [0, 2]$. Pock et al. show that these elements can be used for a large class of convex optimisation problems and they are easy to compute. Furthermore, they show that the preconditioned primal dual algorithm converges faster than the primal dual with estimated σ and τ as described in section 4.1.4.3

4.2 ROF Superresolution

4.2.1 From Primal to Primal-Dual

The discret version of ROF superresolution problem has the form

$$\min_{f_j^{(u)}} = \left\{ \left(\|\nabla f_j^{(u)}\|_{2,1} + \frac{\lambda}{2} \|DBf_j^{(u)} - f_j^{(d)}\|_2^2 \right) \right\}, \forall j = 1 \ldots S \tag{4.10}$$

Since the convex conjugate $f^*(p)$ from the norm $\|.\|_{2,1}$ is the indicator function $\delta_\Sigma(p)$ where $\Sigma = \{p : ||p||_\infty \leq 1\}$, we can turn the primal problem in equation (4.10) with the Legendre-Fenchel transform defined in equation (4.4) into the

following primal-dual formulation.

$$\min_{f_j^{(u)}} \max_{p_j} = \left\{ \Big(\underbrace{\langle p_j, \nabla f_j^{(u)} \rangle}_{\langle y, Kx \rangle} - \underbrace{\delta_\Sigma(p_j)}_{F^*(y)} + \underbrace{\frac{\lambda}{2} \| DB f_j^{(u)} - f_j^{(d)} \|_2^2}_{G(x)} \Big) \right\}, \forall j = 1 \dots S \tag{4.11}$$

This formulation corresponds to the primal-dual formulation in equation (4.7), which allows us to solve the problem with a primal-dual algorithm.

4.2.2 ROF Proximal Operators

As described in section 4.1.4.2, for the terms $F^*(y)$ and $G(x)$ the proximal operators have to be calculated.

4.2.2.1 Proximity of $F^*(y)$

For the ROF primal-dual formulation, the proximity of $F^*(y)$ is

$$p_j = (I + \sigma \partial F)^{-1}(\overline{p_j}) = \arg\min_{p_j} \left\{ \frac{\| p_j - \overline{p_j} \|^2}{2\sigma} + \delta_\Sigma(p_j) \right\} \tag{4.12}$$

which leads to

$$p_j = \frac{\overline{p_j}}{\max(1, |\overline{p_j}|)} \tag{4.13}$$

which is described by Chambolle et. al. in [7] as a reduction on pointwise Euclidean projectors onto L^2 balls.

4.2.2.2 Proximity of $G(x)$

The proximity of $G(x)$ is

$$f_j^{(u)} = (I + \tau \partial G)^{-1}(\overline{f_j^{(u)}}) = \arg\min_{f_j^{(u)}} \left\{ \frac{\| f_j^{(u)} - \overline{f_j^{(u)}} \|^2}{2\tau} + \frac{\lambda}{2} \| DB f_j^{(u)} - f_j^{(d)} \|_2^2 \right\} \tag{4.14}$$

The partial derivative with respect to $f_j^{(u)}$ is calculated and set to zero.

$$\frac{\partial G}{\partial f_j^{(u)}} = f_j^{(u)} - \overline{f_j^{(u)}} + \tau \lambda (DB)^T (DB f_j^{(u)} - f_j^{(d)}) = 0 \tag{4.15}$$

Through rearranging the above, we come to the following equation

$$\underbrace{(I + \tau \lambda (DB)^T (DB))}_{A} \underbrace{f_j^{(u)}}_{x} = \underbrace{\overline{f_j^{(u)}} + \tau \lambda (DB)^T f_j^{(d)}}_{b} \tag{4.16}$$

The equation (4.16) can be seen as a linear system of equation of form $Ax = b$. For such a system, several methods exist to solve it for x. For our implementation, we choose the conjugate gradient method, which is described in section 4.1.3. As we can see in algorithm 1, the conjugate gradient method uses an initial guess for x_0 as starting point for solving the system. In the case of solving the proximal operator for $G(x)$, we use the conjugate gradient method in each iteration of the ROF superresolution primal-dual algorithm, which is shown in algorithm 3. The purpose of the primal-dual algorithm is to find a "good" superresolution image $f_j^{(u)^{n+1}}$, which becomes "better" at every iteration. The purpose of the conjugate gradient method is to solve the linear equation system $Ax = b$ to x shown in equation (4.16), where $x = f_j^{(u)^{n+1}}$ in our case, which is shown in algorithm 3. So, we have for every computation of the proximal operator of $G(x)$ an initial guess $f_j^{(u)^{n}}$, which becomes "better" at every iteration of the primal-dual algorithm, which leads to a fast computation of the proximal operator of $G(x)$.

4.2.3 Iterative ROF Superresolution Primal-Dual Algorithm

The equation 4.20 can be implemented as an iterative Primal-Dual algorithm where the pseudocode is shown in algorithm 3

4.2.4 Step Size

As described in section 4.1.4.3, choosing the right values for τ and σ is a crucial factor which affects the convergence of the algorithm. As pointed out in equation (4.11), the operator $K = \nabla$ for the ROF primal dual algorithm. With the

Algorithm 3 Iterative ROF Superresolution Primal-Dual Algorithm

Initialization:

$\tau_j, \sigma_j > 0\ \{\forall j, j \in \mathbb{N} | j \leq S\};\ \theta \in [0,1];\ {f_j^{(u)}}^0 \in X;\ {p_j}^0 \in Y;$
$maxIterPD, maxIterCG \in \mathbb{N}^+;\ toleranceCG > 0$

Algorithm:

$\overline{{f_j^{(u)}}^0} = {f_j^{(u)}}^0$
$A = (1 + \tau\lambda(DB)^T(DB))$
for $j = 1$ to S **do**
  while $n \leq maxIterPD$ **do**
    $\overline{p_j} = {p_j}^n + \sigma_j \nabla \overline{{f_j^{(u)}}^n}$
    ${p_j}^{n+1} = \frac{\overline{p_j}}{\max(1, |\overline{p_j}|)}$
    $b = {f_j^{(u)}}^n + \tau\lambda(DB)^T f^{(d)}$
    ${f_j^{(u)}}^{n+1} = CG(A, b, maxIterCG, toleranceCG, {f_j^{(u)}}^n)$
    $\overline{{f_j^{(u)}}}^{n+1} = {f_j^{(u)}}^{n+1} + \theta({f_j^{(u)}}^{n+1} - {f_j^{(u)}}^n)$
  end while
end for

definition of our discrete ∇ in section 3.1.2.1, K has the following structure.

$$K = \begin{pmatrix}
-\frac{1}{h_1} & 0 & \dots & 0 & \frac{1}{h_1} & \dots & \dots & 0 \\
0 & -\frac{1}{h_1} & & 0 & 0 & \frac{1}{h_1} & & 0 \\
\vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\
0 & \dots & 0 & -\frac{1}{h_1} & 0 & \dots & 0 & \frac{1}{h_1} \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
\vdots & & & & & & & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
-\frac{1}{h_2} & \frac{1}{h_2} & 0 & \dots & \dots & \dots & \dots & 0 \\
0 & -\frac{1}{h_2} & \frac{1}{h_2} & 0 & \dots & \dots & \dots & 0 \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & 0 & -\frac{1}{h_2} & \frac{1}{h_2} & 0 & \dots & \dots & 0 \\
0 & 0 & 0 & -\frac{1}{h_2} & \frac{1}{h_2} & 0 & \dots & 0 \\
\vdots & & & & & & & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & \dots & \dots & \dots & 0 & -\frac{1}{h_2} & \frac{1}{h_2} & 0 \\
0 & \dots & \dots & \dots & \dots & 0 & -\frac{1}{h_2} & \frac{1}{h_2} \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0
\end{pmatrix} \tag{4.17}$$

By ignoring the border pixel in operator K, and we can calculate the values for the constants τ and σ, by using formula (4.9) with $\alpha = 1$, which leads to

$$\tau = \frac{h_1 h_2}{2(h_1 + h_2)},\ \sigma = \frac{h_1}{2} \tag{4.18}$$

4.3 TVL_1 Superresolution Primal-Dual

4.3.1 From Primal to Primal-Dual

In order to find an implementation for the TVL_1 superresolution, we use the discrete version of the optimisation problem which has the form

$$\min_{f_j^{(u)}} = \left\{ \left(\|\nabla f_j^{(u)}\|_{2,1} + \lambda \|DBf_j^{(u)} - f_j^{(d)}\|_{2,1} \right) \right\}, \forall j = 1 \dots S \tag{4.19}$$

We know that $f^*(p)$ from the norm $\|.\|_{2,1}$ is the indicator function $\delta_{\Sigma_1}(p)$ where $\Sigma_1 = \{||p||_\infty \leq 1\}$ and and that $f^*(p)$ from $\lambda\|.\|_{2,1}$ is $\delta_{\Sigma_\lambda}(p)$ where $\Sigma_\lambda = \{||p||_\infty \leq \lambda\}$, we can formulate the primal dual problem as

$$\min_{f_j^{(u)}} \max_{p_j, q_j} \left(\langle p_j, \nabla f_j^{(u)} \rangle - \delta_{\Sigma_1}(p_j) + \langle q_j, DBf_j^{(u)} - f_j^{(d)} \rangle - \delta_{\Sigma_\lambda}(q_j) \right), \forall j = 1 \dots S \tag{4.20}$$

As we can see, the above formula has one primal variable $f_j^{(u)}$ and two dual variables p_j, q_j. In order to have an expression in the form of equation (4.7) to use a primal dual algorithm for the TVL_1 superresolution problem, the dual variables p_j, q_j are stacked into a vector. This leads to

$$\min_{f_j^{(u)}} \max_{p_j, q_j} \left(\left\langle \underbrace{\begin{pmatrix} p_j \\ q_j \end{pmatrix}}_{y}, \underbrace{\begin{pmatrix} \nabla \\ DB \end{pmatrix}}_{K} \underbrace{\left(f_j^{(u)} \right)}_{x} \right\rangle - \underbrace{\begin{pmatrix} \delta_{\Sigma_1}(p_j) \\ q_j^T f_j^{(d)} + \delta_{\Sigma_\lambda}(q_j) \end{pmatrix}}_{F^*(y)} + \underbrace{0}_{G(x)} \right) \tag{4.21}$$

For the sake of completeness, a zero has been added which has of course no effects but it represents the term $G(x)$ from equation (4.7).

4.3.2 TVL_1 Proximal Operators

As stated in algorithm 2, we have to calculate the proximal operators for $F^*(y)$ and $G(x)$ in equation (4.21)

4.3.2.1 Proximity of $F^*(y)$

For the TVL_1 superresolution primal dual formulation, the proximity of $F^*(y)$ is

$$p_j = (I + \sigma_p \partial F)^{-1}(\overline{p_j}) = \arg\min_{p_j} \left\{ \frac{\|p_j - \overline{p_j}\|^2}{2\sigma_p} + \delta_{\Sigma_1}(p_j) \right\} \tag{4.22}$$

$$q_j = (I + \sigma_q \partial F)^{-1}(\overline{q_j}) = \arg\min_{q_j} \left\{ \frac{\|q_j - \overline{q_j}\|^2}{2\sigma_q} + q_j^T f_j^{(d)} + \delta_{\Sigma_\lambda}(q_j) \right\} \tag{4.23}$$

which leads to

$$p_j = \frac{\overline{p_j}}{\max(1, |(\overline{p_j}|)}, \; q_j = \max(-\lambda, \min(\lambda, (\overline{q_j} + \sigma(-f_j^{(d)})))) \tag{4.24}$$

4.3.2.2 Proximity of $G(x)$

For the TVL_1 superresolution primal dual formulation the proximity of $G^*(x)$ is

$$f_j^{(u)} = (I + \tau \partial G)^{-1}(\overline{f_j^{(u)}}) = \arg\min_{f_j^{(u)}} \left\{ \frac{\|f_j^{(u)} - \overline{f_j^{(u)}}\|^2}{2\tau} + 0 \right\} = \overline{f_j^{(u)}} \tag{4.25}$$

4.3.3 Iterative L_1 Superresolution Primal-Dual Algorithm

The equation 4.20 can be implemented as an iterative primal-dual algorithm

4.3.4 Step Size

As described in section 4.1.4.3, there are two methods which guarantee the convergence of the primal-dual algorithm. For both ways, the constants τ_j, σ_j^q and σ_j^p have to be chosen in a correct way, which requires the calculation of the linear operator K. K, shown in equation (4.21), is a compounded matrix of ∇ and DB with size $2M^{(u)}N^{(u)} + M^{(d)}N^{(d)} \times M^{(u)}N^{(u)}$. We decided to calculate the matrices T_j, Σ_j^q and Σ_j^p via the diagonal preconditioners, which is described in section 4.1.4.3, to ensure the convergence of the algorithm.

Algorithm 4 Iterative TVl1 Superresolution Primal-Dual Algorithm

Initialisation:

$\tau_j, \sigma_j^q, \sigma_j^p > 0 \ \{\forall j, j \in \mathbb{N} | j \leq S\};\ \theta \in [0,1];\ f_j^{(u),0} \in X;\ p_j^0, q_j^0 \in Y;$
$maxIterPD \in \mathbb{N}^+$

Algorithm:

for $j = 1$ to S **do**
  while $n \leq maxIterPD$ **do**
    $\overline{p_j} = p_j^n + \sigma_j^p \nabla f_j^{(u),n}$
    $p_j^{n+1} = \frac{\overline{p_j}}{\max(1, |\overline{p_j}|)}$
    $\overline{q_j} = q_j^n + \sigma_j^q (DB f_j^{(u),n} - f_j^{(d)})$
    $q_j^{n+1} = \max(-\lambda, \min(\lambda, \overline{q_j} + \sigma_j^q (-f_j^{(d)})))$
    $\overline{f_j^{(u),n+1}} = \overline{f_j^{(u),n}} - \tau_j (-\mathrm{div} p_j^{n+1} + B^T D^T q_j^{n+1})$
    $f_j^{(u),n+1} = \overline{f_j^{(u),n+1}} + \theta (\overline{f_j^{(u),n+1}} - \overline{f_j^{(u),n}})$
  end while
end for

Chapter 5

Experimental Results

Contents

5.1 Environment **64**

5.2 Synthetic Data Evaluation **65**

5.3 Real Data Evaluation **69**

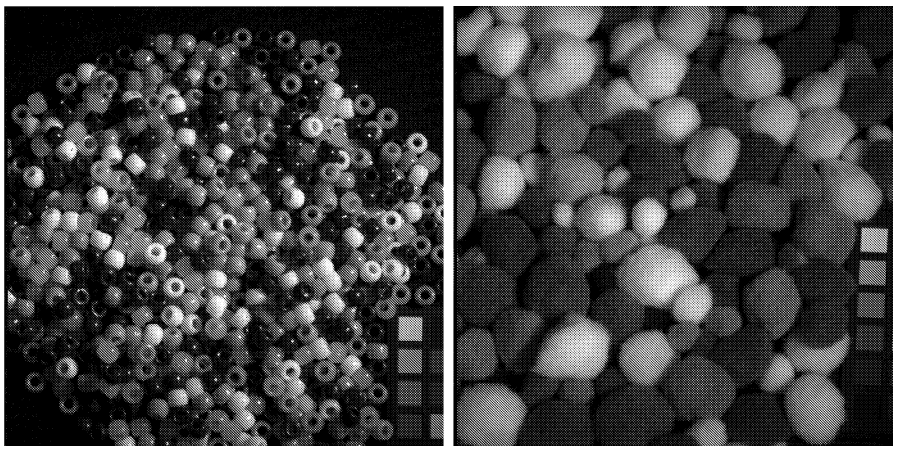
5.4 Summary **82**

5.1 Environment

For our experimental results, we chose two several experimental environments to analyse the performance of our proposed algorithms. For the first environment we used two data sets from the Columbia CAVE Laboratory Multispectral Image Database [8]. These sets are used as ground truth images which are subsequently blurred and downsampled with known parameter to create a synthetic data set. As we know the exact burring and downsamping operator we are able to calculate the reconstruction quality of the high resolution image of our approach by using the metrics described in section 5.2.1. For the second environment we used a real data set where the image comes from a HELIOS camera based hyperspectral line scan imaging system to show performance of the approach in a real environment. In this scenario no exact information of the blurring or the downsampling operator are available, thus, we have to approximate the blurring and the downsampling operator to use our superresolution approach.

5.2 Synthetic Data Evaluation

For the synthetic data evaluation, we chose the "Beads" and "Pompoms" data sets from the the Columbia CAVE Laboratory Multispectral Image Database [8]. Each of these data sets has thirty-one grayscale images with 512×512 pixels. For our synthetic data sets, we use only the first 240×120 pixels of each image. Together, these thirty-one image sections form our synthetic three-dimensional high resolution image $f^{(u)}$. Each spectral plane $f_j^{(u)}$ where $j = 1..31$ is then blurred with a two-dimensional Gaussian kernel with size 7×5 and $\sigma = 2.5$ and based on these several samplings and shiftings are processed to create several low resolution images with different sizes and shifting factors. These low resolution images are then used as input images for our superresolution approach and the results are compared with the reference image to determine the performance of our algorithm. As we have no noise in this setting, we use $\lambda = 100$ for the ROF and TVL1 algorithms to attach great importance to the data term of the optimisation problem (3.11) in order to achieve good results.



(a) **(b)**

Figure 5.1: This figure shows the 20^{th} spectral band images from the "Beans" (a) and the "Pompoms" (b) data sets from Columbia CAVE Laboratory Multispectral Image Database [8].

5.2.1 Evaluation Metrics

To evaluate the performance of our proposed ROF and TVL1 methods, we use the common image-processing metrics Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SIMM), which are extended to the third-dimension so that we are able to measure the reconstruction quality of three-dimensional spectral images.

5.2.1.1 Peak Signal to Noise Ratio (PSNR)

The PSNR is a widely used quality metric which can be calculated in the following way:

$$MSE(X,Y) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\Big(X(i,j) - Y(i,j)\Big)^2 \tag{5.1}$$

$$PSNR(X,Y) = 10\log_{10}\Big(\frac{MAX_i^2}{MSE(X,Y)}\Big) \tag{5.2}$$

where MAX_i denotes the maximum possible input value from a pixel.

5.2.1.2 Structural Similarity (SIMM)

In contrast to the PSNR, the SIMM quality metric measures the similarity of structural information in two images. The SIMM and MSIMM as described by Wang et al. in [45] are calculated through the formulas

$$SIMM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{5.3}$$

and

$$MSIMM(X,Y) = \frac{1}{M}\sum_{j=1}^{M} SSIM(x_j, y_j) \tag{5.4}$$

5.2.2 Hyperspectral Metrics

As we describe in section 5.2.1.1 and 5.2.1.2, the MSSIM and PSNR indexes are calculated from a ground-truth image X and superresolution image Y. In our environment, there is such a pair for each spectral dimension, so we calculate the MSSIM and PSNR values for each band and average them at the end, which leads to the mean MSSIM (MMSSIM) and mean PSNR (MPSNR) index.

$$MPSNR(X,Y) = \frac{1}{S}\sum_{i=1}^{S} PSNR(X_i, Y_i) \tag{5.5}$$

$$MMSIMM(X,Y) = \frac{1}{S}\sum_{i=1}^{S} MSSIM(X_i, Y_i) \tag{5.6}$$

5.2.3 Performance with varying Shifting Factor

To show the effects through different shifting factors, we create four low resolution images from the same ground truth image with the shifting factors 0.25, 0.5, 0.75 and 1 and a sampling rate from 1. A sampling rate from 1 means that the spatial resolution of the input image is equal to the spatial resolution of the ground truth image. So, in this case, no resolution enhancement is done. Figure 5.2 shows the MPSNR and MMSIMM indexes for the four images with varying shifting factor. As we expected, if the shifting factor increase the perfor-

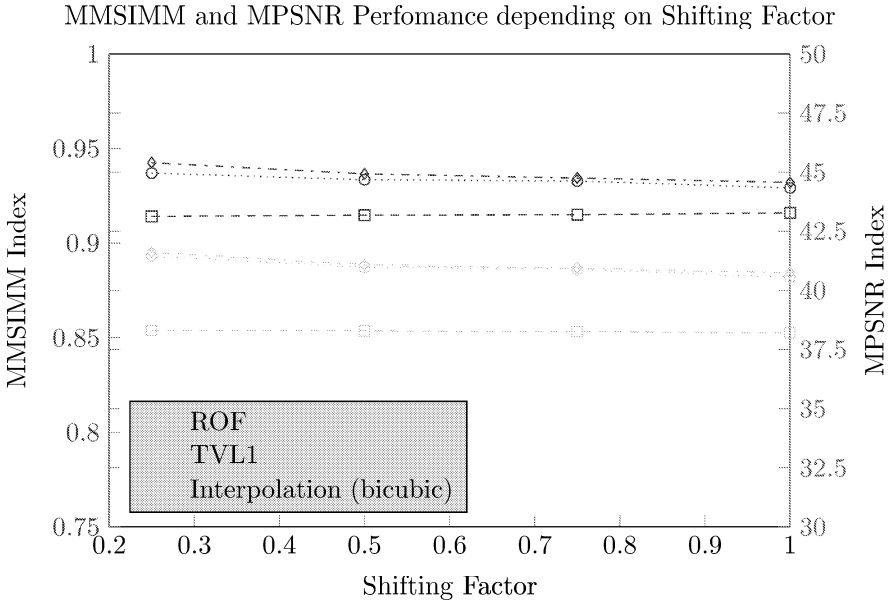


Figure 5.2: The figure shows the influence of the shifting factor to the performance of the ROF, TVL1 and bicubic interpolation approaches. In this scenario, $\lambda = 100$.

mance of the ROF and TVL1 models decreases. This is because the overlapping area between adjacent pixel decreases if the shifting factor increases. But the influence of the different shifting factors s in the range [0.25 1] to the performance of the algorithms is relatively small. But fine structures can be reconstructed

Nevertheless, we can see that our proposed approaches outperform the bicubic interpolation approach in both metrics and all tested shifting factors.

5.2.4 Performance with varying Sampling Rate

To show the performance of the proposed algorithm with different sampling rates, we choose a shifting factor 0.5 and sampling rates of 0.25, 0.5 and 1. A sampling rate of 0.25 means that the spatial resolution of the ground truth image $f^{(u)}$ is four times higher than spatial resolution of the input image $f^{(d)}$. So, in the case of sampling factor 1, the input and the ground truth image have the same spatial resolution. Figure 5.3 shows that ROF and TVL1 outperforms

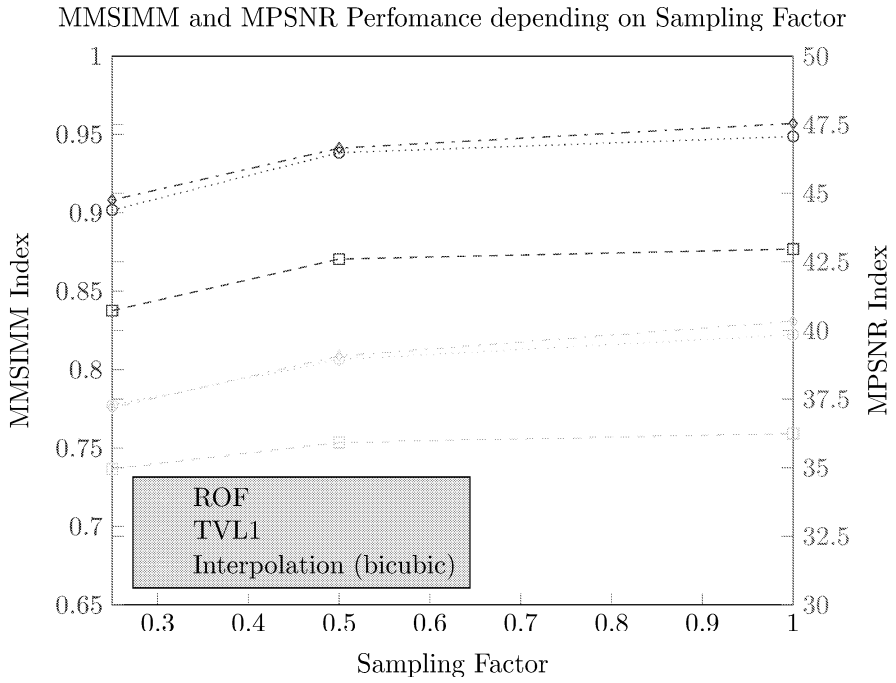


Figure 5.3: The figure shows the influence of the sampling factor to the performance of the ROF, TVL1 and bicubic interpolation approaches. In this scenario, $\lambda = 100$.

the bicubic interpolation approch in both metrics at every chosen sampling rate. Furthermore, as we expected, with increasing sampling rate the performance of the approaches increases. This is obviously because by increasing the sampling

rate, we receive more information for the same area, and, hence, the results are better. Compared to the shifting factor, we can see that the sampling rate has more influence on the performance than the shifting factor.

5.2.4.1 Summary

In the tables 5.1 and 5.2, the resulting images of different sampling and shifting factors are shown. As we know the exact blurring and downsampling operator and as there is no noise, we achieve a very good reconstruction quality with the ROF and TVL1 superresolution methods for the synthetic data sets. When the sampling rate is 1 and the shifting factor is 0.25 fine details can be reconstructed by the ROF and TVL1 method. This is shown in the first and second row of table 5.1 where the texture of the pompoms can be seen in the images which are calculated by ROF and TVL1. In contrast to this, the texture details can not be reconstructed by the bicubic interpolation approach. The same is true for the beads data set where also fine details can be reconstructed by the ROF and TVL1 method but not by the bicubic interpolation approach. If the shifting factor increases and the sampling factor decreases the reconstruction quality also decreases, but again the ROF and TVL1 methods outperfom the bicubic interpolation approach. The evaluation support our expectation of the performance behaviour where the sampling rate has much more influence on the reconstruction quality than the shifting factor. Nevertheless, with a smaller shifting factor fine details can be better reconstructed. Finally, we can say that the ROF and TVL1 methods outperforms the bicubic interpolation approach for the used synthetic data sets in all cases.

5.3 Real Data Evaluation

For the real data evaluation, we use test data from the EVK HELIOS test environment to show the performance of our superresolution approach in real environments.

5.3.1 Camera System Setup

Before we start with capturing images from the system, we have to calibrate it. The first parameter is the recording frequency of the camera, which is set to its

Table 5.1: This table shows the results for the noise free pompoms data set with known blurring and downsampling. $\lambda = 100$ in all cases. In the first row there is no downsampling and the shifting factor $s = 0.25$. In the second row an image section from row number one is shown in detail. In the third row, the shifting factor $s = 1$, hence, no overlap exists and the downsampling factor in x and y-direction is 4. In the forth row an image section from row number three is shown in detail.

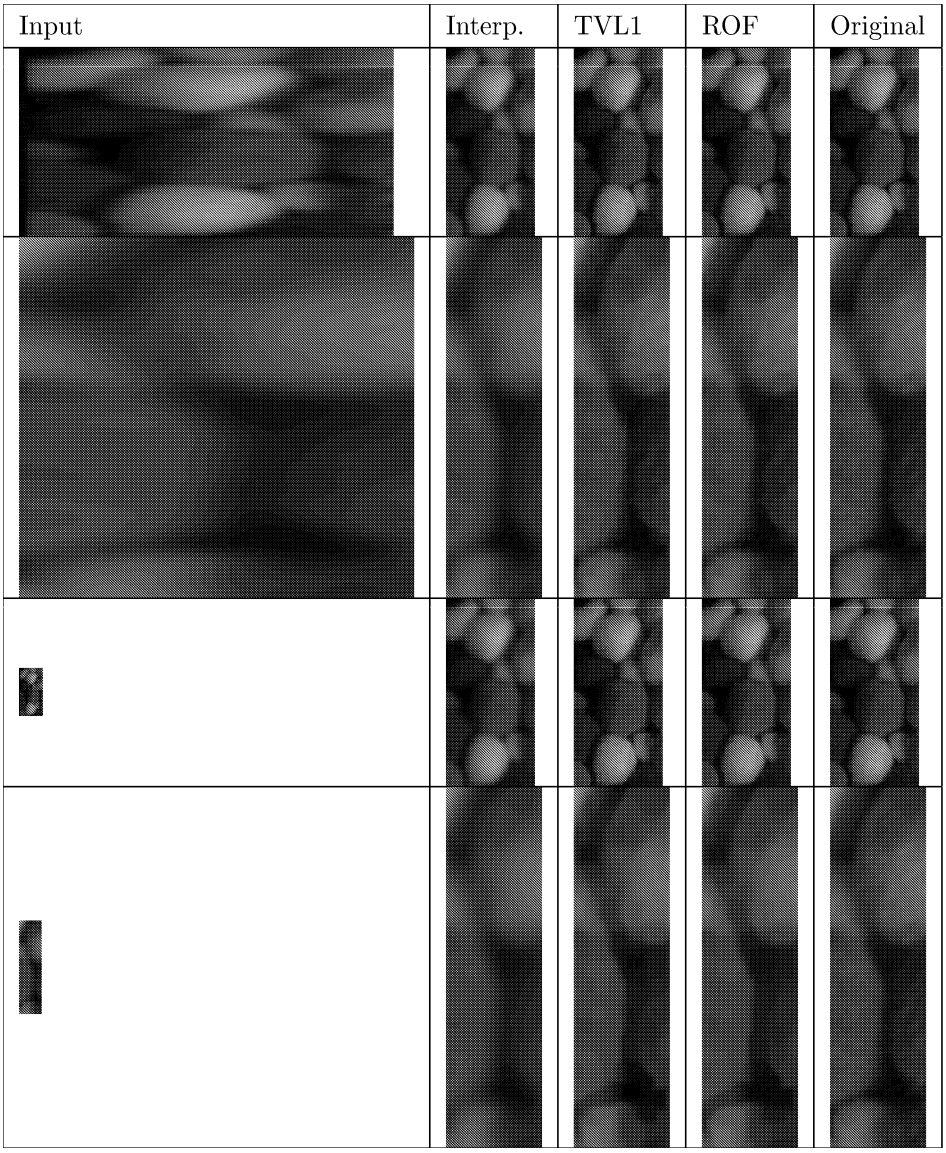
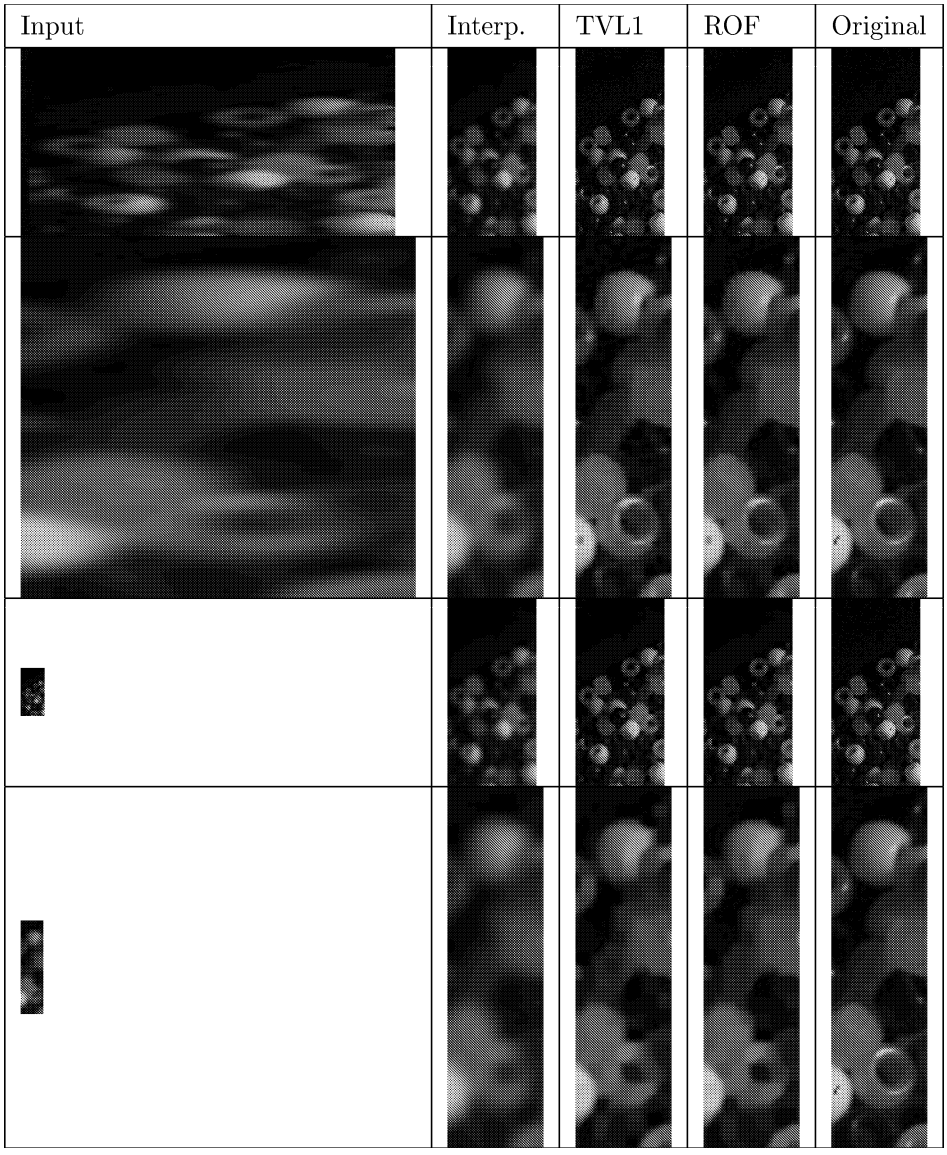


Table 5.2: This table shows the results for the noise free beads data set with known blurring and downsampling. $\lambda = 100$ in all cases. In the first row there is no downsampling and the shifting factor $s = 0.25$. In the second row an image section from row number one is shown in detail. In the third row, the shifting factor $s = 1$, hence, no overlap exists and the downsampling factor in x and y-direction is 4. In the forth row an image section from row number three is shown in detail.



maximum, in our case 300 Hz. Next, we manually adjust the halogen headlights to have a maximal illumination in the recording area. Then we manually finding the best focus for the camera and subsequently, we make a white and a black balancing to normalise the input. The last configuring step sets the speed of the conveyor belt. For our test scenario, the conveyor belt speed plays an important role because as described in section 3.2, the overlap between two adjacent lines depends on the conveyor belt speed and the camera frequency and is a critical factor in our assumption about the superresolution. As in our setup the camera frequency is fixed to its maximal value, only the conveyor belt speed controls the size of the overlapping area between the lines. Thus, we choose various conveyor belt speeds to verify behaviour and performance of our superresolution approach with different overlap areas.

5.3.2 Calibration Pattern

For an optimal functioning of our approach, we have to make accurate assumptions about the blurring and the shifting factor. For both of them, we need a calibration pattern which allows us to find correspondences between the calibration image and the captured image. We choose a calibration pattern layout with filled circles of different sizes arranged in a grid. On the one hand, this allows us to measure the distortion caused by the shifting factor and on the other hand, it is easy to determine the center of a circle which provides us with correspondence points for the registration step of the point spread function estimation. In figure 5.5, the used calibration pattern is shown.

5.3.2.1 Pattern Signatures

To use the pattern as a reference image, we have manually created a 3D hyperspectral image based on the 2D pattern where for each pixel which is either black or white the corresponding spectral signature in the 3D hyperspectral image is inserted. Because we have no information about the “right” material signatures, we have to extract it from the captured image. For that reason, we capture the whole grid with the HELIOS camera system and extract the material signature for black and white pixels by selecting an area with connected “pure” pixels of either black or white pixels for each spectral dimension separately. The selected pixels are then averaged, so we obtain a value for each spectral dimension for

black or white pixels in the 2D grid respectively. In figure 5.4, the spectral signatures for black and white are shown.

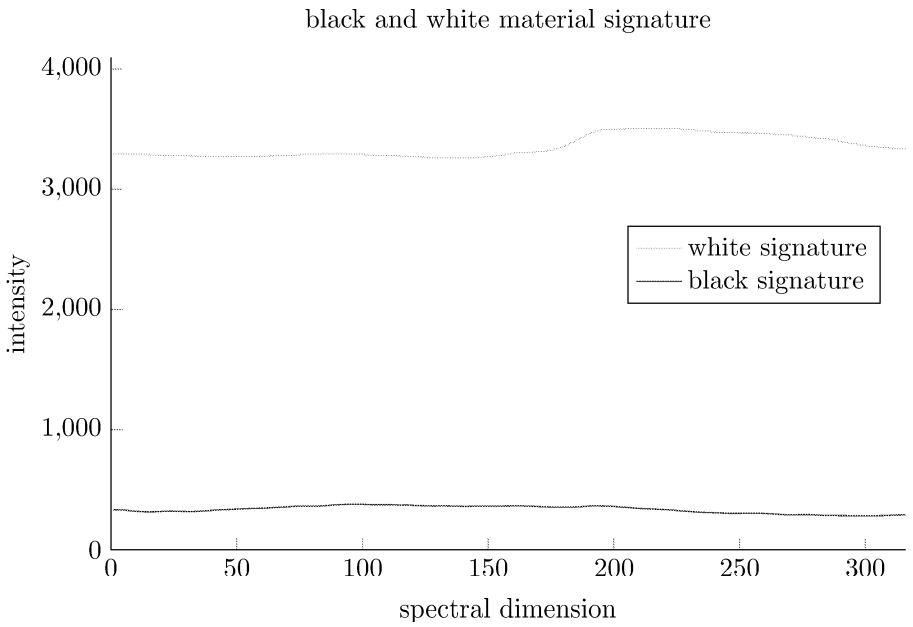


Figure 5.4: Plot of the black and white colour spectral signature from the calibration pattern

5.3.3 Determining Shifting Factor s

As described in section 5.3.1, the conveyor belt speed controls the overlap area, so, for our approach it is necessary to have accurate information about the shifting factor s, which is

$$s = 1 - overlap$$

if an overlap exists. As in our case the speed of conveyor belt is manually set and no information about the exact speed is available, we decide to find the shifting factor by a simple pattern-based measurement method. For this method, we use again the grid which is shown in figure 5.5 and capture the whole grid with the HELIOS camera system. The grayscale image from a captured image with the shifting factor $s = 0.43$ and spectral dimension $d = 200$ is shown in figure 5.6. As we can see in the figure, through the image acquisition process of the system,

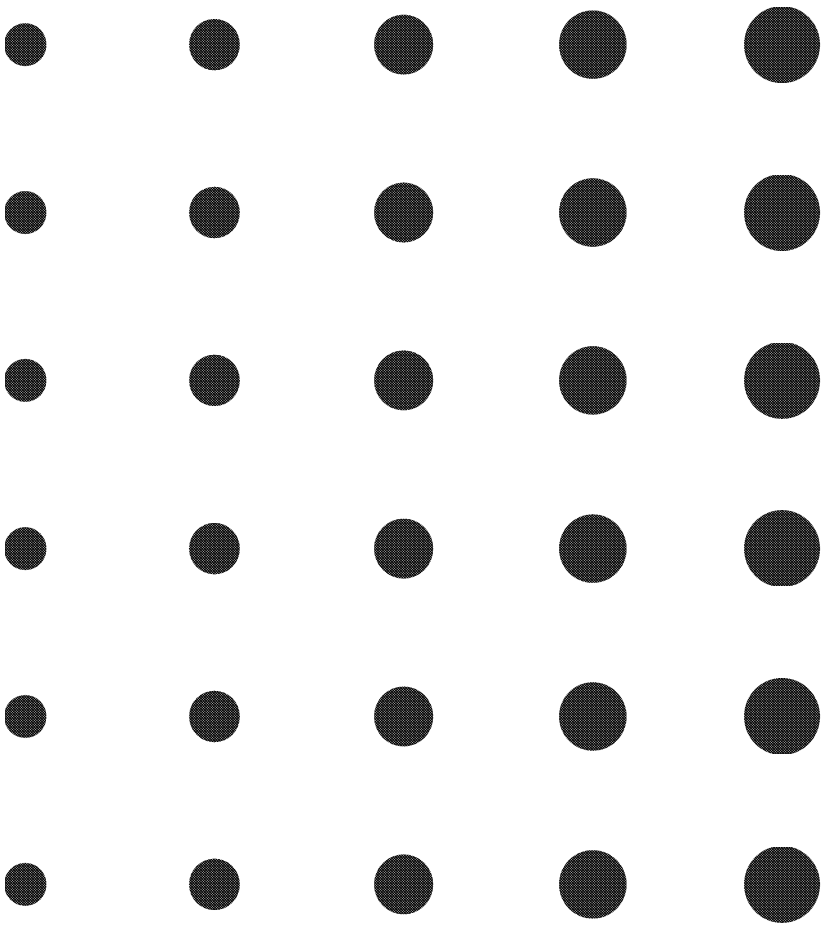


Figure 5.5: This figure shows the layout of the calibration pattern

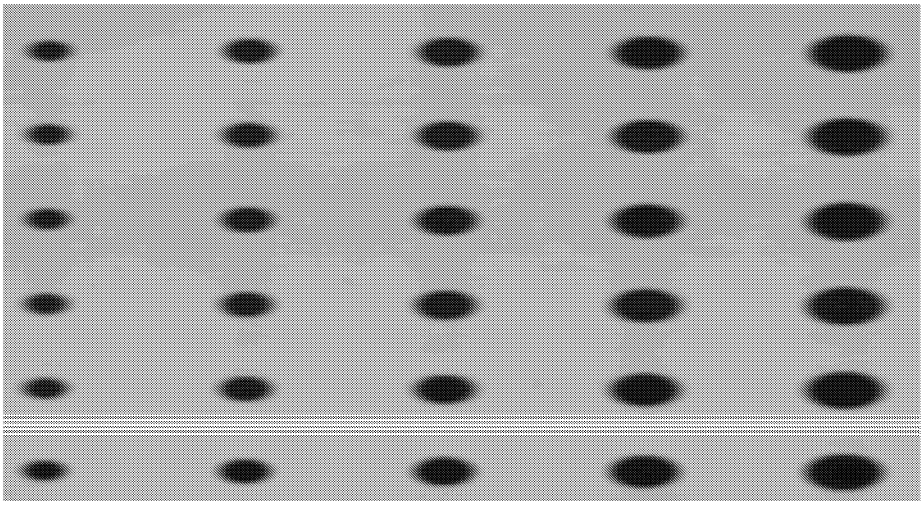


Figure 5.6: This figure shows the grayscale image from the captured pattern with the shifting factor = 0.43 and the spectral dimension 200

a circle on the grid becomes an ellipse on the captured image where the length of the semi-major axis a depends on the shifting factor. As we know that a ellipse is a circle if the lengths of the semi-major and semi-minor axis a and b are equal, we can calculate s by

$$s = \frac{b}{a} \tag{5.7}$$

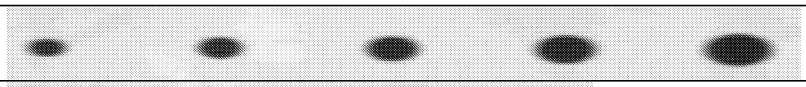
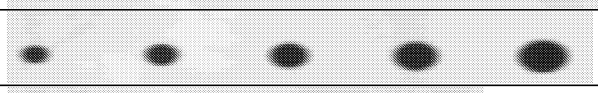
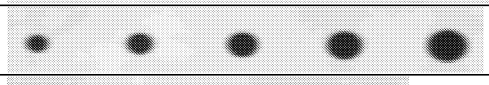
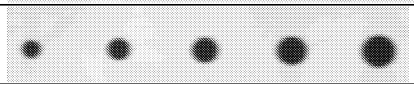
To extract information about the ellipses from the image, we choose an arbitrary spectral band, and treshold it based on the black and white signature information. The treshold value t is calculated by the black intensity b and the white intensity w on the chosen spectral band $d = 200$ in the following way

$$t = \frac{w - b}{2} \tag{5.8}$$

Afterwards, we invert the resulted binary image and select a region with a whole ellipse from the image. Subsequently, we fit an ellipse in this area and calculate the semi-major axis a and semi-minor axis b with the software provided by Ohad Gal [12]. Figure 5.7 shows this process schematically. Based on the extracted ellipse parameter, we calculate s with the equation in (5.7). For more robustness we repeat this process several times and average the calculated shifting factor. In table 5.3, the relationship between the shifting factor s and the captured image grid at various conveyor belt speeds is shown. We can see that the amount of

captured lines for the whole grid increases when the shifting factor decreases. Thus, there is more information about the same area available when the shifting factor is lower. For the real data evaluation, we use all four shifting factors from the table.

Table 5.3: The figure shows the first row of the grid captured by the EVK HELIOS test environment with different conveyor belt speeds respectively shifting factors.

s	first row of captured grid image
0.43	
0.61	
0.78	
0.96	

5.3.4 Determining λ

As we defined λ in section 3.6 as weighting factor between the regularisation term which controls the smoothness of the result and the data term which controls the affiliation to the input data, finding an appropriate λ is also an important process and has strong influence on the results. In the real data set, we expect little noise, so we can attach more importance to the data term, which leads to a relative large λ. We experimentally found out that $\lambda = 150$ works well in combination with the Gaussian blur kernel. In the case of blur kernel estimation by least square, we found out that $\lambda = 30$ works fine.

5.3.5 Determining the Blur Kernel

As we define our hyperspectral image as a series of lines captured by a line scan camera and our approach focuses on increasing the spatial and temporal resolution of the image, we decide to use a two-dimensional kernel. For the real data evaluation, we use different kernel sizes and estimation methods for finding

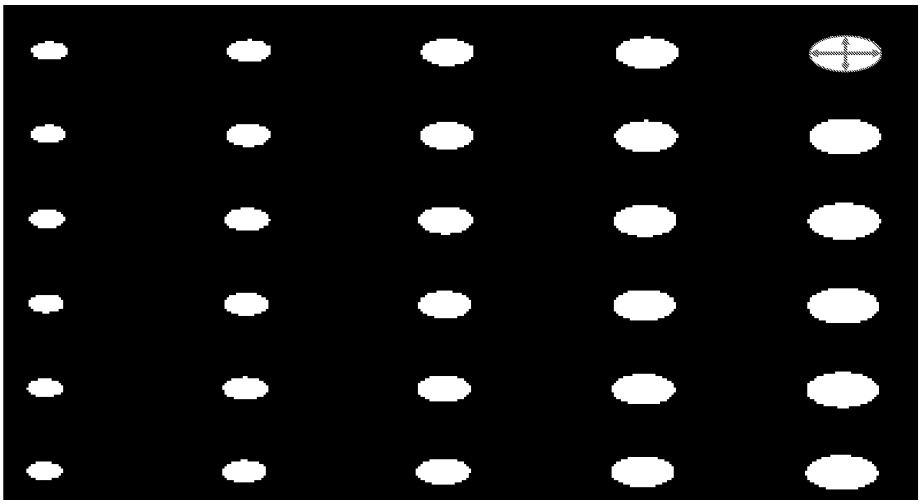


Figure 5.7: By fitting an ellipse in the binary image, the major axis a and b can be calculated, which helps to find the shifting factor s of the current environment

an appropriate kernel, which allows us to reduce the effects caused by the image acquisition process in order to increase the quality of our superresolution results.

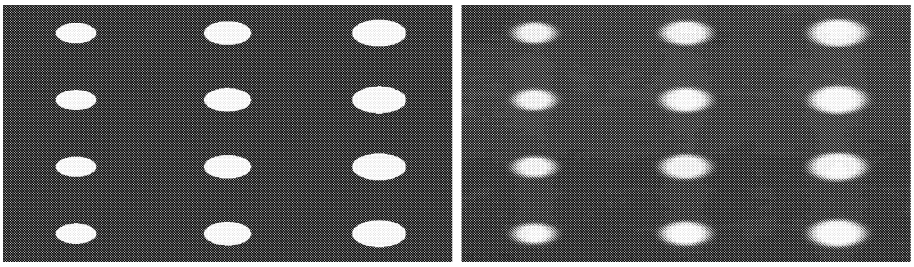
5.3.5.1 Kernel Size

Depending on the spatial resolution of the image, a defined blur kernel size has different influence on the deblurring process. This is obvious because if we use a fixed kernel size for an image and we double up the spatial resolution of the image, the kernel loses the half range of influence on the image. As we compute superresolution images with different spatial and temporal image sizes for real data evaluation, we also have to compute different blur kernel sizes to see which of them works best. We decide to use kernel sizes either 3, 5 or 7 in x or y-direction.

5.3.5.2 Blur Kernel Estimation with Constrainted Least Squere

Once we have estimated the shifting factor s, we calculate the blurring of the camera system. For this, we again use the captured image from the grid shown in figure 5.5 but for the blur kernel estimation, we additionally need a ground truth image. This image is created in the following manner. First, we take the

digital image of the grid and calculate the new image based on the shifting factor s. Subsequently, we register the captured image and the ground truth image and cut off the useless parts created by the registration. The remaining images are then used to estimate the blur kernel with a simple least square problem which is solved by a simple gradient descent update schema, which is described in section 3.6.2.2. In table 5.4, the best results for the CLS blur kernel with



(a) reference image **(b)** blurred image

Figure 5.8: This figure shows the reference and the blurred image which are used to calculate the one dimensional point spread function

increasing factor 1 and 2 are shown.

5.3.5.3 Blur Kernel Estimation with Gaussian

An other method is to assume that the blurring is a Gaussian function. In this case, the blur can be calculated through the Gaussian function, which is described in section 3.6.2.1. For the real data blur kernel estimation, we choose different blur kernel sizes and several σ_x and σ_y. In table 5.5, the best results are shown.

5.3.6 Real Data Evaluation Results

In contrast to the synthetic data, we have no exact ground truth image, so we only can determine the best results by comparing the resulting images. As described before, we have tried several parameter sets to find the best results which are shown in figures 5.9 and 5.10.

Table 5.4: This table shows the least square estimated blur kernels for different increasing factors

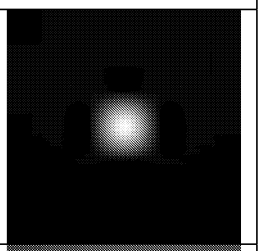
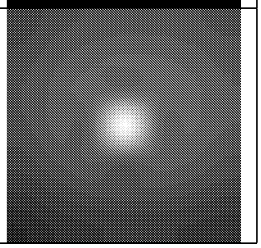
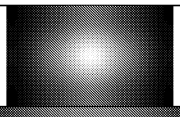
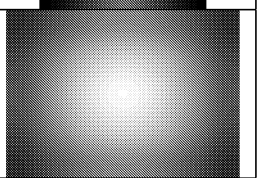
factor	size	blurkernel
1	7×7	
2	7×7	

Table 5.5: This table shows the best estimated Gaussian blur kernels for different increasing factors

factor	size	σ_x	σ_y	blurkernel
1	3×5	0.75	0.85	
2	5×7	1.7	1.9	

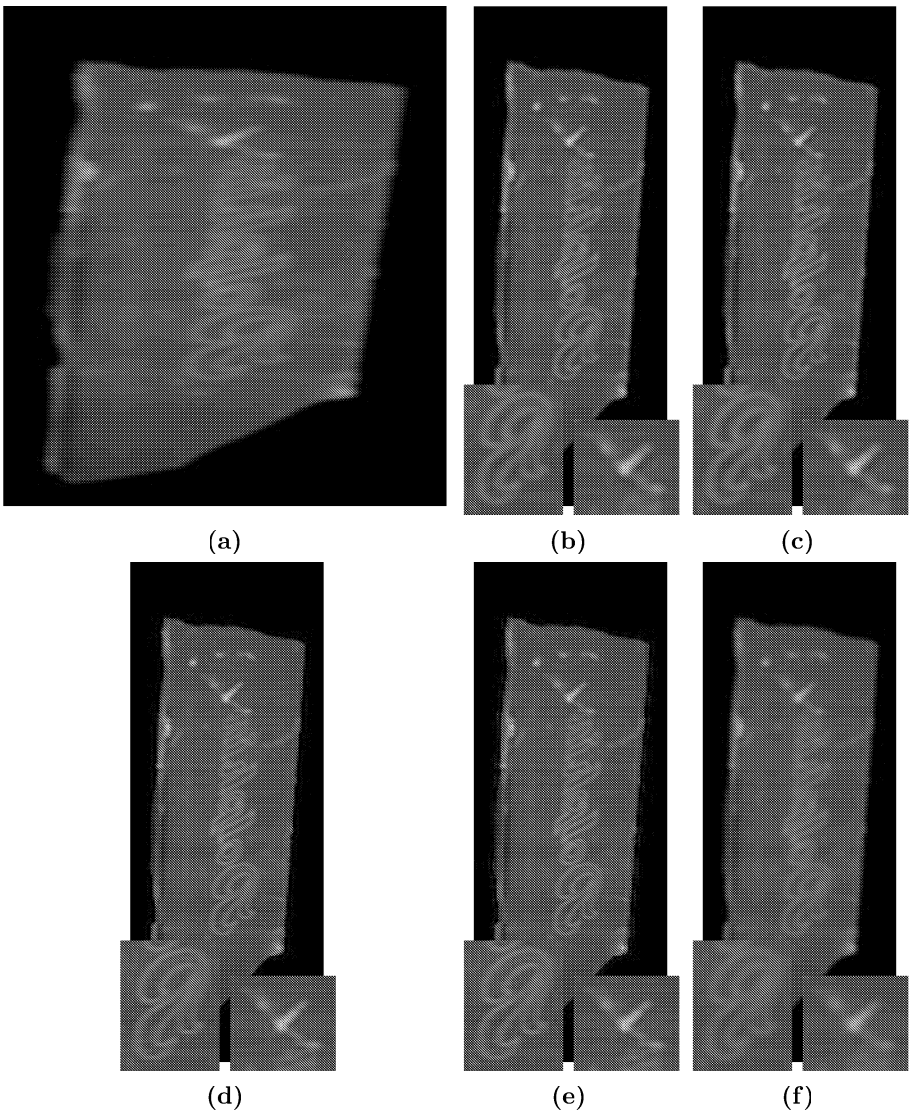


Figure 5.9: This figure shows the best results of proposed ROF and TVL1 method where the input image (a) has a shifting factor $s = 0.43$ and the resulting images (b), (c), (d), (e), (f) have the same spatial resolution as the input image. Image (b) and (c) are the ROF and TVL1 high resolution image with CLS estimated blur kernel, image (d) and (e) are the ROF and TVL1 high resolution images with Gaussian kernel and image (f) is the high resolution image which was calculated by the Matlab bicubic interpolation algorithm. In all images the contrast is adjusted such that the whole range of gray values is used. The both rectangles on the bottom of each image show two image regions in detail.

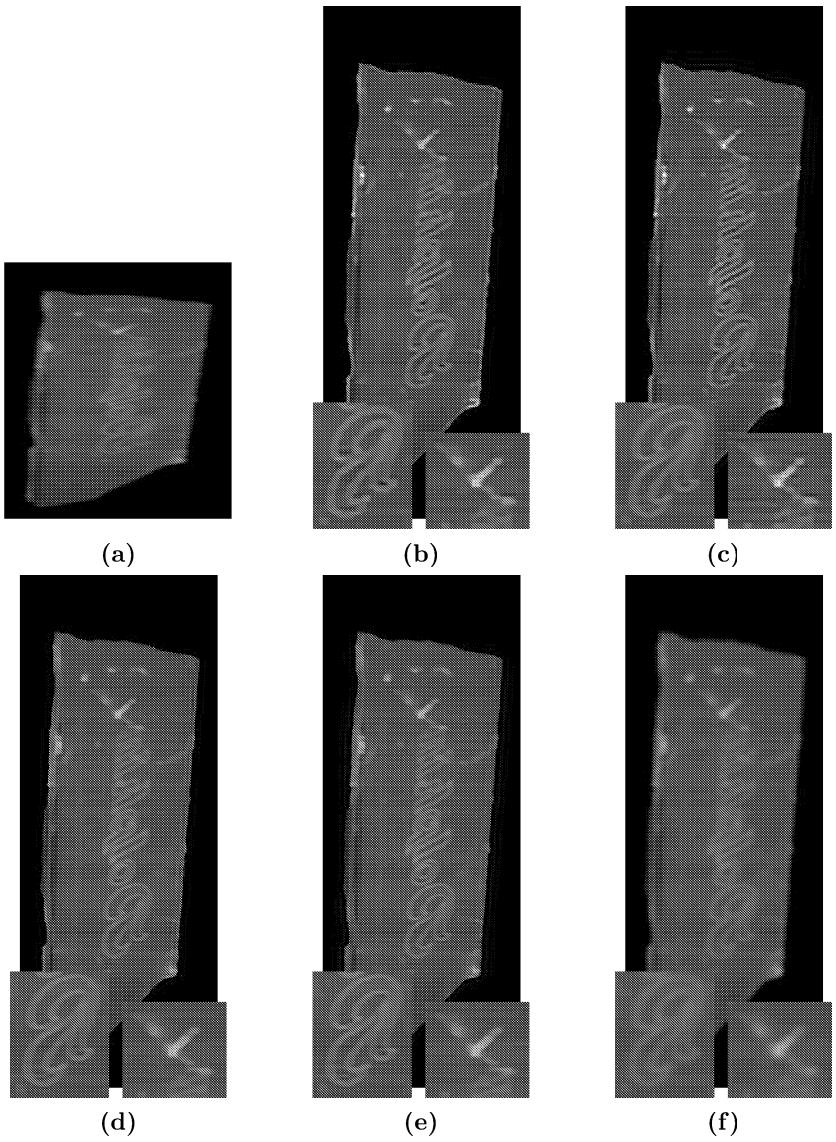


Figure 5.10: This figure shows the best results of proposed ROF and TVL1 method where the input image (a) has a shifting factor $s = 0.43$ and the resulting images (b), (c), (d), (e), (f) have and increased resolution by the factor 2. Image (b) and (c) are the ROF and TVL1 high resolution image with CLS estimated blur kernel, image (d) and (e) are the ROF and TVL1 high resolution images with Gaussian kernel and image (f) is the high resolution image which was calculated by the Matlab bicubic interpolation algorithm.In all images the contrast is adjusted such that the whole range of gray values is used. The both rectangles on the bottom of each image show two image regions in detail.

5.4 Summary

We have found that the least square estimated blur kernels as discussed in 5.3.5.2 do not achieve any improvements compared to Gaussian estimated kernels. Furthermore, we can see that the ROF and TVL1 models are better than interpolation but in real data, there is no great difference. Although no exact evaluation for real data is possible because of no accurate reference images, we can see that the proposed ROF and TVL1 superresolution approach outperform the bicubic interpolation in a real data environment. The high resolution images which are calculated by the ROF or TVL1 approaches are even sharper than calculated by the interpolation approach. Moreover interpolation is not able to correctly recover the distortion caused by the overlap by adjacent pixels in temporal dimension, reduce blur or noise effects. So, finally, we can say that the proposed approaches are more powerful and useful in hyperspectral imaging systems than bicubic interpolation.

Chapter 6

Conclusion and Outlook

Contents

6.1 Conclusion 83
6.2 Outlook 84

6.1 Conclusion

In this master thesis, we have shown how a superresolution reconstruction optimisation problem can be derived from a HELIOS hyperspectral line-scan camera based material separation system in order to achieve a higher hyperspectral image resolution. Chapter 1 provided an overview of spectroscopy, remote imaging and (hyper)spectral imaging. Furthermore, the layout and several components of a HELIOS camera based hyperspectral line-scan imaging system has been explained. And finally, an overview of superresolution, image interpolation and the advantages of high resolution images was given. Chapter 2 showed the fundamental approaches to improve the quality of high resolution images from two ore more subpixel shifted low resolution images and several extension to hyperspectral imaging. In chapter 3, a superresolution approach for a HELIOS camera based system was proposed, which is based on the assumption that such a system delivers hyperspectral images with an overlap in adjacent pixels. With this assumption, a convex superresolution problem was formulated to reconstruct an image with a higher resolution and better quality than the input image. Chapter 4 provided an introduction into to the general primal dual algorithm, which

is a modern method to solve a specific class of problems including the proposed convex superresolution problem from chapter 3. In Chapter 5, we presented the performance evaluations of our proposed approaches and compared it with bicubic interpolation to show the powerfulness and usefulness in synthetic and real environments.

6.2 Outlook

In this master thesis, we analysed the image acquisition process of a HELIOS hyperspectral line-scan camera based material separation system and proposed a ROF and TVL1 superresolution approach which increases the quality and resolution of the hyperspectral image. The performance evaluation has shown that for the real environment the ROF method with an Gaussian estimated blur kernel works best. For future work, we recommend analysing the real-time performance of the proposed algorithms, in particular a CUDA-based GPU implementation because the primal-dual algorithm is particularly qualified for such an implementation. Moreover, some additional constraints and extensions can improve the quality of the high resolution image. Especially incorporating the knowledge of the spectral signature of the endmembers may boost the image quality of the calculated high resolution hyperspectral images.

Bibliography

[1] Ankur Handa, Richard A. Newcombe, A. A., and Davison, A. J. Applications of legendre-fenchel transformation to computer vision problems. Tech. Rep. DTR11-7, Imperial College - Department of Computing, September 2011.

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

[3] Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and der Vorst, H. V. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA, 1994.

[4] Brodzik, A. K., and Mooney, J. M. Convex projections algorithm for restoration of limited-angle chromotomographic images. *J. Opt. Soc. Am. A 16*, 2 (Feb 1999), 246–257.

[5] Chambolle, A. An algorithm for total variation minimization and applications, 2004.

[6] Chambolle, A., Caselles, V., Novaga, M., Cremers, D., and Pock, T. An introduction to Total Variation for Image Analysis, 2009.

[7] Chambolle, A., and Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis. 40*, 1 (May 2011), 120–145.

[8] Columbia University, D. o. C. S. Multispectral image database. `http://www.cs.columbia.edu/CAVE/databases/multispectral/`. [Online; accessed 16-July-2013].

[9] Dong, Y., Hintermüller, M., and Neri, M. An efficient primal-dual method for l1tv image restoration. *SIAM J. Imaging Sciences 2*, 4 (2009), 1168–1189.

[10] Elad, M., and Feuer, A. Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *Image Processing, IEEE Transactions on 6*, 12 (1997), 1646–1658.

[11] Eren, P. E., Sezan, M. I., and Tekalp, A. M. Robust, object-based high-resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing 6*, 10 (1997), 1446–1451.

[12] Gal, O. fit_ ellipse. `http://www.mathworks.com/matlabcentral/fileexchange/3215-fitellipse`. [Online; accessed 16-July-2013].

[13] Geman, S., and Geman, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 6*, 6 (Nov. 1984), 721–741.

[14] Gross, H. N., and Schott, J. R. Application of spatial resolution enhancement and spectral mixture analysis to hyperspectral images. 30–41.

[15] Gu, Y., Zhang, Y., and Zhang, J. Integration of spatial-spectral information for resolution enhancement in hyperspectral images. *IEEE T. Geoscience and Remote Sensing 46*, 5 (2008), 1347–1358.

[16] Guo, Z., Wittman, T., and Osher, S. L1 unmixing and its application to hyperspectral image enhancement. In *Defense, Security, and Sensing* (Apr. 2009), vol. 7334 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, The International Society for Optical Engineering., pp. 73341M+.

[17] Hardie, R. C., Barnard, K. J., Bognar, J. G., Armstrong, E. E., and Watson, E. A. High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system. *Optical Engineering 37*, 1 (1998), 247–260.

[18] Hong, M.-C., Kang, M. G., and Katsaggelos, A. K. Regularized multichannel restoration approach for globally optimal high-resolution video sequence. 1306–1316.

[19] Huang, T. S., and Tsay, R. Y. Multiple frame image restoration and registration. In *Advances in Computer Vision and Image Processing* (Greenwich, 1984), vol. 1, JAI, pp. 317–339.

[20] Irani, M., and Peleg, S. Improving resolution by image registration. *CVGIP: Graph. Models Image Process. 53*, 3 (Apr. 1991), 231–239.

[21] Johnson, C., and Society, A. M. *Matrix Theory and Applications. Proc Held Phoenix, Jan 10-11, 1989.* Proceedings of Symposia in Applied Mathematics Series. Amer Mathematical Society, 1990.

[22] Kim, S. P., Bose, N. K., and Valenzuela, H. M. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech, and Signal Processing 38*, 6 (1990), 1013–1027.

[23] Kim, S. P., and Su, W. Y. Recursive high-resolution reconstruction of blurred multiframe images. *Trans. Img. Proc. 2*, 4 (Oct. 1993), 534–539.

[24] Landgrebe, D. Hyperspectral image data analysis. *Signal Processing Magazine, IEEE 19*, 1 (2002), 17–28.

[25] M. Burgstaller, M. Smart cameras for colouring the invisible. `www.chemical-imaging.com/uploads/4/7/5/6/4756078/colouring_the_invisible_zeiss_jena2010.pdf`. [Online; accessed 16-July-2013].

[26] Manolakis, D., Marden, D., and Shaw, G. A. Hyperspectral image processing for automatic target detection applications, 2003.

[27] Microimages, Inc. Introduction to hyperspectral imaging with TNTmips. Microimages, Inc., Lincoln, NE, USA, 2001.

[28] Mitzel, D., Pock, T., Schoenemann, T., and Cremers, D. Video super resolution using duality based tv-l1 optical flow. In *Proceedings of the 31st DAGM Symposium on Pattern Recognition* (Berlin, Heidelberg, 2009), Springer-Verlag, pp. 432–441.

[29] Moor, J. Application of nir imaging and multivariate data analysis for pharmaceutical products. `http://www.chemical-imaging.com/uploads/4/7/5/6/4756078/20100917graz_nir_spectral_imaging_for_the_pharmaceutical_industry.pdf\\`. [Online; accessed 16-July-2013].

[30] Park, S. C., Park, M. K., and Kang, M. G. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE 20*, 3 (may 2003), 21 – 36.

[31] Patti, A. J., Sezan, M. I., Member, S., Tekalp, A. M., and Member, S. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Transactions on Image Processing 6* (1997), 1064–1076.

[32] Pock, T., and Chambolle, A. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *ICCV* (2011), pp. 1762–1769.

[33] Price, J. C. Combining panchromatic and multispectral imagery from dual resolution satellite instruments. *Remote Sensing of Environment 21*, 2 (1987), 119 – 128.

[34] Robinson, G. D., Gross, H. N., and Schott, J. R. Evaluation of two applications of spectral mixing models to image fusion. *Remote Sensing of Environment 71*, 3 (2000), 272 – 281.

[35] Rockafellar, R. T. *Convex Analysis (Princeton Mathematical Series)*. Princeton Univ Pr.

[36] Rudin, L. I., Osher, S., and Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D 60*, 1-4 (Nov. 1992), 259–268.

[37] Schultz, R. R., and Stevenson, R. L. A Bayesian approach to image expansion for improved definition. *Image Processing, IEEE Transactions on 3*, 3 (1994), 233–242.

[38] Sezan, M. An overview of convex projections theory and its application to image recovery problems. *Ultramicroscopy 40*, 1 (1992), 55 – 67.

[39] Shippert, P. Introduction to hyperspectral image analysis.

[40] Stark, H., and Oskoui, P. High-resolution image recovery from image-plane arrays, using convex projections. *J. Opt. Soc. Am. A 6*, 11 (Nov. 1989), 1715+.

[41] Tekalp, A. M., Ozkan, M. K., and Sezan, M. I. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. In *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on* (Mar. 1992), vol. 3, IEEE, pp. 169–172 vol.3.

[42] Tom, B., Katsaggelos, A., and Galatsanos, N. Reconstruction of a high resolution image from registration and restoration of low resolution images. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference* (1994), vol. 3, pp. 553–557 vol.3.

[43] Trussell, H., and Civanlar, M. The feasible solution in signal restoration. *Acoustics, Speech and Signal Processing, IEEE Transactions on 32*, 2 (1984), 201–212.

[44] Unger, M., Pock, T., Werlberger, M., and Bischof, H. A convex approach for variational super-resolution. In *Pattern Recognition - 32nd DAGM Symposium, Darmstadt, Germany, September 22-24, 2010. Proceedings* (2010), M. Goesele, S. Roth, A. Kuijper, B. Schiele, and K. Schindler, Eds., vol. 6376 of *Lecture Notes in Computer Science*, Springer, pp. 313–322.

[45] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. *IEEE TRANSACTIONS ON IMAGE PROCESSING 13*, 4 (2004), 600–612.

[46] Yang, J., and Huang, T. *Image super-resolution: historical overview and future challenges.* from the book: Super-Resolution Imaging (edited by Peyman Milanfar). CRC Press (Taylor & and amp and Francis Group), 2011.

[47] Yuan, L., Sun, J., Quan, L., and Shum, H.-Y. Image deblurring with blurred/noisy image pairs. In *ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), SIGGRAPH '07, ACM.