

Structure-aware Image Denoising, Super-resolution, and Enhancement Methods

A dissertation submitted towards the degree
Doctor of Engineering (Dr.-Ing.) of the Faculty of
Mathematics and Computer Science of Saarland University

submitted by
Kireeti Bodduna
Saarbrücken 2021



Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Saarbrücken, Germany

Defense Committee Members

Dean: Prof. Thomas Schuster

Chair: Prof. Philipp Slusallek

Reviewers: Prof. Joachim Weickert

Prof. Karen Eguiazarian

Academic Assistant: Dr. Pascal Peter

Date of the Defense: February 22, 2022

Short Abstract

Denoising, super-resolution and structure enhancement are classical image processing applications. The motive behind their existence is to aid our visual analysis of raw digital images. Despite tremendous progress in these fields, certain difficult problems are still open to research. For example, denoising and super-resolution techniques which possess all the following properties, are very scarce: They must preserve critical structures like corners, should be robust to the type of noise distribution, avoid undesirable artefacts, and also be fast. The area of structure enhancement also has an unresolved issue: Very little efforts have been put into designing models that can tackle anisotropic deformations in the image acquisition process. In this thesis, we design novel methods in the form of partial differential equations, patch-based approaches and variational models to overcome the aforementioned obstacles. In most cases, our methods outperform the existing approaches in both quality and speed, despite being applicable to a broader range of practical situations.

Kurzzusammenfassung

Entrauschen, Superresolution und Strukturverbesserung sind klassische Anwendungen der Bildverarbeitung. Ihre Existenz bedingt sich in dem Bestreben, die visuelle Begutachtung digitaler Bildrohdaten zu unterstützen. Trotz erheblicher Fortschritte in diesen Feldern bedürfen bestimmte schwierige Probleme noch weiterer Forschung. So sind beispielsweise Entrauschungs- und Superresolutionsverfahren, welche alle der folgenden Eigenschaften besitzen, sehr selten: die Erhaltung wichtiger Strukturen wie Ecken, Robustheit bezüglich der Rauschverteilung, Vermeidung unerwünschter Artefakte und niedrige Laufzeit. Auch im Gebiet der Strukturverbesserung liegt ein ungelöstes Problem vor: Bisher wurde nur sehr wenig Forschungsaufwand in die Entwicklung von Modellen investiert, welche anisotrope Deformationen in bildgebenden Verfahren bewältigen können. In dieser Arbeit entwerfen wir neue Methoden in Form von partiellen Differentialgleichungen, patch-basierten Ansätzen und Variationsmodellen um die oben erwähnten Hindernisse zu überwinden. In den meisten Fällen übertreffen unsere Methoden nicht nur qualitativ die bisher verwendeten Ansätze, sondern lösen die gestellten Aufgaben auch schneller. Zudem decken wir mit unseren Modellen einen breiteren Bereich praktischer Fragestellungen ab.

Abstract

Structure preservation and enhancement are two of the most important pillars of image processing research. There is a fine but key difference between their definitions. In structure preservation, one is interested in removing undesirable structures but preserving the desirable ones. The latter structures are not significantly altered. In structure enhancement, on the other hand, the desirable structures are significantly altered in order to better visualise them. In this thesis, we consider denoising and super-resolution among the structure preservation applications. We also deal with enhancement of cell structure images obtained using electron microscopy.

Single-frame Denoising. In this application, one tries to obtain a noise-free image from a single raw noisy image. Patch-based methods are approaches which rely on information from a large number of pixels. Consequently, such models produce good denoising results. However, they also tend to form artefacts and require a-priori knowledge about noise statistics. We present a novel approach that avoids the above drawbacks. In particular, we design an iterative non-linear filter that operates on smooth patch reorderings for artefact avoidance. The non-linearity solely encapsulates the properties of a signal but not the noise distribution. Thus, our method is applicable to several synthetic and practical world noise models. Despite excluding noise statistics within the non-linearity, our technique is competitive with state-of-the-art synthetic noise removal approaches which use such information. We also present experiments on practical electron microscopy data where our method outperforms others.

Our above model considers disc-shaped patches. In this thesis, we additionally evaluate sector- and stripe-shaped patches. The sector-shaped modelling in particular uses one-sided derivatives unlike most of the existing filters. The superior structural adaptivity of anisotropic neighbourhoods like sectors and stripes resulted in better preservation of critical structures like edges and corners.

Multi-frame Denoising. Here, we propose three extensions for single-frame patch-based filters when there exist multiple frames of the same scene. The first of these extensions employs reference patches on every frame, thus utilizing the complete available information. The remaining two techniques use a separable spatio-temporal filter to reduce interactions between dissimilar

regions, hence mitigating artefacts. In order to deal with object movement within different frames of a dataset, we combine all our extensions with motion estimation algorithms. Two of our proposed multi-frame filters outperform existing extensions on most occasions by a significant margin. They are also competitive with a state-of-the-art learning-based technique. Moreover, a separable spatio-temporal design makes one of our methods the fastest among several filters.

Multi-frame Super-resolution. In this application, one is interested in computing a high-resolution image from multiple ones with a lower spatial resolution. A so called observational model depicts the relation between a high-resolution image and its low-resolution counterparts. In this context, we present the first systematic evaluation of the observational models for multi-frame super-resolution. This has led to significant improvements in terms of both quality and speed. Moreover, obtaining high-resolution images from data corrupted with clipped noise is algorithmically challenging. So far such issues have hardly been tackled, and the few existing approaches use simplistic models. To this end, we show the usefulness of two techniques which have a better ability to adapt to the structural information within noisy data.

Structure-aware Image Enhancement. Connecting interrupted line-like structures is a frequent problem in image processing. Here we focus on the specific needs that occur in 3D biophysical data analysis in cryo-electron microscopy (cryo-EM). We introduce a powerful framework by combining a specific semi-local Hough transform with a directional evolution equation. The Hough transform allows to find the principal directions in which the structures exist in a robust way. The evolution equation is designed as a partial differential equation that smoothes along these principal directions. We evaluate the structure enhancement performance of our method on both synthetic and cryo-EM data. In contrast to a classical image enhancement model, our method can also handle the anisotropic deformations in the image acquisition process of cryo-EM.

Throughout this thesis, within the applications of denoising, structure enhancement, and super-resolution, we make several extensive evaluations. These are motivated from a point of view of quality as well as computational burden. On both fronts, we have made a significant progress in contrast to existing works.

Acknowledgements

First and foremost, I sincerely thank my supervisor *Prof. Joachim Weickert*, for providing me an opportunity to write a doctoral thesis in his lab. His guidance not only helped me in discovering scientific knowledge but also in learning optimal ways of presenting it to the audience. Several challenging conversations with him saved some crucial time and also led to a better understanding of my own work.

This thesis was funded by four agencies: The German research foundation (through a Gottfried Wilhelm Leibniz prize that was awarded to Prof. Weickert), Saarland University, University of Frankfurt, and the Saarbrücken Graduate School of Computer Science. I especially want to thank *Prof. Achilleas Frangakis* and *Dr. Michelle Carnell*, in this respect. Prof. Frangakis' contribution was also vital in understanding the image processing requirements of the microscopy community. Dr. Carnell was always available for a counselling session at the graduate school.

I also want to thank the members of both Mathematical Image Analysis (MIA) and the Frangakis groups for creating an apt environment for conducting research. Especially the technical discussions I had with my collaborator *Dr. Marcelo Cardenas* were always insightful. Also, the feedback from *Dr. Matthias Augustin*, *Dr. Pascal Peter*, *Dr. Sabine Müller* and *Tobias Alt* on draft versions of my thesis was very helpful. The friendly assistance provided by our administrative staff at MIA, *Ellen Wintringer* and *Peter Franke*, aided me in concentrating solely on research activities.

Prof. Rajesh Siddavatam was my first research supervisor. The motivation he provided and the confidence he had in me were invaluable. My friends and relatives in India and Germany always helped me in recharging my batteries. Last but not least, without the support of my parents *Padmaja* and *Venugopal Bodduna* as well as my brother *Virinchi*, this journey would have been even more difficult.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 15 |
| 1.1 | Goals and Contributions | 18 |
| 1.1.1 | Single-frame Image Denoising | 18 |
| 1.1.2 | Multi-frame Image Denoising | 24 |
| 1.1.3 | Multi-frame Super-resolution | 25 |
| 1.1.4 | Structure-aware Image Enhancement | 28 |
| 1.2 | Thesis Structure | 29 |
| 2 | A Review of Image Acquisition, Noise Models and Error Measures | 32 |
| 2.1 | Image Acquisition Techniques | 33 |
| 2.1.1 | Acquiring Images Through Photon Capture | 34 |
| 2.1.2 | Acquiring Images Through Electron Capture | 34 |
| 2.2 | Noise Models | 36 |
| 2.2.1 | Additive White Gaussian Noise | 36 |
| 2.2.2 | Poissonian Noise | 37 |
| 2.2.3 | Poissonian-Gaussian Mixture Noise | 38 |
| 2.3 | Connections between Image Acquisition Techniques and Noise Models | 39 |
| 2.4 | Statistical Measures of Image Quality | 41 |
| 2.4.1 | Mean Squared Error | 41 |
| 2.4.2 | Peak Signal-to-noise Ratio | 42 |
| 2.4.3 | Fourier Ring Correlation | 43 |
| 3 | A Review of Image Denoising Methods | 45 |
| 3.1 | Gaussian Noise Elimination | 47 |
| 3.1.1 | Cartesian Domain Filters | 47 |
| 3.1.2 | Transformed Domain Filters | 53 |

| | | |
|----------|---|------------|
| 3.1.3 | Dual Domain Filters | 56 |
| 3.2 | Poissonian Noise Elimination | 57 |
| 3.3 | Baseline Denoising Filters Considered | 58 |
| I | Single-frame Image Denoising | 62 |
| 4 | Disc Diffusion - A Non-local Diffusion Model | 63 |
| 4.1 | Modelling and Theory | 64 |
| 4.1.1 | Continuous Modelling of Disc Diffusion | 64 |
| 4.1.2 | Discrete Modelling of Disc Diffusion | 66 |
| 4.2 | Experiments and Discussion | 68 |
| 4.2.1 | Datasets and Methods for Evaluation | 68 |
| 4.2.2 | Parameter Selection | 68 |
| 4.2.3 | Eliminating Synthetically Generated Noise | 70 |
| 4.2.4 | Eliminating Real-world Noise | 71 |
| 4.3 | Conclusions | 72 |
| 5 | Non-linear Filtering on Fast Patch Reorderings | 74 |
| 5.1 | Modelling and Theory | 77 |
| 5.1.1 | Fast Patch Reordering Stage | 78 |
| 5.1.2 | Non-linear Smoothing Stage | 78 |
| 5.1.3 | Correspondences with Disc Diffusion | 81 |
| 5.2 | Experiments and Discussion | 82 |
| 5.2.1 | Datasets and Methods for Evaluation | 82 |
| 5.2.2 | Parameter Selection | 82 |
| 5.2.3 | Synthetic Noisy Data | 83 |
| 5.2.4 | Electron Microscopy Data | 93 |
| 5.3 | Conclusions | 98 |
| 6 | Sector Diffusion - A Corner Preserving Diffusion Model | 100 |
| 6.1 | Modelling and Theory | 102 |
| 6.1.1 | Continuous Modelling of Sector Diffusion | 102 |
| 6.1.2 | Discrete Modelling of Sector Diffusion | 104 |
| 6.2 | Experiments and Discussion | 108 |
| 6.2.1 | Datasets and Methods for Evaluation | 108 |
| 6.2.2 | Parameter Selection | 109 |
| 6.3 | Conclusions | 113 |

| | | |
|------------|--|------------|
| 7 | Importance of Filter Shape in Denoising: Stripe Diffusion | 114 |
| 7.1 | Modelling and Theory | 117 |
| 7.1.1 | Non-symmetric Stripe Diffusion | 117 |
| 7.1.2 | Symmetric Stripe Diffusion | 119 |
| 7.2 | Experiments and Discussion | 120 |
| 7.2.1 | Datasets and Parameter Selection | 120 |
| 7.2.2 | Additive White Gaussian Noise | 121 |
| 7.2.3 | Clipped-Additive White Gaussian Noise | 122 |
| 7.3 | Conclusions | 124 |
| | | |
| II | Multi-frame Image Denoising | 126 |
| | | |
| 8 | Multi-frame Extensions of Patch-based Filters | 127 |
| 8.1 | Modelling and Theory | 130 |
| 8.1.1 | Filters for Single-frame Image Datasets | 130 |
| 8.1.2 | Multi-frame Extensions of Single-frame Filters | 131 |
| 8.1.3 | Optical Flow Methods | 134 |
| 8.2 | Experiments and Discussion | 136 |
| 8.2.1 | Datasets | 136 |
| 8.2.2 | Parameter Selection | 136 |
| 8.2.3 | Perfectly Registered Datasets | 139 |
| 8.2.4 | Non-registered Datasets | 146 |
| 8.3 | Conclusions | 149 |
| | | |
| III | Multi-frame Super-resolution | 151 |
| | | |
| 9 | Evaluating Super-resolution Observational Models and Regularisers | 152 |
| 9.1 | Modelling and Theory | 157 |
| 9.1.1 | Super-resolution Observational Models | 158 |
| 9.1.2 | Optical Flow | 158 |
| 9.1.3 | Edge-enhancing Diffusion | 159 |
| 9.2 | Experiments on Noise-free Datasets | 161 |
| 9.2.1 | Image Datasets | 161 |
| 9.2.2 | Parameter Selection | 162 |
| 9.2.3 | Results and Discussion | 163 |

| | | |
|-----------|--|------------|
| 9.2.4 | More Efficient Model | 166 |
| 9.3 | Experiments on Noisy Images | 167 |
| 9.3.1 | Image Datasets. | 168 |
| 9.3.2 | Parameter Selection. | 168 |
| 9.3.3 | Smoothness Term Evaluation. | 169 |
| 9.3.4 | Data Term Evaluation. | 170 |
| 9.4 | Conclusions | 171 |
| IV | Structure-aware Image Enhancement | 173 |
| 10 | Hough-based Evolutions for Structure Enhancement | 174 |
| 10.1 | Modelling and Theory | 178 |
| 10.1.1 | A General Directional Data Evolution | 179 |
| 10.1.2 | Hough Transform-based Directional Data Evolution | 180 |
| 10.1.3 | Modifying Classical Ideas for Adapting to Cryo-EM Data | 181 |
| 10.1.4 | Numerical Algorithm | 183 |
| 10.2 | Experiments and Discussion | 183 |
| 10.2.1 | Parameter Selection | 183 |
| 10.2.2 | Synthetic Data | 186 |
| 10.2.3 | Real-world Data | 188 |
| 10.3 | Conclusions | 191 |
| 11 | Conclusions and Outlook | 192 |
| 11.1 | Conclusions | 192 |
| 11.2 | Outlook | 193 |
| 12 | Bibliography | 195 |
| 13 | Own Publications | 258 |
| 14 | Abbreviations | 261 |
| 15 | List of Symbols | 263 |

Chapter 1

Introduction

In scientific research, we try to gain new insights about physical and abstract concepts. Several manifestations of these understandings are employed in our everyday life as well as back-channelled for furthering scientific research itself. Imaging applications are very good examples of how we are utilising scientific knowledge to satisfy our daily needs and also to push the limits of our perception. We use these applications in our cell phones besides employing them in microscopy and astronomy (Figure 1.1).

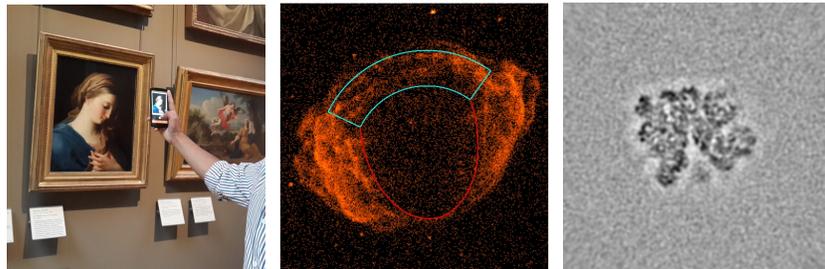


Figure 1.1: Images acquired in various walks of life. **Left:** Cell phone images. **Centre:** Astronomical analysis [1]. **Right:** Microscopic analysis (Courtesy of Lasse Sprankel - Goethe University of Frankfurt).

However, depending on the physical conditions that are present while acquiring desirable imaging data, there is always a possibility of collecting undesirable information. The manifestation or visualisation of such information depends heavily on the specific application. Thus, it is important that we clearly define the application at hand along with the desirable and undesirable data associated with it. For example, in the images correspond-

ing to astronomy and microscopy in Figure 1.1, we can observe unpleasant granular data which one refers to as *unstructured noise*. As the name suggests, such a noise degradation does not have a definite spatial structure. There also exist degradations that have a definite structure, which are referred to as *structured noise*. However, in this thesis, we specifically deal with the former kind which covers a vast majority of imaging applications. Among these unstructured degradations, there exist several subcategories. The formal specification of each one of them is called a *noise model*.

Depending on the type of degradation, we design mathematical techniques that compute noise-free images which give us the best visualisation of the physical world. These methods are called *image denoising filters*. While removing noise from a single image is referred to as *single-frame denoising*, eliminating it from multiple images of the same scene is known as *multi-frame image denoising*.

The above denoising applications produce an output image of the same spatial resolution as the input. *Super-resolution* techniques, on the other hand, have the ability to increase the spatial resolution of images as well as recover lost information due to camera optics. When one tries to produce such high resolution images from multiple low resolution images of the same scene, the specific application is called *multi-frame super-resolution*. The same task when performed under the presence of noise, is referred to as *robust multi-frame super-resolution*.

Ideas from single-frame denoising methods can be extended to multi-frame denoising as well as robust multi-frame super-resolution. This signifies their importance in image processing research. Moreover, all the approaches for above applications are designed from a point of view of preserving image structures while removing noise (Figure 1.2). Such techniques are known as *structure preserving filters*. Astronomy [2–4], microscopy [5–10], seismology [11] and medical imaging [12] are some important examples of areas that require such approaches.

Many methods that focus on preserving structural discontinuities treat all possible directions equally. However, we also encounter situations where removing discontinuities in a particular direction leads to better perception of images. In such cases, we perform a smoothing along a particular local direction which enables structure enhancement (Figure 1.2). In contrast to structure preserving techniques, this is a less explored field and referred to as

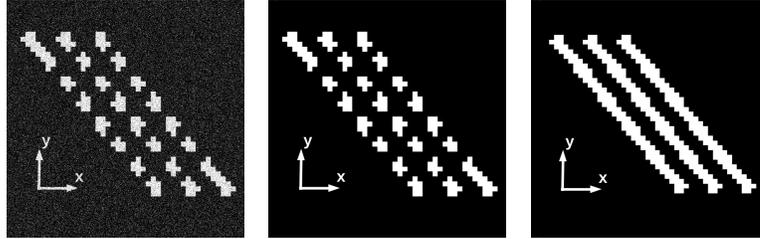


Figure 1.2: **Left:** Noisy image. **Centre:** Ideal output of a structure preserving filter with noisy image on left as input. **Right:** Ideal output of a structure enhancing technique with noise-free image in the centre as input.

structure enhancing image processing. Fingerprint images, artistic paintings, computed tomography scans, and clothing fibre images have been enhanced by such algorithms [13].

Chapter Structure. We divide the rest of this Chapter into two parts. In Section 1.1, we describe our goals and contributions pertaining to both structure preserving and enhancing image processing research. In Section 1.2, we present the organisation of this thesis.

1.1 Goals and Contributions

In this part, we introduce our contributions to four applications: Single-frame denoising, multi-frame denoising, multi-frame super-resolution, and image structure enhancement. We present each of these topics by answering two specific questions: What are the weaknesses in the present state-of-the-art techniques that have been designed for these applications? What are the the main ideas and tools that we use in this thesis to solve these problems? We begin with single-frame filters.

1.1.1 Single-frame Image Denoising

In this application, a single noisy image is used to construct a noise-free one. In order to accomplish this, one generally makes certain assumptions about the properties of noise-free images and noise models associated with the image acquisition process. The non-local Bayes (NLB) [14] and 3D block matching (BM3D) [15] approaches are among the most popular and better performing single-frame filters today. Both these techniques are non-local

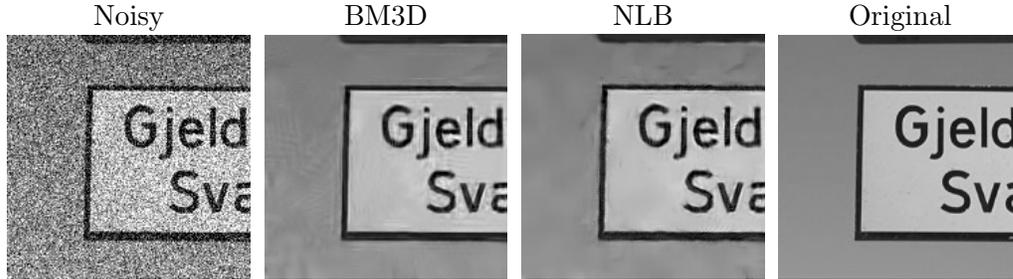


Figure 1.3: NLB and BM3D produce undesirable artefacts in homogenous regions.

and patch-based: “Non-local” because they make use of information from distant regions in an image. “Patch-based” means they assume that similar pixels have similar neighbourhoods around them. The efficient usage of a large amount of information makes such methods very robust. Additionally, NLB and BM3D explicitly model the noise distribution within their frameworks. All these factors lead to the production of superior noise-free images. However, both of these techniques have three problems:

- Information between dissimilar regions in an image might be exchanged.
- Both of these filters can perform an incomplete noise elimination for high noise amplitudes.
- The explicit assumption about a particular noise model sensitises them towards being applicable to other noise distributions.

The first two among the above factors are responsible for existence of undesirable structures in the denoised images, which we refer to as *artefacts* (Figure 1.3). Patch-based methods like NLB and BM3D exist on one side of the spectrum of image denoising approaches. On the other side, we have the diffusion-based methods. These filters are modelled using partial differential equations that control the grey value distribution in an image with respect to time. They have the ability to circumvent artefacts in homogeneous regions of reconstructed images [16] while remaining robust to the kind of noise distribution. However, non-local extensions of these techniques have not yet been thoroughly studied. This brings us to the first goal of this thesis.

Goal 1. Design a non-local diffusion-based approach which is robust to the noise model and also possesses the ability to avoid artefacts in filtered images.

1.1. Goals and Contributions

Our Contribution. We start with a classical diffusion-based idea proposed by Weickert [17]. In this work, the author analysed an anisotropic image evolution process which is capable of preserving edges very well. In order to utilise more information within the filtering process than what is used in [17], we introduce its non-local extension which we call *disc diffusion* (DD).

Even though the iterative nature of the DD framework does not leave any traces of noise in homogeneous regions, it has a couple of drawbacks:

- It is sub-optimal for denoising real world images which contain significant amounts of texture information.
- Although DD has a non-local formulation, it still cannot effectively use information which is very far away like patch-based approaches.

There is another interesting non-local solution existing in the literature which seems promising in avoiding these drawbacks: Ram et al. [18, 19] proposed to employ a smooth patch-based reordering of pixels and subsequently filter them. This idea is powerful enough to produce results of BM3D quality by just using a very basic filter. However, the patch reordering is computationally very expensive and basically requires to solve a travelling salesman problem. Hence, our second goal evolved:

Goal 2. Design a method which combines ideas from sophisticated diffusion-based filtering and simplified patch-based pixel reorderings.

Our Contribution. In order to reduce the computational burden, we employ a simple sort operation for patch reordering. However, this comes at the expense of some disordered pixels which can lead to a loss in reconstruction quality. To compensate for this factor, we use a filtering technique that rewards both patch and pixel similarities in a multiplicative manner. Such a combination results in an approach that is robust in situations when one of the similarity assumptions is violated due to the presence of noise. The specific pixel similarity implementation that we use, is inspired from diffusion-based methods. Thus, this novel filter is robust to the kind of noise distribution and also avoids artefacts. Keeping in mind the ingredients of our model, we name it as *non-linear filtering on fast patch reorderings* (NFPR).

NFPR can be considered as a patch-based extension of the disc diffusion approach. However, there is still scope for improvement in terms of analysing

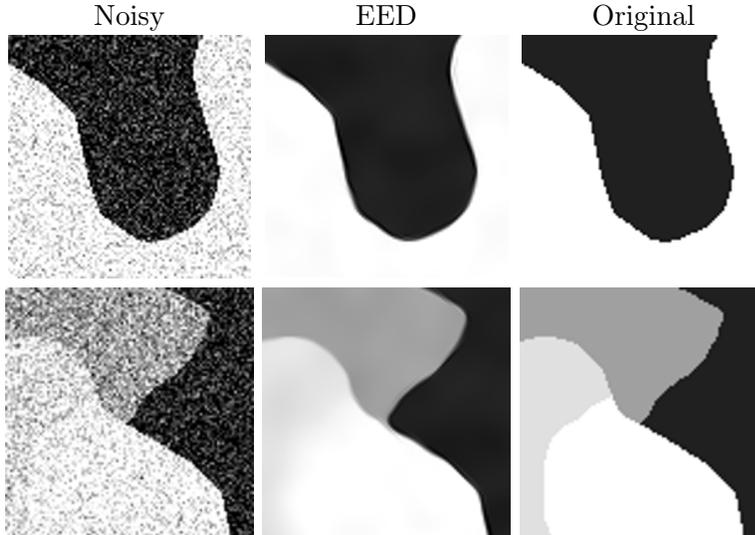


Figure 1.4: **Top:** EED removes fine structures along an edge. **Bottom:** EED also does not properly preserve the corners in comparison to the original image.

shapes other than a disc. This would help in superior performance of the filter in textured regions.

First, let us consider another approach that has seen broader applications, the edge-enhancing diffusion (EED) framework by Weickert [13]. It tries to enhance the quality of an image by smoothing along the edge structures present in it. However, edges represent one-dimensional discontinuities. Corners, on the other hand, are formed when multiple edges intersect. Even though EED is able to preserve edges, its design is not suitable to achieve the same for corners. This can be clearly seen in Figure 1.4, where fine structures along an edge and corner regions are disturbed by EED. In general there is a lack of corner preserving filters, as this is a difficult modelling task in contrast to edge preservation. It is important to fill this gap as corners form a significant part of textured regions in an image.

Goal 3. Design a diffusion-based technique that has the capability to preserve both corners and edges, in the presence of noise.

Our Contribution. We divide a disc-shaped neighbourhood into sectors. This particular shape has a superior ability to adapt to corner regions in the

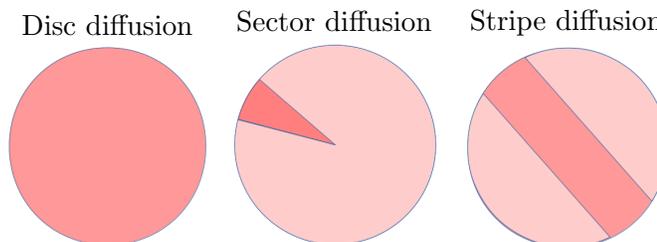


Figure 1.5: Structure element shapes for various diffusion techniques.

image. We then perform smoothing within these sectors for preserving the corners. We call the resulting method *sector diffusion* (SD).

Our SD approach incorporates one-sided derivatives in its continuous model unlike EED. To our knowledge, diffusion filters that are explicitly based on one-sided derivatives have not been described in the literature so far. The division of a disc into sectors is a direct consequence of using one-sided derivatives. Such ideas are mathematically very challenging and could open the door to several new directions of research.

Although sectors are suitable for preserving corners, they have a smaller number of pixels than a disc. The percentage of pixels in an image that belong to corner regions is rather low. Thus, if we adopt a sector-shaped neighbourhood for all regions in the image, this would lead to a sub-optimal result in terms of overall denoising performance. One of the solutions for this is to calculate a separate shape-adaptive neighbourhood for every pixel in the image [20–23]. However, computing a different shape for every pixel is computationally very expensive. This brings us to our next goal.

Goal 4. Design a structure element which acts as common ground in between a disc and a sector.

Our Contribution. The new structure element shall cover a larger area than a sector, but show a better ability to catch anisotropic behaviour than a disc. Stripes satisfy both of these requirements. Thus, we introduce the stripe-shaped structure element in a diffusion-based context. Figure 1.5 shows the disc-, sector-, and stripe-shaped structure elements.

This brings us to the end of our contributions in this thesis pertaining to single-frame denoising. Now, we move on to its multi-frame counterpart.

1.1.2 Multi-frame Image Denoising

Computing a single denoised image from multiple corrupted images of the same scene is called the multi-frame image filtering problem. We encounter this scenario in video denoising as well as when we acquire multiple images in highly noisy applications like microscopy. Most multi-frame filters, unsurprisingly, are extensions of ideas borrowed from the single-frame scenario. The currently best performing multi-frame extensions employ combined spatio-temporal filtering ideas [24–26]. However, separation of spatial from temporal information can lead to artefact avoidance as it reduces interactions between dissimilar regions. Furthermore, a systematic evaluation of multi-frame extensions is also missing.

Goal 5. This particular goal is a combination of two objectives:

- Design a multi-frame extension which separates spatial from temporal information while filtering.
- Complete a comprehensive evaluation of multi-frame denoising techniques.

Our Contribution. In order to achieve the above objectives, we propose the following two solutions, respectively:

- We use filter-then-average and average-then-filter methodologies for separating spatial from temporal data. This leads to artefact reduction.
- Our comprehensive and systematic evaluation includes three kinds of experiments. Firstly, we evaluate our proposed multi-frame extension on three different single-frame methods - NLB, BM3D, and NFPR. This provides evidence in order to consider our technique as a general one for extending single-frame patch-based filters. Secondly, we perform the above experiments for three kinds of frequently encountered noise distributions. This gives an idea of the best available approach for various practical situations. Thirdly, in order to deal with object movement within a multi-frame dataset, we have made use of robust motion compensation methods.

This concludes our contributions that are related to denoising. In the upcoming part, we deal with super-resolution techniques.

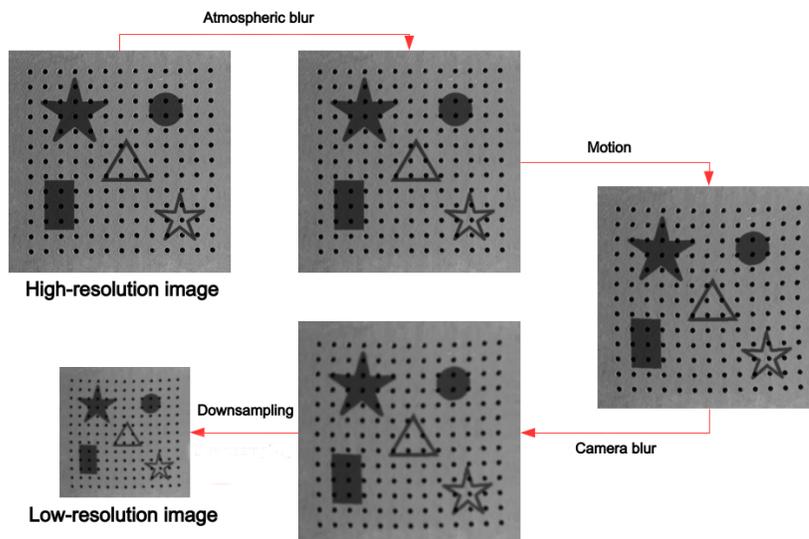


Figure 1.6: The degradations undergone by a high resolution scene in the process of acquiring a low-resolution image.

1.1.3 Multi-frame Super-resolution

The particular difference between multi-frame super-resolution and multi-frame denoising is that, in super-resolution we try to increase the spatial size of the images. We also make an effort to recover information which was lost due to camera optics. Formally, we try to compute a high resolution scene from multiple low resolution realisations of it. The relation between both resolutions is encoded in a so-called *observational model*. The standard observational model for super-resolution (SR) reconstruction was first proposed and represented in a matrix-vector formulation by Elad et al. [27]. It encompasses a sequential modelling of the degradations due to blurring, motion, and downsampling (Figure 1.6).

The physical phenomena mentioned in the above figure are represented by mathematical operators in the observational model. One can see that the low resolution images are blurred versions of the high resolution scene. In order to compute the latter, we try to invert the acquisition process. To this end, we compute the motion on the low resolution blurred images and its up-sampled version is approximated as high resolution motion. Both utilisation of blurred images as well as motion computation itself may lead to errors

which paves way for our next goal.

Goal 6. Find the best order of operators in the observational model to minimise the above mentioned errors while computing a high resolution scene.

Our Contribution. We systematically evaluate all possible permutations of the standard SR observational model in order to identify the one which leads to reduction in errors. To this end, we exploit the flexibility provided by variational models. We encode various combinations of the observational model within these models.

While the above mentioned variant of super-resolution did not take into account the presence of noise, there also exist situations where noise reduction and resolution enhancement are simultaneously required. This is the robust multi-frame super-resolution problem. Both operator permutations and image filtering algorithms need to be investigated in this scenario.

Goal 7. Evaluate the various permutations of the SR observational model in the presence of noise. Additionally, investigate the advantage of a non-local denoising technique in the SR scenario.

Our Contribution. We evaluate all possible combinations of the standard SR observational model in a noisy layout. The particular noise model employed also covers over- and under-exposed imaging conditions. Furthermore, we also use the sector diffusion and edge-enhancing diffusion operators for regularisation purposes.

Denoising and super-resolution strategies are structure preserving algorithms as explained in Figure 1.2. In the following, we discuss their less known structure enhancing counterpart.

1.1.4 Structure-aware Image Enhancement

The classical coherence enhancing diffusion (CED) technique [13, 28] is the best example of a structure enhancing method. It is another partial differential equation-based framework like EED, which enhances coherent structures in an image. An example for fingerprint data is shown on the left side of Figure 1.7. However, CED is unable to account for the special limited angle image acquisition of cryo-electron microscopy (cryo-EM). Here, the acquired

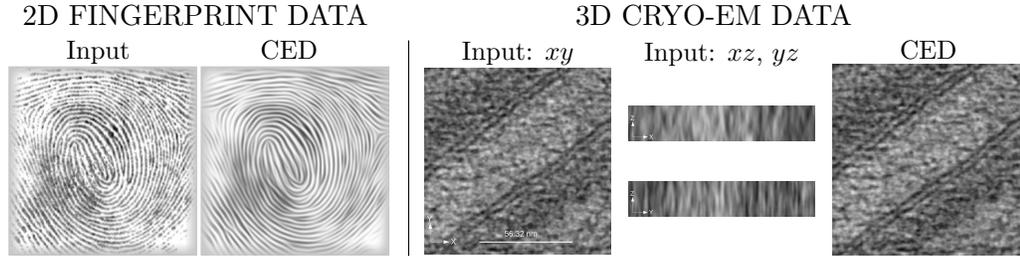


Figure 1.7: **Left:** Coherence enhancement with CED on fingerprint image (Courtesy of Joachim Weickert). **Right:** CED is not suitable for cryo-EM data (Courtesy of Achilleas Frangakis).

data is prone to directional blurring and noise degradations (right side of Figure 1.7). Since CED is misdirected by the blur in the z direction, it is unable to produce the desired structure enhancement in the xy planes. Thus, the final goal of this thesis evolved:

Goal 8. Design a model for processing data acquired through special limited angle cryo-EM techniques: It should be able to enhance image structures like CED and also be robust with respect to noise.

Our Contribution. We introduce the Hough-based image evolution framework. It is a combination of a general directional image evolution process with the classical Hough transform [29–32]. The latter method is generally used to detect complex patterns in images. We use a semi-local version of it to find the local direction in which the image structures exist. This is followed by steering the smoothing process according to the directional image evolution. The flexible and robust nature of the Hough transform is majorly responsible in dealing with the data acquisition problems in cryo-EM.

This ends the brief discussions on our goals and contributions pertaining to both structure preservation and enhancement. In the ensuing section, we present the organisational structure of the thesis.

1.2 Thesis Structure

We begin by reviewing a couple of widely used image acquisition procedures, in Chapter 2. Here, we also introduce our mathematical notations. This

is followed by a formal review of the noise models associated with image acquisition and also the various statistical measures of image quality.

In Chapter 3, we briefly review the huge amount of research on image denoising filters. Our presentation puts a highlight on the particular noise distributions and physical assumptions underlying the design of these filters. The novel results in this thesis are presented in four parts:

Part I - Single-frame Image Denoising. We introduce the disc diffusion model in Chapter 4 as a non-local extension of a classical anisotropic diffusion method. In Chapter 5, we present the non-linear filtering on fast patch reorderings approach. It is basically a patch matching-based extension of disc diffusion. This is followed by the introduction of the sector diffusion technique in Chapter 6, where we showcase the corner preservation ability of this model. Subsequently, we propose the stripe diffusion approach in Chapter 7. This particular chapter will help us understand the importance of filter shapes in image denoising.

Part II - Multi-frame Image Denoising. In Chapter 8, our novel multi-frame extensions of patch-based filters are introduced. We perform a systematic and comprehensive evaluation of these strategies for three different types of synthetic noise distributions.

Part III - Multi-frame Super-resolution. In Chapter 9, we perform an extensive evaluation of the super-resolution observational model. Here, we also evaluate the ability of sector diffusion and EED as SR regularisers in a noisy scenario.

Part IV - Structure-aware Image Enhancement. Chapter 10 contains a description of our Hough-based evolution technique for cryo-EM data processing.

A joint summary of the main conclusions from this thesis is presented in Chapter 11, together with outlooks on further research directions. At the end of the thesis, one can find information regarding bibliography, list of symbols, abbreviations, and own publications.

Chapter 2

A Review of Image Acquisition, Noise Models and Error Measures

In Chapter 1, we briefly introduced the various goals of this thesis along with our contributions in achieving them. It is time that we begin to learn more details about these topics. A natural place to start would be to get acquainted with the procedures through which one acquires digital images. As almost all image acquisition is noisy, this naturally leads to questions about the formal representations of noise models associated with specific acquisition techniques. Hence, in this chapter, we cover topics related to both acquisition and associated noise models. Additionally, we also review three standard statistical error measures that help quantify the quality of images.

Chapter Structure. In Section 2.1, we review a couple of standard image acquisition methodologies. In the ensuing Section 2.2, we formulate the common noise models that are considered in the image processing community. In Section 2.3, we connect noise models with acquisition procedures. Finally, in Section 2.4, we review the statistical measures of image quality.

2.1 Image Acquisition Techniques

In this thesis, we consider digital images acquired using two basic procedures: The well-known photon capture and the less-known electron capture.

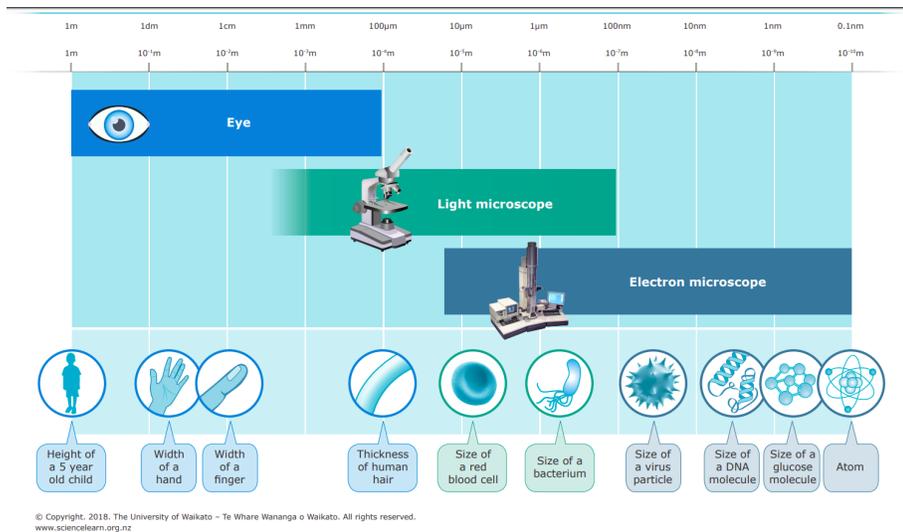


Figure 2.1: Juxtaposition of the resolving power of our eye, light microscopes, and electron microscopes (Courtesy of Science Learning Hub - Pokapū Akoranga Pūtaiao, University of Waikato. www.sciencelearn.org.nz).

Electrons can have a much smaller wavelength when compared to visible radiation. Hence, the resolution of the resulting images is considerably much higher than those acquired using light microscopy or other visible radiation capturing hardware. Figure 2.1 shows the relative resolving power of these two techniques as well as our own eye. However, the high resolving power of electrons often comes at the expense of having a larger amount of noise. Thus, depending on the task at hand, one uses these two techniques accordingly. In the following, we briefly describe both methods.

2.1.1 Acquiring Images Through Photon Capture

The phenomenon by which one acquires a digital image through photon-based capturing, is described by a series of physical processes [33–49]. The scene which is to be imaged, radiates photons (either by itself or by reflecting an already existing beam) which are cast onto the image acquisition device. The photons are then directed by the camera lens onto the semiconductor elements of a charge-coupled device (CCD) sensor or a complementary metal oxide semiconductor (CMOS) sensor. At this stage, the photons are converted into electrical charges which are amplified and further converted into

digital numbers. Each digital number represents the intensity at every pixel, which is the fundamental square element of a sensor. The final raw image one obtains is a collection of intensities at all the pixels. In a CCD sensor, the above analog to digital conversion is performed for every row separately. On the contrary, in a CMOS sensor, the conversion to voltage followed by amplification is calculated at every pixel. It is thus obvious that in both these processes the amount of accumulated charge decides the value of intensity at each pixel in the acquired images. Now we move on to the electron acquisition counterpart of this procedure.

2.1.2 Acquiring Images Through Electron Capture

Gathering imaging data using electron beams is a property of electron microscopy techniques [50, 51]. Here, an electron beam is incident on a sample which is to be imaged. After the beam passes through the sample, there exist both unscattered and scattered electrons. We can categorise the latter type into elastic and inelastic scattering both of which possess information about the atoms they have interacted with, in the sample. This data is encoded in the amplitude and phase factors of the scattered electrons which is converted into digital numbers through CCD, CMOS or hybrid pixel detectors [52–57].

Even though raw images acquired through the above methodologies fall under the discrete data category, in order to process them we also utilise knowledge from continuous mathematics. It is thus evident that we need a notation to clearly differentiate between various elements from both discrete and continuous mathematics. To this end, we use the convention presented in Table 2.1, throughout this thesis.

According to the above notational style, we represent the positions of the pixels in raw images as discrete sampling points Ω in a continuous image domain $\Omega \subset \mathbb{R}^2$. The discrete and continuous collections of intensities throughout this domain are represented by \mathbf{f} and f , respectively. Now, we are ready to formally review various noise models associated with photon- and electron-based image acquisition procedures.

2.2 Noise Models

In this section, we describe three different types of noise models. In particular, we cover categories which do not have a definite spatial structure,

| Type | Examples | Details |
|---------------------------|--------------------------------------|------------------------------|
| Scalars | a, b, c | lower case |
| Vectors | $\mathbf{u}, \mathbf{v}, \mathbf{f}$ | lower case + bold |
| Matrices | \mathbf{D}, \mathbf{Q} | upper case + bold |
| Functions | f, g, h | lower case |
| Statistical Distributions | \mathcal{G}, \mathcal{P} | upper case + calligraphics |
| Functionals | E | upper case |
| Sets | Ω, \mathbb{R} | upper case + special symbols |

Table 2.1: Convention for mathematical symbols.

so-called unstructured noise. This type of degradation covers a vast majority of applications. Information regarding structured noise can be found in [58, 59]. We begin with the ubiquitous Gaussian noise model.

2.2.1 Additive White Gaussian Noise

Let $\Omega \subset \mathbb{R}^2$ denote a rectangle and $\mathbf{x} \in \Omega$ be a position. Consider a noisy image $f : \Omega \rightarrow \mathbb{R}$ acquired through photon or electron capture, or one that is synthetically generated on a computer. Either way, under the *additive noise* model f is considered as an additive combination of the original noise-free image $v : \Omega \rightarrow \mathbb{R}$ and pure noise $n : \Omega \rightarrow \mathbb{R}$. This model [45] can be formulated as

$$f(\mathbf{x}) = v(\mathbf{x}) + n(\mathbf{x}). \quad (2.1)$$

The noise distribution for $n(\mathbf{x})$ can be approximated by various statistical distributions.

The most common additive noise model is *additive white Gaussian noise* (AWGN). Here, one assumes an explicit Gaussian distribution for the pure noise $n(\mathbf{x})$ as a function of time [45]. The resulting probability density function of a Gaussian random variable $n(\mathbf{x})$ is denoted as $n_{\mathcal{G}}$ and is explicitly defined as

$$n_{\mathcal{G}}(n(\mathbf{x}), \mu_{\mathcal{G}}, \sigma_{\mathcal{G}}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{G}}} \cdot e^{-\frac{(n(\mathbf{x})-\mu_{\mathcal{G}})^2}{2\sigma_{\mathcal{G}}^2}}. \quad (2.2)$$

Here, $\mu_{\mathcal{G}}$ and $\sigma_{\mathcal{G}}$ denote the mean and standard deviation of the Gaussian distribution, respectively. Also, $n(\mathbf{x}), \mu_{\mathcal{G}} \in \mathbb{R}$ and $\sigma_{\mathcal{G}}^2 \in \mathbb{R}^+$. The noise variance is independent of the intensities in the input image and thus makes this model a signal-independent one. Moreover, this particular type of degradation has huge practical consequences [49] as the central limit theorem allows

to model several independent noise contributors in an imaging system as a normal distribution.

The “colour” of the noise - for example white, blue, pink, or brown - is named depending upon the properties of the corresponding power spectral density in the Fourier domain. Contrary to others, white noise has a flat power spectral density, i.e., one that does not depend on the frequency. It is the most commonly encountered type and, thus, the only noise color we consider in this thesis. Moreover, in accordance with the general practice in the denoising community, we set μ_G to zero. Hence, we just consider a Gaussian model with zero mean, additive nature, and white color.

2.2.2 Poissonian Noise

Under this non-additive noise model [60], we consider the raw image f to be a collection of independent random Poisson variables whose mean is the underlying signal v to be estimated. The corresponding conditional probability is characterised by a Poissonian distribution and is explicitly defined as follows,

$$P(f(\mathbf{x})|v(\mathbf{x})) = \mathcal{P}(v(\mathbf{x})) = \frac{v(\mathbf{x})^{f(\mathbf{x})} e^{-v(\mathbf{x})}}{f(\mathbf{x})!}. \quad (2.3)$$

Here, $v(\mathbf{x}) \in \mathbb{R}^+$ and $f(\mathbf{x}) \in \mathbb{N} \cup \{0\}$. One can derive that the mean and variance of the Poissonian distribution are the same: $\mu_{\mathcal{P}} = \sigma_{\mathcal{P}} = v(\mathbf{x})$. Thus, they are dependent on the noise-free signal. This is reason why the Poissonian distribution \mathcal{P} , in the above equation, has an argument $v(\mathbf{x})$.

2.2.3 Poissonian-Gaussian Mixture Noise

One can formulate a combination of AWGN and Poissonian noise models as [61]

$$f(\mathbf{x}) = n_{\mathcal{P}}(\mathbf{x}) + n_{\mathcal{G}}(\mathbf{x}). \quad (2.4)$$

The Gaussian component $n_{\mathcal{G}}(\mathbf{x})$ is obtained by considering $\mu_G = 0$ in (2.2). The Poissonian component $n_{\mathcal{P}}(\mathbf{x})$ is characterised as follows,

$$n_{\mathcal{P}}(\mathbf{x}) \sim \mathcal{P}(\chi \cdot v(\mathbf{x})). \quad (2.5)$$

In the above equation, \mathcal{P} denotes a Poisson distribution as already defined in (2.3). The only difference is that we have an additional multiplication factor

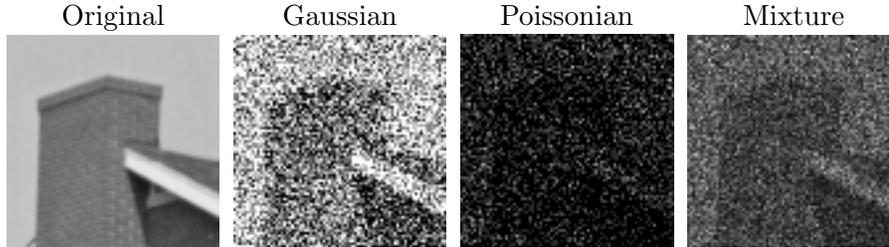


Figure 2.2: House image corrupted with different types of noise. $\sigma_G = 80$ for Gaussian, $\chi = 1.0$ for Poissonian, $\chi = 5.0$ and $\sigma_G = 127.5$ for mixture noise.

χ , which controls the amount of noise in the raw images. The overall variance of the mixture model can be derived by using the elementary properties of Poissonian and Gaussian distributions as [61]

$$\sigma_{\text{noise}}^2 = a \cdot v(\mathbf{x}) + b. \quad (2.6)$$

Here, $a = \frac{1}{\chi}$ and $b = \sigma_G^2$. As one can observe, this noise model is a combination of both signal-dependent and -independent types. Since the variance of the Poissonian component is inversely proportional to χ , a low value of it indicates high noise amplitudes. It is very important to note that, throughout this thesis, we use σ_G and χ to represent the amount of noise. The estimation of the noise characteristics from raw images is also an active area of research [61–67]. Figure 2.2 shows a noise-free image corrupted with different types of degradations. These noise corruptions are generally specific to the image acquisition procedure. We take a more detailed look at these connections in the next section.

2.3 Connections between Image Acquisition Techniques and Noise Models

The degradations of the image signal in photon-based acquisition can be classified into three categories [33, 34, 41, 49]:

- Signal-dependent temporal noise (photon shot noise).
- Signal-independent temporal noise (reset, thermal, flicker, and dark current shot noise).

2.3. Connections between Image Acquisition Techniques and Noise Models

- Fixed-pattern noise (dark FPN, light FPN, defected pixels, and hot spots).

As the names suggest, the first two categories are defined based on observations in the time dimension. The last one is defined according to its spatial structure and falls under the structured noise category. There exist efficient techniques [49, 68] to overcome the degradations due to reset, flicker, and fixed-pattern noise at the hardware stage itself. However, it is difficult to circumvent the signal corruption due to photon shot, thermal, and dark current photon shot noise categories. Thus, we require image denoising algorithms. While photon shot noise is approximated as Poissonian observations of the original signal, the rest of the degradations are usually modelled as AWGN [33, 41, 49].

Similar to capturing photons, the image collection process in cryo-electron microscopy (cryo-EM) - an electron capture-based technique - also has signal corruptions. The raw observations were approximated using Poissonian statistics in [8, 9]. The entire image processing pipeline in cryo-EM, however, is a complex procedure. Noisy datasets within this pipeline were considered as additive and signal independent in [51]. Nevertheless, the body of research regarding details of noise distributions in cryo-EM is substantially smaller than the one regarding the acquisition of normal images. This is due to three major difficulties: Huge amount of noise in high resolution acquisition (pixel size is about 10^{-10} meters), large image sizes (typically 8000×8000 pixels), and our less frequent encounters with such images.

Lately, the mixture noise model in (2.4) has been well-accepted for raw sensor images [61, 69–72] as it encapsulates all possibilities between both extremes of signal-dependent and -independent noise categories. This general noise model is thus valid for several applications such as astronomy [2–4], microscopy [5–10, 73–75], seismology [11], and medical imaging [12]. In [61, 76, 77], an additional noise clipping was also considered in order to take into consideration the lost dynamic range because of over- and under-exposures. Unless and until specifically mentioned, throughout this thesis we do not clip the dynamic range of the corrupted signal.

Both noise models and raw images are used as input for image denoising methods. Before we review these filters themselves in the next chapter, we first present the statistical error measures used to quantify the quality of a reconstructed image. These measures, in principle, are designed such

that their output is coherent with our own visual senses. The comparative evaluation of various image reconstruction algorithms requires such methods.

2.4 Statistical Measures of Image Quality

In this thesis, we use three mathematical measures: Mean squared error (MSE), peak signal-to-noise ratio (PSNR), and Fourier ring correlation (FRC). While the first two measures are designed in the Cartesian/spatial domain, FRC is calculated in the Fourier domain. The structural similarity index measure (SSIM) [78] along with its extensions [79–83] is also a commonly used metric, but is rather complex. Both MSE/PSNR and SSIM have their own advantages and disadvantages [78, 84, 85]. The simplistic nature of the MSE/PSNR is the reason for their ubiquitous usage in contrast to SSIM or other visual perception-based metrics [86]. In the following, we describe MSE, PSNR, and FRC in more detail. All these approaches take as input two images which we want to compare and give as output a number (MSE/PSNR) or a vector (FRC) that gives insights into the similarity of both input images. Thus, if one of these images is a denoised image and another one is a ground truth, these measures allow us to judge the quality of the denoising procedure.

2.4.1 Mean Squared Error

Let $\Omega \subset \mathbb{R}^2$ denote a rectangle. The mean squared error between the restored image $\mathbf{u} : \Omega \rightarrow \mathbb{R}$ and the original noise-free image $\mathbf{v} : \Omega \rightarrow \mathbb{R}$, is defined as

$$\text{MSE}(\mathbf{u}, \mathbf{v}) = \frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} (\mathbf{u}_{(i,j)} - \mathbf{v}_{(i,j)})^2. \quad (2.7)$$

Here, the operator $|\cdot|$ denotes cardinality of the set. It is trivial that MSE is zero if and only if both images are equal and the value is always positive for unequal images. Moreover, a low value indicates high similarity between the images. In the image denoising literature, MSE is frequently used when the images are corrupted by AWGN.

2.4.2 Peak Signal-to-noise Ratio

This particular statistic is just a logarithmic rescaling of the MSE value:

2.4. Statistical Measures of Image Quality

$$\text{PSNR}(\mathbf{u}, \mathbf{v}) = 10 \log_{10} \left(\frac{255^2}{\text{MSE}(\mathbf{u}, \mathbf{v})} \right). \quad (2.8)$$

Here, the assumption is that the range of the image is bounded from above by 255 and below by 0. The logarithmic rescaling and the specification of the bounds makes PSNR an apt measure for Poissonian noise scenarios. Moreover, unlike MSE, a high PSNR value indicates high similarity between the images. However, since logarithm is a monotonous function, MSE and PSNR both give rise to the same order of experimental rankings.

Both MSE and PSNR are generally used when we have the original noise-free image \mathbf{v} , available. In practical scenarios this is not the case. Thus, we can use two strategies here: Firstly, we perform a lot of experiments on data where we add the noise by ourselves. This gives us an impression of the optimal range for different parameters used in the filter. We then use this knowledge along with our own visual cues in the practical scenario. However, if we also want to have a statistical measure in real world situations, FRC is an apt criterion.

2.4.3 Fourier Ring Correlation

FRC is an error assessment method that is widely used in cryo-EM [87]. It measures the similarity between two images in the Fourier space. One first calculates the Fourier coefficients of the two images, which are divided into a number of narrow frequency ranges called rings. For each ring, a corresponding cross-correlation coefficient between the two images is calculated. A curve is thus plotted with ring number on x-axis and corresponding cross-correlation coefficient on the y-axis. This is called the FRC curve. An FRC curve between two completely similar images would be the straight horizontal line $y = 1$. This indicates highest similarity at every frequency.

Thus, for cryo-EM data, we first consider two pre-aligned noisy images of the same sample. In order to judge the quality of a denoising algorithm, we employ it on both of these images. Afterwards, we calculate the FRC curve between the two denoised images. The physical interpretation of this procedure is the following: The denoised image pair must either possess the original structures in the scene or leftover noise and artefacts created due to noise elimination. Since the latter properties vary from one noise realisation of a scene to another, an FRC curve indicates the quality of the former. In

general, one draws a horizontal line at $y = 0.5$ [88] and checks the corresponding highest frequency (ring number) at which this line intersects the FRC curve. The parameters of the denoising model being tested are generally selected such that this particular frequency is high. Once we have the optimal parametric selections, we sum up the two acquired images in order to increase the signal content. The denoising algorithm with the obtained parametric combination is then employed on this summed up image, which yields the final denoised version.

The mathematical definition of the FRC between two images \mathbf{u} and \mathbf{v} can be formulated as

$$\text{FRC}(\mathbf{u}(r), \mathbf{v}(r)) = \frac{\sum_{r_i \in r} \hat{\mathbf{u}}(r_i) \cdot \hat{\mathbf{v}}(r_i)}{\sqrt{\sum_{r_i \in r} |\hat{\mathbf{u}}(r_i)|^2 \cdot \sum_{r_i \in r} |\hat{\mathbf{v}}(r_i)|^2}}. \quad (2.9)$$

Here, r represents the ring number. The Fourier transformed images of \mathbf{u} and \mathbf{v} are denoted by $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, respectively. The summation is over all the coefficients that are present in a particular ring.

Similar to MSE and PSNR, FRC also might sometimes give a result which defies visual perception. Thus, experts in electron microscopy consider both visual analysis and FRC simultaneously.

Chapter 3

A Review of Image Denoising Methods

In the previous chapter, we have acquainted ourselves with image acquisition techniques, noise models, and statistical measures of image quality. We are now ready to learn more about image denoising filters. In this chapter, we thus review the literature on single-frame filters which also form the core of multi-frame denoising and robust super-resolution.

Single-frame noise elimination methods produce the noise-free image u from a single raw input image f . To this end, every such denoising filter employs the following three key steps:

- *Gathering Similar Information* - Natural world noise-free images have several regions which are made up of either repetitive or less varying grey values. Based on this fundamental property, we perceive and differentiate between multiple objects and multiple regions within the same object. A noise realisation of a scene affects this property and thus also our visual perception. In order to denoise a corrupted pixel/reference pixel, one first selects a set of pixels which are similar to the reference one. This selection process depends on certain similarity assumptions which are specific to the type of filter.
- *Combining/Filtering Similar Information* - The goal of this step is to combine the selected set of pixels in the first step, to yield a single value that will replace the noisy reference pixel value. The motivation behind this is to regain back the lost property of possessing repetitive

and less varying grey values in noise-free images. To this end, each of the pixels in the above set are assigned weights that are specific to the type of filter under consideration. These weights decide the amount of contribution of each pixel within the set.

- *Reconstructing Image from Filtered Information* - Some denoising methods yield multiple denoised versions of a reference pixel. This final step is responsible for obtaining one value from these multiple versions.

In this thesis, we will concentrate on only those categories of noise where all pixels are assumed to be corrupted. Such degradations cover a vast majority of imaging applications and hence are the most considered ones in the denoising community. The other category, where only a percentage of the total number of pixels are corrupted, is called *impulse noise*. Examples of filters designed for such degradations can be found in [89, 90]. Based on this bifurcation, various denoising techniques apply the above three filtering steps on either all pixels or just a percentage of them. In order to design these steps, one makes certain assumptions about the properties of both noise-free images and the formal noise model which they are designed to eliminate. Since the former is generally common among all the image denoising methods, one generally bifurcates these techniques based on the latter: Methods that are designed for eliminating additive white Gaussian noise (AWGN) and Poissonian data. We begin with the AWGN model.

3.1 Gaussian Noise Elimination

The assumptions about the desired noise-free image and the noise type can be made in both spatial domain or a transformed domain. Accordingly, AWGN filters are further classified into three parts: Cartesian, transformed, and dual domain techniques.

3.1.1 Cartesian Domain Filters

One utilises three key physical assumptions about noise-free images in order to design a filter in the Cartesian domain:

- Pixels within a small spatial neighbourhood of an image have similar intensities. This is called *spatial assumption*.

3.1. Gaussian Noise Elimination

- Grey values or tonal values of pixels which look alike have minimum intensity variation among them. This is referred to as *tonal assumption*.
- Pixels which are similar to each other also possess almost identical neighbourhoods around them, which is the *neighbourhood assumption*.

In the following, we briefly summarise several types of Cartesian AWGN elimination methods keeping in mind the above physical assumptions. We discuss the inter-category correspondences and equivalences between these different kind of filters, once we complete introducing each one of them. We begin with the classical local smoothing noise elimination methods.

Local Smoothing Filters. These approaches use the spatial assumption to design a filter whose output can be formulated in general, as

$$u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{\Omega(\mathbf{x}, \mathbf{h})} w_s(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) d\mathbf{y}. \quad (3.1)$$

Here, $\Omega(\mathbf{x}, \mathbf{h})$ is a neighbourhood around \mathbf{x} whose size is determined by \mathbf{h} . The weights for each spatial position \mathbf{y} within this neighbourhood are represented by w_s . Further, $C(\mathbf{x}) = \int_{\Omega(\mathbf{x}, \mathbf{h})} w_s(\mathbf{x} - \mathbf{y}) d\mathbf{y}$ is a normalisation factor. The very common Gaussian smoothing [91], the filters of Nadaraya [92] and Watson [93] together with the vast class of local polynomial approximation (LPA) filters [94–99], fall under the category of local smoothing filters. Moreover, several image adaptive versions of LPA filters have also been designed [98–109]. One considers a combination of both tonal and spatial assumptions in order to model image adaptive approaches. Other techniques similar to such adaptive methods go by the names moving least-squares, reproducing, moment, Savitzky-Golay, and kernel regression [49] filters.

Range/Sigma Filters. These are another class of denoising methods which also consider a combination of spatial and tonal weighting functions, similar to the image adaptive LPA filters. An inclusive formula for computing the denoised image using such techniques can be given as

$$u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{\Omega(\mathbf{x}, \mathbf{h})} w_s(\mathbf{y}) w_t(f(\mathbf{x}) - f(\mathbf{y})) f(\mathbf{y}) d\mathbf{y}. \quad (3.2)$$

The Yaroslavsky [110], Lee [111], Susan [112] and bilateral [113, 114] filters fall into this category of range filters. Bilateral filtering, especially, has also

been adopted in various other image processing applications [115–117].

Diffusion Filters. These techniques compute the denoised image as a solution to a partial differential equation (PDE) [118–123]. Similar to range filters, diffusion-based techniques also employ the spatial and tonal assumptions. One can formally encompass these methods using the following general equation:

$$\partial_t u = \operatorname{div}(\mathbf{D} \nabla u). \quad (3.3)$$

The noisy image f is used as the initialisation for the above evolution. The term $\partial_t u$ denotes the temporal derivative of u and ∇u represents the continuous 2D spatial gradient. A so-called diffusivity function determines the entries of the 2-by-2 diffusion tensor \mathbf{D} . Both ∇u and entries of the diffusion tensor together model the spatial and tonal assumptions: In linear isotropic diffusion [91, 124–127], one uses a constant, scalar diffusivity. In the case of nonlinear isotropic diffusion [128–130], one still has a scalar diffusivity, but this is now a function of the gradient magnitude of the involving image. Anisotropic diffusion techniques [131–134] extend these ideas further by constructing a symmetric positive definite diffusion tensor which can no longer be equivalently expressed by a single scalar-valued function. Both non-linearity and anisotropy (preferential directional smoothing) are responsible for preserving image structures.

Variational Filters. Another class of methods which employ the spatial and tonal assumptions are the variational denoising approaches. In these techniques, one calculates the denoised image as a minimiser of a particular energy functional which models the required physical assumptions. An overall formulation of these filters can be written as

$$\begin{aligned} u &= \operatorname{argmin}_m E(m, f), \\ E(m, f) &= E_{\text{data}}(m, f) + E_{\text{smooth}}(m). \end{aligned} \quad (3.4)$$

Here $E(m, f)$ denotes the energy functional which can be decomposed into data part E_{data} and smoothness term E_{smooth} . The data term ensures that the solution is close to the initial noisy image f and the smoothness term allows us to model the various physical assumptions. The huge class of total variation penalty-based filters [135–143] and recent works by Steidl and Teuber [134, 144] are good examples of variational denoising techniques.

3.1. Gaussian Noise Elimination

Through the above Equation (3.4), one is looking for an image that minimises the energy. However, a denoising operator which directly minimises the MSE between the denoised image and the original image can also be computed. The best linear operator of such a kind is the classical Wiener filter [145]. This technique needs information about the noise model in order to approximate the original image using the noisy one. Once it is available, one can compute the Wiener filter in both spatial [146] or transformed domains. A locally adaptive version of the spatial Wiener filter has been studied by Lee [147]. We re-introduce Wiener filtering formally when we review transformed domain methods in a later section of this chapter.

Bayesian Filters. A category of filters that is close in spirit to the variational filters are Bayesian denoising techniques [148–152]. Here, one obtains the denoised image as the maximiser of the posterior probability function, which models the spatial and tonal physical assumptions. The restored image using such a type of filter can be computed as

$$u = \operatorname{argmax}_m p_{\text{post}}(m|f),$$
$$p_{\text{post}}(m|f) = \frac{p_{\text{like}}(f|m)p_{\text{prior}}(m)}{p(f)}. \quad (3.5)$$

Here, the Bayes formula has been employed for calculating the posterior probability p_{post} . The likelihood term p_{like} keeps the denoised image close to the initial noise-free image f , while the prior p_{prior} gives one the freedom to model spatial and tonal assumptions [120, 148, 153–156].

Non-local Averaging Filters. The paper of Efros and Leung [157] which synthesizes textures, is considered as the seminal work that modelled the third assumption - the neighbourhood similarity. However, De Bonet [158] has employed such an idea even earlier in the denoising context itself, as mentioned in [67]. At about the same time, a similar idea was also used in fractal imaging [159]. More importantly, the access to additional information in an image through such non-local formulations has revolutionised image processing. Its first major effect in the image denoising community appeared in the form of the non-local means (NLM) algorithm [160, 161]. The neighbourhood similarity can be considered as a combination of both its spatial and tonal counterparts. It is particularly powerful in a denoising scenario because it takes into account information from multiple pixels.

One cannot solely rely on tonal similarity, which uses information from just one pixel, for optimal results. The denoised image u of NLM thus can be computed using the following formula:

$$u(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{\Omega(\mathbf{x},h)} w_{\text{nbd}}(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}. \quad (3.6)$$

The neighbourhood weights $w_{\text{nbd}}(\mathbf{x}, \mathbf{y})$ in the NLM algorithm are computed using a Gaussian function. An L_2 -distance between square-shaped patches surrounding \mathbf{x} and \mathbf{y} is the argument of this function.

Several data adaptive and automated parameter selection versions of NLM have also been studied [162–167]. A contemporary technique of NLM that was proposed by Awate and Whitaker [168], also exploits the advantages of non-local information. Moreover, when we said non-local formulations have revolutionised image filtering methods, we meant that such ideas have been adopted in variational methods [169–173], Bayesian approaches [14, 174–179], range filters [180, 181] as well as local smoothing filters [18]. These ideas are also being used in various other image processing applications [58, 169, 172, 182–185]. However, the non-local extensions of diffusion-based methods have not been yet thoroughly studied. This is one of the main goals of our work.

Correspondences Between Various Filters in the Cartesian Domain.

Keeping in mind the volume of literature on image denoising, it is not surprising that we find correspondences and equivalences between different filters. There is a significant amount of literature specifically dedicated to uncover such relations. However, we can only find correspondences between different type of filters in the Cartesian domain if they model the same type of assumptions. For example, since both range filters and partial differential equations try to model spatial and tonal assumptions, one can find correspondences between them [123, 186–192]. Similarly, we can also find studies which concentrate on variational, PDE, and Bayesian approaches [120, 193–195]. There also exists a correspondence between the NLM algorithm and the non-local Bayes approach [14, 196], both of which model the neighbourhood assumption. This shows the importance of having a physical perspective: One can come up with ever more complicated techniques. However, if the choice of the underlying physical assumptions are not given adequate importance, there is high chance that the filter has equivalences with existing ones. Thus,

not leading to significant practical progress. In other words, one must carefully and completely review the combination of the assumptions used in the existing literature, to come up with novel methods.

3.1.2 Transformed Domain Filters

The second type of AWGN elimination methods are the transformed domain-based filters. As the name suggests, one eliminates AWGN by using certain transforms of the initial noisy image f , like Fourier or wavelet. After performing a filtering operation in the transformed domain, the images are back transformed to the Cartesian domain for visualisation. The main assumption of these methods is that the basic structure of images can be represented using only a few coefficients (a sparse set) in the respective transformed domains. Attenuating the less informative coefficients of the noisy image in the transformed domain is the core denoising feature of these techniques. Unlike Cartesian domain assumptions, this design is relatively difficult for us to visualise since our visual capacity operates in the Cartesian space.

Let us start by remembering the idea on which the Wiener filter is built. The goal is to find an operator \mathbf{w} which when acted upon the noisy image, gives the least MSE value between the resultant and the original image \mathbf{v} . Formally one can write this operator as

$$\mathbf{w} = \underset{\mathbf{m}}{\operatorname{argmin}} \operatorname{MSE}(\mathbf{m} * \mathbf{f}, \mathbf{v}). \quad (3.7)$$

Here, $\mathbf{m} * \mathbf{f}$ is a discrete convolution operation. Let us assume that the noisy image is approximated as $\mathbf{f} = \mathbf{v} + \mathbf{n}$, with \mathbf{n} representing the noise vector. The solution to the above equation in a discrete Fourier basis thus turns out to be the operator

$$\hat{\mathbf{w}}(\hat{\mathbf{x}}) = \frac{|\hat{\mathbf{v}}(\hat{\mathbf{x}})|^2}{|\hat{\mathbf{v}}(\hat{\mathbf{x}})|^2 + |\hat{\mathbf{n}}(\hat{\mathbf{x}})|^2}. \quad (3.8)$$

Here, $\hat{\mathbf{x}}$ is a 2D frequency vector, $\hat{\mathbf{v}}(\hat{\mathbf{x}})$ represents the 2D discrete Fourier transform of $\mathbf{v}(\mathbf{x})$ and $|\hat{\mathbf{v}}(\hat{\mathbf{x}})|^2$ is its power spectrum. The above solution can also be formulated in the Cartesian domain with the help of convolutions [186]. Such operations attenuate the contributions of the high frequencies which noise is generally associated with. Since this filter requires the noise-free image, it is called an ideal filter or the Fourier-Wiener filter.

We must keep in mind that the above filter is a global operator as it acts on the complete image in the Cartesian domain. This property might give rise to unwanted structures in the filtered image that correspond to the global structure. Thus, in order to avoid such artefacts, one needs to localise the effect of the filter. This has been achieved using wavelets [197–204], discrete cosine transform (DCT), and also a mixture of bases [205].

The other problem in using the ideal Wiener filter is that one needs to estimate the Fourier coefficients of the original signal. Yaroslavsky [206, 207] proposed the estimation of the original signal through the noisy image by using the AWGN model followed by a localised DCT filtering. Such a technique is referred to as the empirical Wiener filter. However, this still does not completely get rid of the artefacts, which can be attributed to attenuating the coefficients that also represent the edges. The probability of such an occurrence with wavelets, on the other hand, is lower as they are optimal for sparse representations of the data [186]. The initial efforts in this direction (Donoho and Johnstone [208], Donoho [209]), only partially got rid of these filtering artefacts. Many advanced ideas [210, 211] were employed to try and solve this problem: Data adaptive threshold selection [212], translation invariance [138], wavelets in combination with variational [213–217] and Bayesian [218–223] filters, and data adaptive basis techniques which are not just restricted to wavelets [22, 109, 224–231]. Among the data adaptive basis techniques there exist strategies [229, 230] where the number of elements in the basis exceeds the dimensionality of the signal. One refers to this as an overcomplete basis and it allows for relatively sparser representations [232].

In a similar fashion to that of Cartesian domain techniques, we observe two things about transformed domain filters. On the one hand, non-local patch-based extensions [15, 23, 233–235] of transformed domain-based filters utilise information from far-away regions. On the other hand, there exist studies which find correspondences and equivalences between wavelet-based, PDE-based and variational methods [193, 194, 236–238].

3.1.3 Dual Domain Filters

While all the above mentioned denoising filters are designed in either the Cartesian or a transformed domain, there also exist techniques which use a combination of both. We can divide these filters into two classes for a better understanding: The non-local patch-based [21, 239, 240] and the semi-

3.2. Poissonian Noise Elimination

non-local [232, 241–243] approaches. Members of the former class perform a post-processing of results obtained using either Cartesian or transformed domain strategies. In particular, they employ a Fourier shrinkage of the shape- and data-adaptive patches considered from the initial denoised results. Methods belonging to the latter class do not adhere to the patch-similarity idea. They assume that the noise statistics are known and estimate the noise variance in both spatial and frequency domains. Once noise is estimated, it is subtracted from the raw image to obtain the noise-free signal.

With the above discussion, we have come to the end of our brief survey of AWGN elimination filters. We close it by mentioning the two general approaches that are frequently used in the image denoising world. These ideas in practice can be applied to almost every denoising method. Firstly, although patch-based methods have a non-local impact, it is still limited to a particular search window. In order to increase the spatial scale of the filter impact, the multi-scale approach is widely used [208, 240, 244–252]. It basically considers a processing of the data at several spatial scales. Secondly, there are also techniques which use patch and image information from external databases as well [175, 179, 240, 253–259].

Now that we have reviewed the various classes of AWGN elimination methods, we move on to the less prominent Poissonian noise removal filters.

3.2 Poissonian Noise Elimination

In this particular section, we cover methods for elimination of both Poissonian noise and Poissonian-Gaussian mixture noise. These methods can be divided into the ones which deal directly with Poissonian noise and the ones which use schemes that enable the application of AWGN elimination algorithms. We review these techniques in the same order.

Let us remember that, in order to design a denoising algorithm, we need to make assumptions about the noise-free image and the noise model under consideration. These assumptions can be made in both Cartesian and transformed domains. Under the spatial domain category, surprisingly very few filters have been designed for Poissonian noise elimination. A couple of works in this direction are the following: The first of them proposes a method which uses the patch similarity assumption through the utilisation of an external database of patches [260]. The second one among them is

an energy minimisation approach which models tonal and spatial assumptions [261]. Otherwise, the three Cartesian domain physical assumptions about the noise-free image are somehow not as frequently used for Poissonian observations as they are employed for AWGN elimination. Most filters in this class are built in a transformed domain. They exploit the fact that the reproducing property of Poissonian distribution exists across all scales for just a particular Haar wavelet transform. In case of AWGN, this is true for all orthogonal wavelet transforms. This difference between both noise distributions [262] has led to a huge amount of research which uses wavelet-based techniques to remove Poissonian noise as well as Poissonian-Gaussian noise [4, 263–273]. Other sparsity-based worth mentioning approaches which directly deal with Poissonian noise can be found in [274–276].

Since it is practically difficult to design image denoising methods for every type of image degradation, the second type of Poissonian denoising algorithms searches for a way to make use of AWGN filters. To this end, the variance stabilising transformations (VSTs) [2, 60, 66, 67, 277–287] convert a Poissonian distribution to its Gaussian counterpart. Once the noise is removed with an AWGN algorithm, a suitable inverse transformation is applied to get back to the original dynamic range. Another filter which applies a similar philosophy is the plug-and-play Bayesian prior framework [288, 289].

With the above information on Poissonian noise elimination techniques, we have completed our review of both AWGN and Poissonian noise removal approaches.

3.3 Baseline Denoising Filters Considered

In this section, we consider five non-local filters from the above huge amount of literature for a further discussion. They have been frequently used for a comparative analysis with the models designed in this thesis:

- 3D block matching (BM3D) [15] - a transformed domain approach.
- non-local Bayes (NLB) [14] and linear filtering on smooth patch reorderings (LFSPO) [18] - two Cartesian domain techniques.
- dual domain image denoising (DDID) [241] and non-local dual domain denoising (NLDD) [239] - two dual domain filtering solutions.

3.3. Baseline Denoising Filters Considered

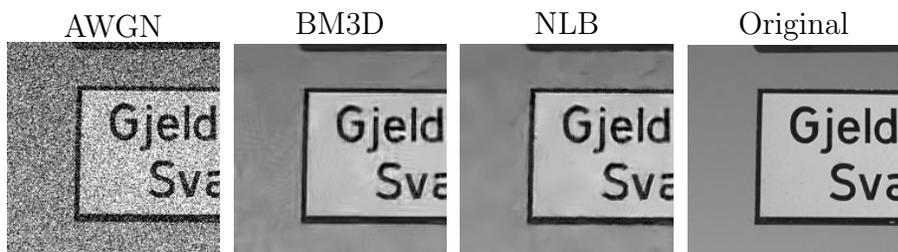


Figure 3.1: Artefacts in noise-free outputs of NLB and BM3D ($\sigma_G = 50$).

Although we already made a small discussion about the artefacts produced by BM3D and NLB (Figure 3.1) in the first Chapter, we explore this topic a bit further now. These unwanted structures can be understood from two different points of view:

- Both the above methods filter a 3D group of similar patches. However, there can also exist dissimilar regions within this group. In order to stop information exchange between such regions and to completely remove the noise, they make use of specific non-linear filtering ideas. Despite this, such a modelling can lead to just a partial fulfillment of the two goals. Thus, leading to the presence of artefacts.
- BM3D in particular, which is the better performing among these two filters, is a widely-accepted single-frame image denoising technique. It makes assumptions about the noise-free image in the Fourier domain. We need to be very careful while making assumptions in such domains as this is not where we visualise images. A local change in this transformed domain has a non-local effect in the Cartesian domain. This could further lead to exchange of information between regions of dissimilar grey values and thus appearance of artefacts. The NLB approach also has a correspondence with Wiener filtering [290], which is a transform domain-based solution. It is also not artefact free. Thus, modern extensions of the classical Wiener filtering are still prone to artefacts.

Along with NLB and BM3D, we also employ NLDD, DDID, and LFSPO techniques within our comparative evaluations. The latter three are better solutions for circumventing the above artefact problem. We will learn more about them in Chapter 5.

Considering the huge amount of literature that we have surveyed in this chapter, it is not surprising that there have also been works which studied the theoretical bounds of quality that patch-based methods can achieve [175,291]. The gist of these studies is that while BM3D almost reaches optimal performance in texture like scenarios, there is definitely scope for improvement in piecewise constant images. The reason behind this is the artefacts created by BM3D.

One can notice that throughout this chapter we concentrated on the physical assumptions underlying image denoising filters. In model-driven approaches, we thus decide on the arithmetics within a filter based on our physical intuition. There also exist data-driven methods which model the required arithmetics based on a given set of training data. Such approaches are relatively less intuitive in contrast to model-based ones. Learning, deep learning and neural networks are some of the terms that are used to coin data-driven techniques [292–295]. There also exist works that uncover relations between model- and data-driven methods [296–298] and works with hybrid models that utilise ideas from both worlds [299–302]. We frequently analyse the strengths and weaknesses of both model-based and data-driven approaches in the upcoming chapters.

This finishes our survey of single-frame image denoising methods. We are now ready to learn more about our specific contributions that push the knowledge limits of this field.

Part I

Single-frame Image Denoising

Chapter 4

Disc Diffusion - A Non-local Diffusion Model

Many processes in image processing exhibit a non-local behaviour. That is the case, for instance in bilateral filters [112, 113, 303] which are widely used for image enhancement and denoising. Moreover, non-local patch-based methods [14, 15, 58, 161, 295, 304–308] have been among the leading noise elimination methods for more than a decade. Lately, such techniques have also been modified to solve correlated noise and deblurring problems efficiently [58, 59].

Classical PDEs and variational methods have also been generalised to non-local processes [170, 309–317]. Image inpainting [170, 315], irregularity detection [170], image decomposition [317], denoising [312, 315], super-resolution [313], and reconstruction [310] are some of the many applications where such PDEs and variational methods have been employed. It is thus quite evident that these non-local processes make their presence in several applications. Diffusion-based techniques can be robust to the type of noise and can also avoid artefacts produced by frequency domain-based filtering. Despite this, it is surprising that very little effort has been put into studying non-local counterparts of these approaches in a denoising context.

Our Goal. Model a non-local diffusion-based approach that is robust to the kind of noise degradation and also avoids artefacts in filtered images.

Our Contribution. We begin with an anisotropic diffusion model devel-

oped by Weickert [17] two decades ago. This approach has the property to preserve edges but does not make use of non-local image information which is vital in image denoising. In this chapter, we extend this method by including non-local information within the filtering process.

Chapter structure. This chapter is outlined as follows: We introduce the ideas behind the modelling of our filter in Section 4.1. The experiments and discussions on the results are presented in Section 4.2. We consolidate our conclusions from the experiments along with an outlook on future work in Section 4.3.

4.1 Modelling and Theory

4.1.1 Continuous Modelling of Disc Diffusion

Let us start by introducing some background. In [17], Weickert analysed the anisotropic image evolution process

$$\partial_t u(\mathbf{x}, t) = \frac{1}{\pi} \int_0^\pi \partial_\theta (g(\partial_\theta u_\sigma) \partial_\theta u) d\theta. \quad (4.1)$$

The main idea behind that model is to construct an anisotropic filter by integrating over one-dimensional diffusion processes. Here ∂_θ is the directional derivative in direction $(\cos \theta, \sin \theta)^\top$, g is a diffusivity function, and u_σ denotes a convolved version of u with a Gaussian of standard deviation σ . This idea gives rise to an image evolution process that is capable of preserving edges very well. Our goal here is to adapt (4.1) to the non-local setting. To this end, we model a new evolution equation

$$\partial_t u(\mathbf{x}, t) = \int_\Omega J(|\mathbf{y} - \mathbf{x}|) g \left(\frac{u_\sigma(\mathbf{y}) - u_\sigma(\mathbf{x})}{|\mathbf{y} - \mathbf{x}|} \right) (u(\mathbf{y}) - u(\mathbf{x})) d\mathbf{y}, \quad (4.2)$$

where the integration over the image domain Ω reflects the original idea in (4.1) of considering information along all possible orientations. However, it also incorporates non-locality into the filter: Given a positive parameter ρ , the function $J(s)$ is defined as a slightly Gaussian-smoothed version of $F(s) := \frac{1}{s^2}$, for $0 \leq |s| \leq \rho$. It is also assumed to decrease fast but smoothly to zero for $|s| > \rho$. This last assumption motivates the name disc diffusion, since the integration is essentially done within a disc of radius ρ . The slight

Gaussian smoothing of F is needed to avoid the singularity of J as $s \rightarrow 0$. We choose the diffusivity g to be [133]

$$g(s) = 1 - \exp\left(\frac{-3.31488}{\left(\frac{s}{\lambda}\right)^8}\right) \quad (4.3)$$

with a positive parameter λ . With these choices, (4.2) can also be seen as an example of a theoretical framework developed in [318].

4.1.2 Discrete Modelling of Disc Diffusion

We now discuss the space and time discretisations of (4.2). If u_1, u_2, \dots, u_M represent the grey values of the M pixels in the image grid, the space-discrete formulation of (4.2) is given by

$$\frac{d}{dt}u_i = \sum_{1 \leq j \leq M} J(|\mathbf{x}_j - \mathbf{x}_i|) g\left(\frac{u_{\sigma j} - u_{\sigma i}}{|\mathbf{x}_j - \mathbf{x}_i|}\right) (u_j - u_i). \quad (4.4)$$

Here, the variables \mathbf{x}_i and \mathbf{x}_j denote the positions of the pixels i and j . The quantities $u_{\sigma i}$ and $u_{\sigma j}$ represent the corresponding Gaussian-smoothed values.

As already mentioned, in (4.2) we have chosen J to be a Gaussian-smoothed version of $F(s) := \frac{1}{s^2}$, in order to avoid the singularity as $s \rightarrow 0$. In the space-discrete setting, however, this situation is never met, hence we may use F directly. Moreover, since we also assumed that $J(s)$ decreases fast towards zero for $s > \rho$, here we simply consider points within a disc of radius ρ . With these choices, the explicit time discretisation of (4.4) with a time step τ is given by

$$u_i^{k+1} = u_i^k + \tau \left(\sum_{j \in B_{i,\rho}} g\left(\frac{u_{\sigma j}^k - u_{\sigma i}^k}{|\mathbf{x}_j - \mathbf{x}_i|}\right) \frac{u_j^k - u_i^k}{|\mathbf{x}_j - \mathbf{x}_i|^2} \right). \quad (4.5)$$

Here, $B_{i,\rho}$ is the disc centered at the position of pixel i with radius ρ . The above equation can be written in a more compact way using a matrix $\mathbf{Q}(\mathbf{u}^k) \in \mathbb{R}^{M \times M}$:

$$\mathbf{u}^{k+1} = \mathbf{Q}(\mathbf{u}^k)\mathbf{u}^k. \quad (4.6)$$

Here the entries of the matrix $q_{i,j}(\mathbf{u}^k)$ are given by

$$q_{i,j}(\mathbf{u}^k) = \begin{cases} 1 - \sum_{j \in B_{i,\rho}} \frac{\tau \cdot g_{i,j}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2}, & \text{if } j = i, \\ \frac{\tau \cdot g_{i,j}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2}, & \text{if } j \in B_{i,\rho} \text{ and } j \neq i, \\ 0, & \text{otherwise,} \end{cases} \quad (4.7)$$

with $g_{i,j}(\mathbf{u}^k) := g\left(\frac{u_{\sigma_j}^k - u_{\sigma_i}^k}{|\mathbf{x}_j - \mathbf{x}_i|}\right)$. Also, it is not difficult to check that the row sums of the above matrix are 1. According to the discrete diffusion theory of Weickert [133], for the maximum-minimum principle to hold, we also require all the entries of the matrix $\mathbf{Q}(\mathbf{u}^k)$ to be non-negative. Hence, the following condition must hold:

$$1 - \sum_{j \in B_{i,\rho}} \tau \cdot g\left(\frac{u_{\sigma_j}^k - u_{\sigma_i}^k}{|\mathbf{x}_j - \mathbf{x}_i|}\right) \frac{1}{|\mathbf{x}_j - \mathbf{x}_i|^2} \geq 0. \quad (4.8)$$

With prior knowledge that the maximum value of the diffusivity is 1, we have the final bounds on the time step size:

$$0 \leq \tau \leq \frac{1}{\sum_{j \in B_{i,\rho}} \frac{1}{|\mathbf{x}_j - \mathbf{x}_i|^2}}. \quad (4.9)$$

If we interchange i and j in (4.7), the values of the matrix $\mathbf{Q}(\mathbf{u}^k)$ do not change. Hence, it is symmetric. Since the row sums of this matrix are 1 and $\mathbf{Q}(\mathbf{u}^k)$ is symmetric, the columns also have a unit sum. Thus, the average grey value is preserved during the image evolution (see [133]). Now we move to the image denoising experiments.

4.2 Experiments and Discussion

4.2.1 Datasets and Methods for Evaluation

In order to check the robustness of our method with respect to the noise model, we consider two different types of noisy data. First, we corrupt the Texmos¹ (synthetic) and Svalbard (real-world) images with additive white

¹<http://sipi.usc.edu/database/>

Gaussian noise (AWGN with standard deviation σ_G). We also use noisy cryo-electron microscopy data. Unlike for normal cameras, the noise model in data acquired using an electron microscope is not well understood. Hence such images are very good test cases for checking the robustness of an image denoising filter with respect to the kind of noise distribution.

We compare our proposed disc diffusion (DD) framework with the non-local Bayes (NLB) [14, 290] and 3D block matching (BM3D) [15, 319] filters.

4.2.2 Parameter Selection

We have five parameters in the disc diffusion model: The Gaussian smoothing parameter σ , the contrast parameter λ , the radius ρ of the disc shaped neighbourhood, the total number of iterations k_{\max} and the time step size τ .

Influence of the Disc Radius. The plots in Figure 4.1 show the influence of the disc radius on the denoising performance of DD. Three things can be observed from the graphs:

- Firstly, the reconstructed images have a high mean squared error (MSE) value for low radii. This is can be attributed to the fact that the disc does not contain enough pixels.
- The graph is almost constant for a small range of radii. From (4.5), we can see that information from neighbouring pixels is weighted by $\frac{1}{|\mathbf{x}_j - \mathbf{x}_i|^2}$. When $|\mathbf{x}_j - \mathbf{x}_i|$ starts to increase, these weights drop drastically, which decreases the influence of such neighbours on the diffusion process to a negligible amount. This explains the second observation.
- Thirdly, we can see that the quality of the reconstructed image deteriorates when the radius is very large. Even though the contribution from pixels that are far away is small, it is still non-zero. Consequently, a large number of small error prone contributions that come from areas of dissimilar grey values deteriorate the reconstructed image.

All these observations are irrespective of the amount of noise and the image under consideration. Thus, we choose $\rho = 7$ in all our further experiments, as it is a value in the region where we have least MSE.

Once we have fixed the radius, we can calculate the time step size τ using the upper bound in (4.9). In our experiments, we have chosen τ to be 95%

4.2. Experiments and Discussion

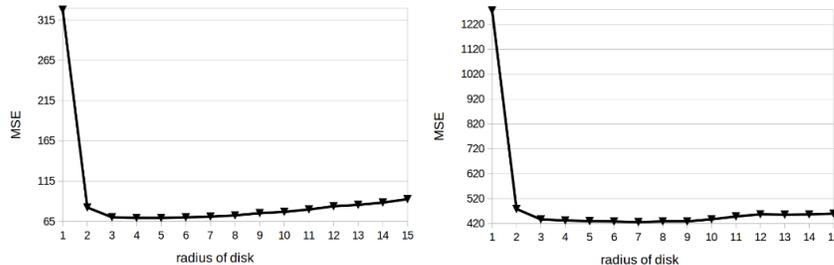


Figure 4.1: Influence of disc radius on denoising performance of DD. **Left:** Svalbard image with $\sigma_G = 25$, $\sigma = 0.8$, and $\lambda = 3.9$. **Right:** Texmos image with $\sigma_G = 75$, $\sigma = 2.3$, and $\lambda = 1.1$. Image size: 512×512 .

| Image (σ_G) | σ | λ | k_{\max} | DD | NLB | BM3D |
|----------------------|----------|-----------|------------|--------------|--------|---------------|
| Texmos (25) | 1.1 | 1.7 | 52 | 7.28 | 11.41 | 11.81 |
| Texmos (50) | 1.7 | 1.4 | 92 | 22.72 | 31.76 | 34.07 |
| Texmos (75) | 2.3 | 1.1 | 165 | 40.31 | 61.59 | 59.25 |
| Svalbard (25) | 0.8 | 3.9 | 16 | 20.27 | 22.58 | 22.92 |
| Svalbard (50) | 1.0 | 4.6 | 20 | 54.01 | 60.07 | 61.47 |
| Svalbard (75) | 1.2 | 5.1 | 17 | 108.68 | 118.20 | 107.98 |

Table 4.1: Denoising evaluation of various methods using MSE.

of this upper bound. Thus, we only have to optimise for σ , λ and k_{\max} . In case of AWGN, we optimise the parameters of DD with respect to the MSE. In the case of BM3D and NLB, we utilise the optimal parametric values as suggested in [319] and [290], respectively. Now we are ready to present our denoising results. First, we show experiments on images corrupted by synthetic noise (AWGN) which are then followed by experiments on cryo-electron microscopy (cryo-EM) data.

4.2.3 Eliminating Synthetically Generated Noise

Table 4.1 and Figure 4.2 showcase the MSE values and denoised images, respectively. In terms of MSE as well as visual quality, DD gives the best results. BM3D and NLB produce artefacts in homogeneous regions.

The reason behind the above observation is that the structure preserving ability of DD minimises information exchange between regions of dissimilar greyvalues. Moreover, BM3D and NLB are just two step procedures. They

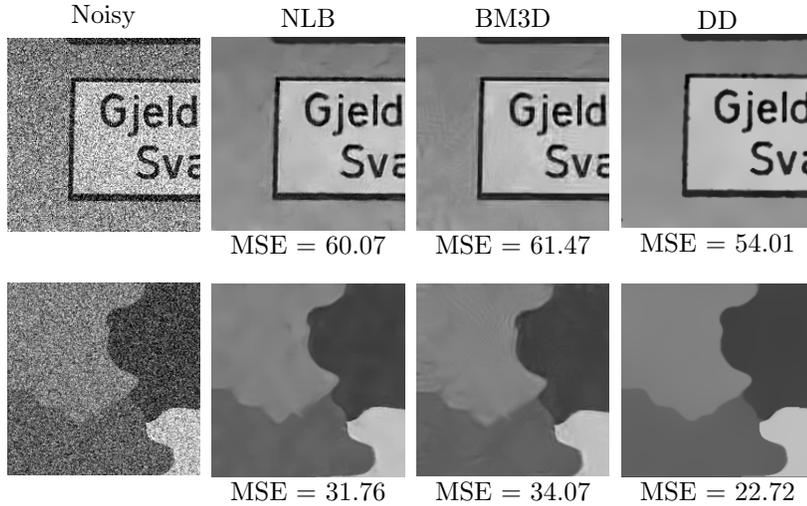


Figure 4.2: **Top and Bottom:** Zoom into Svalbard and Texmos images ($\sigma_G = 50$), respectively.

might lead to an incomplete elimination of noise. Since we perform several iterations in DD, it completely removes the noise.

4.2.4 Eliminating Real-world Noise

We have considered noisy data consisting of a pre-ribosomal particle from yeast cells, acquired using an electron microscope. The denoised images are presented in Figure 4.3. Figure 4.4 shows the corresponding Fourier ring correlation (FRC) plots. We can clearly see that DD gives better results in terms of FRC. The type of electron microscopy data we have considered was approximated with a signal independent additive noise model in [51]. However, NLB and BM3D additionally assume an explicit Gaussian distribution for the noisy signal. Practical data need not comply with this analytical assumption, in which case NLB and BM3D are at a disadvantage. Diffusion-based models on the other hand are relatively robust to the type of noise.

Having discussed the positive aspects of DD, we also need to mention its limitations. Our experience suggests that such a basic diffusion technique cannot compete with frequency domain methods when it comes to preserving texture-like structures. Thus, DD definitely needs modifications to compete with BM3D and NLB for textured images.

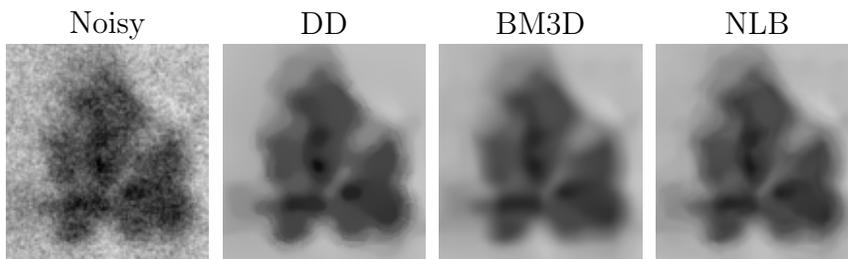


Figure 4.3: Zoom into cryo-EM denoising results. Data courtesy of Utz Ermel - Goethe University of Frankfurt. DD parameters: $\sigma = 2.0$, $\lambda = 1.0$, $k_{\max} = 17$.

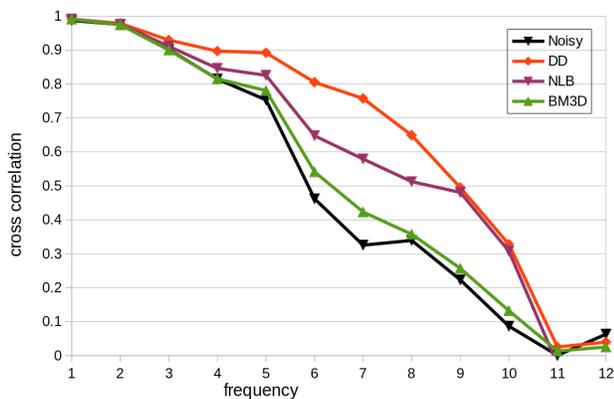


Figure 4.4: Zoom into FRC curves near a cross correlation of 0.5 where the values can be trusted [88]. The unknown parameter $\sigma_{\mathcal{G}}$ for NLB and BM3D was optimised using FRC.

4.3 Conclusions

In this chapter, we have introduced a non-local extension of a classical diffusion-based filter. It is surprising that a non-local counterpart of a technique that was developed two decades ago is competitive with complex techniques like BM3D and NLB for piecewise constant images. Moreover, our proposed disc diffusion model is robust to the type of noise and also avoids artefacts.

Patch-based filters generally seem to have a better texture reconstruction capability. In the upcoming chapter, we thus construct a filter which combines ideas from disc diffusion and patch-based approaches.

Chapter 5

Non-linear Filtering on Fast Patch Reorderings

As detailed in Chapter 3, the field of image denoising filters is very rich. The present day state-of-the-art standard in this area is set by patch-based methods [14, 15, 18, 19, 21, 160] as well as neural networks [292–294]. The former techniques have a transparent design. In order to yield the final denoised image, they explicitly model both signal and noise characteristics. Moreover, their combination with neural networks has brought the best out of both worlds [295]. Since the development of transparent designs is also very crucial, in this chapter we contribute one such novel filter.

Although popular patch-based methods produce state-of-the-art results, they also have a couple of drawbacks:

1. *The Artefact Problem.* There are two reasons behind the presence of artefacts. Information exchange between regions of dissimilar grey values and incomplete removal of noise within regions of similar grey values. Although patch-based procedures [14, 15] benefit from non-linear filtering for dealing with these phenomena, they are not completely solved. The smooth patch reordering approaches of Ram et al. [18, 19] seem promising in this aspect. However, they are computationally very expensive. In [21, 239, 241–243], dual-domain image denoising algorithms have been suggested for solving the artefact problem. We will discuss the results of these techniques in a later section of this chapter.
2. *Sensitivity to Noise Type.* Almost all the above methods assume an additive white Gaussian noise (AWGN) model. Thus, they require variance

stabilising transformations [60, 277, 281] in order to be employed for Poissonian and Poissonian-Gaussian (mixture) noise scenarios. These transformations can have their own problem of creating a bias while stabilising the variance. Ideas which exclude the characteristics of noise within the respective non-linear models are surprisingly rare. Such a design would allow us to forgo the bias problem of variance stabilising transformations (VSTs) as well as the errors obtained while fitting practical data with analytical noise distributions [61–64, 66].

The artefact problem is especially severe for large noise amplitudes. In applications like electron microscopy one largely acquires pure noise and a faint signal along with it. Despite the longstanding efforts in designing both deep learning and traditional methods, high noise levels and varying types of noise distributions are not often dealt with. For example, the considered noise levels in some of the recent deep learning-based papers [292–295] as well as model-based filters [14, 18, 19, 21, 239, 241–243], are not large enough for the above imaging applications. In fact, one of the latest model-based methods, which is competitive with neural networks, also considers just an AWGN distribution [320]. To our knowledge, the 3D block matching (BM3D) [15] approach is the only technique which has been extensively evaluated for both severe degradations as well as varying noise distributions [279–281]. However, it forms artefacts and is sensitive to the noise model.

Our Goal. Design a method which combines ideas from diffusion-based filtering and patch reorderings to achieve the following objectives: Preserve image structures, avoid artefacts, remain robust to the kind of noise distribution, retain sharpness in the image, and consume less time.

Our Contribution. In order to accomplish the above goals, we introduce a method that possesses the following features:

1. *Artefact Avoidance and Sharpness Retainment:* Our design builds on the non-local means (NLM) approach [160] since it has the property to circumvent artefacts [14]. However, the NLM solution also smooths over edges and loses sharpness [14]. We propose a combination of patch similarity assumption in NLM and pixel similarity modelling in diffusion-based methods. The latter idea is responsible for retaining the sharpness.
2. *Multiplicative Combination:* Unlike any of the previous denoising filters, we utilise a multiplicative combination of patch and pixel similarity assump-

tions. This makes our solution robust in situations when one of the assumptions might be violated due to the presence of noise.

3. *Smoothing Within Fast Patch Reorderings*: Obtaining smoothly reordered patches for efficient denoising can be very expensive [18, 19]. In order to simplify this, we just utilise a sort operation on the inter-patch distances. Any disordered pixels due to this simplification are dealt with by our robust multiplicative modelling. Considering the less burdensome computational power we call our technique *non-linear filtering on fast patch reorderings* (NFPR).

4. *Collaborative Filtering at Pixel Level*: In collaborative filtering [15], each patch to be denoised collaborates with several others. However, this also allows for interactions between regions of dissimilar greyvalues due to intra-patch smoothing. In our case, we just employ an inter-patch smoothing between the central pixels of similar patches.

5. *Robust to the Kind of Noise*: Our modelling excludes the presence of any noise statistics within the non-linearity. Hence, NFPR can be applied to Gaussian, Poissonian, mixture, clipped, and impulse noise distributions. These are unstructured degradations which do not possess a definite spatial structure. Heavily structured noise [58, 59] has spatial properties similar to that of noise-free data. We do not consider such data in this chapter.

Chapter Structure. In Section 5.1 we introduce the modelling of our proposed NFPR filter. In Section 5.2 we present extensive experimental evaluations for different noise types along with detailed discussions. Finally, in Section 5.3 we conclude with an overview and an outlook on future research.

5.1 Modelling and Theory

Our filter consists of two parts. In the first, we derive a set of pixels with similar grey values. In the second step, these pixels undergo a non-linear smoothing process which yields the final denoised image. In the following, we describe these two steps in more detail.

5.1.1 Fast Patch Reordering Stage

The motive of this part of the filter is to congregate pixels with grey values that belong to the same region in an image. To this end, we employ the following patch-based solution: Let us consider a pixel u_i in the 2D image

domain and a disc-shaped patch B_{search} around it, with radius ρ_{search} . We now compute the Euclidean norm d_{ij} between two disc-shaped patches of radius ρ_{sim} around u_i and u_j for all $u_j \in B_{\text{search}}$. This is followed by constructing a set \mathbf{P}_i of N nearest pixels to u_i in B_{search} according to the computed Euclidean norms. This particular set constitutes the desired combination of pixels that arise from regions of similar grey values. However, there exist two problems with such a combination. In order to completely avoid artefacts, we need to average noisy versions of the same original grey value. In \mathbf{P}_i , we have an assembly of noisy versions of similar grey values, which is the first problem. In [18], Ram et al. solve the computationally expensive travelling salesman problem for achieving a smooth pixel reordering. This was followed by the use of a simple linear smoothing for excellent results. In our case, we just perform a sort operation for the pixel reordering. This saves a lot of time, however, it also comes at the cost of certain disordered pixels in \mathbf{P}_i which is the second problem. In order to deal with both problems, we adhere to a robust structure preserving filter in the second step of NFPR.

5.1.2 Non-linear Smoothing Stage

In this step we design weighting functions that decide the amount of contribution from each of the pixels in \mathbf{P}_i towards the denoised pixel. In particular, we model two assumptions: Similar pixels have minimum grey value variation among themselves (pixel similarity). They are also surrounded by similar neighbourhoods (patch similarity). In order to have a robust denoising, we use a multiplicative combination of both assumptions for better performance when one of them can be violated due to noise. A formulation of such a model through a discrete image evolution process is given by

$$\begin{aligned} \frac{u_i^{k+1} - u_i^k}{\tau} = & a_i^k \cdot \left(\sum_{j \in \mathbf{P}_i^k} g(u_{\sigma_j}^k - u_{\sigma_i}^k) h(d_{ij}^k) (u_j^k - u_i^k) \right. \\ & \left. + \sum_{j \in \mathbf{P}_i^{\text{add},k}} g(u_{\sigma_j}^k - u_{\sigma_i}^k) h(d_{ij}^k) (u_j^k - u_i^k) \right). \end{aligned} \quad (5.1)$$

Note the two terms on the right hand side of the above image evolution. The first term is responsible for information exchange between u_i and the other

pixels within \mathbf{P}_i . However, every pixel in the image other than u_i has its own reordered set \mathbf{P} . It can happen that u_i is present within these sets. Thus, the symbol $\mathbf{P}_i^{\text{add}}$ denotes the set of pixels for which u_i is a closest neighbour in a Euclidean norm sense.

The second term in the above equation represents data exchange between u_i and pixels within $\mathbf{P}_i^{\text{add}}$. Together, both the terms constitute a smoothing process in a space defined by inter-patch distances, unlike diffusion [128,129] or bilateral filters [112,303,321]. Moreover, it can also be seen as a pixel-level collaborative filtering, unlike a patch-level counterpart in [15].

We now describe the individual components of the above two terms in detail: The patch and pixel similarity assumptions are modelled using the functions h and g , respectively. The argument of g is one which is computed on an initial denoised image \mathbf{u}_σ . Such a two step filtering process is inspired from a diffusion-based technique [129]. In particular, we use a collaborative non-local means [160] approach for computing \mathbf{u}_σ :

$$u_{\sigma i}^k = b_i^k \cdot \left(\sum_{j \in \mathbf{P}_i^k} h(d_{ij}^k) u_j^k + \sum_{j \in \mathbf{P}_i^{\text{add},k}} h(d_{ij}^k) u_j^k \right). \quad (5.2)$$

The term on the left hand side of (5.1) is generally associated with Gaussian noise [195]. Similar to the original non-local means filter [160], we do not include the noise standard deviation σ_G as an argument of the non-linear function h . However, later modified versions of NLM, like the ones in [14,322] as well as BM3D, include such knowledge for better performance. We do not adhere to this idea as we aim at a filter which is robust to the kind of noise distribution. The weighting functions g [133] and h in (5.1) are chosen as

$$\begin{aligned} g(s) &= 1 - \exp\left(\frac{-3.31488}{\left(\frac{s}{\lambda}\right)^8}\right), \\ h(s) &= \exp\left(\frac{-s^2}{2\sigma^2}\right). \end{aligned} \quad (5.3)$$

The normalisation constants for the two image denoising processes in (5.1) and (5.2) are represented by a_i and b_i , respectively. We use the symbol M_i to denote the combined number of elements in \mathbf{P}_i and $\mathbf{P}_i^{\text{add}}$. By choosing $a_i = \frac{b_i}{M_i}$ and $\tau \leq 1$, we do not violate the maximum-minimum principle. The

two NFPR steps are computed for a maximum of k_{\max} iterations. The fast patch reordering step is initialised with a Gaussian smoothed (with standard deviation σ_{init}) version of the raw noisy image \mathbf{f} . The non-linear smoothing process is directly initialised with \mathbf{f} .

5.1.3 Correspondences with Disc Diffusion

At the end of the previous chapter, we claimed that NFPR is going to be designed using a combination of ideas from disc diffusion (DD) and patch-based approaches. In this part of the present chapter, we are going to analyse the correspondences and differences between NFPR and DD.

Let us first revisit the fully discrete image evolution equation for disc diffusion from Chapter 4:

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{j \in B_{i,\rho}} g \left(\frac{u_{\sigma_j}^k - u_{\sigma_i}^k}{|\mathbf{x}_j - \mathbf{x}_i|} \right) \frac{u_j^k - u_i^k}{|\mathbf{x}_j - \mathbf{x}_i|^2}. \quad (5.4)$$

The NFPR image evolution process in (5.1) has resemblances with both sides of the above equation. The correspondence with the right hand side is non-trivial: Firstly, the basic filter shape in NFPR and DD is a disc B . However, the set of pixels with which the center pixel i interacts is different in both methods. In DD, this set contains all the pixels within the disc $B_{i,\rho}$ that is centered at pixel i and has radius ρ . In NFPR, however, we completely get rid of this spatial context fixed by $B_{i,\rho}$. Instead, we choose all pixels which are closest to the center pixel in terms of inter-patch distances. These pixels are contained in sets $\mathbf{P}_i^{\text{add}}$ and \mathbf{P}_i in (5.1). As a consequence of this, instead of computing a regular Gaussian smoothing u_{σ_i} like the one in DD, in NFPR such a smoothing is computed in a space defined by inter-patch distances.

Moreover, in DD, we just model pixel similarity. In NFPR, our design contains both pixel and patch similarities. This brings us to the end of the modelling part. In the next section, we present our experimental evaluation.

5.2 Experiments and Discussion

5.2.1 Datasets and Methods for Evaluation

We have considered the classical Lena, Peppers, House, Bridge, Cameraman images (<http://sipi.usc.edu/database/>) as well as 24 images from the

Kodak dataset [323], for our experiments. We corrupt these images with Gaussian, Poissonian and mixture degradations. This serves as example data for synthetic noise. In order to complete a comprehensive evaluation, we also consider real world data acquired using an electron microscope.

In order to have a fair comparison between our filter and others, we need to consider methods which have been tested on multiple kinds of noise. Since BM3D is one technique which has been extensively evaluated for several noise degradations, we use it for our core set of experiments. In the remaining ones, we also use various other state-of-the-art model-based approaches.

5.2.2 Parameter Selection

We have fixed the following NFPR parameters: $\rho_{\text{search}} = 10$, $\rho_{\text{sim}} = 10$, $\sigma_{\text{init}} = 2.5$, $\tau = 0.95$ and $N = 35$. In order to have an easier selection of the parameter σ , we perform an affine rescaling of the Euclidean norms corresponding to the set \mathbf{P}_i to $[0, 255]$. Furthermore, in order to save time, we have employed the patch-reordering step for just two iterations. Finally, we optimise the parameters σ , λ , k_{max} with respect to the peak signal to noise ratio (PSNR) for synthetic data.

In case of electron microscopy images, the ground truth is not available. Hence we use a popular frequency domain measure called Fourier ring correlation (FRC) [87, 88], to quantify the quality of the images. Here, one measures correlation coefficients between two images at several frequency levels. These images are obtained after denoising two different noise realisations of the same scene, using a particular algorithm. Higher correlation indicates a better preservation of structures.

We have now prepared the experimental setup with the help of the above information. From the next paragraph onwards we begin with experiments on synthetically corrupted datasets. Afterwards, we present results on a couple of real-world datasets acquired using an electron microscope.

5.2.3 Synthetic Noisy Data

We begin with AWGN and employ non-local Bayes (NLB) [14, 290], BM3D [15, 319] and our novel NFPR filters for a comparative evaluation. In the left

5.2. Experiments and Discussion

| AWGN | | | | POISSONIAN | | | MIXTURE | | |
|-------|--------------|-------|--------------|------------|--------------|--------------|---------|--------------|--------------|
| Image | NFPR | NLB | BM3D | Image | NFPR | BM3D | Image | NFPR | BM3D |
| L40 | 29.44 | 29.72 | 29.79 | L1.0 | 22.92 | 22.71 | L5.0 | 26.35 | 26.34 |
| L60 | 27.94 | 27.75 | 27.93 | L0.9 | 22.96 | 22.70 | L4.0 | 25.85 | 25.74 |
| L80 | 26.69 | 26.24 | 26.57 | L0.8 | 22.45 | 22.25 | L2.0 | 24.38 | 24.05 |
| L100 | 25.96 | 25.15 | 25.49 | L0.7 | 22.09 | 21.87 | L0.9 | 22.71 | 22.33 |
| L120 | 25.19 | 24.07 | 24.54 | L0.6 | 21.78 | 21.51 | L0.7 | 22.19 | 21.77 |
| L140 | 24.49 | 23.18 | 23.77 | L0.5 | 21.56 | 21.21 | L0.5 | 21.52 | 21.15 |
| B40 | 24.07 | 24.45 | 24.45 | B1.0 | 19.99 | 19.87 | B5.0 | 22.00 | 22.05 |
| B60 | 22.90 | 22.90 | 23.14 | B0.9 | 19.88 | 19.77 | B4.0 | 21.73 | 21.73 |
| B80 | 22.10 | 21.84 | 22.21 | B0.8 | 19.76 | 19.63 | B2.0 | 20.83 | 20.67 |
| B100 | 21.52 | 21.26 | 21.57 | B0.7 | 19.66 | 19.48 | B0.9 | 19.90 | 19.78 |
| B120 | 21.09 | 20.71 | 21.03 | B0.6 | 19.38 | 19.31 | B0.7 | 19.62 | 19.52 |
| B140 | 20.68 | 20.23 | 20.55 | B0.5 | 19.27 | 19.19 | B0.5 | 19.23 | 19.08 |
| H40 | 30.44 | 30.19 | 30.65 | H1.0 | 21.62 | 21.05 | H5.0 | 26.11 | 25.95 |
| H60 | 29.03 | 27.92 | 28.49 | H0.9 | 21.31 | 20.90 | H4.0 | 25.47 | 25.17 |
| H80 | 27.46 | 26.17 | 26.97 | H0.8 | 21.01 | 20.50 | H2.0 | 23.05 | 22.93 |
| H100 | 26.28 | 24.73 | 25.40 | H0.7 | 20.68 | 20.29 | H0.9 | 21.15 | 20.67 |
| H120 | 25.29 | 23.52 | 24.41 | H0.6 | 20.53 | 19.87 | H0.7 | 20.59 | 20.00 |
| H140 | 24.44 | 22.61 | 23.35 | H0.5 | 19.90 | 19.51 | H0.5 | 20.05 | 19.39 |
| P40 | 30.52 | 30.35 | 30.43 | P1.0 | 23.00 | 23.03 | P5.0 | 26.56 | 26.49 |
| P60 | 28.90 | 28.34 | 28.53 | P0.9 | 22.77 | 22.71 | P4.0 | 26.09 | 25.94 |
| P80 | 27.62 | 27.03 | 27.19 | P0.8 | 22.48 | 22.50 | P2.0 | 24.40 | 24.24 |
| P100 | 26.70 | 25.78 | 26.07 | P0.7 | 22.31 | 22.14 | P0.9 | 22.74 | 22.48 |
| P120 | 25.88 | 24.79 | 25.11 | P0.6 | 22.01 | 21.79 | P0.7 | 22.32 | 22.04 |
| P140 | 25.15 | 23.87 | 24.31 | P0.5 | 21.61 | 21.44 | P0.5 | 21.61 | 21.37 |
| K40 | 27.50 | 27.79 | 27.91 | K1.0 | 22.37 | 22.10 | K5.0 | 25.01 | 25.06 |
| K60 | 26.17 | 26.04 | 26.28 | K0.9 | 22.22 | 21.91 | K4.0 | 24.63 | 24.64 |
| K80 | 25.22 | 24.87 | 25.15 | K0.8 | 22.06 | 21.67 | K2.0 | 23.37 | 23.23 |
| K100 | 24.48 | 23.95 | 24.26 | K0.7 | 21.86 | 21.45 | K0.9 | 22.25 | 21.77 |
| K120 | 23.86 | 23.22 | 23.53 | K0.6 | 21.64 | 21.18 | K0.7 | 21.81 | 21.42 |
| K140 | 23.35 | 22.57 | 22.92 | K0.5 | 21.40 | 20.86 | K0.5 | 21.36 | 20.76 |

Table 5.1: PSNR values for various noise distributions. L40 stands for Lena image corrupted with AWGN of $\sigma_G = 40$. L0.5 in the case of Poissonian noise stands for Lena image with $\chi = 0.5$. In the case of mixture noise, we have added AWGN with $\sigma_G = (0.10 * \chi * 255)$ to the Poissonian noise component. B, H, P denote Bridge, House and Peppers, respectively. The House image is 256×256 sized, while the remaining three are of size 512×512 . K denotes the Kodak dataset for which we have presented the average PSNR of the 24 images. We have employed NLB and BM3D implementations in [290] and [319], respectively.

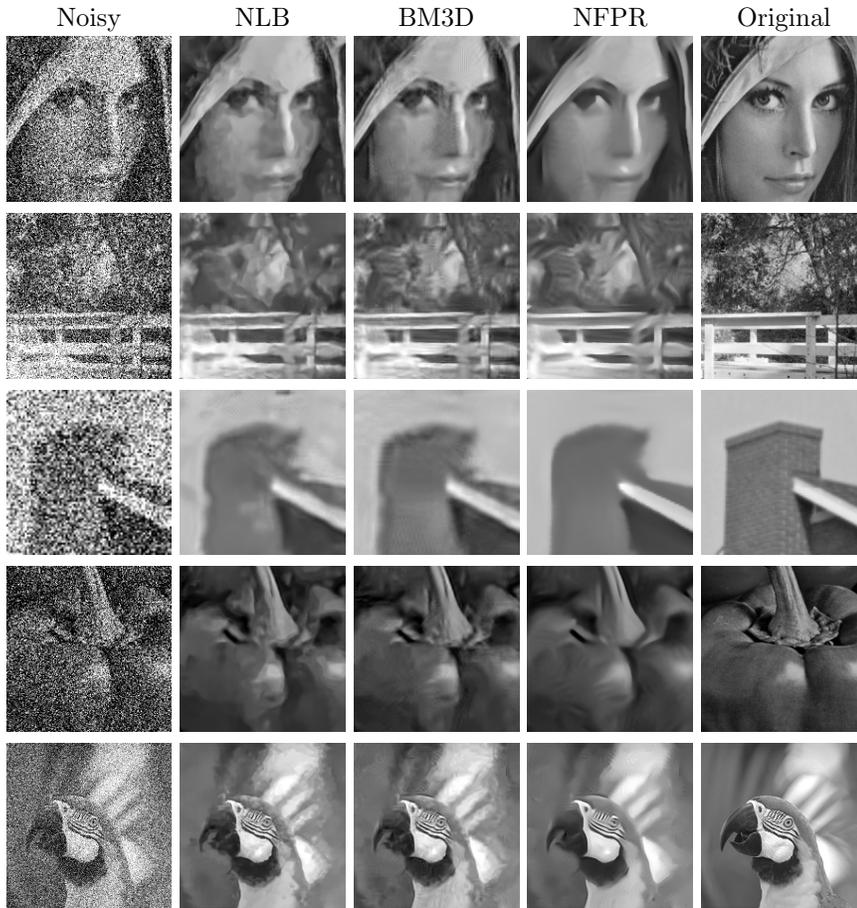


Figure 5.1: **Top to Bottom:** Zoom into Lena, Bridge, House, Peppers and a Kodak image, respectively (AWGN with $\sigma_G = 80$).

column of Table 5.1, we present the PSNR values of denoised images. Figure 5.1 shows some corresponding images. Both visual results and PSNR values suggest that our filter can denoise the images very well, while avoiding the artefacts. The images in particular show sharper and visually pleasing edges. Moreover, in a medical imaging scenario, creation of artefacts is considered a severe drawback as they might be mistaken as original structures. Hence, our technique can be a crucial tool in such situations.

Other methods [21, 239, 241–243] which have also tried to solve the artefact problem can be divided into two categories: While methods of the first type perform a raw processing of the data like we do [241–243], the others carry

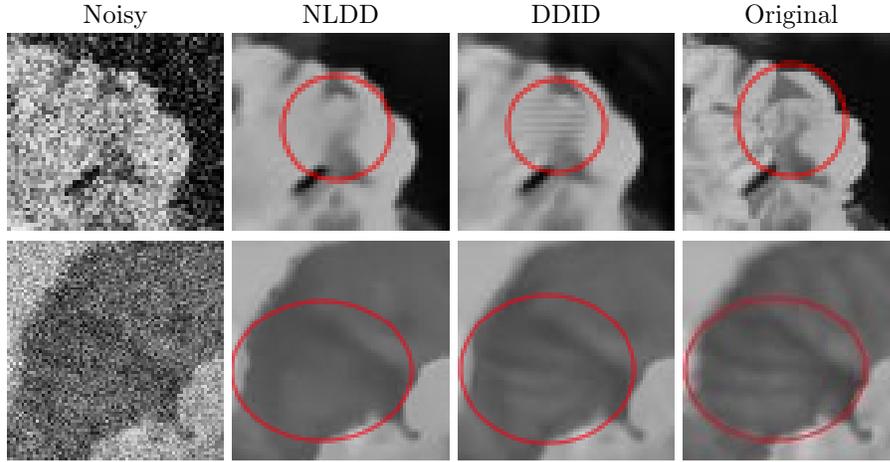


Figure 5.2: **Top:** NLDD avoids artefacts that are present in the DDID output. **Bottom:** However, it also removes original structures on the leaves unlike DDID, as a trade off [239].



Figure 5.3: The artefacts corresponding to accidental noisy patterns are not spatially constant across different noisy realisations ($\sigma_G = 80$).

out a post-processing of the artefacts created by approaches like non-local Bayes and 3D block matching [21, 239].

We consider one technique from each of the above two categories: The dual-domain image denoising (DDID) [241] and the non-local dual-domain (NLDD) [239] filtering approaches. Although Gaussian noise is an unstructured degradation, it can form accidental spatial patterns. DDID preserves these patterns, assuming that they are structures belonging to the original signal, thus creating artefacts. NLDD removes these accidental patterns but only at the risk of suppressing original structures. This is a trade off between the two categories of approaches and can be clearly seen for the Flowers image in Figure 5.2. In Figure 5.3 one can see that even our filter preserves accidental noise patterns. However, they are present at varying spatial positions for different realisations of noise. Thus, the correct way to

| Image | NFPR | Best | DDID2 | Image | NFPR | REC |
|-------|--------------|--------------|--------------|-------|---------------|---------------|
| PS25 | 28.55 | 28.10 | 28.10 | L50 | 91.36 | 82.90 |
| PS40 | 30.49 | 30.46 | 30.46 | L75 | 126.90 | 123.75 |
| L25 | 30.97 | 32.27 | 32.27 | L100 | 164.85 | 163.05 |
| L40 | 29.44 | 30.22 | 30.22 | H50 | 70.94 | 74.45 |
| | | | | H75 | 108.71 | 120.41 |
| | | | | H100 | 153.27 | 167.52 |

Table 5.2: PSNR values for AWGN experiments. Image abbreviations as in Table 5.1 and PS means a smaller version of Peppers with size 256×256 . Whenever we use a smaller version of an image, we adhere to such an abbreviation. Best - largest PSNR among 11 different filters ([14,15,23,161,231,241–243,249,324,325]) evaluated with DDID2 in [243], REC - Ram et al. [18].

deal with these accidental patterns is to collect multiple noise realisations of a scene and then denoise each one them with either NFPR or DDID. This way, we can at least differentiate between original structures and artefacts created due to accidental noise patterns. This might not be possible with post-processing approaches like NLDD. They tend to remove structures irrespective of whether they are originally present in the image or accidentally created due to noise.

Keeping in mind the above observations, we have performed a short comparative analysis of NFPR with DDID2 [243]. This is also accompanied by a brief comparison with the method of Ram et al. [18], which has algorithmic correspondences with our approach. Table 5.2 shows the results of these experiments. The takeaway lessons in the AWGN scenario are the following:

- In case of almost every noise amplitude in highly homogeneous images like Peppers and House, NFPR gives the best results as we avoid artefacts of frequency domain filtering.
- In Lena and Bridge images, which have some amount of texture, the competing methods are superior for low noise amplitudes. This can be attributed to our simpler patch reordering technique and better texture preservation of frequency domain methods.
- In high noise amplitude scenarios, the texture information is completely lost and hence our approach also produces superior results for Lena and Bridge images.

5.2. Experiments and Discussion

| Image | NFPR | BM3D | Best | BM3D++ [279] |
|-------|--------------|--------------|--------------|--------------|
| PS4.0 | 22.95 | 24.20 | 24.25 | 24.04 |
| PS2.0 | 22.20 | 22.36 | 22.26 | 21.93 |
| PS1.0 | 20.90 | 20.69 | 20.57 | 20.44 |
| PS0.5 | 19.70 | 19.35 | 19.21 | 19.05 |
| PS0.2 | 17.71 | 17.68 | 17.60 | 17.54 |
| PS0.1 | 16.64 | 16.03 | 16.52 | 16.45 |
| BS4.0 | 21.62 | 21.72 | 21.71 | 21.71 |
| BS2.0 | 20.70 | 20.58 | 20.69 | 20.69 |
| BS1.0 | 19.49 | 19.25 | 19.86 | 19.86 |
| BS0.5 | 18.57 | 18.47 | 19.08 | 19.08 |
| BS0.2 | 17.62 | 17.16 | 18.13 | 18.13 |
| BS0.1 | 16.70 | 15.85 | 17.30 | 17.30 |

Table 5.3: PSNRs for Poissonian noise experiments. Image abbreviations as in Table 5.1 and 5.2. BM3D - our proposed $\sigma_{\mathcal{G}}$ optimisation for removing the VST bias, Best - largest PSNR among 13 different methods ([18, 22, 23, 60, 73, 160, 229, 260, 275, 276, 279, 289, 326]) evaluated in [279], BM3D++ [279] - with an iterative VST which is the best available framework.

- While dealing with accidental noise patterns, it is safer to collect multiple noise realisations and then filter with NFPR/DDID.

In the next part of the comparative evaluation, we consider Poissonian data. Here, we evaluate our filter with BM3D in combination with a VST [278]. Generally, after application of a VST on Poissonian data, the noise distribution is assumed to be converted to an AWGN with $\sigma_{\mathcal{G}} = 1.0$. This is however not the case as VSTs can induce a bias. In order to get rid of this bias and have a fair comparison with our method, we have optimised $\sigma_{\mathcal{G}}$ after application of the VST w.r.t. PSNR. In the center column of Table 5.1, we present the results of this evaluation. It is clear that despite not requiring the usage of a VST, NFPR outperforms BM3D. We have also carried out a short evaluation with another well-performing VST framework [279] in Table 5.3. Our approach is better for the homogeneous Peppers image. For the Bridge image which has textured regions, BM3D is superior to NFPR. These results thus follow a similar trend as in the AWGN case.

Now we move to the mixture noise scenario. Here, we employed the same optimisation strategy as in the Poissonian noise scenario, to eliminate the VST induced bias. In the right column of Table 5.1 we present the comparison of NFPR with BM3D + VST [327] which shows that our denoising

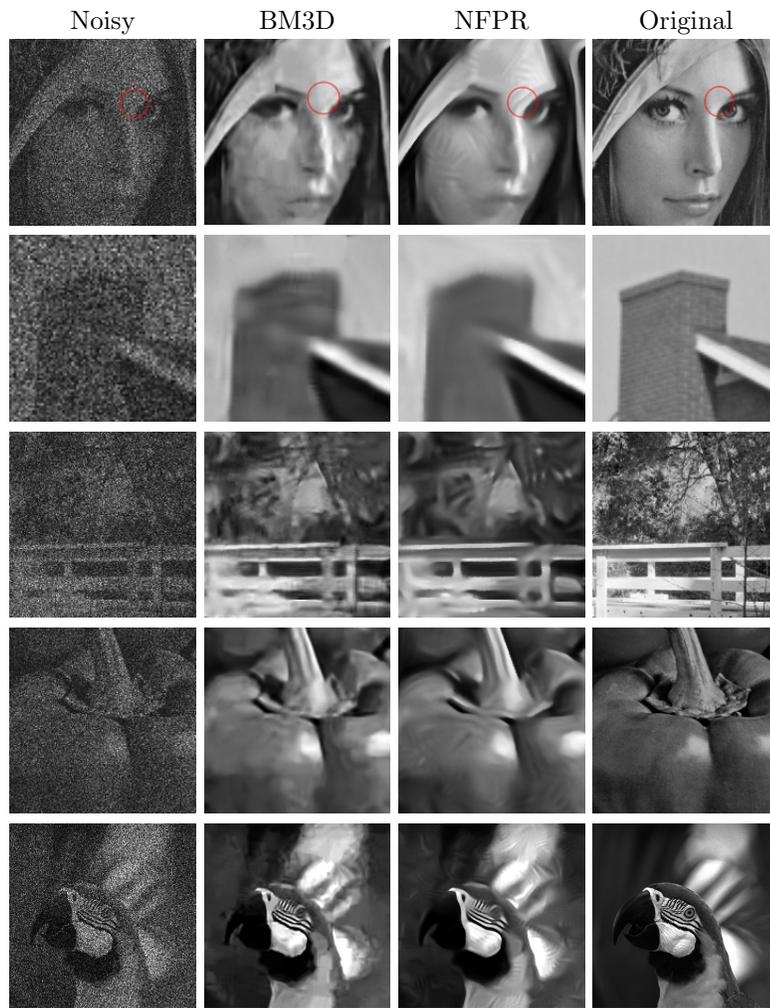


Figure 5.4: **Top to Bottom:** Zoom into Lena, Bridge, House, Peppers and a Kodak image, respectively ($\chi = 5.0$).

5.2. Experiments and Discussion

| Image | NFPR | BM3D | Best | BM3D+ [281] |
|-------|--------------|-------|--------------|--------------|
| L5.0 | 26.35 | 26.34 | 26.18 | 26.18 |
| L2.0 | 24.38 | 24.05 | 24.34 | 24.34 |
| L1.0 | 22.95 | 22.59 | 22.83 | 22.59 |
| CS5.0 | 23.32 | 23.73 | 24.09 | 24.09 |
| CS2.0 | 21.21 | 21.71 | 21.93 | 21.93 |
| CS1.0 | 20.15 | 20.05 | 20.35 | 20.23 |

Table 5.4: PSNRs for Poissonian-Gaussian distribution. Image abbreviations as in Tables 5.1, 5.2. C denotes Cameraman image. BM3D - our proposed $\sigma_{\mathcal{G}}$ optimisation for removing the VST bias, Best - largest PSNR among all the methods ([6, 15, 326]) evaluated in [281], BM3D+ [281] - optimal VST which is the best available framework for mixture noise.

method outperforms BM3D. Figure 5.4 presents some corresponding images of this evaluation. Concerning the artefacts, similar observations can be made as in the AWGN scenario: As highlighted by the red circle for Lena image, noise can form accidental spatial patterns that are not present in the original image. Our filter preserves these patterns similar to DDID. BM3D almost completely removes them, analogous to NLDD in the AWGN scenario. The pictures after removing these accidental structures might look pleasant. However, this comes at the risk of removing original structures itself. Such an observation is also in agreement with the one in [241]. Moreover, in Table 5.4 we showcase a short comparison of NFPR with the best available framework [281] for Poissonian-Gaussian degradation. The practical world data is generally approximated with the mixture model. A competitive performance by NFPR in this scenario is thus a significant step forward.

It is worth noticing that despite excluding the noise parameter within the arguments of the non-linear functions in our filter, NFPR can outperform those which include this knowledge. One might argue that our approach is at an advantage as we are optimising its parameters w.r.t. PSNR. However, BM3D and NLB are operating at the theoretical optimum as they utilise the ideal $\sigma_{\mathcal{G}}$ value for AWGN. In practice, this needs to be estimated which might also lead to errors. In the case of Poissonian and mixture noise models, we optimise the noise variance w.r.t. PSNR in order to minimise VST induced bias. This ensures a fair comparison. For reproducibility, the parametric values used to generate the results in all the above mentioned experiments are presented in Tables 5.5 - 5.6. With this, we end the synthetic noise experiments. Our philosophy, which does not require a-priori knowledge of

| AWGN | | | | POISSONIAN | | | | | MIXTURE | | | | |
|------|----------|-----------|------------|------------|----------|-----------|------------|-----------------------|---------|----------|-----------|------------|-----------------------|
| Data | σ | λ | k_{\max} | Data | σ | λ | k_{\max} | σ_{est} | Data | σ | λ | k_{\max} | σ_{est} |
| L40 | 150 | 11.5 | 16 | L1.0 | 250 | 1.0 | 12 | 1.4 | L5.0 | 180 | 0.5 | 11 | 1.1 |
| L60 | 160 | 15.5 | 16 | L0.9 | 260 | 0.5 | 12 | 1.5 | L4.0 | 180 | 0.5 | 11 | 1.1 |
| L80 | 175 | 20.0 | 14 | L0.8 | 270 | 0.5 | 13 | 1.6 | L2.0 | 200 | 0.5 | 10 | 1.2 |
| L100 | 175 | 23.5 | 16 | L0.7 | 290 | 1.0 | 13 | 1.5 | L0.9 | 250 | 0.5 | 13 | 1.6 |
| L120 | 190 | 27.0 | 15 | L0.6 | 290 | 0.5 | 14 | 1.6 | L0.7 | 290 | 1.0 | 14 | 1.6 |
| L140 | 195 | 31.5 | 15 | L0.5 | 330 | 0.5 | 16 | 1.7 | L0.5 | 300 | 1.0 | 15 | 1.6 |
| B40 | 130 | 15.0 | 13 | B1.0 | 250 | 0.5 | 9 | 1.3 | B5.0 | 210 | 0.5 | 6 | 1.0 |
| B60 | 160 | 20.5 | 8 | B0.9 | 260 | 1.0 | 9 | 1.4 | B4.0 | 200 | 0.5 | 6 | 1.0 |
| B80 | 165 | 26.0 | 9 | B0.8 | 250 | 1.0 | 10 | 1.4 | B2.0 | 200 | 0.5 | 7 | 1.0 |
| B100 | 180 | 30.5 | 9 | B0.7 | 300 | 1.0 | 11 | 1.4 | B0.9 | 260 | 0.5 | 9 | 1.3 |
| B120 | 180 | 34.5 | 10 | B0.6 | 290 | 1.0 | 12 | 1.4 | B0.7 | 280 | 0.5 | 11 | 1.5 |
| B140 | 190 | 36.5 | 11 | B0.5 | 330 | 0.5 | 13 | 1.4 | B0.5 | 300 | 1.0 | 13 | 1.3 |
| H40 | 140 | 10.5 | 27 | H1.0 | 230 | 0.5 | 9 | 1.3 | H5.0 | 190 | 0.5 | 11 | 1.1 |
| H60 | 160 | 12.0 | 27 | H0.9 | 230 | 0.5 | 10 | 1.3 | H4.0 | 180 | 0.5 | 11 | 1.1 |
| H80 | 180 | 15.0 | 23 | H0.8 | 250 | 0.5 | 12 | 1.5 | H2.0 | 180 | 0.5 | 10 | 1.1 |
| H100 | 185 | 17.5 | 17 | H0.7 | 240 | 0.5 | 12 | 1.3 | H0.9 | 230 | 1.0 | 11 | 1.3 |
| H120 | 205 | 22.5 | 17 | H0.6 | 260 | 1.0 | 13 | 1.5 | H0.7 | 240 | 0.5 | 13 | 1.6 |
| H140 | 200 | 25.0 | 19 | H0.5 | 290 | 1.0 | 14 | 1.7 | H0.5 | 260 | 0.5 | 14 | 1.5 |
| P40 | 155 | 11.5 | 16 | P1.0 | 250 | 0.5 | 9 | 1.3 | P5.0 | 190 | 0.5 | 9 | 1.1 |
| P60 | 160 | 16.0 | 17 | P0.9 | 250 | 1.0 | 9 | 1.5 | P4.0 | 190 | 0.5 | 9 | 1.1 |
| P80 | 185 | 18.5 | 15 | P0.8 | 270 | 0.5 | 10 | 1.4 | P2.0 | 200 | 0.5 | 8 | 1.2 |
| P100 | 195 | 21.0 | 15 | P0.7 | 300 | 0.5 | 11 | 1.5 | P0.9 | 260 | 0.5 | 10 | 1.4 |
| P120 | 205 | 25.0 | 15 | P0.6 | 300 | 1.0 | 11 | 1.3 | P0.7 | 280 | 0.5 | 11 | 1.5 |
| P140 | 205 | 29.5 | 15 | P0.5 | 350 | 0.5 | 12 | 1.4 | P0.5 | 300 | 0.5 | 12 | 1.4 |

Table 5.5: Parameter values used for generating the results in Table 5.1. σ_{est} denotes optimised $\sigma_{\mathcal{G}}$ for BM3D.

| Data | σ | λ | k_{\max} | Data | σ | λ | k_{\max} | Data | σ | λ | k_{\max} | σ_{est} | Data | σ | λ | k_{\max} | σ_{est} |
|------|----------|-----------|------------|------|----------|-----------|------------|-------|----------|-----------|------------|-----------------------|-------|----------|-----------|------------|-----------------------|
| PS25 | 125 | 9.5 | 21 | L50 | 150 | 14.5 | 17 | PS4.0 | 180 | 0.5 | 11 | 1.1 | L5.0 | 180 | 0.5 | 11 | 1.1 |
| PS40 | 145 | 12.5 | 16 | L75 | 170 | 19.0 | 15 | PS2.0 | 180 | 0.5 | 11 | 1.1 | L2.0 | 200 | 0.5 | 10 | 1.2 |
| L25 | 130 | 9.0 | 19 | L100 | 175 | 23.5 | 16 | PS1.0 | 200 | 0.5 | 10 | 1.2 | L1.0 | 240 | 0.5 | 12 | 1.4 |
| L40 | 150 | 11.5 | 16 | H50 | 160 | 11.0 | 23 | PS0.5 | 250 | 0.5 | 13 | 1.6 | CS5.0 | 160 | 0.5 | 12 | 1.0 |
| | | | | H75 | 165 | 15.5 | 23 | PS0.2 | 290 | 1.0 | 14 | 1.6 | CS2.0 | 170 | 0.5 | 10 | 0.9 |
| | | | | H100 | 185 | 17.5 | 17 | PS0.1 | 300 | 1.0 | 15 | 1.6 | CS1.0 | 200 | 0.5 | 9 | 0.8 |
| | | | | | | | | BS4.0 | 210 | 0.5 | 6 | 1.0 | | | | | |
| | | | | | | | | BS2.0 | 200 | 0.5 | 6 | 1.0 | | | | | |
| | | | | | | | | BS1.0 | 200 | 0.5 | 7 | 1.0 | | | | | |
| | | | | | | | | BS0.5 | 260 | 0.5 | 9 | 1.3 | | | | | |
| | | | | | | | | BS0.2 | 280 | 0.5 | 11 | 1.5 | | | | | |
| | | | | | | | | BS0.1 | 300 | 1.0 | 13 | 1.3 | | | | | |

Table 5.6: **Left to Right:** Parameters corresponding to results in Tables 5.2 - 5.3, respectively.

the noise distribution, proves useful in multiple ways. We will catch sight of these strengths in the upcoming experiments on practical data.

5.2.4 Electron Microscopy Data

Here, we have considered two datasets of ribosome particles in yeast cells. In each of them, we have two pre-aligned noisy images of the same scene. This enables quantifying the quality of denoised images using the FRC criterion. In Figure 5.5 we show the resulting denoised images along with the corresponding FRC curves. These were computed for our approach as well as AWGN versions of NLB and BM3D.

Four important features can be observed in the noise degradations of an electron microscope:

- These images are obtained after averaging a few hundreds of very noisy images which do not show any sort of signal to the naked eye.
- The appearance of noise is different in each of the datasets. Moreover, these degradations are also different from both Gaussian and mixture kinds of noise considered in the synthetic images. Hence, the noise model is not well understood.
- Although the category of noise distribution in the first dataset seems to be an unstructured one, the other is definitely spread over relatively larger spatial scales. Thus, it is a slightly structured degradation.
- There is no guarantee that the noise distribution is spatially invariant.

These observations signify the need for a filter that is robust to the type of noise and also works for highly degraded images. For the first dataset, in contrast to NLB and BM3D, our method produces sharper images with clearly visible structures. In the second one, NLB is not able to remove the noise completely while BM3D produces a less sharper output than NFPR. While these are visual interpretations which might differ from person to person, FRC curve is a quantitative objective measure. The noise distribution along with its accidental spatial patterns are presumably different in two different noise realisations of the same scene. Hence, the FRC criterion quantifies the ability of a filter to preserve original image structures. The FRC curves in Figure 5.5 show that our filter outperforms NLB and BM3D.

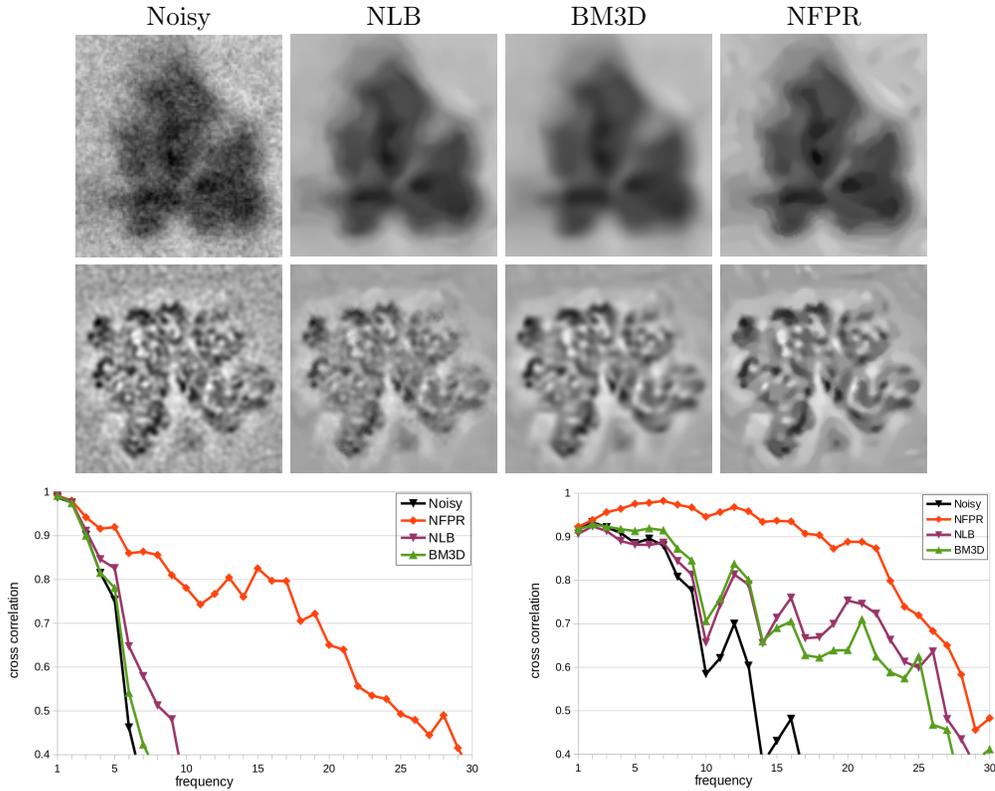


Figure 5.5: **Top:** Zoom into ribosome image of a yeast cell. **Middle:** Zoom into ribosome image of another yeast cell. Images courtesy of Utz Ermel and Lasse Sprankel from the Goethe University of Frankfurt. Original size of both datasets is 256×256 . **Bottom Left and Bottom Right:** Zoom into the $[0.4-1.0]$ correlation range of the corresponding FRC plots for both images, respectively. NFPR parameters used: $\sigma = 170$, $\lambda = 2.5$, $k_{\max} = 35$. In case of NLB and BM3D we have optimised the unknown σ_G with respect to FRC.

5.2. Experiments and Discussion

In the synthetic noise scenario we had four takeaway lessons from the AWGN experiments. Some of these observations were also true for Poissonian and mixture noise layouts. In a similar spirit to these remarks, the following observations stand out for practical microscopy data:

- NFPR has a better structure preserving performance.
- It is directly applicable to unstructured as well as slightly structured degradations.
- Spatially variant noisy data can also be dealt with using NFPR.

The BM3D and NLB filters employ data adaptive parametric choices. Thus, one could definitely achieve further quality gain for NFPR in this direction. However, in this work we concentrate on confirming the capability of our technique as a robust solution with respect to the noise model. Autonomous and adaptive parameter selection for NFPR, similar to the non-local means filter [322, 328], requires careful and separate studies. Our experience indicates that NFPR is also applicable for the clipped-AWGN model which mimics over-exposed/under-exposed conditions. In the impulse noise [89, 90] scenario, only a percentage of pixels are corrupted. Here, our strategy needs to be combined with noise detection algorithms.

Having now looked at an extensive evaluation between BM3D, DDID and ours, it is vital that we also comprehend the modelling differences between them: BM3D performs both inter- and intra-patch smoothing. Despite the usage of non-linear filtering, intra-patch smoothing cannot completely stop interactions between regions of dissimilar grey values. On the other hand, in our filter we only perform an inter-patch smoothing. This is one of the main reasons why NFPR avoids artefacts despite using less than 1% of the pixels in contrast to BM3D: We use just 35 pixels in a reordered set, while BM3D considers 32 patches with 144 pixels each.

The DDID class of methods also avoids artefacts. These approaches assume that the noise model is known a-priori and estimate the noise distribution according to this assumption. These estimates are then subtracted from the raw data to obtain the noise-free signal. We have a diametrically opposite philosophy. NFPR assumes that the signal, but not the noise, can be modelled using patch and pixel similarities. The advantage over DDID is that, we are robust to the noise model.

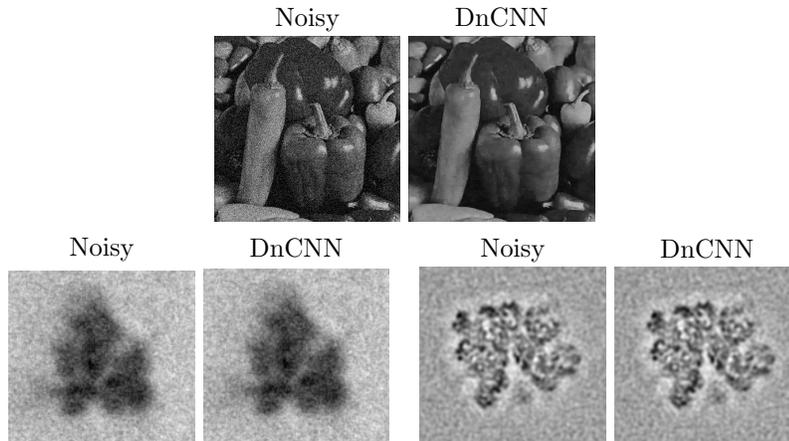


Figure 5.6: **Top:** Synthetic noise. **Bottom:** Electron microscopy data.

Deep learning approaches have been setting the state-of-the-art standard in image denoising recently. Let us consider the denoising convolutional neural network (DnCNN) method [292], which is one of the most widely accepted learning-based solutions for removing AWGN. One can clearly observe from the Figure 8.2 that DnCNN indeed works for a synthetic Gaussian noise model. However, it fails to denoise the image in practical scenarios. Such strategies cannot cover a broader spectrum of noise models as they are uni-dimensionally trained for a very specific kind of data. Other learning approaches which aim at solving this problem have not completely been successful. They either require selection of the loss function which is deterioration dependent [293] or need analytical noise models [329].

In conclusion, deep learning methods are perfectly suitable for applications where the type of noise is well known and the amplitudes of noise are small. However, for highly degraded images, in scenarios where training data is not easily available, and in situations when the data cannot be easily approximated with an analytical noise model, our method is better suited. Electron microscopy is one such application because the process of acquiring raw data to obtaining the final 3D structure of a specimen is a very complex pipeline. One might require a denoising technique at any point in this pipeline. After almost every step, the noise distribution changes. Thus, it is highly recommended to have a filter at hand which is robust to the type of degradation. We have experimentally demonstrated this very property of NFPR in this chapter.

The most time consuming part of our filter is the patch reordering stage. Experiments on an NVIDIA GeForce GTX 970 graphics card using C++ and CUDA, indicated that our method took 2 seconds for denoising the 256×256 sized image ($\sigma_G = 100$). Recently, BM3D was implemented in real-time on a GPU [330]. Thus, we can further speed up NFPR with faster versions of a naive patch-matching implementation.

5.3 Conclusions

The crux of this chapter takes the very first steps in the direction of a universal denoising filter. Such a filter must possess all-round qualities like preserving structural image information, avoiding artefacts, consuming less computational time and not requiring a-priori knowledge about noise statistics. Although image denoising research is decades-old, only a handful of works have studied this general direction. Most works in both learning-based and model-based worlds create a noise degradation synthetically on a computer and then utilise this very knowledge within the design of a filter. Practical data need not comply with this idea. In this work, we have concentrated solely on modelling the noise-free signal without including any noise parameters within the non-linearity. Although such a formulation might reduce the filter performance on a specific synthetic noise distribution, we have shown that it gives a clear advantage on practical world data. Our technique can thus be seen as a successful attempt in the pursuit of a universal filter. It can be directly employed for eliminating unstructured as well as slightly structured kinds of noise which cover a vast spectrum of imaging applications.

It is now clear from the experiments conducted in this chapter that there are several advantages of combining the patch-based similarity assumption with its pixel-based counterpart. However, we just scrutinised a disc shaped neighbourhood. In the upcoming two chapters we examine a couple of anisotropic filter shapes.

Manuscript details of research content from this Chapter:

K. Bodduna and J. Weickert. Image Denoising with Less Artefacts: Novel Non-linear Filtering on Fast Patch Reorderings. *arXiv:2002.00638 [eess.IV]*, February 2020.

Chapter 6

Sector Diffusion - A Corner Preserving Diffusion Model

In Chapter 4, we saw that the disc diffusion (DD) filter has the ability to avoid artefacts in homogenous regions. However, it cannot compete with well-performing methods like 3D block matching (BM3D) [15] and non-local Bayes (NLB) [14] for textured images. Then we combined the pixel similarity assumption in disc diffusion with the patch similarity counterpart in non-local means (NLM) [160, 161] algorithm in Chapter 5. This resulted in the non-linear filtering on fast patch reorderings (NFPR) approach. We found that, unlike DD, NFPR was able to compete with BM3D and NLB on textured images while also retaining the ability to prevent artefacts in homogenous regions. Nevertheless, there is still scope for designing diffusion-based techniques that perform better in texture regions. In particular, one must investigate filters that catch anisotropic behaviour in an efficient manner. In this chapter, we concentrate on one such solution.

We have seen that non-linear diffusion models possess the ability to preserve edges. Corners, on the other hand, are regions where multiple edges exist in a local neighbourhood. These structures form a significant part of textured regions in an image. There is still room for improvement when it comes to processing corners. This is a less explored and relatively difficult research area. One idea that helps in preserving corners is to determine a-priori whether a particular location is a corner or not and perform anisotropic smoothing accordingly [331–333]. An energy minimisation method was also employed for processing corners [134, 144]. Here, the idea is to first find

the two dominant directions in a local neighbourhood using a structure tensor. This is followed by modelling an energy that incorporates these directions and is responsible for performing a smoothing along them. There are, however, many situations where more than two edges intersect. Here, one needs to employ a smoothing in multiple directions unlike the assumptions in [134, 144, 334, 335]. Shape adaptive filters [20–23] have also been proposed for dealing with corner like structural image information. They generally compute a data adaptive structure element around every pixel. This is computationally very intensive.

Our Goal. Introduce a diffusion model which has good performance in terms of corner preservation.

Our Contribution. The sector diffusion (SD) model we propose has multiple ingredients. The most important one among them is the novel idea of incorporating one-sided derivatives in its modelling. As a consequence of this feature, SD is able to smooth in multiple directions instead of just two, does not require to determine a-priori whether a particular location belongs to a corner region, and also has a pre-computable filter shape.

Chapter structure. This chapter is outlined as follows: We introduce the motivation behind modelling of our sector diffusion filter in Section 6.1. The experiments and discussions on the results pertaining to corner preservation and denoising are presented in Section 6.2. We summarise our conclusions from the experiments along with an outlook on future work in Section 6.3.

6.1 Modelling and Theory

6.1.1 Continuous Modelling of Sector Diffusion

Let us revisit the integration model of Weickert [17] (equation (4.1) from Chapter 4):

$$\partial_t u(\mathbf{x}, t) = \frac{1}{\pi} \int_0^\pi \partial_\theta (g(\partial_\theta u_\sigma) \partial_\theta u) d\theta. \quad (6.1)$$

This model considers each direction θ separately and is thus capable of diffusing along edges, but not across them. We used the above image evolution as a starting point for disc diffusion. However, instead of directly consid-

ering the non-local extension of the above filter, sector diffusion adheres to a different idea. *In order to improve the structure adaptation even further, we replace the directional derivatives by one-sided directional derivatives and integrate over $[0, 2\pi]$ instead of $[0, \pi]$:*

$$\partial_t u(\mathbf{x}, t) = \frac{1}{2\pi} \int_0^{2\pi} \partial_\theta^+ (g(\partial_\theta^+ u_\sigma^\theta) \partial_\theta^+ u) d\theta. \quad (6.2)$$

Here, u_σ^θ represents a one-sided smoothing of u in the orientation given by the angle θ , and ∂_θ^+ denotes a one-sided derivative in the same orientation. In contrast to the usual Gaussian smoothing applied in (6.1), this one-sided smoothing allows the filter to distinguish two different derivatives for a given direction: One in the orientation of θ , and the other in the orientation of $\theta + \pi$. A formal definition of these concepts can be realised by considering the restriction of u to the corresponding ray starting at \mathbf{x} , in the orientation of each θ . Namely, for fixed \mathbf{x}, t, θ , we consider $u(h; \mathbf{x}, t) := u(\mathbf{x} + h(\cos(\theta), \sin(\theta))^T, t)$, for $h \in [0, \infty]$. Then, the one-sided directional derivative ∂_θ^+ is formally defined as

$$\partial_\theta^+ u := \lim_{h \rightarrow 0^+} \frac{u(h; \mathbf{x}, t) - u(\mathbf{x}, t)}{h}. \quad (6.3)$$

To our knowledge, diffusion filters that are explicitly based on one-sided directional derivatives have not been described in the literature so far. In fact, we have also not come across theories about image evolutions that involve one-sided derivatives in general. They are mathematically challenging and could open the door to several new directions of research.

In order to introduce a second alteration of model (6.1), we supplement one-sided derivatives with concept of non-locality. This leads to

$$\partial_t u(\mathbf{x}, t) = \int_{B_{\mathbf{x}, \rho}} J(|\mathbf{y} - \mathbf{x}|) g\left(\frac{u_\sigma(\mathbf{y}; \mathbf{y} - \mathbf{x}) - u_\sigma(\mathbf{x}; \mathbf{y} - \mathbf{x})}{|\mathbf{y} - \mathbf{x}|}\right) (u(\mathbf{y}) - u(\mathbf{x})) d\mathbf{y}. \quad (6.4)$$

In contrast to the disc diffusion modelling in (4.2), the above non-local formulation consists of the value $u_\sigma(\mathbf{z}; \mathbf{y} - \mathbf{x})$. It corresponds to a one-dimensional Gaussian smoothing of u inside the segment

$$\lambda_{\mathbf{x}\mathbf{y}}(s) := \left\{ \mathbf{x} + s \frac{(\mathbf{y} - \mathbf{x})}{|\mathbf{y} - \mathbf{x}|} : s \in [0, \rho] \right\}, \quad (6.5)$$

evaluated at \mathbf{z} . This idea of making the diffusivity dependent on values inside an orientation dependent segment determines the structure preservation

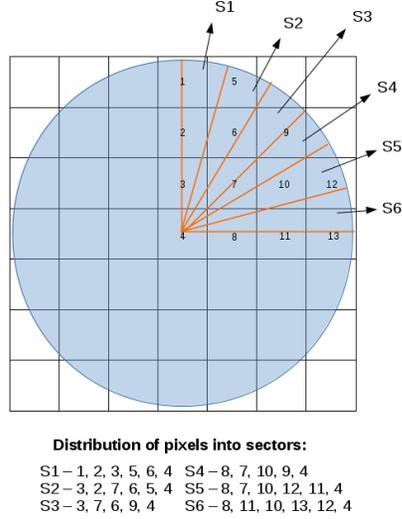


Figure 6.1: Mapping of pixels to sectors

capabilities of the model. With these design choices, (6.4) can also be considered as an example of a general non-local theoretical framework developed in [318]. We will now discuss how to translate the above non-local filter into space-discrete and fully discrete versions by dividing the disc $B_{x,\rho}$ into sectors. This explains the name sector diffusion.

6.1.2 Discrete Modelling of Sector Diffusion

The following ideas form the core of the sector diffusion model:

- In order to properly align our filter with corners in an image, we first divide a disc shaped neighborhood $B_{i,\rho}$ of radius ρ centered around pixel i into M sectors. With the objective of reducing interactions between regions of dissimilar grey values we employ robust smoothing within these sectors. This mirrors the continuous modelling idea of smoothing within the segments λ_{xy} .
- The second key idea is that we map a particular pixel i within this disc to a sector S_ℓ , if they have a non-zero area of intersection. This would give rise to a distribution where each pixel can be mapped to multiple sectors. Thus, we have more pixels in each sector in contrast to the situation that would arise when we map each pixel to just one

sector. Figure 6.1 shows the distribution of the pixels in a disc shaped neighbourhood to their respective sectors, as an example.

- The final design objective is that we employ one-sided derivatives instead of central derivatives for discretisation purposes. The latter have a property of smoothing over the central pixel, thus destroying the corners. This idea is again a direct consequence of considering orientations rather than directions in the continuous model.

With the above motivations in mind, we now go ahead to define the space-discrete formulation of the sector diffusion model as

$$\frac{d}{dt}u_i = \sum_{\ell=1}^M \sum_{j \in S_\ell} g_{i,j} \cdot \frac{u_j - u_i}{|\mathbf{x}_j - \mathbf{x}_i|^2}. \quad (6.6)$$

One can immediately see a double summation in the above equation which arises due to the division of a disc into sectors. *In the case of disc diffusion, we encounter just a single summation in its evolution equation.* Also, in the above equation, $g_{i,j} = g\left(\frac{u_{\sigma j\ell} - u_{\sigma i\ell}}{|\mathbf{x}_j - \mathbf{x}_i|}\right)$ and S_ℓ is the set of pixels within a particular sector ℓ . The symbols \mathbf{x}_i and \mathbf{x}_j denote the position of the pixels i and j in the image grid. The sector-restricted smoothing is defined as

$$u_{\sigma j\ell} = \frac{1}{c} \sum_{k \in S_\ell} h_1(k, j, \sigma) u_k. \quad (6.7)$$

Here, c is a normalisation constant and

$$h_1(k, j, \sigma) = \exp\left(\frac{-|\mathbf{x}_k - \mathbf{x}_j|^2}{2\sigma^2}\right). \quad (6.8)$$

The value of $u_{\sigma i\ell}$ in (6.6) is also computed using (6.7), but with a weighting function h_2 which is slightly different from h_1 :

$$h_2(k, i, \sigma) = \exp\left(\frac{-|\mathbf{x}_k - \mathbf{x}_i|^2}{2\sigma^2}\right) - \exp\left(\frac{-|\mathbf{x}_k - \mathbf{x}_i|^2}{2\sigma_1^2}\right). \quad (6.9)$$

By choosing a very small value for σ_1 ($\ll \sigma$), we downweight the contribution of the weight corresponding to the case $k = i$. This gives more a robust smoothing estimate $u_{\sigma i\ell}$ [336]. Such a slightly different Gaussian smoothing is required to compensate for the asymmetric shape of a sector. *In disc diffusion, however, we always use the function h_1 .*

The space-time discrete evolution counterpart of (6.6) with a time step τ can be written as

$$\mathbf{u}_i^{k+1} = \mathbf{u}_i^k + \tau \cdot \left(\sum_{l=1}^M \sum_{j \in S_l} g \left(\frac{u_{\sigma_{jl}}^k - u_{\sigma_{il}}^k}{|\mathbf{x}_j - \mathbf{x}_i|} \right) \frac{u_j^k - u_i^k}{|\mathbf{x}_j - \mathbf{x}_i|^2} \right). \quad (6.10)$$

The above equation can be formulated in a more compact way using a matrix $\mathbf{Q}(\mathbf{u}^k) \in \mathbb{R}^{M \times M}$:

$$\mathbf{u}^{k+1} = \mathbf{Q}(\mathbf{u}^k) \mathbf{u}^k. \quad (6.11)$$

The entries of the matrix $q_{i,j}(\mathbf{u}^k)$ are given by

$$q_{i,j}(\mathbf{u}^k) = \begin{cases} 1 - \sum_{l=1}^M \sum_{j \in S_l} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2}, & \text{if } j = i, \\ \sum_{l \in P_j} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2}, & \text{if } j \in B_\rho(i) \text{ and } j \neq i, \\ 0, & \text{otherwise,} \end{cases} \quad (6.12)$$

with diffusivity weight $g_{ij}(\mathbf{u}^k) = g \left(\frac{u_{\sigma_{jl}}^k - u_{\sigma_{il}}^k}{|\mathbf{x}_j - \mathbf{x}_i|} \right)$. Remember that each pixel within the disc can be present in multiple sectors. Thus, we use the symbol P_j to represent the set of sectors in which pixel j is present. *The existence of the sets S_ℓ and P_j is the main difference between the matrices corresponding to disc and sector diffusion techniques.* Moreover, in the above matrix, the values of $u_{\sigma_{jl}}$ and $u_{\sigma_{il}}$ would be different in the case where u_j is the center pixel and u_i is the center pixel. This is due to the fact that we have a different set of neighbours for computing the Gaussian smoothing in both cases. It is a direct consequence of the intrinsic shape of a sector. We also have slightly different smoothing functions, h_1 and h_2 , for computing $u_{\sigma_{jl}}$ and $u_{\sigma_{il}}$, respectively. *Thus, unlike in the case of disc diffusion, $\mathbf{Q}(\mathbf{u}^k)$ is not symmetric in this approach.*

We choose the time step τ in such a way that a maximum-minimum principle is valid. In other words we want to guarantee that after one time step the resulting image \mathbf{u}^{k+1} has lower and upper bounds given by the maximum and minimum values of the image \mathbf{u}^k . The discrete diffusion theory of Weickert [133] was proposed for evolution equations where $\mathbf{Q}(\mathbf{u}^k)$ is symmetric. However, it mentions that the maximum-minimum principle can still be met without this requirement of symmetry: First of all we need that the row

sums of the matrix $\mathbf{Q}(\mathbf{u}^k)$ are equal to 1. Due to the area-based mapping from pixels to sectors, we have

$$\sum_{\{j \in B_{i,\rho} : j \neq i\}} \sum_{l \in P_j} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2} = \sum_{l=1}^M \sum_{j \in S_l} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2}. \quad (6.13)$$

The left hand side of the above equation denotes the sum of the weights pertaining to all pixels within the disc $B_{i,\rho}$, other than the center pixel i . The right hand side denotes the sum of the corresponding weights for i itself. Using the above equation, we can prove that the row sums are equal to 1:

$$1 - \sum_{l=1}^M \sum_{j \in S_l} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2} + \sum_{\{j \in B_{i,\rho} : j \neq i\}} \sum_{l \in P_j} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2} = 1. \quad (6.14)$$

In order to satisfy the maximum-minimum principle, we additionally need that the entries of the matrix $\mathbf{Q}(\mathbf{u}^k)$ are non-negative. This is fulfilled if the following condition holds:

$$1 - \sum_{l=1}^M \sum_{j \in S_l} \frac{\tau \cdot g_{ij}(\mathbf{u}^k)}{|\mathbf{x}_j - \mathbf{x}_i|^2} \geq 0. \quad (6.15)$$

Given that the maximum value of diffusivity is 1, the final limits for the time step size simplify to

$$0 \leq \tau \leq \frac{1}{\sum_{l=1}^M \sum_{j \in S_l} \frac{1}{|\mathbf{x}_j - \mathbf{x}_i|^2}}. \quad (6.16)$$

This concludes the modelling part of sector diffusion. Now we evaluate its corner preservation and denoising capability.

6.2 Experiments and Discussion

6.2.1 Datasets and Methods for Evaluation

In the interest of evaluating the corner preservation ability of SD, we have corrupted the synthetic Texmos test image with additive white Gaussian noise (AWGN). We also consider real-world test images Lena, House, Peppers, and Bridge, corrupted with the clipped-AWGN (grey values clipped to a dynamic range $[0, 255]$). In our comparative analysis, we have considered the well known edge-enhancing diffusion (EED) method [132] which has been discretised according to [337].

6.2. Experiments and Discussion

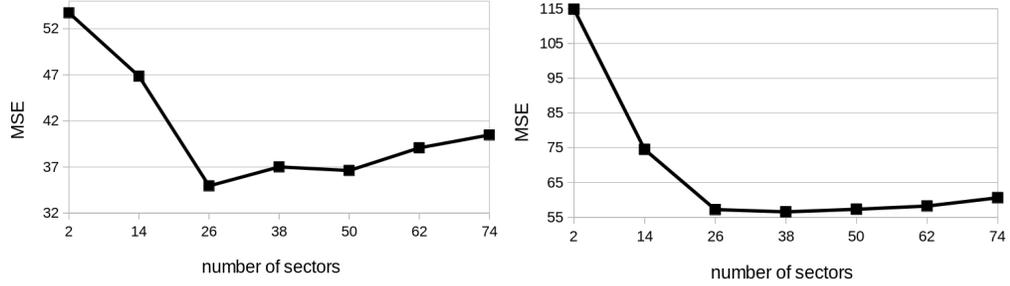


Figure 6.2: Influence of the number of sectors on the denoising output for Texmos image. **Left:** $\sigma_G = 50$, parameters used: $\sigma = 0.5$, $\rho = 7$, and $\lambda = 1.0$. **Right:** $\sigma_G = 75$, parameters used: $\sigma = 0.5$, $\rho = 7$, and $\lambda = 1.2$. Image size: 512×512 .

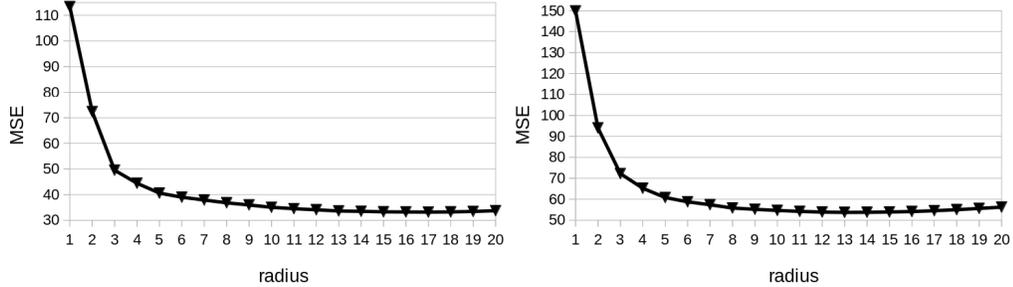


Figure 6.3: Influence of the radius on the denoising output for Texmos image. **Left:** $\sigma_G = 50$, parameters used: $\sigma = 0.5$, $M = 36$, and $\lambda = 1.2$. **Right:** $\sigma_G = 75$, parameters used: $\sigma = 0.5$, $M = 36$, and $\lambda = 1.2$.

6.2.2 Parameter Selection

We have seven parameters in sector diffusion (SD): The smoothing parameters σ, σ_1 , contrast parameter λ , radius of the disc ρ , total number of iterations k_{\max} , number of sectors M and the time step size τ . However, we will shortly see that we do not have to optimise ρ , M and τ for denoising purposes.

Influence of the Number of Sectors on Denoising. Figure 6.2 shows the mean squared error (MSE) values of the denoised Texmos image, as a function of the number of sectors chosen. We can make three crucial observations from the plot: We obtain high errors for low number of sectors ($M < 26$). This is due to the fact that the adaptivity to the image structures is not optimal for sectors with larger areas. We also observe large MSE val-

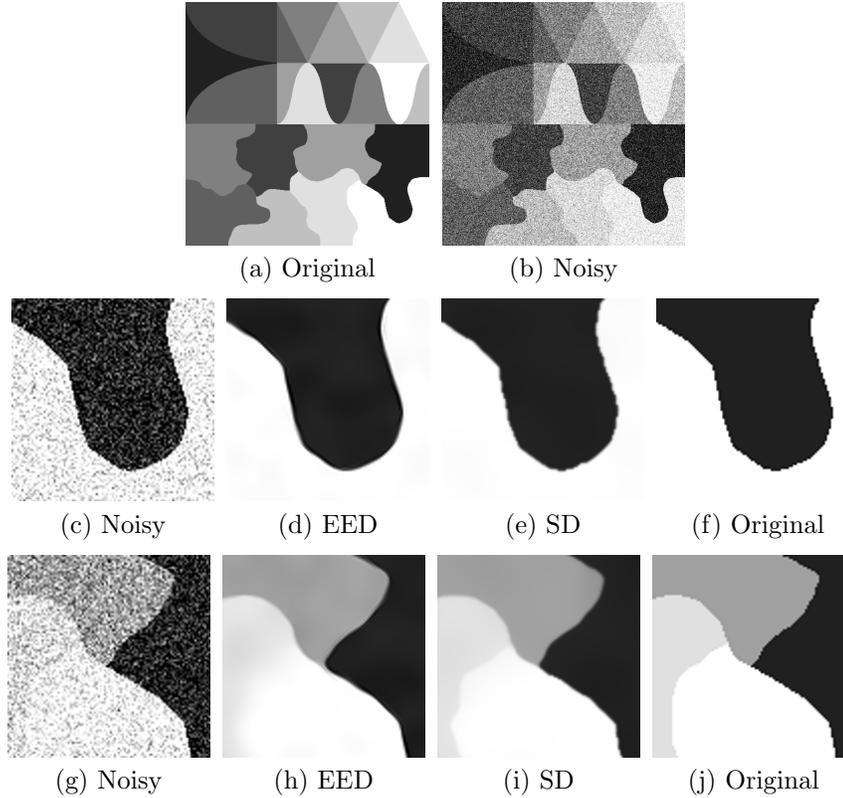


Figure 6.4: Denoised Texmos images using various methods ($\sigma_G = 50$). **Top Row:** Full images. **Middle Row:** Zoom into the images for visualising fine structures along an edge. **Bottom Row:** Zoom into the images for visualising two particular corner regions. SD Parameters: $\sigma = 0.5$, $\lambda = 1.0$ and $k_{\max} = 65$.

ues for very high number of sectors ($M > 50$) because we have less number of pixels in each sector which leads to sub-optimal smoothing. Finally, least MSE is achieved for a moderate range ($26 < M < 50$). Within this range we avoid both the above disadvantages. We can additionally notice that these observations are common for two different standard deviations of noise, thus signifying the importance of sector-based modelling.

Influence of the Disc Radius on Denoising. Figure 6.3 presents the MSE values of the denoised Texmos image as a function of the disc radius. We can make observations similar to the above graphs: We achieve better MSE values for a moderately large radius. When the radius is low there are

6.2. Experiments and Discussion

| Image | EED | | | | SD | | | |
|-------|----------|-----------|------------|--------------|----------|-----------|------------|---------------|
| | σ | λ | k_{\max} | MSE | σ | λ | k_{\max} | MSE |
| L40 | 1.2 | 7.5 | 34 | 98.67 | 0.6 | 3.1 | 7 | 92.99 |
| L60 | 1.8 | 5.0 | 63 | 156.24 | 0.6 | 3.3 | 11 | 138.48 |
| L80 | 2.0 | 4.6 | 87 | 230.28 | 0.6 | 2.9 | 18 | 180.66 |
| B40 | 0.9 | 14.4 | 12 | 294.32 | 0.5 | 3.3 | 4 | 261.62 |
| B60 | 1.6 | 20.5 | 8 | 333.37 | 0.5 | 4.1 | 6 | 360.87 |
| B80 | 1.4 | 10.4 | 28 | 514.23 | 0.6 | 4.0 | 9 | 436.60 |
| H40 | 0.9 | 11.1 | 34 | 96.62 | 0.7 | 2.6 | 9 | 104.31 |
| H60 | 1.1 | 12.1 | 33 | 167.72 | 0.7 | 2.7 | 14 | 152.24 |
| H80 | 1.8 | 5.8 | 72 | 247.09 | 0.6 | 2.7 | 19 | 207.65 |
| P40 | 1.2 | 8.1 | 28 | 102.97 | 0.6 | 2.1 | 10 | 86.57 |
| P60 | 1.7 | 5.6 | 51 | 200.31 | 0.6 | 1.8 | 19 | 133.19 |
| P80 | 1.9 | 5.1 | 68 | 353.61 | 0.6 | 1.7 | 30 | 188.86 |

Table 6.1: MSE values of denoised images including parameters used. L40 stands for Lena image with $\sigma_G = 40$. B, H, P denote Bridge, House and Peppers images, respectively.



Figure 6.5: Zoom into Lena images ($\sigma_G = 60$).

too few pixels within the disc which does not result in a robust smoothing estimate. A very high radius is also sub-optimal since this gives rise to a situation where grey values from different regions in the image can interact with each other.

Based on the above observations, we have chosen $M = 36$ and $\rho = 7$. Such a choice was made keeping in mind both computational time and the quality of the denoised images. We use 95% of the upper bound for the time step τ in (6.16). The small positive constant value σ_1 is set to 0.0005. Finally, we have to optimise only three parameters: σ , λ , and k_{\max} .

Corner Preservation Experiment. Figure 6.4 shows the denoised versions of the Texmos test image obtained using both SD and EED filters.

One can observe the superior corner and fine structure preserving characteristics of SD over EED.

Real-world Image Denoising Experiment. Table 6.1 presents the mean squared error (MSE) values of denoised images. It is clear from the table that SD gives the best results. This can also be verified visually from Figure 6.5, where the eye of Lena is better preserved by SD.

We attribute the better performance of SD to four specific model features: Using one-sided derivatives for the diffusion process, division of a disc-shaped neighbourhood into sectors, usage of an area-based mapping between pixels within the disc and sectors, and employment of a robust smoothing within these sectors.

6.3 Conclusions

We have seen that corner preservation with a diffusion-based filter requires sophisticated modelling ideas. Each of these ingredients is vital for the efficient performance of SD. Sector diffusion is in fact the first diffusion method that consequently uses only one-sided directional derivatives. In its local formulation, this is a model that offers also structural novelties from a mathematical perspective, since it cannot be described in terms of a partial differential equation. From a practical perspective, the non-local sector diffusion possesses a higher structural adaptivity and a better denoising performance than simpler diffusion models. Thus, it appears promising to study its usefulness in applications beyond denoising. We investigate such an idea in Chapter 9.

We began this chapter by pursuing the direction of evaluating anisotropic filter shapes other than a disc. Although these efforts proved to be fruitful, in the upcoming chapter we will analyse another neighbourhood shape that can adapt to anisotropy within data. This will ensure that we complete a comprehensive evaluation of several filter shapes.

Chapter 7

Importance of Filter Shape in Denoising: Stripe Diffusion

Every image denoising method tries to make three crucial decisions: Choosing the combination of pixels on which the filter needs to be applied, deciding the type of non-linearity for filtering purposes, and finally the process through which the denoised image is generated from the filter outputs.

In this chapter, we concentrate on the first step of the above mentioned sequential process. The most important question in this step is about the spatial shape of the filter, which is also commonly known as a neighbourhood or a structure element. Based on current research, most models can be divided into two categories: Firstly, filters which use fixed shapes defined by a square [14, 15, 18, 160] or a disc [336, 338] (also see Chapters 4, 5). The latter methods generally lead to rotationally invariant models. The other category contains models where the shape of the filter adapts to the image structure [20–23]. However, computing a different shape for every pixel is a computationally expensive process. There is a lack of research when it comes to carefully evaluating fixed shapes other than disc or a square which offer anisotropy. Such a contribution is also useful for structure enhancement (Chapter 10) and robust super-resolution imaging applications (Chapter 9).

In the previous chapter we introduced a sector-shaped neighbourhood (see Figure 7.1b) which falls under the category of fixed filter shapes. The anisotropy offered by a sector lead to better preservation of corners. Also, in Chapter 4, a disc-shaped element was used in a similar non-local diffusion setting as sector diffusion. While a disc contains more pixels in it than a

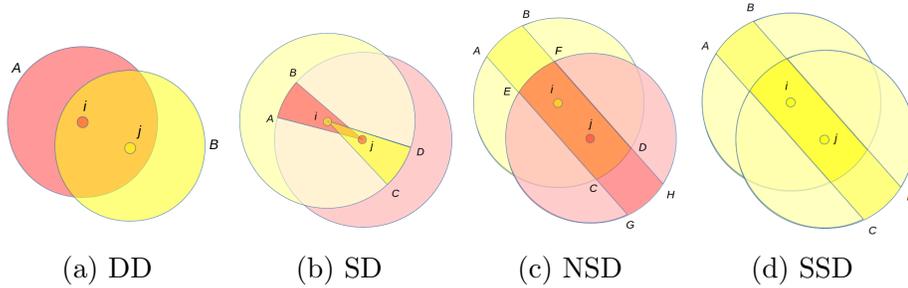


Figure 7.1: Structure elements for various diffusion techniques.

sector, it does not offer as much anisotropy. There has not been any research which combines the advantages of both sector and a disc.

Our Goal. Design a structure element which combines the strengths of both disc and sector shapes.

Our Contribution. We introduce the stripe-shaped diffusion technique. A stripe covers a larger area than a sector, but shows a better ability to catch anisotropic behaviour than a disc. Moreover, we introduce two versions of stripe diffusion - non-symmetric stripe diffusion (NSD) and symmetric stripe diffusion (SSD) which lead to non-symmetric and symmetric image evolutions, respectively. Figures 7.1c and 7.1d offer a first look at the stripe-shaped neighbourhood. Furthermore, we evaluate disc-diffusion (DD), sector diffusion (SD), NSD and SSD in two different image denoising scenarios: Additive white Gaussian noise (AWGN) and clipped-AWGN. While the former depicts a common noise model, the latter represents situations when images are acquired in over- or under-exposed conditions. Such a simultaneous evaluation reveals the robustness of a filter to the kind of noise. This study is highly ignored in the present research works.

Designing fixed shape neighbourhoods that offer anisotropy is still an open problem. However, even before filling this void completely, patch-matching extensions of square or disc shaped neighbourhoods have already been designed [14,15,18]. Our evaluation of different filter shapes thus becomes even more important. This is due to the fact that any improvements from our semi-non-local study could have direct repercussions in a non-local scenario.

Finally, symmetric and non-symmetric image evolutions have their own

advantages and disadvantages in terms of quality and speed. Our SSD and NSD frameworks contribute to both these fields.

Chapter structure. This chapter is structured as follows: We introduce the ideas behind modelling of various structure elements in Section 7.1. The experiments and discussions on the strengths and weaknesses of different filter shapes are presented in Section 7.2. We put forward our conclusions from the experiments along with an outlook on future work in Section 7.3.

7.1 Modelling and Theory

Diffusion-based methods rely on pixel similarity within a spatial neighbourhood. In disc diffusion (Chapter 4), as the name suggests, this neighbourhood is a disc. There exist two denoising steps in every iteration of the image evolution under DD. The first is a simple Gaussian smoothing. The second is a diffusion-based filtering with weights computed on the Gaussian smoothed image. Such a modelling makes the weights robust to the presence of noise [129]. In sector diffusion (Chapter 6), we divide the disc into sectors and then perform the two denoising steps within every sector. The motivation behind such a strategy is that sectors are superior in adapting to corner-like image structures. However, we have less number of pixels in every sector, which is not optimal for image denoising.

In this chapter, we use a stripe-shaped structure element. In contrast to a disc and sector, a stripe offers better anisotropy and more pixels within the structure element, respectively. In this section, we first present the non-symmetric version of stripe diffusion which is closer in spirit to sector diffusion. We then introduce the symmetric variant by combining ideas from both DD and NSD mechanisms.

7.1.1 Non-symmetric Stripe Diffusion

In order to compute the denoised image using NSD, we first consider a disc $B_{i,\rho}$ of radius ρ around every pixel u_i in the 2D image domain Ω . We then divide this disc into M stripes centered around u_i . Each of these stripes spans a different angle with the x-axis. These angles are equally spaced and range between 1 and 180 degrees, such that the entire disc is covered. In Figure 7.1c we can visualise two stripes $ABDC$ and $EFHG$: One each for pixels u_i

and u_j (labelled as i and j for simplicity), respectively. Both these stripes span the same angle with the x-axis. Once we have divided the disc into stripes, we then designate every pixel j within $B_{i,\rho}$ to a stripe ℓ , if they have a non-zero area of intersection. Thus, each pixel in the disc can belong to multiple stripes. Similar to sector diffusion such a design leads to a relatively higher number of pixels in each stripe in contrast to the situation where one pixel is mapped to a single stripe. Now we are ready to define the discrete image evolution process with time step τ as

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{\ell=1}^M \sum_{j \in S_\ell} g \left(\frac{u_{\sigma j\ell} - u_{\sigma i\ell}}{|\mathbf{x}_j - \mathbf{x}_i|} \right) \frac{u_j - u_i}{|\mathbf{x}_j - \mathbf{x}_i|^2}. \quad (7.1)$$

The above equations looks completely similar to the sector diffusion-based image evolution. However, there are two differences: *The set S_ℓ contains the addresses of pixels in stripe ℓ instead of those within a sector.* Moreover, both $u_{\sigma i\ell}$ and $u_{\sigma j\ell}$ are computed using the same Gaussian smoothing process:

$$u_{\sigma j\ell} = \frac{1}{c} \sum_{n \in S_\ell} a(j, n, \sigma) u_n. \quad (7.2)$$

Thus, unlike SD, we do not employ two different weighting functions for computing $u_{\sigma i\ell}$ and $u_{\sigma j\ell}$.

Now, we need to understand why exactly NSD leads to a non-symmetric image evolution. Let us revisit Figure 7.1c. The pixel u_j is present within the disc around u_i and also inside the stripe $ABDC$. When we are computing the image evolution equation for u_i , the Gaussian smoothed estimates for both u_i and u_j are calculated using pixels within the stripe $ABDC$. However, when we compute the same equation for u_j , the pixels which contribute to the Gaussian smoothing come from the stripe $EFHG$. This change in the combination of pixels, leads to a non-symmetric image evolution.

The above explanation holds if we replace stripes $ABDC$ and $EFHG$, with sectors CDi and ABj , respectively (see Figure 7.1b). The inherent shape of these structure elements makes these processes non-symmetric. A detailed explanation for this in terms of a matrix-vector formulation is presented in Chapter 6.

In disc diffusion, the pixel combinations that contribute to the Gaussian smoothing do not change. Thus, it leads to a symmetric image evolution.

| Image | AWGN | | | | Clipped-AWGN | | |
|-------|---------------|--------|---------------|--------|--------------|---------------|---------------|
| | SSD | DD | NSD | SD | NSD | NSD-M | SD |
| L40 | 97.72 | 95.16 | 91.55 | 92.51 | 93.50 | 93.99 | 92.67 |
| L60 | 137.61 | 135.55 | 129.23 | 132.57 | 146.53 | 140.72 | 136.39 |
| L80 | 176.52 | 175.56 | 167.73 | 172.61 | 226.83 | 197.01 | 182.05 |
| L100 | 209.47 | 201.05 | 199.47 | 201.67 | 322.39 | 253.64 | 218.50 |
| B40 | 265.15 | 261.27 | 256.10 | 256.55 | 260.56 | 259.53 | 259.64 |
| B60 | 347.53 | 351.88 | 342.57 | 344.17 | 367.70 | 359.56 | 356.39 |
| B80 | 421.67 | 425.56 | 418.10 | 420.76 | 487.12 | 450.92 | 442.32 |
| B100 | 478.69 | 480.31 | 474.95 | 477.74 | 615.16 | 534.56 | 514.94 |
| H40 | 86.09 | 89.32 | 82.92 | 104.27 | 84.59 | 86.69 | 103.80 |
| H60 | 134.73 | 143.49 | 130.25 | 158.01 | 141.68 | 140.58 | 156.20 |
| H80 | 185.56 | 200.00 | 187.46 | 210.73 | 227.07 | 212.34 | 210.50 |
| H100 | 247.96 | 258.45 | 247.55 | 263.25 | 348.72 | 301.63 | 266.71 |
| P40 | 71.22 | 71.18 | 66.79 | 73.53 | 93.62 | 83.11 | 84.82 |
| P60 | 103.47 | 102.78 | 98.30 | 104.55 | 180.19 | 138.68 | 132.34 |
| P80 | 137.92 | 135.40 | 132.15 | 136.76 | 321.83 | 217.23 | 186.23 |
| P100 | 165.03 | 160.61 | 159.86 | 166.38 | 501.27 | 307.38 | 243.20 |

Table 7.1: MSE values of denoised images. L40 stands for Lena image with $\sigma_G = 40$. B, H, P denote Bridge, House and Peppers respectively.

However, we can turn a non-symmetric stripe diffusion process into a symmetric one by choosing a common ground between DD and NSD.

7.1.2 Symmetric Stripe Diffusion

The SSD image evolution equation is designed in a similar spirit to DD:

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{j \in B_{i,\rho}} g \left(\frac{u_{\sigma_j} - u_{\sigma_i}}{|\mathbf{x}_j - \mathbf{x}_i|} \right) \frac{u_j - u_i}{|\mathbf{x}_j - \mathbf{x}_i|^2}. \quad (7.3)$$

The only difference between SSD and DD is the manner in which we choose the pixels that contribute to the Gaussian smoothing \mathbf{u}_σ . We combine all the pixels in the two stripes $ABDC$ and $EFHG$ shown in Figure 7.1c. This results in a bigger stripe $ABDC$ presented in Figure 7.1d. Irrespective of whether we are computing the image evolution equation for u_i or u_j , such a construction would not change the pixel combination that contributes to the Gaussian smoothing. This leads to a symmetric image evolution.

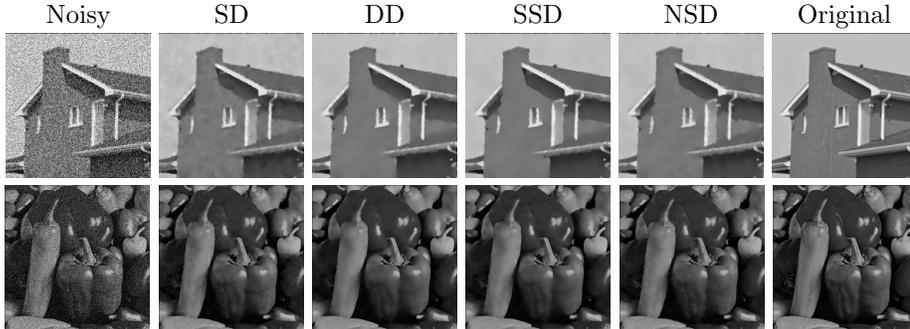


Figure 7.2: **Top and Bottom:** House and Peppers images ($\sigma_G = 40$).

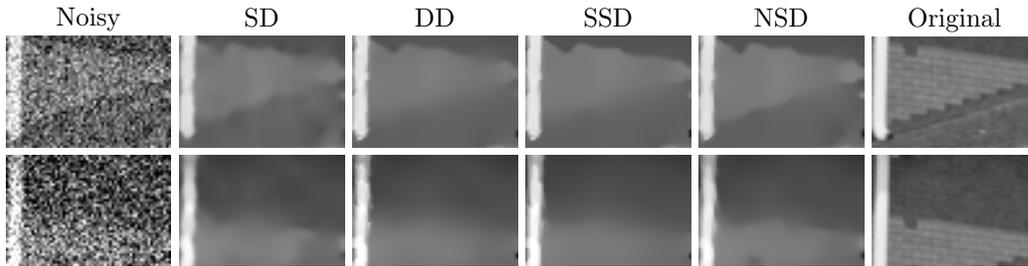


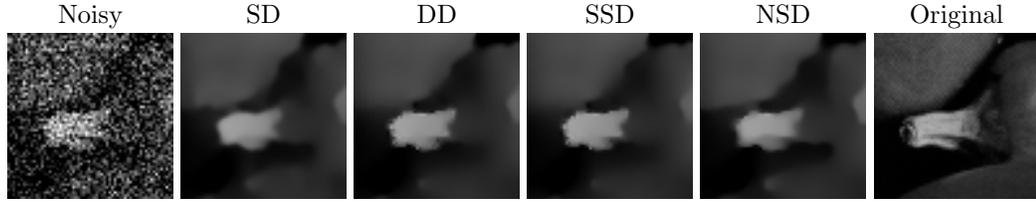
Figure 7.3: **Top and Bottom:** Zoom into House image with $\sigma_G = 40$ and 60 , respectively.

7.2 Experiments and Discussion

7.2.1 Datasets and Parameter Selection

We have corrupted Lena, Bridge, House, and Peppers images with AWGN and clipped-AWGN. We have chosen a radius of $\rho = 7$ for the disc-shaped neighbourhood. The number of stripes M is selected as 18 for NSD. In SSD, the number of stripes is equal to the number of pixels within the disc. We choose τ as 95% of the upper bound value which would satisfy the maximum-minimum principle of the image evolution.

The symbol ρ_t is used to represent half of the thickness of each stripe. We select this value such that every pixel in the disc is mapped to atleast one stripe. In the case of NSD, choosing $\rho_t = 1.21$ would satisfy this requirement. In SSD, $\rho_t = 0.5$ would suffice due to a design consequence: Every pixel pair $i - j$ within the disc already has its own corresponding stripe. Thus, we just optimise $\sigma, \lambda, k_{\max}$ w.r.t. the mean squared error (MSE).

Figure 7.4: Zoom into Peppers image ($\sigma_G = 60$).

7.2.2 Additive White Gaussian Noise

On the left side of Table 7.1, we can see the MSE values of denoised images for the AWGN layout. We can observe from these errors that NSD outperforms other methods. This can be attributed to the stripe-shaped neighbourhood. SSD also has the same shape for the structure element. However, it considers too many pixels in a stripe which is required for its symmetric modelling. This increases the chances of interaction with pixels having dissimilar greyvalues. Thus, its performance is inferior than NSD. Figure 7.2 shows a few denoised images of this evaluation. At first glance we do not observe too many differences between the methods. However, the zoomed images in Figures 7.3 and 7.4 confirm that in contrast to other methods the edge preservation capability of NSD is superior. In Table 7.2, we present the parameteric values that were used for obtaining the results in Table 7.1.

7.2.3 Clipped-Additive White Gaussian Noise

In our second type of experiments, the test images were first corrupted with AWGN and the resulting dynamic range was then clipped to [0-255] range. On the right side of Table 7.1, we can see the MSE values corresponding to clipped noise elimination. We can see that the performance of NSD and SD has deteriorated after clipping the noise. This can be attributed to the information lost due to clipping: Imagine that there were two pixels with greyvalues 220 and 240, in the original image. After adding noise, let us assume that their greyvalue has become larger than 255. Since we clip the noise, both the pixels now have the same value of 255. This depicts the lost information. Now, it seems plausible that the structure elements which use a smaller number of pixels will be more robust to information lost due to clipping. This can be attributed to the lesser probability of interaction between pixels from different regions. In contrast to NSD, this is one of the

reasons why SD has better performance in such a layout.

However, we must remember that SD also employs a slightly robust version of Gaussian smoothing (Chapter 6), in order to compensate for the asymmetric shape of a sector. Thus, in order to have a fair comparison between SD and NSD, we used a modified version NSD-M, which employs the same version of Gaussian smoothing as SD. One can see that NSD-M performs better than NSD for clipped noise. However, it still does not reach the performance of SD. This experiment signifies that fact that the filters benefit from having less pixels in the structure element for a clipped noise scenario.

We have performed our experiments on a GPU - NVIDIA GeForce GTX 970 graphics card - using C++ and CUDA. They show that each iteration of DD, SD, SSD and NSD consume 8, 22.56, 30.39 and 140.5 milliseconds, respectively for a 256×256 sized image. The stripe restrictive Gaussian smoothing is the reason behind the higher computational times of SSD and NSD. The former is faster than the latter, as we can exploit the symmetric nature of Gaussian smoothing of SSD while implementing it. However, this comes at a trade-off for gain in edge preservation capability of NSD, as we saw in the results. This is a perfect example of the strengths and weaknesses of symmetric and non-symmetric processes.

7.3 Conclusions

As already mentioned in Chapter 3, using a similar kind of assumptions could lead to several correspondences and equivalences between image denoising filters [120, 187–190, 193, 194]. This leads to indistinguishable performances. In this work, we instead evaluated several fixed filter shapes. As a first consequence, we will never be able to find an equivalence between them. Secondly, such a study has also led to practical progress in terms of both quality and speed.

Our evaluation can be considered as a solid foundation for designing patch matching-based extensions of stripe and sector-like filter shapes. This might be fruitful since the stripe- and sector-based filtering techniques are best suited for AWGN and clipped-AWGN models, respectively. They outperform the commonly used disc shape.

7.3. Conclusions

| SSD | | | | DD | | | NSD | | | SD | | |
|------|----------|-----------|------------|----------|-----------|------------|----------|-----------|------------|----------|-----------|------------|
| Data | σ | λ | k_{\max} |
| L40 | 1.3 | 4.0 | 7 | 1.1 | 5.6 | 4 | 1.1 | 4.3 | 9 | 0.6 | 3.4 | 7 |
| L60 | 1.4 | 6.7 | 5 | 1.3 | 5.5 | 6 | 1.4 | 3.6 | 15 | 0.5 | 3.2 | 12 |
| L80 | 1.5 | 7.6 | 6 | 1.9 | 3.6 | 12 | 1.7 | 3.3 | 19 | 0.5 | 3.2 | 18 |
| L100 | 2.0 | 6.0 | 9 | 2.6 | 2.6 | 23 | 1.9 | 3.4 | 21 | 0.3 | 3.7 | 19 |
| B40 | 1.3 | 12.8 | 2 | 1.1 | 7.2 | 2 | 1.2 | 8.6 | 2 | 0.2 | 6.3 | 3 |
| B60 | 1.4 | 10.7 | 2 | 1.3 | 7.1 | 3 | 1.3 | 9.9 | 3 | 0.2 | 6.3 | 5 |
| B80 | 1.5 | 10.4 | 3 | 1.4 | 7.6 | 4 | 1.4 | 10.5 | 4 | 0.3 | 6.5 | 7 |
| B100 | 1.9 | 10.7 | 3 | 1.7 | 8.0 | 4 | 1.5 | 9.6 | 6 | 0.2 | 6.7 | 9 |
| H40 | 1.4 | 3.3 | 11 | 1.1 | 4.1 | 8 | 1.1 | 3.0 | 18 | 0.7 | 3.0 | 8 |
| H60 | 1.6 | 3.7 | 14 | 1.1 | 5.6 | 8 | 1.2 | 3.4 | 22 | 0.6 | 3.1 | 13 |
| H80 | 1.8 | 3.3 | 26 | 1.3 | 5.7 | 9 | 1.3 | 4.2 | 21 | 0.5 | 2.6 | 20 |
| H100 | 2.0 | 3.2 | 30 | 1.6 | 4.8 | 12 | 1.4 | 3.5 | 33 | 0.5 | 2.7 | 26 |
| P40 | 1.3 | 5.0 | 5 | 1.1 | 4.5 | 6 | 1.2 | 3.6 | 11 | 0.6 | 3.0 | 9 |
| P60 | 1.4 | 6.1 | 6 | 1.3 | 5.5 | 6 | 1.4 | 3.4 | 17 | 0.5 | 3.0 | 14 |
| P80 | 1.9 | 4.5 | 11 | 1.7 | 4.6 | 9 | 1.6 | 3.4 | 21 | 0.5 | 3.0 | 20 |
| P100 | 2.1 | 4.8 | 12 | 2.0 | 4.2 | 12 | 1.7 | 3.5 | 25 | 0.3 | 3.7 | 20 |

| NSD | | | | NSD-M | | | SD | | |
|------|----------|-----------|------------|----------|-----------|------------|----------|-----------|------------|
| Data | σ | λ | k_{\max} | σ | λ | k_{\max} | σ | λ | k_{\max} |
| L40 | 1.1 | 4.4 | 8 | 0.8 | 5.3 | 7 | 0.6 | 3.3 | 7 |
| L60 | 1.3 | 3.6 | 14 | 0.7 | 5.5 | 13 | 0.6 | 3.2 | 12 |
| L80 | 1.3 | 3.9 | 17 | 0.6 | 7.8 | 12 | 0.6 | 3.2 | 17 |
| L100 | 1.2 | 5.1 | 16 | 0.6 | 8.2 | | 0.6 | 3.0 | 24 |
| B40 | 1.1 | 9.0 | 2 | 0.7 | 8.5 | 3 | 0.1 | 5.8 | 3 |
| B60 | 1.2 | 9.0 | 3 | 0.5 | 9.0 | 6 | 0.5 | 5.2 | 5 |
| B80 | 1.1 | 9.0 | 5 | 0.5 | 11.9 | 6 | 0.6 | 4.2 | 9 |
| B100 | 1.0 | 9.5 | 7 | 0.5 | 12.9 | 7 | 0.6 | 3.7 | 14 |
| H40 | 1.1 | 3.0 | 16 | 1.1 | 2.7 | 18 | 0.7 | 2.7 | 9 |
| H60 | 1.2 | 3.6 | 17 | 1.1 | 3.3 | 18 | 0.7 | 3.1 | 13 |
| H80 | 1.2 | 4.1 | 18 | 0.8 | 4.6 | 21 | 0.7 | 3.3 | 17 |
| H100 | 1.2 | 4.3 | 20 | 0.6 | 7.1 | 16 | 0.6 | 2.7 | 26 |
| P40 | 1.1 | 3.5 | 11 | 0.5 | 4.9 | 10 | 0.6 | 2.1 | 11 |
| P60 | 1.2 | 3.5 | 16 | 0.5 | 6.2 | 12 | 0.6 | 2.3 | 16 |
| P80 | 1.1 | 4.7 | 17 | 0.5 | 7.6 | 13 | 0.6 | 2.2 | 25 |
| P100 | 1.1 | 5.7 | 15 | 0.5 | 10.0 | 11 | 0.6 | 2.1 | 34 |

Table 7.2: Parametric values for AWGN on top and clipped noise at bottom. Abbreviations as in Table 7.1.

Part II

Multi-frame Image Denoising

Chapter 8

Multi-frame Extensions of Patch-based Filters

In Chapters 4-7, we dealt with the single-frame noise elimination scenario. Therein, we tried to produce a noise-free image from a single raw frame. BM3D [15, 319] and NLB [14, 290] are among the best performing filters in this layout. Both of them are non-local patch-based methods which utilise the similar information available at distant regions in the image. More precisely, they filter a 3D group of similar patches. BM3D in particular is a quasi-standard for modern denoising algorithms. It is used as a benchmark in articles that involve both neural network-based techniques [295] and traditional approaches [14].

Multi-frame filters [24–26, 74, 75, 339–352], on the other hand, utilise information from multiple frames of the same scene to compute the final denoised image. In this chapter, we concentrate on the fundamental problem of finding general approaches that can optimally extend single-frame patch-based methods such as NLB and BM3D to the multi-frame scenario.

There already exist two types of extensions [24–26, 347, 348] for BM3D and NLB. Methods from the first category search for similar 2D patches from all the available frames. However, they use just one reference frame for filtering purposes, thus making limited use of the available information [347, 348]. Extensions from the other category take privilege of having more data in 3D spatio-temporal patches [24–26]. Nevertheless, techniques which utilise 2D patches on multiple reference frames and those which perform separable spatio-temporal filtering have not been studied. The latter can reduce unde-

sirable interactions between regions of dissimilar greyvalues. Furthermore, a careful and systematic evaluation of these extensions is also missing.

Our Goals. The following are our three main objectives:

1. Model an approach that filters 2D patches in all the available frames.
2. Design a technique which employs separable spatio-temporal filtering.
3. Conduct a comprehensive evaluation of these extensions.

Our Contribution. In order to achieve the the first among the above objectives, we employed the 2D patch similarity approach of [347, 348] but using every frame as reference. This ensures that we make use of the complete available information.

Secondly, we introduce two extensions which benefit from separately filtering the different types of data in temporal and spatial dimensions. The first one performs a simple temporal averaging followed by a single-frame spatial filtering, while the other reverses this order.

Finally, in order to conduct a systematic and comprehensive evaluation of all the five extensions, we perform several tasks: We consider both registered and non-registered data. In the former scenario, we consider AWGN, Poissonian as well as mixture noise categories. For the latter scenario, we just evaluate AWGN filters. Here, we also utilise robust optical flow methods for dealing with inter-frame motion. Such a study is interesting as the utilisation of motion compensation was avoided in a state-of-the-art method [26] for circumventing motion estimation errors. We also pay special attention to parameter optimisation of the optical flow approaches. Such an analysis provides valuable additional insights into the importance of well optimised motion estimation in multi-frame denoising.

Furthermore, we provide the first comprehensive examination of general strategies how to extend single-frame filters to multi-frame ones. We apply all our extensions to both BM3D, NLB and our non-linear filtering on fast patch reorderings (NFPR) approach from Chapter 5. This provides solid evidence such that we can apply our extensions to many other patch-based methods as well. Our evaluation also includes very high noise levels. Such large amplitudes of noise, which are consistently ignored in the literature, are very relevant for microscopic and medical imaging applications. For a comparative analysis, we include three state-of-the-art extensions [24, 26, 353].

The neural network-based approach presented in [353, 354] is one among the many learning-based filtering strategies [355, 356] adopted nowadays.

Chapter Structure. In Section 8.1, we first review the central ideas behind the design of NLB and BM3D filters. We then present the five multi-frame extensions under evaluation along with optical flow methods. In Section 8.2, we showcase the results of several denoising experiments along with detailed explanations behind the observed rankings. Finally, in Section 8.3 we conclude the chapter with a summary and an outlook.

8.1 Modelling and Theory

8.1.1 Filters for Single-frame Image Datasets

NLB [14, 290] and BM3D [15, 319] are non-local patch-based denoising methods which consider similar information from distant regions in the image. Both single-frame filters are two step approaches. Furthermore, both steps are split into three sub-steps each, namely grouping, collaborative filtering and aggregation.

Grouping. In order to exploit the advantage of having more information, for every noisy reference patch considered, one forms a 3D group of similar patches using Euclidean distance.

Collaborative Filtering. The term "collaborative" has a literal meaning here: Each patch in a group collaborates with the rest of them for simultaneous and efficient filtering. In NLB, one uses Bayesian filtering (in both main steps) to denoise the 3D groups. In BM3D, a hard thresholding (first main step) and Wiener filtering (second main step) are employed.

Aggregation. In order to derive the final denoised image, one computes a weighted averaging of the several denoised versions of every pixel.

8.1.2 Multi-frame Extensions of Single-frame Filters

In this section, we describe five multi-frame extensions for the above mentioned single-frame filters in detail. For a better comprehension, we arrange

all five of them in an increasing order of design complexity.

In the multi-frame scenario, there exist slightly different types of data in the temporal and spatial dimensions. Thus, in order to filter them carefully the first two extensions employ separable spatio-temporal filtering.

Proposed Extension - Average then Filter (AF). First, we average all the registered frames. Then we employ a single-frame filter for removing the remaining noise in the averaged frame.

Proposed Extension - Filter then Average (FA). Here, we first denoise every registered frame by using a single-frame filter and then average the denoised frames.

The above two approaches differ from the methods in [350, 351] in the following fundamental aspect: Irrespective of the quality of registration, we utilise a temporal average *and* spatial filter strategy. This is different from a temporal average *or* spatially filter technique in [350, 351] that depends on the registration error.

While the first two extensions FA and AF perform a separable spatio-temporal filtering, the subsequent three employ combined filtering ideas. Let us discuss them in more detail now.

Existing Extension - Single Reference Frame Filtering (SF) [347–349]. Here, a single frame among all available ones is considered as the reference frame. One selects reference patches from just this frame. For every reference patch, a group of similar patches is formed using information from all the frames but not just one.

Proposed Extension - Multiple Reference Frame Filtering (MF). The fourth extension differs from SF in three different aspects. Firstly, in order to make complete use of the available information we consider all frames for reference patches. Secondly, we perform an aggregation of denoised pixels in such a way that after the first main step we have as many denoised frames as there are initial ones. This paves the way for the final difference: The above modified aggregation process enables us to consider reference patches from all frames in the second main step too. We cannot do this in the second main step using SF because it has considered reference patches from just one frame initially. We can thus formulate the final denoised image $\mathbf{u}^{\text{final}}$ which

is obtained from a combination of the registered noisy data \mathbf{f} and the initial denoised image $\mathbf{u}^{\text{initial}}$, as

$$\mathbf{u}^{\text{final}}(\mathbf{x}) = \frac{\sum_{\ell} \sum_{P_{\ell}} w_{P_{\ell}}^{\text{wien}} \sum_{Q \in \mathcal{P}(P_{\ell})} \chi_Q(\mathbf{x}) \mathbf{u}_{Q, P_{\ell}}^{\text{wien}}(\mathbf{x})}{\sum_{\ell} \sum_{P_{\ell}} w_{P_{\ell}}^{\text{wien}} \sum_{Q \in \mathcal{P}(P_{\ell})} \chi_Q(\mathbf{x})}. \quad (8.1)$$

Here, \mathbf{x} denotes the 2D position vector. We represent the set of most similar patches to the reference patch P_{ℓ} belonging to frame ℓ , using $\mathcal{P}(P_{\ell})$. For every patch Q in the set $\mathcal{P}(P_{\ell})$, we have $\chi_Q(\mathbf{x}) = 1$ if $\mathbf{x} \in Q$ and 0 otherwise. The symbol $\mathbf{u}_{Q, P_{\ell}}^{\text{wien}}(\mathbf{x})$ denotes the estimation of the value at pixel position \mathbf{x} , belonging to the patch Q . We derive this estimation through Wiener filtering (with coefficients $w_{P_{\ell}}^{\text{wien}}$) a combination of \mathbf{f} and $\mathbf{u}^{\text{initial}}$. In similar spirit to (8.1), we can formulate the NLB aggregation process:

$$\mathbf{u}^{\text{final}}(\mathbf{x}) = \frac{\sum_{\ell} \sum_{P_{\ell}} \sum_{Q \in \mathcal{P}(P_{\ell})} \chi_Q(\mathbf{x}) \mathbf{u}_{Q, P_{\ell}}^{\text{bayes}}(\mathbf{x})}{\sum_{\ell} \sum_{P_{\ell}} \sum_{Q \in \mathcal{P}(P_{\ell})} \chi_Q(\mathbf{x})}. \quad (8.2)$$

Here, the superscript **bayes** implies Bayesian filtering as detailed in [14, 290]. By restricting the total number of frames to one in (8.1) and (8.2), we obtain the original single-frame BM3D and NLB algorithms. This implies that MF encompasses the single-frame filters.

While grouping and filtering stages produce noise-free patches, aggregation computes the final denoised image from them. Employing 3D spatio-temporal patches gives an advantage of having more information at the patch denoising steps itself, even before employing the aggregation process. This exact idea is employed by the final extension.

Existing Extension - Combined Filtering (CF) [24–26]. One fixes 3D spatio-temporal patches and searches for similar volumes instead of patches. Then, a 4D filtering technique is employed, which removes noise using all the considered similar volumes. Such ideas are in accordance with the single-frame NLB and BM3D filters, where one considers a 2D similarity measure combined with a 3D denoising technique.

This finishes the brief discussion of all the multi-frame extensions we are going to deal with in this chapter. Table 8.1 serves as a look up table for these

| Method | Characteristics |
|--------|--|
| AF | 1. separable spatio-temporal filtering 2. average registered frames and then filter |
| FA | 1. separable spatio-temporal filtering 2. filter each registered frame and then average |
| SF | 1. combined spatio-temporal filtering 2. considers 2D reference patches from a single frame |
| MF | 1. combined spatio-temporal filtering 2. considers 2D reference patches from multiple frames |
| CF | 1. combined spatio-temporal filtering 2. considers 3D reference patches across frames |

Table 8.1: The main characteristics of the multi-frame extensions.

| | |
|------------------------|--|
| Input: | Noisy non-registered dataset \mathbf{f}^{nr} |
| Main Algorithm: | <ol style="list-style-type: none"> We employ an optical flow technique for obtaining registered data \mathbf{f} from \mathbf{f}^{nr}. Options for the optical flow methods include SOF-1, SOF-2 or SOF-3. We utilise a combination of single-frame denoising filters with their multi-frame extensions for producing the final denoised output $\mathbf{u}^{\text{final}}$ using registered data \mathbf{f}. Options for the single-frame filters are NLB or BM3D. They can be combined with extensions AF, FA, SF or MF. |
| Output: | Denoised data $\mathbf{u}^{\text{final}}$ |

Table 8.2: A general algorithm of the proposed denoising scheme.

five methods and presents the chief characteristics of each one of them. By combining the five multi-frame extensions and the two single-frame filters, we have ten filters in total. As an example, we will abbreviate one of these combined techniques as BM3D-MF, if it is a combination of single-frame BM3D with extension MF. Due to space constraints, within the experimental tables that are going to be presented in the upcoming subsections, we sometimes use shortforms for NLB-MF as NL-MF and BM3D-MF as BM-MF.

8.1.3 Optical Flow Methods

We perform experiments on both perfectly registered and non-registered datasets. In the latter scenario, we need to first register the images before applying the above multi-frame extensions. Thus, we have employed three robust discontinuity preserving optical flow methods [357–359]. These motion estimation techniques perform better than some classical strategies [360,361].

In all the three approaches, one minimises a similar energy functional to determine the motion vector $\mathbf{w} = (w_1, w_2, 1)^\top$ between frames f_1 and f_2 :

$$\begin{aligned}
E(\mathbf{w}) = \int_{\Omega} & \left(\Psi(|f_2(\mathbf{x} + \mathbf{w}) - f_1(\mathbf{x})|^2) + \right. \\
& \gamma (\Psi(|\nabla f_2(\mathbf{x} + \mathbf{w}) - \nabla f_1(\mathbf{x})|^2) + \\
& \left. \alpha (\Psi(\Phi(\nabla f_1(\mathbf{x})) \cdot (|\nabla w_1|^2 + |\nabla w_2|^2))) \right) d\mathbf{x}.
\end{aligned} \tag{8.3}$$

Here, $\mathbf{x} = (x, y, t)^T$ denotes the spatio-temporal location, Ω is the 2D image domain and ∇ is the spatio-temporal gradient.

The above energy penalises deviations in both gray values and gradients. One enables interactions in between neighboring pixels through the smoothness term. The parameters γ and α represent the gradient and smoothness term weights, respectively.

Moreover, applying $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ results in a robust convex energy functional with $\epsilon = 0.001$ ensuring strict convexity of Ψ . The smoothness function $\Phi(\nabla f_1, \lambda)$ with parameter λ specifies the regularisation strategy. The three optical flow methods that we use in this work differ in the choice of this particular function. We abbreviate these three techniques as SOF-1, -2 and -3 (SOF means sub-optimal flow).

In SOF-1, one employs a decreasing scalar function $\Phi(\nabla f_1, \lambda)$ to preserve image driven flow discontinuities. The second and third optical flow strategies try to avoid blob like artefacts: SOF-2 performs a minimum isotropic diffusion even when the gradient is very large. In SOF-3, one utilises an automatic selection strategy for λ . The same numerical procedure is adopted to compute the solution in all the three methods.

We use the aforementioned optical flow strategies for the first four extensions. The algorithm in Table 8.2 describes the main ideas behind the denoising framework of these approaches. The fifth method CF (author implementations available in the form of commonly known V-BM4D [24] and V-NLB [26]) uses its own motion compensation technique. The differences in motion estimation approaches should not be an issue as we also perform experiments on perfectly registered data. This concludes the modelling and theory part. Now, we move on to the experimental demonstrations.

| Data | α | γ | λ | Best Method | Data | α | γ | Data | α | γ |
|------|----------|----------|-----------|-------------|------|----------|----------|-------|----------|----------|
| G10 | 15 | 1.5 | 0.1 | SOF-2 | S10 | 25 | 1.5 | BH10 | 100 | 0.5 |
| G20 | 25 | 1.5 | 0.1 | SOF-2 | S20 | 75 | 2.5 | BH20 | 130 | 0.5 |
| G40 | 35 | 1.5 | 0.1 | SOF-2 | S40 | 95 | 1.5 | BH40 | 135 | 1.0 |
| G60 | 35 | 1.5 | 0.1 | SOF-2 | S60 | 110 | 0.5 | BH60 | 135 | 0.5 |
| G80 | 45 | 2.5 | 0.1 | SOF-2 | S80 | 85 | 0.5 | BH80 | 130 | 1.5 |
| G100 | 110 | 1.0 | - | SOF-3 | S100 | 95 | 0.5 | BH100 | 100 | 1.5 |
| G120 | 95 | 1.0 | - | SOF-3 | S120 | 90 | 0.5 | BH120 | 90 | 1.5 |

Table 8.3: Optical flow parameter values used for different datasets. G stands for Grove2, S for Shoe, and BH for Bird House. G80 denotes Grove2 dataset with noise standard deviation $\sigma_G = 80$. **Left:** The best among SOF-1, SOF-2 and SOF-3 motion compensation methods for Grove2 dataset. We have considered the tenth frame as the reference frame since ground truth flow information was available between frames 10 and 11. **Middle:** Shoe dataset with SOF-3 approach. **Right:** Bird House dataset with SOF-3 technique. We have utilised the fifth frame as the reference frame for Shoe and Bird House datasets and then employed frames 4-6 for optimizing the optical flow parameters. Also, we have used BM3D-MF and BM3D-FA as denoising filters while optimizing SOF parameters for these two datasets, respectively.

8.2 Experiments and Discussion

8.2.1 Datasets

For creating perfectly registered data, we have considered multiple AWGN realisations of the classical Lena, House, Peppers and Bridge images with fourteen datasets each. They are obtained by a combination of $\sigma_G = 10, 20, 40, 60, 80, 100, 120$ with five- and ten-frame datasets. In a similar spirit, we have also created non-registered data by corrupting the Grove2 [362], Shoe and Bird House [363] images with AWGN.

8.2.2 Parameter Selection

Optical Flow Parameters. For the Grove2 dataset, we have optimised the optical flow parameters with respect to the ground truth flow for all three methods. We then choose the best method to register every dataset. In Shoe and Bird House datasets we have optimised the SOF-3 parameters with respect to the final denoised image directly as the ground truth flow was not available. Table 8.3 shows more details.

Denoising Parameters. Various studies [14, 15, 290, 319, 364] have contributed in making the single-frame filters BM3D and NLB non-parametric,

while retaining the quality of the denoised images as much as possible. In a similar spirit to the above works, in this chapter we use better versions of two extensions AF and MF.

Firstly, at the time of application of the filter in the first extension AF, the noise distribution has already changed due to temporal averaging. Since we are using an AWGN model, we know that the standard deviation of noise is reduced by a factor \sqrt{L} for a dataset with L frames. We can improve the performance of type-AF extensions if we select the filter parameters corresponding to the new standard deviation.

The second improvement is to optimise the number of patches in a 3D group. The threshold parameter on Euclidean distance controls the total number of patches one employs for filtering purposes. Over several noise amplitude levels, our experience suggests that the gain in quality due to the presence of this parameter is relatively lot less when compared to the deterioration because of it. Hence, for simplicity reasons we refrain from using the threshold parameter in any of the first four BM3D extensions.

Moreover, in the multi-frame scenario we have more similar patches, in contrast to the single-frame layout. We thus check in the upcoming sections, whether we can improve the performance of BM3D-MF by doubling the number of patches in a 3D group. We label this particular parametric choice as BM3D-MFO, where O stands for an optimised version.

In case of the denoised results for perfectly registered noisy data using SF and CF techniques, we have always presented the best peak signal to noise ratio (PSNR) value among all frames. This ensures a fair comparison with the remaining three extensions.

For experiments on non-registered datasets, we have calculated the PSNR value by leaving out a border of fifty pixels on all sides of the reference frame at which different frames were registered. We do this in order to mitigate the ill-effects due to unavailable information at the borders of registered images. This also makes sense for several multi-frame imaging applications where we capture the region of interest in the centre of the frame.

8.2.3 Perfectly Registered Datasets

Tables 8.4 and 8.5 showcase the PSNR values of denoised images in the AWGN scenario. Figure 8.1 displays the corresponding visual results, after

| Data | NL-AF | NL-FA | NL-SF | NL-MF | NL-CF | BM-AF | BM-FA | BM-SF | BM-MF | BM-CF |
|------|--------------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|
| B10 | 36.54 | 35.00 | 33.22 | 36.11 | 35.75 | 36.53 | 34.16 | 32.54 | 34.91 | 35.95 |
| B20 | 31.83 | 29.79 | 28.31 | 30.55 | 31.01 | 31.88 | 28.94 | 28.20 | 29.79 | 31.04 |
| B40 | 27.92 | 25.87 | 24.84 | 26.17 | 26.98 | 27.90 | 25.65 | 24.89 | 26.21 | 26.50 |
| B60 | 26.04 | 23.98 | 23.11 | 23.86 | 25.07 | 25.99 | 24.27 | 23.50 | 24.73 | 24.50 |
| B80 | 24.87 | 22.86 | 22.24 | 22.82 | 23.69 | 24.83 | 23.51 | 22.75 | 23.88 | 23.45 |
| B100 | 24.00 | 22.34 | 21.76 | 22.35 | 22.87 | 24.08 | 22.93 | 22.17 | 23.25 | 22.75 |
| B120 | 23.30 | 21.99 | 21.29 | 21.98 | 22.24 | 23.48 | 22.49 | 21.69 | 22.75 | 22.20 |
| P10 | 38.64 | 37.28 | 36.04 | 37.23 | 37.81 | 38.72 | 36.90 | 36.04 | 37.11 | 37.47 |
| P20 | 35.80 | 35.02 | 33.77 | 35.09 | 35.29 | 35.88 | 34.87 | 33.97 | 35.24 | 34.81 |
| P40 | 33.23 | 32.49 | 31.10 | 32.66 | 33.00 | 33.54 | 32.61 | 31.43 | 33.09 | 32.16 |
| P60 | 31.99 | 30.75 | 29.19 | 30.88 | 31.28 | 32.10 | 31.13 | 29.79 | 31.63 | 30.41 |
| P80 | 30.71 | 29.25 | 28.00 | 29.51 | 29.33 | 30.84 | 30.34 | 28.53 | 30.41 | 29.12 |
| P100 | 29.76 | 28.32 | 26.97 | 28.65 | 28.28 | 29.86 | 29.00 | 27.52 | 29.43 | 28.07 |
| P120 | 28.84 | 27.60 | 26.04 | 27.78 | 27.37 | 28.99 | 28.14 | 26.69 | 28.61 | 27.16 |
| L10 | 38.84 | 37.80 | 36.38 | 37.89 | 37.96 | 38.87 | 37.54 | 36.41 | 37.83 | 37.87 |
| L20 | 35.87 | 34.91 | 33.43 | 34.93 | 35.38 | 36.00 | 34.75 | 33.75 | 35.27 | 34.83 |
| L40 | 32.95 | 31.85 | 30.44 | 32.07 | 32.67 | 33.23 | 31.90 | 30.71 | 32.53 | 31.60 |
| L60 | 31.38 | 30.00 | 28.47 | 30.03 | 30.71 | 31.56 | 30.30 | 29.06 | 30.93 | 29.65 |
| L80 | 30.11 | 28.41 | 27.22 | 28.35 | 28.78 | 30.15 | 29.04 | 27.83 | 29.59 | 28.28 |
| L100 | 29.13 | 27.34 | 26.17 | 27.50 | 27.75 | 29.24 | 28.14 | 26.81 | 28.66 | 27.21 |
| L120 | 28.24 | 26.77 | 25.41 | 26.96 | 26.87 | 28.33 | 27.38 | 26.08 | 27.82 | 26.33 |
| H10 | 39.92 | 38.13 | 36.60 | 37.78 | 39.28 | 40.12 | 38.15 | 37.23 | 38.75 | 38.79 |
| H20 | 36.36 | 35.20 | 34.02 | 35.32 | 36.33 | 36.83 | 35.30 | 34.45 | 35.83 | 35.17 |
| H40 | 33.22 | 32.58 | 31.22 | 33.23 | 33.46 | 33.92 | 32.77 | 31.64 | 33.42 | 32.06 |
| H60 | 31.97 | 30.37 | 28.81 | 31.30 | 31.51 | 32.49 | 30.96 | 29.77 | 31.77 | 29.94 |
| H80 | 30.52 | 28.26 | 27.23 | 29.20 | 29.49 | 30.96 | 29.41 | 28.30 | 30.16 | 28.38 |
| H100 | 29.38 | 26.79 | 26.03 | 27.77 | 28.38 | 29.85 | 28.43 | 27.20 | 29.07 | 27.14 |
| H120 | 28.46 | 25.66 | 25.08 | 26.71 | 27.35 | 29.16 | 27.47 | 26.29 | 28.28 | 26.10 |

Table 8.4: PSNR values after denoising five-frame datasets with various methods. Abbreviations: B80 - Bridge with $\sigma_G = 80$, P - Peppers, L - Lena, H - Bridge. Sizes: H - 256×256 , rest - 512×512 .

8.2. Experiments and Discussion

| Data | NL-AF | NL-FA | NL-SF | NL-MF | NL-CF | BM-AF | BM-FA | BM-SF | BM-MF | BM-CF |
|------|--------------|-------|---------|-------|-------|--------------|-------|-------|-------|-------|
| B10 | 39.08 | 35.84 | 33.6700 | 38.40 | 38.41 | 39.06 | 34.73 | 33.05 | 37.34 | 37.71 |
| B20 | 34.11 | 30.19 | 28.51 | 32.13 | 33.32 | 34.13 | 29.19 | 28.76 | 31.72 | 32.27 |
| B40 | 29.83 | 26.11 | 24.90 | 26.95 | 28.55 | 29.80 | 25.83 | 25.03 | 27.15 | 26.96 |
| B60 | 27.63 | 24.14 | 23.13 | 24.07 | 26.07 | 27.61 | 24.45 | 23.58 | 25.46 | 24.75 |
| B80 | 26.31 | 23.00 | 22.31 | 23.01 | 24.38 | 26.24 | 23.69 | 22.81 | 24.49 | 23.69 |
| B100 | 25.42 | 22.50 | 21.82 | 22.54 | 23.37 | 25.34 | 23.14 | 22.26 | 23.85 | 23.02 |
| B120 | 24.64 | 22.17 | 21.39 | 22.18 | 22.60 | 24.63 | 22.70 | 21.79 | 23.29 | 22.51 |
| P10 | 40.50 | 37.58 | 36.14 | 37.71 | 39.16 | 40.64 | 37.12 | 36.26 | 37.96 | 38.03 |
| P20 | 37.12 | 35.36 | 33.89 | 35.61 | 36.13 | 37.16 | 35.14 | 34.17 | 35.94 | 35.20 |
| P40 | 34.69 | 32.87 | 31.23 | 33.25 | 33.75 | 34.72 | 33.00 | 31.57 | 33.93 | 32.59 |
| P60 | 33.06 | 31.14 | 29.29 | 31.48 | 32.03 | 33.40 | 31.55 | 29.93 | 32.52 | 30.87 |
| P80 | 32.12 | 29.63 | 28.18 | 30.02 | 30.18 | 32.26 | 30.47 | 28.70 | 31.36 | 29.62 |
| P100 | 31.32 | 28.83 | 27.18 | 29.23 | 29.10 | 31.38 | 29.64 | 27.74 | 30.46 | 28.60 |
| P120 | 30.41 | 28.19 | 26.30 | 28.55 | 28.13 | 30.52 | 28.82 | 27.01 | 29.69 | 27.76 |
| L10 | 40.53 | 38.18 | 36.50 | 38.62 | 39.26 | 40.59 | 37.86 | 36.61 | 38.81 | 38.59 |
| L20 | 37.32 | 35.30 | 33.55 | 35.54 | 36.45 | 37.38 | 35.09 | 33.96 | 36.27 | 35.42 |
| L40 | 34.54 | 32.21 | 30.52 | 32.71 | 33.70 | 34.64 | 32.26 | 30.85 | 33.52 | 32.12 |
| L60 | 32.66 | 30.39 | 28.52 | 30.52 | 31.70 | 32.99 | 30.68 | 29.20 | 31.87 | 30.20 |
| L80 | 31.62 | 28.73 | 27.35 | 28.76 | 29.73 | 31.72 | 29.53 | 28.01 | 30.62 | 28.90 |
| L100 | 30.69 | 27.79 | 26.38 | 27.98 | 28.59 | 30.72 | 28.65 | 27.13 | 29.63 | 27.89 |
| L120 | 29.81 | 27.26 | 25.68 | 27.53 | 27.64 | 29.87 | 27.95 | 26.35 | 28.84 | 27.05 |
| H10 | 41.72 | 38.41 | 36.70 | 38.12 | 40.59 | 41.89 | 38.38 | 37.42 | 39.82 | 39.53 |
| H20 | 38.17 | 35.53 | 34.12 | 35.79 | 37.32 | 38.48 | 35.56 | 34.61 | 36.66 | 35.63 |
| H40 | 34.96 | 33.02 | 31.32 | 34.03 | 34.32 | 35.41 | 33.16 | 31.91 | 34.37 | 32.57 |
| H60 | 33.14 | 30.86 | 29.00 | 32.18 | 32.60 | 33.87 | 31.45 | 29.99 | 32.77 | 30.48 |
| H80 | 32.14 | 28.79 | 27.65 | 30.08 | 30.61 | 32.62 | 30.02 | 28.64 | 31.31 | 28.89 |
| H100 | 31.20 | 27.28 | 26.28 | 28.74 | 29.39 | 31.57 | 29.05 | 27.55 | 30.14 | 27.75 |
| H120 | 30.35 | 26.19 | 25.32 | 27.99 | 28.32 | 30.85 | 28.17 | 26.60 | 29.37 | 26.76 |

Table 8.5: PSNR values after denoising ten-frame datasets with various methods. Abbreviations as in Table 8.4.

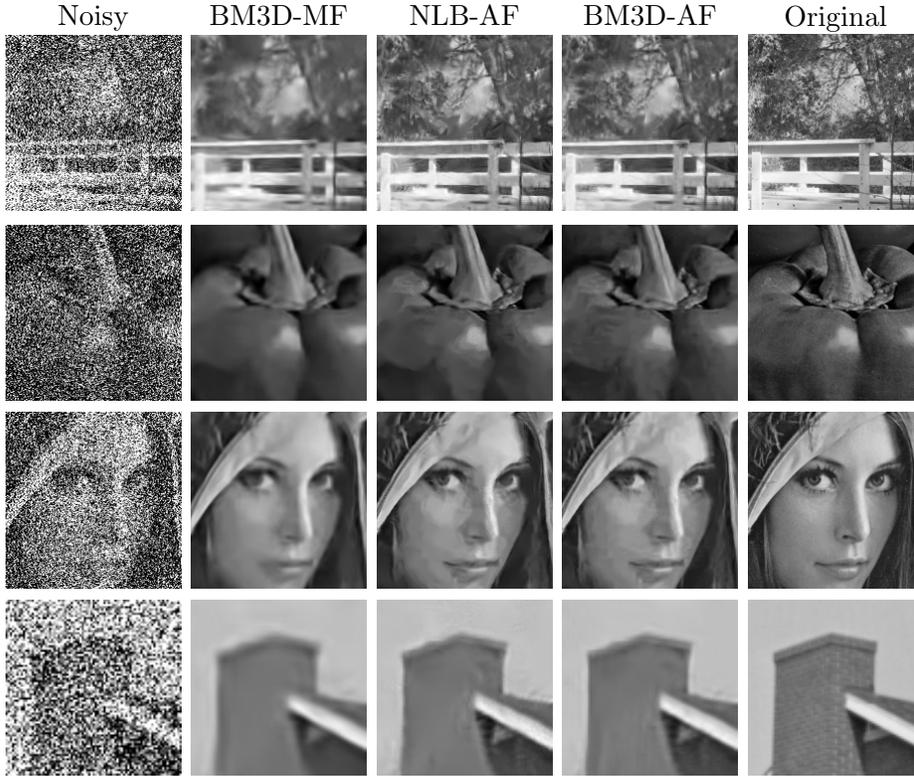


Figure 8.1: Denoised ten-frame datasets using the three best filters ($\sigma_G = 120$). **Top to Bottom:** Zoom into the Bridge, Peppers, Lena, and House images, respectively. The AF-type filters produce better results visually.

we have applied all ten methods. It is clear from these results that extensions of type-AF outperform all other techniques. The two best performing extensions in these experiments are MF and AF. We test these inter-frame connectivity strategies in Poissonian and mixture scenarios as well. Here, we evaluate NLB, BM3D as well as our novel NFPR filters. Results from Table 8.6 show that irrespective of the kind of noise model and the type of single-frame filter under consideration, type-AF extensions give the best results. Now, let us understand the reasons behind this occurrence, carefully.

In the category-FA extensions, we directly apply the single-frame filters on every frame. This is a sub-optimal solution because we do not have enough signal on each of the frames. Techniques belonging to type-SF do not make use of the complete available information as they just consider a

8.2. Experiments and Discussion

| POISSONIAN | | | | MIXTURE | | | |
|------------|--------------|---------|---------|---------|--------------|---------|---------|
| Data | NFPR-AF | BM3D-AF | BM3D-MF | Data | NFPR-AF | BM3D-AF | BM3D-MF |
| L0.9 | 25.37 | 25.16 | 24.52 | L0.9 | 25.17 | 24.86 | 24.24 |
| L0.7 | 24.55 | 24.42 | 23.61 | L0.7 | 24.53 | 24.18 | 23.40 |
| L0.5 | 22.63 | 22.30 | 21.84 | L0.5 | 22.70 | 22.40 | 22.01 |
| L0.1 | 20.62 | 19.98 | 19.42 | L0.1 | 20.75 | 20.21 | 19.77 |
| B0.9 | 21.25 | 21.19 | 20.76 | B0.9 | 21.28 | 21.14 | 20.74 |
| B0.7 | 21.01 | 20.90 | 20.36 | B0.7 | 20.93 | 20.77 | 20.28 |
| B0.5 | 19.87 | 19.76 | 19.45 | B0.5 | 19.84 | 19.70 | 19.47 |
| B0.1 | 18.43 | 18.22 | 17.80 | B0.1 | 18.55 | 18.30 | 18.06 |
| H0.9 | 24.61 | 24.27 | 23.45 | H0.9 | 24.57 | 24.26 | 23.42 |
| H0.7 | 23.74 | 23.45 | 22.30 | H0.7 | 23.89 | 23.55 | 22.37 |
| H0.5 | 21.23 | 20.87 | 20.20 | H0.5 | 21.24 | 20.74 | 20.21 |
| H0.1 | 19.23 | 18.64 | 18.18 | H0.1 | 19.27 | 18.46 | 18.06 |
| P0.9 | 25.48 | 25.38 | 24.59 | P0.9 | 25.36 | 25.19 | 24.40 |
| P0.7 | 24.86 | 24.75 | 23.84 | P0.7 | 24.83 | 24.57 | 23.64 |
| P0.5 | 22.75 | 22.66 | 21.98 | P0.5 | 22.66 | 22.49 | 21.90 |
| P0.1 | 20.45 | 20.23 | 19.57 | P0.1 | 20.70 | 20.45 | 20.04 |

Table 8.6: PSNR values for different noise types. Each dataset contains three noisy realisations of the original image. L0.5 in the case of Poissonian noise stands for Lena image with noise peak $\chi = 0.5$. In the case of mixture noise, we have added AWGN with $\sigma_G = (0.10 * \chi * 255)$, to the Poissonian component. Abbreviations are as in Table 8.4.

single reference frame.

In the MF and CF filters, we avoid the disadvantages of both FA and SF. However, they fall behind type-AF methods for two reasons: Firstly, we separate out temporal and spatial filtering in category-AF techniques. This is advantageous since we have noisy versions of the same original gray value in the temporal dimension for perfectly registered images. In the spatial dimensions we have noisy versions of approximately equal gray values in general. This outperforms simultaneous non-linear filtering of the MF and CF techniques, where we combine the information in all dimensions at once. Such a strategy proves to be inferior even though we use a non-linear filtering in the temporal dimension when compared to the linear temporal averaging of category-AF filters.

Interestingly, a similar result was observed in a single-frame scenario in the work of Ram et al. [18]. By adopting a simple linear filtering on a smoothly reordered set of pixels they could produce results almost equivalent to the sophisticated BM3D filtering. The reason behind such observations is that linear averaging of different noisy versions of the same pixel intensity does not create artefacts like a non-linear combination of dissimilar inten-

sities does. This is also the reason why averaging is preferred in electron microscopy (Chapter 11 of [51]). Moreover, the linear nature of temporal averaging helps in computing the new standard deviation of noise after temporal filtering through theoretical knowledge. The second reason why MF and CF types fall behind category-AF is the following: The latter extension computes the initial grouping on the less noisy averaged image. In all the other four categories we do this on the highly noisy initial images, which makes the grouping error-prone.

The overall better performance of type-AF filters does not mean we can immediately reject the next best MF and CF categories as we assumed perfect registration. We examine the non-registered layout in the upcoming section.

Furthermore, BM3D-AF is superior to NLB-AF (from Tables 8.4-8.5 and Figure 8.1) because BM3D is a better single-frame denoising method than NLB for gray value images. We infer that the usage of the discrete cosine transform and the bi-orthogonal spline wavelet transform in the two main steps of BM3D, respectively, leads to superior anisotropic modelling.

8.2.4 Non-registered Datasets

Tables 8.7, 8.8 and 8.9 display the PSNR values of the denoised images while Figure 8.2 showcases the visual results. It can be clearly seen that NLB-AF and BM3D-AF outperform other approaches several times. However, for low amplitude noise situations, NLB-CF is competitive with the category-AF extensions and even superior to them at certain occasions.

Let us explore the above results a bit further. For all the three datasets, we have performed experiments on two kinds of data: One with less number of frames and the other with more of them. In the latter case it is highly probable that there exists large motion between the reference frame and others which can lead to high errors in motion estimation. Hence, if a particular approach is able to produce better quality results for a high number of frames, this indicates that it is robust to motion estimation errors.

From Tables 8.7, 8.8 and 8.9, we can observe that CF is the only technique which does not even have a single instance where the PSNR value has decreased when more number of frames have been utilised. AF, MF, FA and category-SF filters could produce enough quality improvement for perfectly registered data. However, in the present non-registered layout we can find at

8.2. Experiments and Discussion

| Data | NL-AF | NL-FA | NL-SF | NL-MF | NL-CF | BM-AF | BM-FA | BM-SF | BM-MF | BM-CF |
|------|--------------|-------|-------|-------|--------------|--------------|-------|-------|-------|-------|
| G10 | 33.10 | 31.80 | 32.23 | 32.16 | 34.14 | 32.89 | 31.50 | 31.93 | 31.80 | 33.22 |
| G20 | 30.24 | 28.62 | 28.27 | 28.75 | 30.58 | 30.09 | 28.20 | 28.14 | 28.64 | 29.74 |
| G40 | 27.26 | 25.02 | 24.37 | 24.72 | 27.06 | 27.03 | 25.52 | 25.12 | 25.82 | 26.15 |
| G60 | 25.32 | 23.75 | 23.20 | 23.60 | 25.22 | 25.32 | 24.36 | 23.82 | 24.54 | 24.42 |
| G80 | 24.05 | 23.07 | 22.66 | 23.06 | 23.93 | 24.39 | 23.68 | 23.17 | 23.85 | 23.45 |
| G100 | 23.21 | 22.65 | 22.15 | 22.65 | 23.17 | 23.60 | 23.13 | 22.52 | 23.27 | 22.79 |
| G120 | 22.76 | 22.41 | 21.81 | 22.41 | 22.58 | 23.10 | 22.76 | 22.15 | 22.87 | 22.28 |
| G10 | 33.21 | 31.39 | 32.41 | 32.40 | 35.46 | 33.04 | 31.11 | 32.22 | 32.17 | 33.29 |
| G20 | 30.83 | 28.44 | 28.40 | 29.27 | 31.90 | 30.74 | 28.04 | 28.60 | 29.45 | 29.88 |
| G40 | 27.97 | 24.85 | 24.39 | 24.77 | 28.23 | 27.83 | 25.46 | 25.24 | 26.31 | 26.22 |
| G60 | 26.18 | 23.61 | 23.22 | 23.57 | 26.10 | 26.04 | 24.34 | 23.87 | 24.93 | 24.48 |
| G80 | 24.97 | 23.01 | 22.64 | 23.07 | 24.61 | 24.97 | 23.70 | 23.20 | 24.13 | 23.51 |
| G100 | 23.99 | 22.73 | 22.19 | 22.76 | 23.71 | 24.12 | 23.24 | 22.56 | 23.57 | 22.88 |
| G120 | 23.25 | 22.51 | 21.87 | 22.51 | 22.76 | 23.48 | 22.89 | 22.18 | 23.12 | 22.40 |

Table 8.7: PSNR values of denoised Grove2 images after using a combination of denoising methods and optical flow. **Top:** Four-frame datasets (frames 9-12). **Bottom:** Eight-frame datasets (frames 7-14). Frame size: 640×480 .

| Data | NL-AF | NL-FA | NL-SF | NL-MF | NL-CF | BM-AF | BM-FA | BM-SF | BM-MF | BM-MFO | BM-CF |
|------|-------|-------|-------|-------|--------------|--------------|-------|-------|-------|--------|-------|
| S10 | 37.49 | 36.34 | 35.94 | 36.38 | 37.89 | 37.67 | 36.84 | 36.51 | 36.98 | 36.79 | 37.38 |
| S20 | 34.63 | 33.32 | 32.63 | 33.32 | 35.02 | 35.02 | 34.10 | 33.39 | 34.35 | 34.16 | 34.28 |
| S40 | 31.71 | 30.17 | 29.51 | 30.27 | 32.08 | 32.20 | 31.37 | 30.46 | 31.74 | 31.63 | 31.16 |
| S60 | 30.39 | 28.58 | 27.77 | 28.77 | 30.26 | 30.90 | 29.84 | 28.80 | 30.25 | 30.18 | 29.34 |
| S80 | 29.07 | 27.46 | 26.64 | 27.66 | 28.52 | 29.65 | 28.71 | 27.59 | 29.05 | 29.06 | 28.01 |
| S100 | 28.27 | 26.80 | 25.88 | 27.08 | 27.52 | 28.88 | 27.88 | 26.65 | 28.09 | 28.26 | 26.95 |
| S120 | 27.61 | 26.35 | 25.26 | 26.64 | 26.70 | 28.14 | 27.18 | 25.90 | 27.35 | 27.51 | 26.08 |
| S10 | 37.55 | 35.95 | 35.90 | 36.28 | 38.09 | 37.66 | 36.48 | 36.56 | 37.03 | 36.83 | 37.42 |
| S20 | 35.19 | 33.22 | 32.67 | 33.39 | 35.26 | 35.45 | 34.02 | 33.51 | 34.74 | 34.58 | 34.44 |
| S40 | 32.47 | 30.19 | 29.61 | 30.45 | 32.38 | 32.87 | 31.56 | 30.63 | 32.12 | 32.29 | 31.49 |
| S60 | 31.28 | 28.65 | 27.84 | 29.05 | 30.61 | 31.79 | 30.10 | 28.97 | 30.93 | 30.97 | 29.77 |
| S80 | 30.13 | 27.56 | 26.75 | 28.02 | 29.01 | 30.62 | 29.06 | 27.75 | 29.78 | 29.90 | 28.53 |
| S100 | 29.34 | 27.01 | 25.98 | 27.49 | 28.02 | 29.93 | 28.27 | 26.85 | 28.87 | 29.11 | 27.57 |
| S120 | 28.68 | 26.66 | 25.37 | 27.10 | 27.19 | 29.27 | 27.65 | 26.04 | 28.01 | 28.35 | 26.77 |

Table 8.8: PSNR values of denoised Shoe images after using a combination of denoising methods and optical flow. **Top:** Five-frame datasets (frames 3-7). **Bottom:** Ten-frame datasets (frames 1-10). Frame size: 1280×720 . **Abbreviation:** BM-MFO uses twice the number of patches as in BM-MF.

| Data | NL-AF | NL-FA | NL-SF | NL-MF | NL-CF | BM-AF | BM-FA | BM-SF | BM-MF | BM-CF |
|-------|--------------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|
| BH10 | 36.63 | 35.00 | 34.86 | 35.61 | 35.03 | 36.63 | 34.99 | 34.84 | 35.37 | 35.07 |
| BH20 | 33.46 | 31.21 | 30.67 | 31.73 | 31.57 | 33.55 | 31.19 | 30.88 | 31.84 | 31.60 |
| BH40 | 30.07 | 27.02 | 26.43 | 27.11 | 28.22 | 30.11 | 27.84 | 27.36 | 28.48 | 27.79 |
| BH60 | 28.15 | 25.16 | 24.52 | 24.87 | 26.52 | 28.26 | 26.36 | 25.71 | 26.80 | 25.95 |
| BH80 | 26.71 | 24.50 | 24.05 | 24.38 | 25.34 | 26.95 | 25.53 | 24.88 | 25.89 | 24.90 |
| BH100 | 25.73 | 24.18 | 23.73 | 24.13 | 24.62 | 26.03 | 24.97 | 24.29 | 25.24 | 24.19 |
| BH120 | 24.96 | 23.96 | 23.43 | 23.94 | 24.08 | 25.27 | 24.55 | 23.82 | 24.76 | 23.66 |
| BH10 | 36.13 | 34.36 | 34.90 | 35.27 | 35.62 | 36.12 | 34.32 | 34.94 | 35.23 | 35.19 |
| BH20 | 33.79 | 30.94 | 30.74 | 31.97 | 31.97 | 33.84 | 30.81 | 31.25 | 32.39 | 33.29 |
| BH40 | 30.89 | 26.88 | 26.54 | 27.42 | 28.47 | 31.00 | 27.72 | 27.60 | 29.16 | 27.99 |
| BH60 | 29.18 | 25.06 | 24.55 | 24.92 | 26.66 | 29.24 | 26.37 | 25.79 | 27.33 | 26.11 |
| BH80 | 27.80 | 24.50 | 24.11 | 24.44 | 25.49 | 27.91 | 25.40 | 24.99 | 26.42 | 25.05 |
| BH100 | 26.71 | 24.28 | 23.80 | 24.25 | 24.75 | 26.84 | 24.70 | 24.39 | 25.76 | 24.39 |
| BH120 | 25.86 | 24.12 | 23.53 | 24.08 | 24.19 | 25.91 | 24.66 | 23.96 | 25.23 | 23.91 |

Table 8.9: PSNR values of denoised Bird House images after using a combination of denoising methods and optical flow. **Top:** Five-frame datasets (frames 3-7). **Bottom:** Ten-frame datasets (frames 1-10). Frame size: 1280×720 .

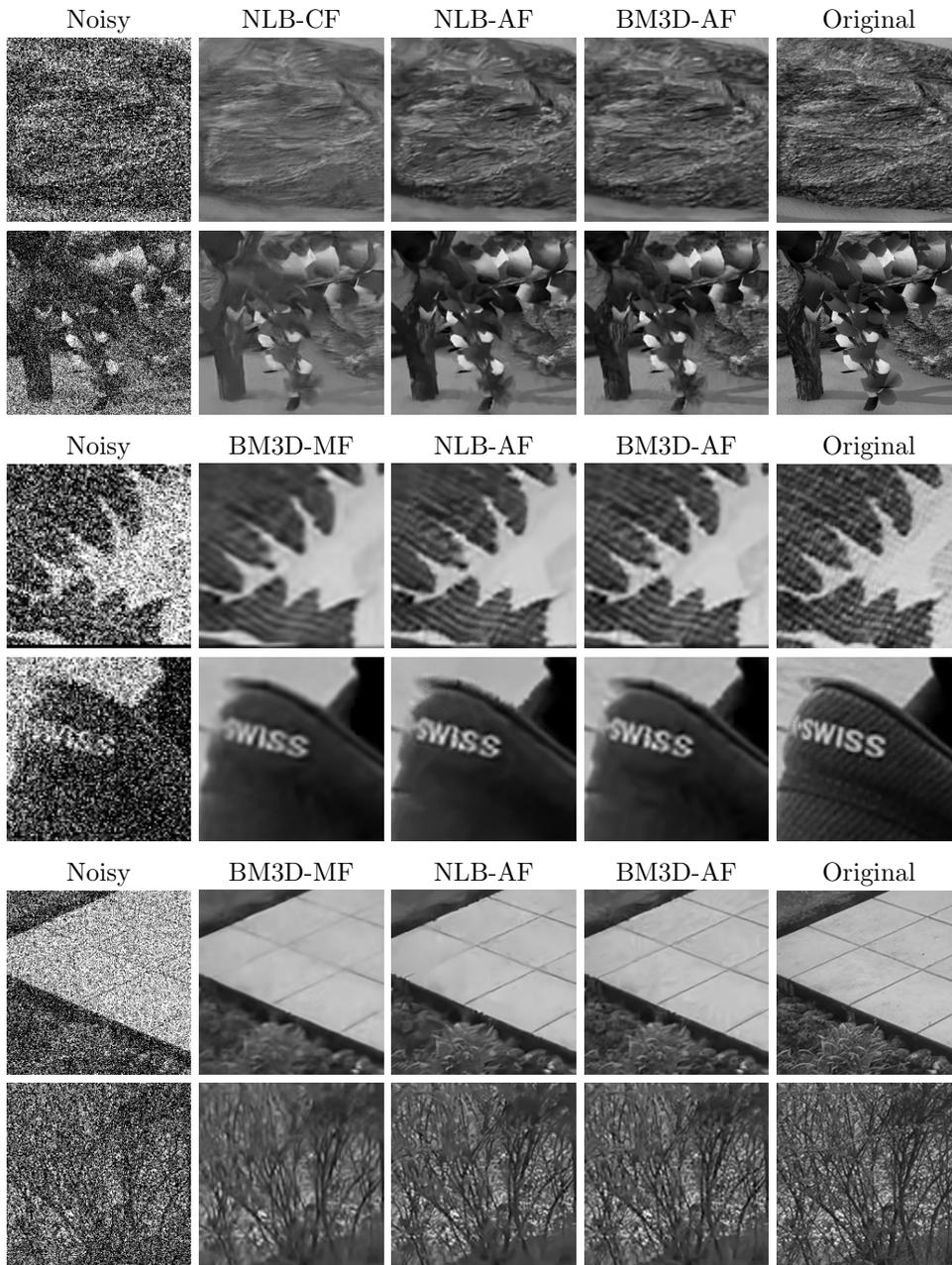


Figure 8.2: **Top to Bottom:** Different regions of Grove2, Shoe and Bird House datasets, respectively, using the three best extensions ($\sigma_G = 80$). We can see that the AF-type filters produce visually pleasing results and are also superior in preserving texture information.

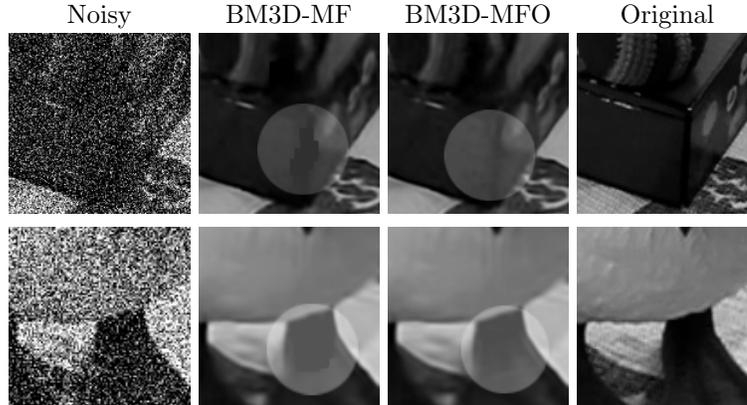


Figure 8.3: Different regions of Bird House dataset ($\sigma_G = 80$). One can observe that by using twice the number of patches than BM3D-MF, BM3D-MFO gets rid of the dark square shaped patches.

least one instance for each of these extensions where the quality has deteriorated with an increase in number of frames. The only explanation behind this is the robustness of category-CF extensions with respect to motion. However, at regions where the motion registration is correct, the performance of AF-type techniques is so high that they can outperform category-CF approaches. Nevertheless, optical flow methods will continue to improve in the future. Thus, the philosophy of our proposed category-AF extensions will benefit from these advancements.

The BM3D-MFO variant employs twice the number of patches than BM3D-MF. The increase in PSNR values from BM3D-MF to BM3D-MFO in Table 8.8 for high noise amplitudes and visual results in Figure 8.3 indicate the following: The black patches in darker regions of the image can be eliminated using BM3D-MFO. However, we must use the above strategy of increasing the number of patches only if we encounter black patches. Having too many of them in a 3D group gives rise to an undesirable blurring.

We can draw two conclusions for non-registered datasets: Optical flow methods are capable of helping type-AF filters generate superior results. Secondly, separable spatio-temporal filtering is very productive.

In recent years, learning-based denoising models have gained a lot of attention. In order to finish a comprehensive evaluation of our proposed technique, we compare its performance with a state-of-the-art neural network-

| Data | NL-AF | BM-AF | VNLNET | Data | NL-AF | BM-AF | VNLNET |
|------|--------------|--------------|--------|------|--------------|--------------|--------------|
| B10 | 39.08 | 39.06 | 38.35 | G10 | 33.21 | 33.04 | 34.64 |
| B20 | 34.11 | 34.13 | 33.36 | G20 | 30.83 | 30.74 | 31.02 |
| B40 | 29.83 | 29.80 | 29.00 | G40 | 27.97 | 27.83 | 27.73 |
| P10 | 40.50 | 40.64 | 38.64 | S10 | 37.55 | 37.66 | 38.39 |
| P20 | 37.12 | 37.16 | 35.88 | S20 | 34.63 | 35.45 | 35.36 |
| P40 | 34.69 | 34.72 | 33.30 | S40 | 31.71 | 32.87 | 32.37 |
| L10 | 40.53 | 40.59 | 39.36 | BH10 | 36.13 | 36.12 | 37.10 |
| L20 | 37.32 | 37.38 | 36.38 | BH20 | 33.79 | 33.84 | 33.78 |
| L40 | 34.54 | 34.64 | 33.25 | BH40 | 30.89 | 31.00 | 30.14 |
| H10 | 41.72 | 41.89 | 40.50 | | | | |
| H20 | 38.17 | 38.48 | 36.76 | | | | |
| H40 | 34.96 | 35.41 | 34.01 | | | | |

Table 8.10: PSNRs after denoising 10-frame datasets with various methods. **Left:** Perfectly registered datasets. **Right:** Non-registered layout. Abbreviations are as in Table 8.4. Moreover, G stands for Grove2, S for Shoe, and BH for Bird House.

based filter - video non-local network (VNLNET) [353,354]. Table 8.10 shows the PSNR values of this evaluation. The results show that our strategy outperforms VNLNET in the perfectly registered scenario and is competitive with it in the non-registered layout.

All the above results show that type-AF filters are among the best performing methods irrespective of whether there is any motion or not in the image dataset, what criteria have been used to optimise the optical flow, and what kind of optical flow technique has been employed. In the future, type-AF filters can be designed with occlusion handling [352], deflickering, sharpening [24], better optical flow and denoising strategies.

The AF-type frameworks are the fastest among all extensions as they employ separable spatio-temporal filtering. Since temporal averaging can be performed in real time, their net complexity is just a combination of the optical flow method and the 2D single-frame filter employed on the temporally averaged frame. All the experiments in this chapter were performed using a CPU (Intel(R) Core(TM) i7-6700 CPU @3.4 GHz using C++ and OpenMP) implementation. We also have a GPU (NVIDIA GeForce GTX 1070 graphics card using ANSI C and CUDA) version of BM3D-MF. We have already shown that BM3D-MF encompasses the original single-frame BM3D algorithm mathematically. Thus, the same GPU implementation can also be employed for BM3D-AF. With such an approach, we have observed that BM3D-AF is 7.25 times faster than BM3D-MF for a $4 \times 640 \times 480$ sized

dataset. It consumes just 1.82 seconds for the filtering process after motion compensation. Also, the CPU implementation of BM3D-AF is over 50 times faster than NLB-CF, which is a current state-of-the-art technique.

8.3 Conclusions

We have optimised the usage of NLB, BM3D and NFPR filters for the multi-frame scenario. We can conclude from the experiments that our proposed following sequential process gives the best results in most cases: They register the images with robust optical flow methods, temporally average the registered noisy images, and then apply the single-frame filters. Such an observation has surprisingly not been recognised for many years. This reaffirms the fact that sometimes simpler solutions are the most powerful ones. They can also be competitive with sophisticated neural network architectures. Furthermore, we achieve a significant improvement at the cost of zero additional parameters and far less computational time. Our technique preserves a large amount of detail in high amplitude noise scenarios. It is also effective for all the three commonly considered synthetic noise models. Thus, the category-AF extensions in combination with robust optical flow methods can be employed in practice for many patch-based multi-frame applications.

Manuscript details of research content from this Chapter:

- K. Bodduna and J. Weickert. Enhancing patch-based methods with inter-frame connectivity for denoising multi-frame images. In *Proc. IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 2414-2418.
- K. Bodduna and J. Weickert. Poisson Noise Removal Using Multi-Frame 3D Block Matching. In *Proc. 8th IEEE European Workshop on Visual Information Processing (EUVIP)*, Rome, Italy, Oct. 2019, pp. 58-63.
- K. Bodduna and J. Weickert. Removing Multi-frame Gaussian Noise by Combining Patch-based Filters with Optical Flow. *Journal of Electronic Imaging*, 30(3): 033031, June 2021.

Part III

Multi-frame Super-resolution

Chapter 9

Evaluating Super-resolution Observational Models and Regularisers

Generating high resolution (HR) images is one of the main objectives of photography. These images show more details of the scene which is crucial for many applications such as satellite [365], aerial [366], medical [367], mosaic [368], infrared [369], facial [370], text [367], compressed video recovery [371], number plate [372], and fingerprint [373] imaging. Instead of opting for expensive high precision optics, we can enhance the resolution in these images through usage of mathematical techniques. Super-resolution (SR) is one such technique designed to overcome the resolution limits of cameras. Generating a high resolution (HR) image from one single low resolution (LR) image is referred to as single-frame super-resolution [374–376]. In this work, we concentrate on multi-frame super-resolution, where information from multiple LR images is fused into a single HR image [375, 377–391].

In bio-medical and bio-physical applications we encounter images that possess a significant amount of noise. Multi-frame super-resolution in the presence of noise is thus practically relevant and also a very challenging research field. Algorithms that are designed to solve this problem compute derivative information on noisy data which showcases the ill-posed nature of the problem. In view of these algorithmic challenges, it is not surprising that very little efforts have been put into obtaining high resolution images from noisy low resolution data.

Deep learning-based methods are less suitable for bio-physical applications like electron microscopic imaging due to three main reasons: Firstly, there is very little ground truth data available. Secondly, the raw noise type is not well understood, unlike normal cameras. Finally, this imaging pipeline employs a huge amount of steps to obtain the final structure of the specimen under observation. After each one of these steps, the noise type changes. This makes deep-learning models that are trained for a particular kind of noise, sub-optimal.

In the present chapter we specifically aim at reconstructing a HR image from its LR versions. We consider both noise free LR images as well as those which have been corrupted by clipped-additive white Gaussian noise (AWGN).

Formalisation of the Problem. For multi-frame super-resolution, we want to find a HR image \mathbf{u} of resolution $N_H = H_1 \times H_2$ from N low resolution images $\{\mathbf{f}_i\}_{i=1}^N$ of resolution $N_L = L_1 \times L_2$. The low resolution images are assumed to be degraded versions of the real world HR scene. The seminal work on multi-frame super-resolution goes back to Tsai et al. [392]. The widely followed [382, 393–405] formulation of the relation between the high resolution scene and its LR realisations is [27]

$$\mathbf{f}_i = \mathbf{D}\mathbf{B}\mathbf{W}_i\mathbf{u} + \mathbf{e}_i. \quad (9.1)$$

In this observational model, we express the motion of the objects in the image using \mathbf{W}_i (size: $N_H \times N_H$). The operator \mathbf{B} (size: $N_H \times N_H$) denotes the blur due to the point spread function of the camera. We represent the downsampling of the HR scene by the camera detector system using \mathbf{D} (size: $N_L \times N_H$). The vector \mathbf{e}_i depicts the noise (error) acquired due to the imaging system. The operators \mathbf{B} and \mathbf{D} do not have an index i as we assume the same camera conditions for all images.

The standard model (9.1), however, has a disadvantage: The operator \mathbf{W}_i acts on the high-resolution scene \mathbf{u} . Hence, the model assumes that we have motion information at the high-resolution scale. In practice, we just have the downsampled and blurred images \mathbf{f}_i at our disposal. Motion at high-resolution must be approximated by upsampling the one computed on a lower resolution. Thus, the following question arises: Can one improve the practical performance of the SR approach by permuting the order of the operators? This is the first question that we address in this chapter.

Moreover, there is a second problem: For super-resolution of noisy data, the ideal observational model in (9.1) should be stabilised with a regulariser. In most cases, this is done by embedding it into the following quadratic energy minimisation framework:

$$E(\mathbf{u}) = \frac{1}{2} \sum_{i=1}^N |\mathbf{D}\mathbf{B}\mathbf{W}_i\mathbf{u} - \mathbf{f}_i|^2 + \frac{1}{2} \alpha_{\text{SR}} |\mathbf{A}\mathbf{u}|^2. \quad (9.2)$$

Here, \mathbf{A} is a discrete approximation of the continuous gradient operator, α_{SR} is the regularisation constant, and $|\cdot|$ denotes the Euclidean norm. The first term is the data term that encapsulates the observational model. The second one serves as smoothness term which eliminates noise. Minimising (9.2) by setting its gradient to zero gives

$$\sum_{i=1}^N \mathbf{W}_i^\top \mathbf{B}^\top \mathbf{D}^\top (\mathbf{D}\mathbf{B}\mathbf{W}_i\mathbf{u} - \mathbf{f}_i) - \alpha_{\text{SR}} \mathbf{A}_{\text{HD}}\mathbf{u} = \mathbf{0}, \quad (9.3)$$

where $\mathbf{A}_{\text{HD}} = \mathbf{A}^\top \mathbf{A}$ is the discrete approximation of the continuous Laplacian operator. In this chapter, we use a Gaussian blur kernel with standard deviation σ_{SR} , such that \mathbf{B}^\top equals \mathbf{B} . We denote the upsampling and down-sampling matrices by \mathbf{D}^\top and \mathbf{D} , respectively. The operator \mathbf{W}_i represents forward warping, while \mathbf{W}_i^\top encodes backward registration. The explicit gradient descent scheme with parameters τ (the time step size) and k_{max} (the number of iterations) to solve Equation (9.3) is given by

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau \left(\alpha_{\text{SR}} \mathbf{A}_{\text{HD}}\mathbf{u}^k - \sum_{i=1}^N \mathbf{W}_i^\top \mathbf{B}^\top \mathbf{D}^\top (\mathbf{D}\mathbf{B}\mathbf{W}_i\mathbf{u}^k - \mathbf{f}_i) \right). \quad (9.4)$$

In this evolution equation, \mathbf{A}_{HD} acts as the denoiser. However, such a noise elimination scheme uses a simple homogeneous diffusion process that also blurs important structures. As far as the usage of diffusion-based regularisers for super-resolution is concerned, only a few papers with simplistic models are available. Thus, it is highly desirable to introduce more advanced structure preserving regularisers. This is our second challenge.

Our Goal. Irrespective of whether the LR images are corrupted with noise or not, we want to find the order of operators in the observational model which gives least error. Moreover, we also want to evaluate the usage of

advanced structure preserving regularisers for effectively dealing with noise.

Our contribution. SR observational models that deviate from the imaging physics are mathematically still plausible. Variational techniques are known for the flexibility they provide while modelling various problems. Thus, to address the first challenge in this chapter, we exploit this liberty to evaluate six permutations arising from three operators in the standard observational model. Very few works [396, 406, 407] have focused their research in this particular direction. In contrast to these works, we complete an evaluation of all possible permutations of the standard model. Moreover, our experiments are not restricted to a specific type of motion. We also employ a better motion compensation technique than those utilised in these works. Such a comprehensive evaluation also enables us to make advancements from an algorithmic complexity viewpoint.

The second challenge in this chapter is to incorporate structure preserving regularisers. We start with replacing the homogeneous diffusion operator by the classical model of edge-enhancing anisotropic diffusion [132]. Although this model is around since a long time, its performance for super-resolution has not been examined so far. Moreover, we also make use of the sector diffusion (SD) model of Chapter 6, for superior structure preservation purposes. We deliberately avoid popular denoising methods such as 3D block matching [15] and non-local Bayes [14]: Most of these techniques rely heavily on a correct noise model, which renders them inferior for clipped noise, in particular with large amplitudes.

Chapter structure. The outline of this chapter is as follows: We propose various SR observational models in Section 9.1. Here, we also review an optical flow method as well as the EED- and SD-based SR reconstruction schemes. The SR experiments on noise-free and noisy images are presented in Sections 9.2 and 9.3, respectively. We conclude with a summary about multi-frame SR reconstruction as well as an outlook on future work, in Section 9.4.

9.1 Modelling and Theory

In this section, we first review the various possible permutations of the super-resolution observational model in Equation (9.1). Afterwards, we review a

| Model | Equation |
|-------|---|
| M1 | $\mathbf{DBW}_i\mathbf{u} + \mathbf{e}_i = \mathbf{f}_i$ |
| M2 | $\mathbf{DW}_i\mathbf{Bu} + \mathbf{e}_i = \mathbf{f}_i$ |
| M3 | $\mathbf{BDW}_i\mathbf{u} + \mathbf{e}_i = \mathbf{f}_i$ |
| M4 | $\mathbf{W}_i\mathbf{DBu} + \mathbf{e}_i = \mathbf{f}_i$ |
| M5 | $\mathbf{BW}_i\mathbf{Du} + \mathbf{e}_i = \mathbf{f}_i$ |
| M6 | $\mathbf{W}_i\mathbf{BDu} + \mathbf{e}_i = \mathbf{f}_i$ |
| M2.1 | $\mathbf{Bu} + \mathbf{e}_i = \mathbf{W}_i^\top \mathbf{D}^\top \mathbf{f}_i$ |

Table 9.1: The seven SR observational models.

classical optical flow method that has been employed for motion compensation purposes. This is followed by introducing the EED- and SD-based SR reconstruction schemes.

9.1.1 Super-resolution Observational Models

Table 9.1 shows the various permutations of the original observational model M1. While models M2-M6 depict the five other possible permutations, M2.1 represents a technique that is derived from M2. The motivation behind the modelling of M1-M6 is quality reasons. M2.1, on the other hand, is designed to exploit the precomputable nature of the term on the right hand side of the corresponding equation. Such a design is faster than any of the other models.

9.1.2 Optical Flow

The warping matrix in Equation 9.1 represents the displacements that the objects in the HR scene have undergone before being captured as a LR image by the camera. We make use of a simplified version of the popular optical flow method by Brox et al. [360] to estimate this matrix. In particular, we omit gradient constancy and just consider grey value constancy. By using a theory of multi-scale warping, this method has the ability of handling large displacements. Also, it does not assume a particular type of motion and hence it is a very good fit for estimating the warping matrix. We consider one of the LR images to be the reference one. The warping matrix is calculated

for every LR image with respect to this reference image. In the following we briefly sketch the main ideas behind the employed optical flow method.

Let $\mathbf{x} := (x, y, t)^T$ denote the position vector and $\mathbf{w} := (w_1, w_2, 1)^T$ the unknown displacement vector field. Penalising the deviations from the grey value constancy and enabling interaction between pixels can be modelled by the following continuous energy functional:

$$E(w_1, w_2) = \int_{\Omega} \left(\Psi(|f(\mathbf{x} + \mathbf{w}) - f(\mathbf{x})|^2) + \alpha_{\text{OF}} (\Psi(|\nabla w_1|^2) + |\nabla w_2|^2) \right) d\mathbf{x}, \quad (9.5)$$

where $\Omega \subset \mathbb{R}^3$ is the image domain, $f : \Omega \rightarrow \mathbb{R}$ denotes the image sequence and ∇ represents a spatio-temporal gradient. The function $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ mitigates the effects of outliers in a quadratic energy. A small positive constant ϵ ensures that the energy functional is strictly convex as well as differentiable. Moreover, α_{OF} is the regularisation parameter. The goal is to find a \mathbf{w} which minimises the above energy functional. The multi-scale warping approach is integrated in the Euler-Lagrange equations of the above energy functional. More specific details about the parameters and the optical flow method itself can be found in the paper by Brox et al. [360].

For best HR reconstruction results, one could use more complex optical flow models, anisotropic blurring kernels, robust data terms, and sophisticated interpolation strategies. However, the aim of our work is to evaluate the performance of the data and smoothness terms. To this very end, we keep things simple while selecting these factors. This is also why we omitted the gradient constancy assumption in the optical flow approach.

9.1.3 Edge-enhancing Diffusion

Edge-enhancing diffusion was proposed by Weickert [132] with the goal to enhance smoothing along edges while inhibiting it across them. To achieve this, one designs a diffusion tensor \mathbf{D} with eigenvectors \mathbf{v}_1 and \mathbf{v}_2 that are parallel and perpendicular to a Gaussian smoothed image gradient. This is followed by setting the eigenvalue corresponding to the eigenvector perpendicular to the gradient to one, indicating full flow. The eigenvalue corresponding to the eigenvector parallel to the gradient is determined by a diffusivity function. Using this idea, one can inhibit smoothing across edges. The following is the continuous mathematical formulation of the evolution of image u under

9.2. Experiments on Noise-free Datasets

EED:

$$\partial_t u = \operatorname{div}(\mathbf{D}(\nabla u_\sigma) \nabla u), \quad (9.6)$$

$$\mathbf{D}(\nabla u_\sigma) = g(|\nabla u_\sigma|^2) \cdot \mathbf{v}_1 \mathbf{v}_1^T + 1 \cdot \mathbf{v}_2 \mathbf{v}_2^T, \quad (9.7)$$

$$\mathbf{v}_1 \parallel \nabla u_\sigma, \quad |\mathbf{v}_1| = 1 \quad \text{and} \quad \mathbf{v}_2 \perp \nabla u_\sigma, \quad |\mathbf{v}_2| = 1. \quad (9.8)$$

Here, div is the 2D divergence operator and ∇u the spatial gradient. The Gaussian-smoothed image is u_σ . Computing the gradient on u_σ makes the diffusion process robust under the presence of noise. Both EED and SD evolution equations are initialised with the noisy image \mathbf{f} . Finally, the diffusivity function $g(x)$ is chosen as [133]

$$g(x) = 1 - \exp\left(\frac{-3.31488}{\left(\frac{x}{\lambda}\right)^8}\right). \quad (9.9)$$

Thus, by replacing the Laplacian \mathbf{A}_{HD} in (9.4) with the space discrete version \mathbf{A}_{EED} of the EED operator in (9.6), we arrive at the EED-based scheme for reconstructing the high resolution scene:

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau \left(\alpha_{\text{SR}}(\mathbf{A}_{\text{EED}}(\mathbf{u}^k)) - \sum_{i=1}^N \mathbf{W}_i^\top \mathbf{B}^\top \mathbf{D}^\top (\mathbf{D} \mathbf{B} \mathbf{W}_i \mathbf{u}^k - \mathbf{f}_L^i) \right). \quad (9.10)$$

Details regarding discretisation of the operator \mathbf{A}_{EED} can be found in [337]. In a similar fashion to EED, the SD-based SR reconstruction scheme can be written as

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau \left(\alpha_{\text{SR}}(\mathbf{A}_{\text{SD}}(\mathbf{u}^k)) - \sum_{i=1}^N \mathbf{W}_i^\top \mathbf{B}^\top \mathbf{D}^\top (\mathbf{D} \mathbf{B} \mathbf{W}_i \mathbf{u}^k - \mathbf{f}_L^i) \right). \quad (9.11)$$

We refer Chapter 6 for details regarding the sector diffusion model.

9.2 Experiments on Noise-free Datasets

9.2.1 Image Datasets

To evaluate the performance of the methods mentioned in Table 9.1, we have generated two LR test image sequences by corrupting two ground truth

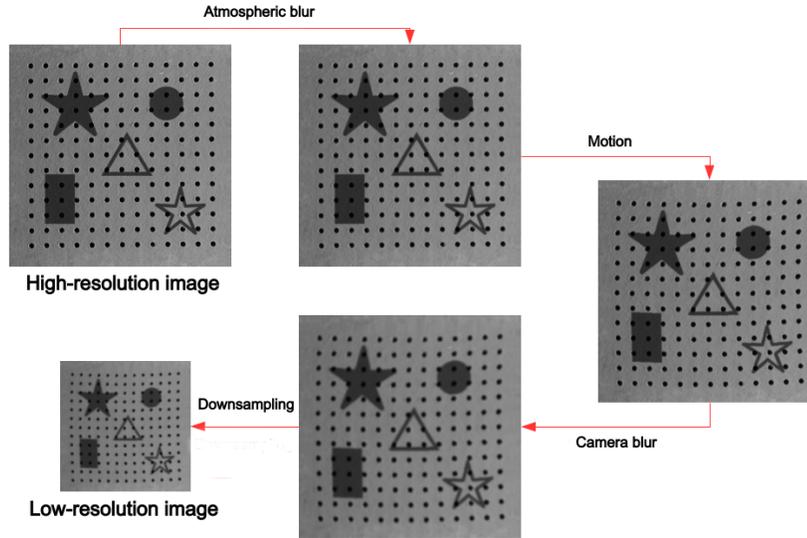


Figure 9.1: Generation of Shapes LR sequence.

HR images: The ‘Shapes’ image in Figure 9.1 has a texture background¹. The second image sequence is composed of self generated text (Figure 9.2). The ground truth of both image sequences is a 512×512 greyscale image. Figure 9.1 shows the generation of the Shapes image sequence from its ground truth. We have specifically simulated a deformation type of motion with subpixel displacements. It is well known that subpixel displacements are a requirement for super-resolution [408]. The ground truth image in ‘Text’ dataset undergoes a similar degradation process but without atmospheric blur. Both image sequences have 13 images each with a zoom factor of $z = 3$. We have used the last image as the reference one in both cases.

9.2.2 Parameter Selection

Optical Flow Parameters. The numerical parameters are the downsampling factor for the warping scheme η , inner fixed point iterations η_1 , outer fixed point iterations η_2 , and the successive over-relaxation parameter ω . We select $\eta_1 = 10$, $\eta_2 = 10$ and $\omega = 1.95$. More details regarding η , α_{OF} , and Gaussian pre-smoothing parameter σ_{OF} , will be given in later sections.

¹<http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>

SR Parameters. We have optimised the two model parameters α_{SR} (for regularisation) and σ_{SR} (standard deviation of the Gaussian blur kernel) with respect to the mean squared error (MSE). The numerical parameters are k_{max} and τ . A decay of the norm of the residue by a factor of 10^{-5} was used as the stopping criterion for k_{max} . We also utilise a fast explicit diffusion (FED) scheme [409] to accelerate the explicit gradient descent scheme. We observed that $\tau = 0.1$ was a stable time step size for all the proposed observational models through backtracking search. An area upsampled reference image was used as an initialisation.

9.2.3 Results and Discussion

To emphasise the importance of optical flow computation for super-resolution, we cover a large spectrum of optical flow qualities: We use both ground truth flow (GTF) as well as sub-optimal flow (SOF). The SOF is estimated using the above mentioned optical flow method.

Table 9.2 shows the MSE values of the reconstructed HR scenes while Figure 9.2 shows the images. We can conclude that the standard observational model gives the best results for the ground truth optical flow. For SOF, M2 gives the best results. Moreover, irrespective of whether we are using GTF or SOF, Models 1 and 2 are the two best ones among the six possible permutations. While the standard observational model is not a surprising winner, one needs to understand the better performing nature of M2 in the case of SOF. We will state the reason behind this in the next paragraph. Furthermore, for all the permutations we observe that the error using GTF is much smaller than that obtained using SOF. This reinforces the critical nature of motion estimation in SR reconstruction.

Zhang et al. [406] and Rockefort et al. [396] have discussed the application of both observational Models 1 and 2 but only when affine motion is assumed. Wang and Qi [407] also discuss Models 1 and 2 but without any constraints on the type of motion, similar to our work. They specifically show that SR reconstruction using M1 introduces a systematic error. In other words, if one looks closely at M1, we are warping the high resolution scene. However, the entries of this warping operator are calculated using blurred images. This induces a systematic error. It only vanishes when we use the ground truth optical flow. Thus, M1 outperforms M2 for GTF.

| # | Observational Model | MSE (GTF) | MSE (SOF) | σ_{SR} (GTF) | σ_{SR} (SOF) | α_{SR} (GTF) | α_{SR} (SOF) |
|----|--------------------------------------|-----------|-----------|----------------------------|----------------------------|----------------------------|----------------------------|
| M1 | $DBW_i \mathbf{u} = \mathbf{f}_L^i$ | 10.85 | 173.18 | 0.34 | 0.42 | 0.0002 | 0.003 |
| M2 | $DW_i B \mathbf{u} = \mathbf{f}_L^i$ | 23.42 | 162.86 | 0.35 | 0.45 | 0.0005 | 0.002 |
| M3 | $BDW_i \mathbf{u} = \mathbf{f}_L^i$ | 77.92 | 248.38 | 0.50 | 0.55 | 0.0009 | 0.005 |
| M4 | $W_i DB \mathbf{u} = \mathbf{f}_L^i$ | 250.35 | 294.00 | 0.31 | 0.33 | 0.001 | 0.0006 |
| M5 | $BW_i D \mathbf{u} = \mathbf{f}_L^i$ | 422.82 | 451.70 | 0.42 | 0.43 | 0.002 | 0.001 |
| M6 | $W_i BD \mathbf{u} = \mathbf{f}_L^i$ | 423.52 | 451.91 | 0.42 | 0.43 | 0.002 | 0.002 |

Table 9.2: Proposed SR observational models along with the MSE values of the obtained reconstructed SR image for Text sequence. Results for both ground truth (GT) and sub-optimal flow (SOF) are presented along with the selected parametric values.

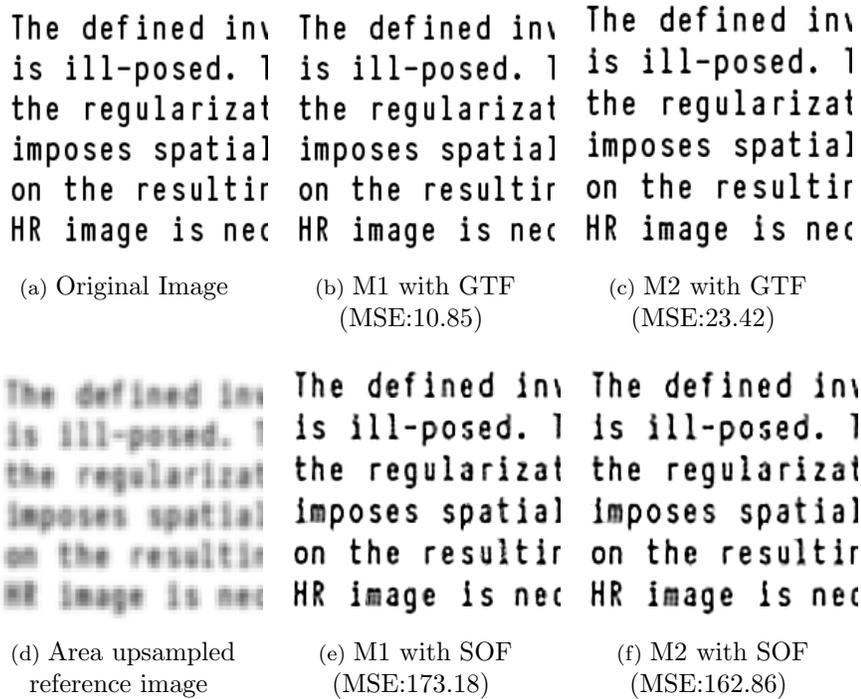


Figure 9.2: SR reconstructions of Text dataset along with MSE and flow.

On the other hand, in M2, we are applying a motion that is estimated using blurred images on a blurred HR scene. This relieves M2 from the systematic error of M1. However, M2 still has an operator commutability

9.2. Experiments on Noise-free Datasets

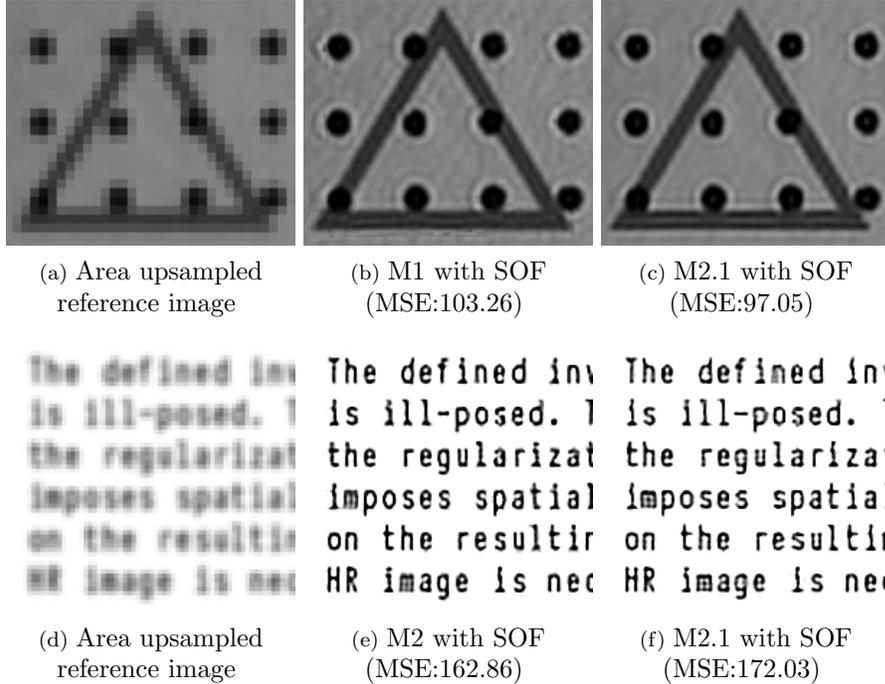


Figure 9.3: SR reconstructions of Shapes and Text datasets using M2.1. We select $\eta = 0.9, \alpha_{\text{OF}} = 15.0, \sigma_{\text{OF}} = 0.3$, as the SOF parameters for the Text sequence. Similarly, we use $\eta = 0.5, \alpha_{\text{OF}} = 8.0, \sigma_{\text{OF}} = 0.3$ as SOF parameters for the Shapes sequence.

error as it interchanges the warp and blur matrices. Such an error vanishes only when the motion is affine, which is not the case for our datasets.

This must not be confused with the fact that we are using a relatively simple motion compensation technique by dropping the gradient constancy assumption in the optical flow model. On the contrary, we have employed a more advanced motion estimation method than the one [410] adopted by Wang and Qi [407]. However good our optical flow might be, if the motion within the dataset is not affine, we are bound to have a commutability error. Which error among the systematic and commutability types is higher is totally dependent on the specific image dataset. This will finally decide the order in the ranking of Models 1 and 2. In our specific case while using SOF, since M2 outperforms M1, we conjecture that commutability error is lower than the systematic error.

Keeping in mind the above conclusions, when we retrospect the reported

works on super-resolution after [407] was published, it is surprising that in most of them the standard observational model has been used [382, 397, 399, 400, 402, 403, 405]. M2 has not been considered as a possible alternative. Thus, in the following, we perform a careful further analysis on this strategy.

9.2.4 More Efficient Model

Now we propose another mathematically plausible model derived from M2 and discuss what could be the advantages of using it. The following is the representation of what we denote as M2.1: $\mathbf{B}\mathbf{u} = \mathbf{W}_i^T \mathbf{D}^T \mathbf{f}_L^i$. It is clear that it is derived from M2 as the ordering of operators is the same. However, it is definitely different from M2 itself as warping and downsampling are interpolation operations. Interpolation in general is not an invertible operation.

The gradient descent of the energy which uses M2.1 in the data term, is given by

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau(\alpha_{\text{SR}} \mathbf{A}_{\text{HD}} \mathbf{u}^k - (N \mathbf{B}^T \mathbf{B} \mathbf{u}^k - \mathbf{C})), \quad (9.12)$$

where $\mathbf{C} = \sum_{i=1}^N \mathbf{B}^T \mathbf{W}_i^T \mathbf{D}^T \mathbf{f}_L^i$ can be precomputed. Such a precomputation is not possible with M2.

Figure 9.3 shows the reconstructed images using M2.1. The parameters $\sigma_{\text{SR}} = 0.64, \alpha_{\text{SR}} = 0.006$ were selected for the Shapes sequence and $\sigma_{\text{SR}} = 0.59, \alpha_{\text{SR}} = 0.002$ for the Text dataset. These settings were obtained after optimising the parameters w.r.t. MSE.

We can conclude from Figure 9.3 that the reconstructed HR images obtained using M2.1 are not far off from Models 1 and 2 in terms of image reconstruction quality. However, since we can precompute \mathbf{C} , the gradient descent of M2.1 is twenty-five times faster for the Shapes dataset and eighteen times faster for the Text sequence in contrast to Models 1 and 2, respectively. This can be a decisive advantage in time critical applications.

9.3 Experiments on Noisy Images

In the previous section we did not consider any kind of noise in the LR images. Here, we corrupt the images with clipped-AWGN. We first evaluate EED and SD in terms of their SR regularisation capability. Then we choose the best of the two as regulariser for data term evaluation.

9.3. Experiments on Noisy Images

| Dataset | σ_{OF} | α_{OF} |
|---------|----------------------|----------------------|
| Text1 | 2.6 | 13.3 |
| Text2 | 1.0 | 15.6 |
| Text3 | 2.3 | 6.3 |
| House1 | 3.8 | 13.5 |
| House2 | 1.2 | 17.0 |
| House3 | 2.7 | 16.5 |

Table 9.3: Parameter settings for optical flow calculation. We have two model parameters: α_{OF} (smoothness parameter) and σ_{OF} (Gaussian pre-smoothing). Numerical parameters are chosen as $\eta = 0.95$ (downsampling factor), $\eta_1 = 10$ (inner fixed point iterations), $\eta_2 = 10$ (outer fixed point iterations) and $\omega = 1.95$ (successive over-relaxation parameter).

9.3.1 Image Datasets.

We have considered two high-resolution scenes in the form of ‘Text’² and ‘House’ images. The ground truth HR images have been warped (randomly generated deformation motion), blurred (Gaussian blur with standard deviation 1.0), downsampled (with bilinear interpolation), and degraded by noise (clipped-AWGN with $\sigma_{\mathcal{G}}=40$).

9.3.2 Parameter Selection.

To account for a large spectrum of optical flow qualities, we have used both the ground truth flow as well as a simplified approach of Brox et al. [360] without gradient constancy assumption. The parameters for different datasets are shown in Table 9.3. We optimise these parameters just once, but not after every super-resolution iteration. For SR reconstruction, we optimise the parameters α_{SR} (smoothness), σ_{SR} (Gaussian blur), σ (Gaussian smoothing in EED and SD) and λ (contrast parameter in EED and SD) with respect to MSE. The grid size is 1. As time step we choose $\tau = 0.05$ for EED and $\tau = 0.012$ for SD, giving experimental stability and convergence to a plausible reconstruction. We initialise \mathbf{u} with a bilinearly upsampled image.

9.3.3 Smoothness Term Evaluation.

The SR reconstruction quality of the two regularisers is evaluated using equations (9.10) and (9.11). From Table 9.4 and Figures 9.4, 9.5 we observe that

²<https://pixabay.com/en/knowledge-book-library-glasses-1052014/>

| Dataset | EED | | | | | | SD | | | | | |
|---------|----------|----------------------|-----------|----------------------|------------------|--------|----------|----------------------|-----------|----------------------|------------------|---------------|
| | σ | σ_{SR} | λ | α_{SR} | k_{max} | MSE | σ | σ_{SR} | λ | α_{SR} | k_{max} | MSE |
| H1 | 0.6 | 0.8 | 11.0 | 118.0 | 37 | 110.45 | 0.9 | 1.0 | 2.0 | 1.6 | 17 | 83.12 |
| H2 | 0.7 | 0.5 | 12.0 | 120.0 | 9 | 162.64 | 0.8 | 0.7 | 1.7 | 5.3 | 17 | 133.35 |
| H2-S | 0.7 | 0.5 | 13.0 | 115.0 | 9 | 172.94 | 0.9 | 0.8 | 1.8 | 4.5 | 17 | 141.62 |
| H3 | 0.6 | 0.4 | 14.0 | 127.0 | 48 | 201.91 | 0.6 | 0.8 | 2.3 | 2.9 | 49 | 161.26 |
| T1 | 1.0 | 1.1 | 9.0 | 14.0 | 136 | 164.72 | 0.6 | 1.1 | 3.0 | 0.3 | 48 | 158.50 |
| T2 | 1.3 | 0.9 | 7.0 | 18.0 | 11 | 397.09 | 0.6 | 1.0 | 2.7 | 0.6 | 34 | 378.72 |
| T2-S | 1.3 | 1.0 | 7.0 | 18.0 | 14 | 510.80 | 0.6 | 1.1 | 2.9 | 0.5 | 32 | 499.60 |
| T3 | 1.2 | 0.4 | 7.0 | 14.0 | 13 | 674.65 | 0.6 | 0.6 | 2.3 | 0.6 | 49 | 657.82 |

Table 9.4: MSE values of SR reconstructed images including parameters used. T2 stands for Text2 dataset with ground truth optical flow, while T2-S was computed using sub-optimal calculated flow. Ground truth image size for Text: 512×512 . T1-T3 represent images downsized by factors 1, 2 and 3, respectively. Image size for House: 256×256 . H1-H3 represent images downsized by factors 1, 1.5 and 2, respectively. Every dataset has 30 images each, with the last of them being the reference frame for registration.

| Model | Left | | | | | | Right | | | | | |
|-------|----------|----------------------|-----------|----------------------|------------------|---------------|----------|----------------------|-----------|----------------------|------------------|---------------|
| | σ | σ_{SR} | λ | α_{SR} | k_{max} | MSE | σ | σ_{SR} | λ | α_{SR} | k_{max} | MSE |
| M1 | 0.6 | 1.0 | 2.7 | 0.6 | 34 | 378.72 | 0.6 | 1.1 | 2.9 | 0.5 | 32 | 499.60 |
| M2 | 0.6 | 1.0 | 2.7 | 0.6 | 34 | 382.63 | 0.6 | 1.1 | 3.4 | 0.4 | 33 | 502.78 |
| M3 | 0.6 | 0.6 | 2.9 | 1.2 | 20 | 381.75 | 0.3 | 0.8 | 3.8 | 1.0 | 21 | 500.50 |
| M4 | 0.6 | 0.8 | 2.8 | 0.6 | 35 | 392.66 | 0.6 | 0.9 | 3.0 | 0.5 | 32 | 511.09 |
| M5 | 0.6 | 0.5 | 2.7 | 0.7 | 33 | 391.91 | 0.4 | 0.6 | 4.6 | 0.3 | 43 | 513.29 |
| M6 | 0.3 | 0.5 | 4.1 | 0.4 | 60 | 403.50 | 0.4 | 0.6 | 4.6 | 0.3 | 48 | 518.04 |
| M2.1 | 0.6 | 1.5 | 3.3 | 0.2 | 55 | 394.85 | 0.6 | 1.6 | 3.5 | 0.2 | 56 | 523.34 |

Table 9.5: Data term evaluation. **Left:** Text2 with ground truth flow. **Right:** With sub-optimal flow.

SD outperforms EED consistently. This holds both for ground truth and sub-optimal optical flow, over all downsampling factors. The superior structural adaptivity of SD is the reason behind this observation.

9.3.4 Data Term Evaluation.

Since we have observed a superior performance of SD for regularisation purposes, we also use it in the smoothness term while evaluating the data term. Table 9.5 shows the MSE values of the reconstructed high resolution scene with all observational models from Table 9.1. For ground truth flow, the observational model M1 performs best. This is in accordance with our results

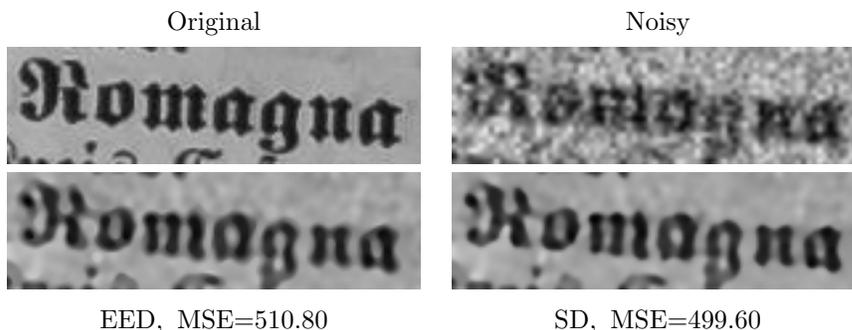


Figure 9.4: Zoom into SR reconstructions for the Text2 dataset with sub-optimal flow.

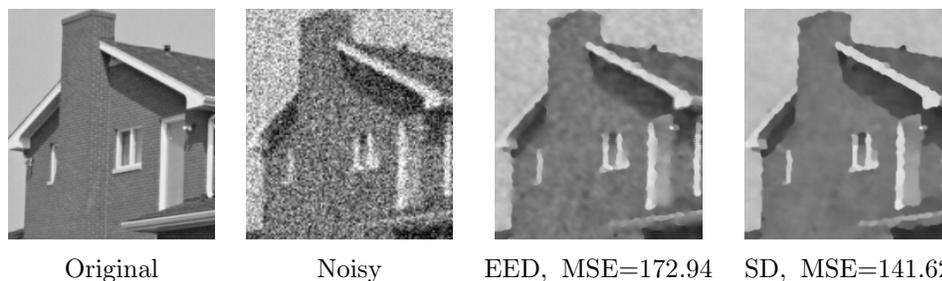


Figure 9.5: Zoom into SR reconstructions for the House2 dataset with sub-optimal flow.

in Section 9.2 for a noise-free scenario.

For SOF, M1 also outperforms M2. Interestingly, this is contradictory to the findings in a noise-free scenario, where M2 gave superior results. We thus conjecture that the commutability error is higher than the systematic error in a noisy scenario.

In Section 9.2, model M2.1 was much faster than M2 with only little loss in reconstruction quality. However, this model becomes irrelevant in the noisy scenario, as M1 outperforms M2. We also encounter a further quality loss when replacing M2 by M2.1.

9.4 Conclusions

Super-resolution requires to model three physical phenomena: Blur, warp, and downsample. In this chapter, we have performed the first systematic evaluation of the influence of the order of these three operators on the result of a variational super-resolution model. This has led to the surprising

result that it is not always the physically most plausible and most widely used model which performs best in a practical setting. Thus, it is worthwhile to consider also alternative models. Moreover, we saw that closely related models can lead to algorithms with strongly differing efficiency: By reformulating the blur-warp-downsample model we managed to come up with a novel model that was 18–25 times more efficient. These insights emphasise the fundamental importance of careful model design.

Our work also belongs to the scarce amount of literature that ventures to investigate super-resolution models in the practically relevant scenario of substantial amounts of clipped noise. In contrast to classical least squares approaches with homogeneous diffusion regularisation we have paid specific attention to structure preserving regularisers such as edge-enhancing anisotropic diffusion (EED). Interestingly, EED has not been used for super-resolution before, in spite of the fact that alternatives such as BM3D and NLB are less suited for data with highly corrupted clipped noise. More importantly, we have also evaluated the performance of the sector diffusion technique as a super-resolution regulariser. Its higher structural adaptivity enabled it to outperform EED.

Manuscript details of research content from this chapter:

- K. Bodduna and J. Weickert. Evaluating data terms for variational multi-frame super-resolution. In F. Lauze, Y. Dong, A. B. Dahl, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 10302 of *Lecture Notes in Computer Science*, pages 590-601, Springer, Cham, 2017.
- K. Bodduna, J. Weickert, and M. Cárdenas. Multi-frame Super-resolution from Noisy Data. In A. Elmoataz, J. Fadili, Y. Quéau, J. Rabin, L. Simon, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 12679 of *Lecture Notes in Computer Science*, pages 565-577, Springer, Cham, 2021.

Part IV

Structure-aware Image Enhancement

Chapter 10

Hough-based Evolutions for Structure Enhancement

In the introduction of this thesis, we mentioned two types of filters: Structure preserving methods (Chapters 4-9) and the less researched structure enhancing techniques. In this chapter, we study about a method belonging to the latter kind. We also specifically concentrate on cryo-electron microscopy (cryo-EM) data.

Enhancing oriented structures is a classical problem in image processing. Fingerprint images, artistic paintings, computer tomography scans, and clothing fibre images are specific areas where we encounter such a problem [13]. On a broader scale, this application is encountered in fluid dynamics, meteorology, forensic studies, computer vision, biomedical and biophysical image analysis.

In 2D, analysing and processing oriented structures has a long tradition. The structure tensor [411] and its equivalent concepts play a prominent role in this context. While early work by Kass and Witkin [412] as well as Rao and Schunck [413] apply it as a pure analysis tool, Weickert et al. [13, 28] use it to steer a so-called coherence-enhancing anisotropic diffusion (CED) process. This has triggered several follow up works that employ diffusion-based ideas to enhance oriented structures [144, 334, 414]. Mühlich et al. [415] have studied the presence of multiple orientations in a local neighbourhood. Stochastic models [416] have also been used for analysis of contour shapes in images. More recently, template matching-based on orientation scores [417] has been proposed for detecting combined orientation and blob patterns.

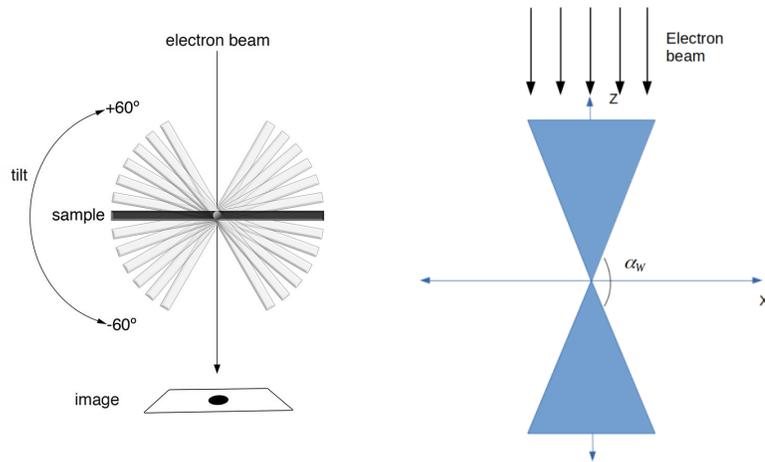


Figure 10.1: **Left:** Image acquisition in cryo-EM. Courtesy of Eikosi, 2015. http://commons.wikimedia.org/wiki/File:Electron_Tomography.tif. **Right:** The triangles represent the region where the data is smeared at every pixel in Cartesian space due the presence of the missing wedge.

In 3D, the first technique for coherence enhancement was proposed by Weickert et al. [28]. Related works to this anisotropic diffusion technique in 3D include papers by Krissian et al. [418] and by Payot et al. [419] for medical imaging applications. Methods based on partial differential equations continue to be important for enhancing 3D data sets that are difficult and expensive to acquire, such that deep learning approaches are less suited. However, in order to achieve optimal quality, these methods should be well adapted to the imaging process.

All the above mentioned methods might not be directly applicable to data acquired through cryo-EM: One challenge in designing filters for cryo-EM data is the limited angle tomography problem, also known as missing wedge problem [50,51]. It arises from the geometric design of data acquisition using the electron microscope. One cannot acquire data from all orientations of the sample (left side of Figure 10.1), which leads to presence of the missing wedge in the Fourier space. When one reconstructs the 3D data in the Cartesian space from individual projections, it is blurred/smeared in the directions where the missing wedge exists. On the right side of Figure 10.1, we can see the directions in which the data is blurred at every pixel in the Cartesian space. The two triangles represent these directions where the data

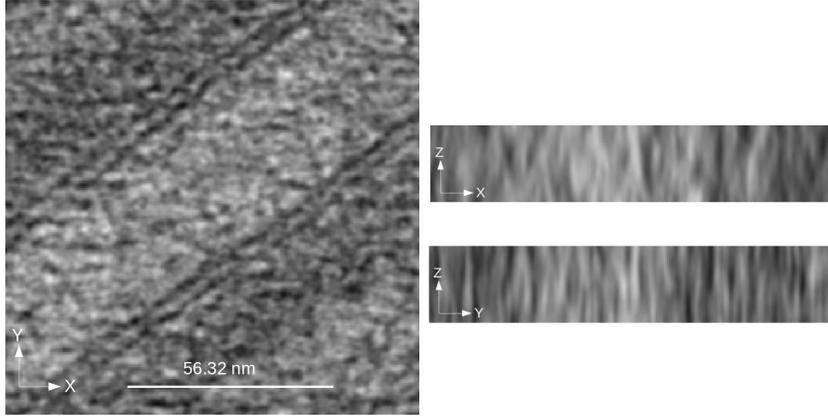


Figure 10.2: **Left:** Noisy xy slice. **Right:** Directional blur in z direction. Images have been obtained through courtesy of Achilleas Frangakis.

is smeared at every pixel. Here, we depict a particular xz plane. This region extends throughout the y direction which is perpendicular to the plane in which the figure exists. Due to the smearing effect present in the data, classical formulas based on gradient calculations cannot be considered for enhancing the image structures. Moreover, the high resolution nature of image acquisition in cryo-EM also gives rise to a lot of noise. Figure 10.2 shows the noisy and directional blur nature of rat liver cell data acquired through cryo-EM techniques.

Another challenge in cryo-EM is the following: The relation between an object O and its corresponding cryo-EM projection image I in the Fourier space is given by $I(\hat{x}) = O(\hat{x})H(\hat{x})$. Here, \hat{x} represents the 2D frequency vector. The contrast transfer function H (CTF) is the Fourier transform of the point spread function of the camera. Retrieving the object O from the image I is called CTF correction. However, this process can lead to the creation of directional discontinuities in the acquired data. Figure 10.3 shows such an effect in the protinaceous desmosome networks of rat liver cells.

Our Goal. The main goal in this chapter is to design a filter that extends the application of oriented structure enhancement in 3D to the specific needs in cryo-EM. More precisely, the filter needs to account for the limited angle tomography problem, noisy nature of image acquisition, and also the discontinuities that might be created due to CTF data correction.

Our Contribution. We consider a combination of a general directional

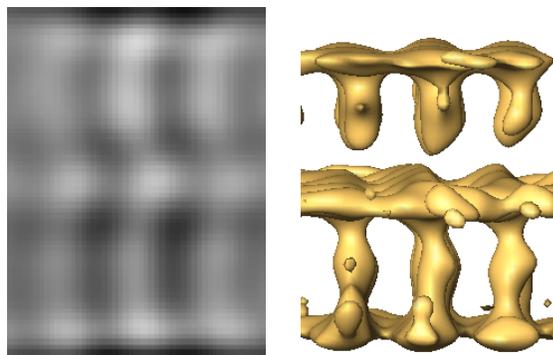


Figure 10.3: **Left:** 2D orthoslice. **Right:** Corresponding 3D isosurface. Data has been obtained through courtesy of Achilleas Frangakis.

image evolution process with the classical Hough transform [29–32]. The latter technique is generally used to detect complex patterns in images. We use a semi-local version of it to find the local direction in which the image structures exist. The smoothing process is then steered in this direction using the directional image evolution process. The flexible and robust nature of the Hough transform helps the *Hough-based evolution* (HE) method introduced in this chapter to deal with all the above problems in cryo-EM effectively.

Chapter Structure. The organisation of this chapter is as follows: In Section 10.1 we introduce the HE method. We also mention the specific changes that have to be made to the classical ideas used for structure enhancement, in order to adapt them to cryo-EM data. In Section 10.2, we compare the performance of the HE method with the popular CED approach on synthetic and cryo-EM data. We also discuss how the ideas used in the modelling of the HE algorithm lead to the desired enhancement of oriented structures in the images. In Section 10.3, we conclude with a summary and future outlook.

10.1 Modelling and Theory

The structure enhancing technique that we introduce in this section can be divided into two parts: First, we explain the core image evolution process that we adopt. This is followed by a review of the Hough transform which helps detect the local direction in which the smoothing is steered. We end this section with the modifications of classical ideas that are needed to adapt

to the specific requirements in cryo-EM.

10.1.1 A General Directional Data Evolution

Let $\Omega \subset \mathbb{R}^3$ denote a cuboidal image domain and consider some 3D data set $f : \Omega \rightarrow \mathbb{R}$. We can obtain a family $\{u(\cdot, t) \mid t \geq 0\}$ of smoothed versions of f by regarding f as initial value of a 3D directional image evolution that satisfies the following partial differential equation:

$$\partial_t u = \partial_{\boldsymbol{\eta}} u = \boldsymbol{\eta}^\top \text{Hess}(u) \boldsymbol{\eta} \quad (10.1)$$

with reflecting boundary conditions. Here, $\text{Hess}(u)$ denotes the spatial Hessian of u . The smoothing direction $\boldsymbol{\eta}$ is space-variant and is characterised by its angles $\theta(\mathbf{x})$ and $\phi(\mathbf{x})$ in a spherical coordinate system, i.e.

$$\boldsymbol{\eta} = \boldsymbol{\eta}(\mathbf{x}) = \begin{pmatrix} \sin(\theta(\mathbf{x})) \cos(\phi(\mathbf{x})) \\ \sin(\theta(\mathbf{x})) \sin(\phi(\mathbf{x})) \\ \cos(\theta(\mathbf{x})) \end{pmatrix}. \quad (10.2)$$

Equations of type (10.1) have a long tradition, in particular in the 2D setting. For instance, for 2D mean curvature motion [420], one chooses $\boldsymbol{\eta}(\mathbf{x}, t) \perp \nabla u(\mathbf{x}, t)$. Obviously such a choice – which smoothes along isophotes – cannot close interrupted structures. To this end, one needs more advanced local structure descriptors than the gradient, e.g. the Gabor transform-based methods that are used in the evolution equation of Carmona and Zhong [421]. The Carmona–Zhong approach, however, is designed for processing 2D images and has not been adapted to tackle 3D cryo-EM data. We prefer another local structure descriptor that is better suited for our specific needs. It is based on a semi-local Hough transform and is discussed next.

10.1.2 Hough Transform-based Directional Data Evolution

The novelty of our work is that we choose the smoothing direction $\boldsymbol{\eta}(\mathbf{x})$ mentioned above using a semi-local Hough transform [29, 31, 32] on the original data f . This direction corresponds to the line segment in a local neighbourhood along which the image structure to be enhanced is present. It has to be mentioned that the Hough transform is very robust in detecting

Pseudo Code: Hough-based Evolution

Input: Original discrete data set \mathbf{f} and the following parameters:

ρ_1 - radius of the sphere shaped neighbourhood,

ρ_2 - half of the length of the line segments,

T - threshold for the Hough transform,

τ - time step size of the explicit scheme,

k_{\max} - total number of iterations.

Modified Hough Algorithm for Computing $\eta(x)$:

1. Select a ball B of radius ρ_1 around each pixel of \mathbf{f} .
2. Line segments of half length ρ_2 centered at every pixel within B are considered by discretising the angles θ and ϕ .
3. The θ and ϕ values of the line segment which has the largest percentage of pixels with grey values larger than T represent the local dominant direction η . To tackle the missing wedge, only line segments within the angle α_W from Figure 10.1 are considered in the voting process.

Main Algorithm:

1. Initialisation: $\mathbf{u}^0 = \mathbf{f}$
2. **For** $k = 0, 1, 2, \dots, k_{\max} - 1$:

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau \cdot \eta^\top \text{Hess}(\mathbf{u}^k) \eta,$$
 where $\text{Hess}(\mathbf{u}^k)$ is approximated with central finite differences.

Output: Image with enhanced structures $\mathbf{u}^{k_{\max}}$.

the local dominant direction because a small relative majority in the voting process suffices for obtaining the dominant direction. In this sense, for noise of isotropic nature, the Hough-based image evolution is more robust than structure tensor-based methods such as CED. Moreover, as we will see in the upcoming section, our modified Hough-based selection of dominant directions is able to tackle the missing wedge effect in cryo-EM.

The above pseudo-code explains the complete algorithm and its parameters in detail. Note that we compute the Hough transform only for the initial data set f , not for its evolution $u(., t)$. This saves computational time and leads to a linear method. We have not noticed qualitative differences

compared to a non-linear variant where we adapt the dominant direction to the evolving image $u(., t)$.

10.1.3 Modifying Classical Ideas for Adapting to Cryo-EM Data

Let us now discuss the exact modifications we have made to classical concepts in order to adapt them better to the scenario of limited angle tomography.

Restricting the Search Space of Dominant Directions. It was already explained in the beginning of this chapter that the missing wedge problem arises due to the design of the data acquisition process: The specimen whose images are acquired cannot be tilted above a certain angle (generally 60°). Thus, we do not have projections from all angles. This leads to missing information while reconstructing the 3D data from the available projections. It creates a smearing effect of the reconstructed data in the directions where the data cannot be collected. In order to tackle this, the modified Hough algorithm mentioned above just considers the line segments which are outside the wedge represented in Figure 10.1. In other words, since our aim is to enhance the structures in the directions where we do not encounter the smearing effect, the search space of Hough directions is restricted by choosing $\alpha_W = 30^\circ$. One can notice that this value is smaller than 60° . The reason behind this is the discrete nature of pixels. It does not allow for a smooth transfer of the continuous real-world geometry into the captured data.

Avoiding the Usage of Gradients. Due to the smearing effect of the cryo-EM data, classical formulas for gradient calculation can no longer be used for processing. Generally, gradients calculated on a Gaussian-smoothed image (with standard deviation σ) are used in the voting process of the Hough transform. We instead use the grey values.

If dark structures are to be enhanced, we choose as Hough direction in 3D the one that contains the largest percentage of pixels with greyscale value below a certain threshold. For enhancing bright structures, we consider pixels above the threshold. Most of the previously designed filters mentioned in introductory part of this chapter are based on the structure tensor [411]. The structure tensor averages directional information over a local neighbourhood using gradient formulation. Thus, as we will see in the upcoming section, both gradient-based Hough algorithm and CED (which is built upon the

structure tensor) are not successful. On the other hand, we will also see that the usage of the grey value-based Hough algorithm in the directional image evolution produces the desired results.

10.1.4 Numerical Algorithm

For discretising the 3D evolution in (10.1), we use a straightforward explicit scheme as mentioned in the main algorithm of the pseudo-code. We use central derivative approximations to calculate the spatial derivatives in the Hessian. For a spatial grid size of 1 and $\tau < \frac{1}{6}$, we observed a convergence to a plausible image reconstruction with enhanced structures. A more detailed theoretical study of the Hough-based image evolution behaviour will be a topic for future research.

10.2 Experiments and Discussion

We have assessed the performance of our algorithm on both synthetic and real-world data. We first present details regarding parameter selection for HE before we move on to the results.

10.2.1 Parameter Selection

There are five parameters which need to be selected, out of which two are critical and need to be adapted to the specific data set.

The radius of the sphere ρ_1 , half of the length of the line segment ρ_2 , and the threshold T for the voting process in the Hough transform are the model parameters. The time step size τ and the number of iterations k_{\max} are numerical parameters to reach a desired stopping time.

We suggest $\rho_1 = 3$ for the radius of the sphere shaped neighbourhood. This allows for searching the dominant orientation in a small neighbourhood instead of just around a specific pixel. The selection of the parameters ρ_2 and T is important. They must be adapted to the data set. The parameter ρ_2 , which specifies the length of the line segment, must be greater than the length of the discontinuities in order to detect and remove them. The threshold parameter T must be selected according to the greyscale range at which the structures are present in the image.

As already mentioned, τ has to obey a stability criterion caused for the explicit scheme. Fixing τ implies that the stopping time is proportional to k_{\max} . We obtain smoother data for a larger number of iterations. Also, if the gaps to be closed are large, we need to increase the number of iterations until the structures get connected.

For discretizing the spherical polar coordinate angles θ and ϕ , we use stepsizes that produce 18 samples within each of their respective ranges.

Also, a Gaussian smoothing is only necessary for a gradient-based Hough transform, due to the ill-posedness of differentiation. It is not required for our grey value-based variant. Furthermore, we have affine rescaled the real-world datasets to $[0, 255]$ before the algorithms were applied. This facilitates the reproducibility of results while selecting the threshold parameter. It does not have any other effect on the model itself.

10.2.2 Synthetic Data

The Shepp–Logan phantom data set [422, 423] is a popular synthetic data set used for testing 3D reconstruction algorithms. However, it is not suited for testing the capability of enhancing line-like structures in the presence of a missing wedge. A synthetic image for testing this particular capability of methods is simply missing in the image processing community.

Thus, we have created a 3D image which mimics the effect of the missing wedge and also has discontinuous structures that need to be connected while enhancing them. Figure 10.4 (a) shows different slices of the 3D data set we have created. We can clearly see the disconnected structures in the xy slice and the elongated/smeared structures in the other slices. This mimics the effect of the missing wedge. In Figure 10.4 (b), we present the original structures without the discontinuities. Figure 10.4 (c) depicts the results of our grey value-based HE method. It is able to connect the disconnected structures. We observe that this approach outperforms both the gradient-based HE method and CED whose results are presented in Figures 10.4 (d) and 10.4 (e), respectively.

10.2.3 Real-world Data

Figure 10.5 (a),(d) shows a reconstructed 3D cellular region acquired from an electron microscope. One can see that the data in the z direction is smeared

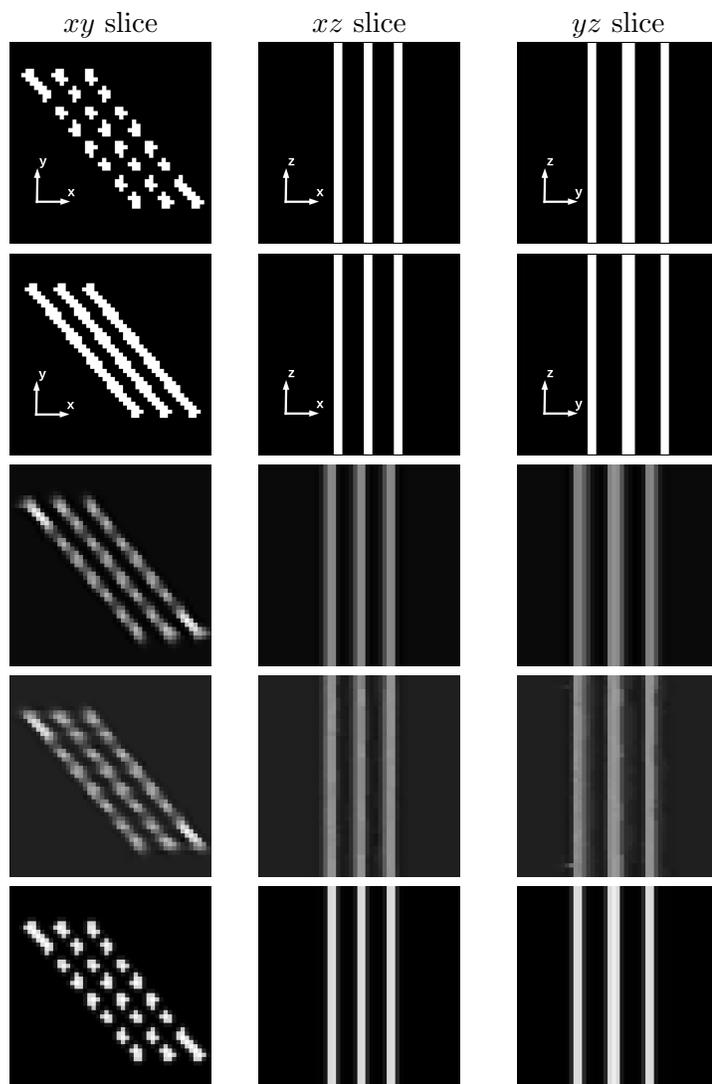


Figure 10.4: Performance of the different approaches for a synthetic data set of size $49 \times 49 \times 49$. **(a) First Row:** Corrupted data. **(b) Second Row:** Original data. **(c) Third Row:** Grey value-based HE ($\rho_1 = 3$, $\rho_2 = 21$, $T = 10$, $\tau = 0.1$, $k_{\max} = 10$). **(d) Fourth Row:** Gradient-based HE ($\rho_1 = 3$, $\rho_2 = 21$, $T = 5$, $\sigma = 0.5$, $\tau = 0.1$, $k_{\max} = 10$). **(e) Fifth Row:** CED ($\lambda = 1.0$, $\sigma = 0.5$, $\rho = 5.0$, $\alpha = 0.001$, $\tau = 0.1$, $k_{\max} = 1000$). We have shown complement images of the original ones, for better viewing. One can observe a better enhancement using grey value-based HE.

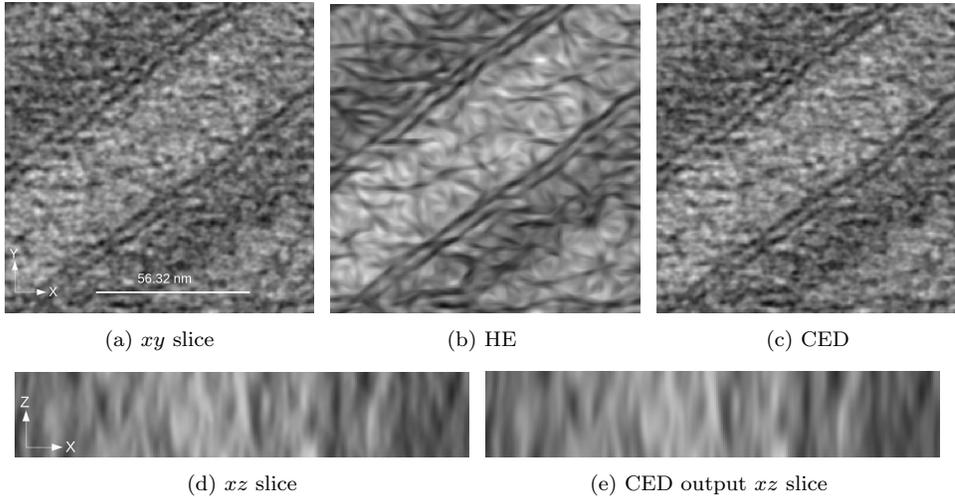


Figure 10.5: Cellular regions of rat liver enhanced using 3D HE. In Figure 10.5(b) we can see the enhanced double walled cell membranes of neighbouring cells. We can also see the desmosomes (structures made of proteins) in directions perpendicular to the cell membrane. HE parameters: $\rho_1 = 3$, $\rho_2 = 21$, $T = 200$, $\tau = 0.1$, $k_{\max} = 50$. CED parameters: $\lambda = 1.0$, $\sigma = 0.5$, $\rho = 5.0$, $\alpha = 0.001$, $\tau = 0.1$, $k_{\max} = 100$. Data set size: $256 \times 256 \times 50$. The xz slice for HE has not been presented as it looks similar to the input version.

and resembles the above mentioned synthetic data set. Figure 10.5(b) displays the resulting enhanced cell structures using the HE algorithm. This image allows better visualisations than the original image in Figure 10.5(a). The enhanced image (Figure 10.5(b)) contains two double walled cell membranes of neighbouring cells. The structures arising in directions perpendicular to the cell membrane are the desmosome networks which are made of proteins. These networks bind neighbouring cells together. The structures are more evident in the enhanced image than in the original image (Figure 10.5(a)). Further, since the original image has more signal in the intra-cellular region than the extra-cellular region, the enhanced image also has clear desmosome networks in the intra-cellular region.

Figure 10.5(c) depicts the result using CED with a straightforward explicit scheme. This structure tensor-based enhancement method fails to enhance the structures in the presence of the missing wedge, as was explained in Section 10.1.3. Figure 10.5(e) shows the xz slice after applying CED. In the presence of a missing wedge, this method always smooths in the z -direction

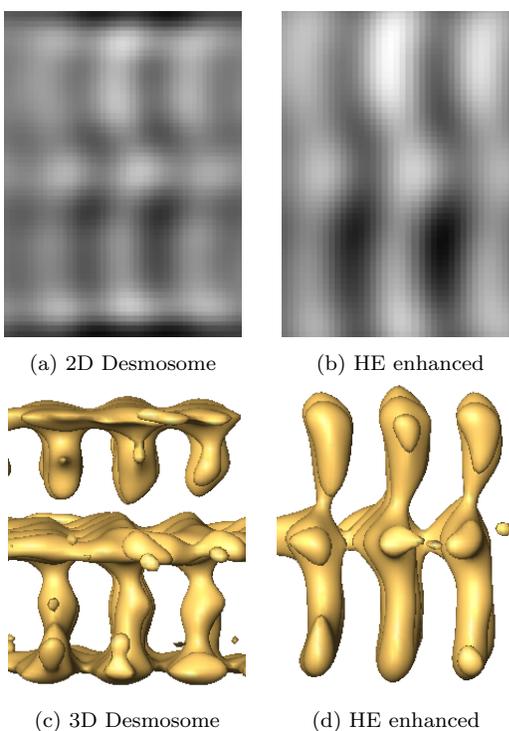


Figure 10.6: Connecting disconnected desmosome structures. Visualisation threshold for both structures is a grey value of 140. This is required for graphical rendering and is different from the threshold parameter T in the Hough-transform. The length of each vertical desmosome is around 28 nanometers. Full data set size: $128 \times 128 \times 128$ pixels. Parameters used: $\tau = 0.025$, $k_{\max} = 300$. The other model parameters are not relevant as we do not employ the Hough transform for this experiment, but fix the smoothing to one particular direction.

as CED detects the coherent structures in the z -direction. Consequently, we do not observe structure enhancement in the xy slices; see Figure 10.5 (c). In the grey value-based HE method, this is avoided by restricting the search space of angles.

Another application of HE is presented in Figure 10.6. The vertical structures in white in Figure 10.6(a) are desmosomes. We infer that the presence of discontinuities in the horizontal direction is due to the CTF correction of the data after acquisition. We see the desmosome structures in the extracellular regions clearly in contrast to the images in the previous experiment. This is due to the fact that several similar structures were averaged to get this

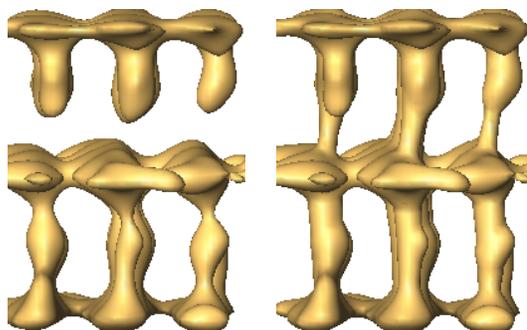


Figure 10.7: Enhancement without losing the shape of the structures. Grey value threshold parameter for visualisation: 146. Parameters used: $\tau = 0.025$, $k_{\max} = 1200$. The other model parameters are not relevant as we do not employ the Hough transform for this experiment, but fix the smoothing to one particular direction. **Left:** Desmosome in 3D. **Right:** Modified HE output.

final image. As a consequence, the missing wedge effect is minimised. Here, we want to enhance the structures only in the direction perpendicular to the discontinuities. Hence, we need not perform the Hough transform. Smoothing in this required direction is governed by setting $\theta = 0^\circ$. This results in $\boldsymbol{\eta} = (0, 0, 1)^T$. The output after removing the discontinuities using HE is displayed in Figure 10.6(b). The graphical renderings of these structures are also depicted in Figure 10.6. The vertical structures are nicely connected.

However, the shape of the structures seems to have changed. The reason behind this is the following: After the HE technique is used, grey values move from one region of the image to another. Let us think of these grey values as particles. The movement of particles changes the shape of the structures. In order to save the initial shape we need to fill missing gaps with new particles. To this end, we update a pixel value only when it has increased in contrast to its value in the previous iteration. In other words, if $\mathbf{u}^{k+1} < \mathbf{u}^k$ then $\mathbf{u}^{k+1} = \mathbf{u}^k$. In Figure 10.6(a), since larger grey values represent particles and smaller ones represent background, the above modification will create new particles and also not disturb the spatial positions of the old ones. Thus, the initial shape of the structures is preserved. The output of such an algorithm is shown in Figure 10.7.

Another method which can help improve the visualisation of images is the contrast control feature of the Amira software - a typical cryo-EM visualisation software. Figure 10.8 shows an example. The same software has

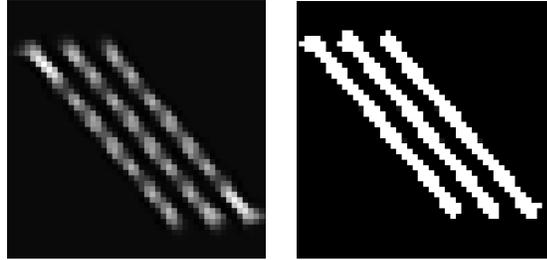


Figure 10.8: Enhancement with contrast control in Amira software. Contrast control parameters - Center: 205.2 and Width: 2.2. **Left:** Default visualisation. **Right:** With contrast control.

also been used for producing the graphical renderings in Figures 10.3, 10.7, and 10.8.

All the experiments in this work have been performed on a GPU (Nvidia Quadro P5000 graphics card with C++ and CUDA). The computational time for the experiments in Figures 10.4-10.6 are 3.45, 84.6 and 10 seconds, respectively.

10.3 Conclusions

We have introduced a method that combines a semi-local Hough transform with a directional image evolution. This approach is designed to enhance oriented structures in 3D data sets from electron microscopy. Our variant of the Hough transform is robust with respect to the unwanted smearing effects produced by the missing wedge in cryo-EM data acquisition. Our model enhances structures that are present within the directions where no smearing occurs. Other methods which are based on derivative information, such as CED and gradient-based HE, fail to overcome this problem. Additionally, our approach is also able to deal with the discontinuities that can occur in cryo-EM data due to CTF correction.

Manuscript details of research content from this chapter:

K. Bodduna, J. Weickert, and A. S. Frangakis. Hough-based evolutions for enhancing structures in 3D electron microscopy. In M. Vento, G. Percannella, editors, *Computer Analysis of Images and Patterns (CAIP)*, volume 11678 of *Lecture Notes in Computer Science*, pages 102-112, Springer, Cham, 2019.

Chapter 11

Conclusions and Outlook

11.1 Conclusions

A significant part of scientific research deals with discovering physical and abstract structures in the very complex world we are living in. The contributions of this thesis follow this philosophy. In all the applications dealt with in this thesis, we faced several complexities: We encountered multiple noise distributions, faced a severe anisotropic deformation in image acquisition process, and needed to design a strategy that improved the performance of not one but several single-frame filters in a multi-frame scenario. Despite such obstacles, our novel approaches were able to reveal physical structures that were otherwise difficult to visualise in the raw input images.

One of the main characteristics of a good technology is that it solves a broader scope of problems than those resolved by previous works. The following summary exemplifies the presence of this attribute in our approaches.

In single-frame denoising, we designed a technique that preserves image structures, is fast, avoids artefacts, and is robust to the type of noise. A filter which possesses such wide-ranging qualities is very rare.

For the multi-frame denoising application, we devised an inter-frame connectivity strategy which improved the performance of a vast variety of single-frame patch-based filters. Moreover, this improvement is irrespective of what kind of noise is under consideration.

In image enhancement, we modelled an evolution equation which is robust to whether or not anisotropic deformations in the data (like the missing wedge

effect in electron microscopy) are present.

Finally, in multi-frame super-resolution we made contributions to both data and smoothness terms. In the former scenario we came up with a model which is more than an order of magnitude faster with a very little compromise in data quality. In the smoothness term study we showed the advantages of using advanced diffusion-based structure preserving regularisers. These contributions help extend the coverage of super-resolution techniques to time critical and noisy image acquisition applications, respectively.

11.2 Outlook

In single-frame denoising, we showed that anisotropic shapes like sector and stripe are more effective than a disc. Thus, in future, one could explore patch-matching extensions of stripe and sector diffusion techniques. Especially due to the fact that a patch-matching extension of a disc in the form of NFPR was highly productive. Moreover, the one-sided directional derivative-based modelling of sector diffusion has not been described in the literature so far. Studying the mathematical properties of such image evolutions is challenging and could be a direction for future research.

In multi-frame denoising, we concluded that a separable spatio-temporal filtering is a better solution than combined denoising. However, we only considered disc and square shaped neighbourhoods enroute this conclusion. In future, one should include sector and stripe shaped anisotropic neighbourhoods along with shape-adaptive filters for the multi-frame study. Moreover, any improvements in single- or multi-frame denoising layouts would also have direct repercussions in a super-resolution scenario.

As far as image enhancement techniques are concerned, one could also study robust anisotropy detection techniques other than the Hough transform, like random sampling consensus algorithms.

Independent of whether it is structure preserving or enhancing models, in this thesis we did not focus on colour images as well as image adaptive and automated parameter selection. These are some other research directions which one could pursue in the future. No matter what the specific direction is, we should not forget one of the chief characteristics of a good technology: Whenever possible, it must be applicable in a broad-ranging number of situations than those dealt with by an existing one.

Chapter 12

Bibliography

Bibliography

- [1] K. J. Borkowski, S. P. Reynolds, D. A. Green, U. Hwang, and R. Petre. Radioactive scandium in the youngest galactic supernova remnant G1.9+0.3. *The Astrophysical Journal Letters*, 724(2):L161–L165, November 2010.
- [2] J. L. Starck, F. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, Cambridge, UK, 1998.
- [3] E. D. Kolaczyk. Nonparametric estimation of gamma-ray burst intensities using Haar wavelets. *The Astrophysical Journal*, 483(1):340–349, July 1997.
- [4] E. D. Kolaczyk and D. D. Dixon. Nonparametric estimation of intensity maps using Haar wavelets and Poisson noise characteristics. *The Astrophysical Journal*, 534(1):490–505, May 2000.
- [5] B. Zhang, M. J. Fadili, J. L. Starck, and J. C. Olivo-Marin. Multiscale variance-stabilizing transform for mixed-Poisson-Gaussian processes and its applications in bioimaging. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 6, pages 233–236, San Antonio, TX, USA, October 2007.
- [6] F. Luisier, T. Blu, and M. Unser. Image denoising in mixed Poisson-Gaussian noise. *IEEE Transactions on Image Processing*, 20(3):696–708, September 2010.
- [7] C. Vonesch, F. Aguet, J. L. Vonesch, and M. Unser. The colored revolution of bioimaging. *IEEE Signal Processing Magazine*, 23(3):20–31, May 2006.

- [8] M. Vulović, R. B. G. Ravelli, L. J. van Vliet, A. J. Koster, I. Lazić, U. Lüken, H. Rullgård, O. Öktem, and B. Rieger. Image formation modeling in cryo-electron microscopy. *Journal of Structural Biology*, 183(1):19–32, July 2013.
- [9] J. E. Evans, C. Hetherington, A. Kirkland, L. Y. Chang, H. Stahlberg, and N. Browning. Low-dose aberration corrected cryo-electron microscopy of organic specimens. *Ultramicroscopy*, 108(12):1636–1644, November 2008.
- [10] J.R. Jinschek, K.J. Batenburg, H.A. Calderon, R. Kilaas, V. Radmilovic, and C. Kisielowski. 3D reconstruction of the atomic positions in a simulated gold nanocrystal based on discrete tomography: Prospects of atomic resolution electron tomography. *Ultramicroscopy*, 108(6):589–604, May 2008.
- [11] D. R. Brillinger. Some wavelet analyses of point process data. In *Proc. IEEE Asilomar Conference on Signals, Systems and Computers (ACSSC)*, volume 2, pages 1087–1091, Pacific Grove, CA, USA, November 1997.
- [12] I. Rodrigues, J. Sanches, and J. Bioucas-Dias. Denoising of medical images corrupted by Poisson noise. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1756–1759, San Diego, CA, USA, October 2008.
- [13] J. Weickert. Coherence-enhancing diffusion filtering. *International Journal of Computer Vision*, 31(2):111–127, April 1999.
- [14] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal Bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, September 2013.
- [15] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, August 2007.
- [16] M. Mainberger, A. Bruhn, J. Weickert, and S. Forchhammer. Edge-based compression of cartoon-like images with homogeneous diffusion. *Pattern Recognition*, 44(9):1859–1873, September 2011.

- [17] J. Weickert. Anisotropic diffusion filters for image processing based quality control. In A. Fasano and M. Primicerio, editors, *Proc. Seventh European Conference on Mathematics in Industry*, pages 355–362. Teubner, Stuttgart, 1994.
- [18] I. Ram, M. Elad, and I. Cohen. Image processing using smooth ordering of its patches. *IEEE Transactions on Image Processing*, 22(7):2764–2774, July 2013.
- [19] I. Ram, M. Elad, and I. Cohen. Image denoising using NL-means via smooth patch ordering. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1350–1354, Vancouver, BC, Canada, May 2013.
- [20] R. Lerallut, E. Decenciere, and F. Meyer. Image filtering using morphological amoebas. *Image and Vision Computing*, 25(4):395–404, April 2007.
- [21] N. Pierazzo, M. E. Rais, J. M. Morel, and G. Facciolo. DA3D: Fast and data adaptive dual domain denoising. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 432–436, Quebec City, QC, Canada, September 2015.
- [22] A. Foi, V. Katkovnik, and K. Egiazarian. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images. *IEEE Transactions on Image Processing*, 16(5):1395–1411, May 2007.
- [23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. BM3D image denoising with shape-adaptive principal component analysis. In *Proc. of the Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, Saint-Malo, France, April 2009.
- [24] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian. Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms. *IEEE Transactions on Image Processing*, 21(9):3952–3966, September 2012.
- [25] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruc-

- tion. *IEEE Transactions on Image Processing*, 22(1):119–133, July 2012.
- [26] P. Arias and J. M. Morel. Video denoising via empirical Bayesian estimation of space-time patches. *Journal of Mathematical Imaging and Vision*, 60:70–93, January 2018.
- [27] M. Elad and A. Feuer. Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Transactions on Image Processing*, 6(12):1646–1658, December 1997.
- [28] J. Weickert, B. M. ter Haar Romeny, A. Lopez, and W. J. van Enk. Orientation analysis by coherence-enhancing diffusion. In *Proc. Real World Computing Symposium*, pages 96–103, Tokyo, Japan, January 1997.
- [29] P. V. C. Hough. Methods and means for recognising complex patterns. *U.S. Patent No. 3069654*, December 1962.
- [30] P. V. C. Hough. Machine analysis of bubble chamber pictures. *Proc. of the International Conference on High Energy Accelerators and Instrumentation*, pages 554–556, September 1959.
- [31] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, January 1972.
- [32] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [33] A. J. P. Theuwissen. *Solid-State Imaging with Charge-Coupled Devices*. Kluwer, Dordrecht, 1995.
- [34] A. J. Blanksby, M. J. Loinaz, D. A. Inglis, and B. D. Ackland. Noise performance of a color CMOS photogate image sensor. In *Proc. IEEE International Electron Devices Meeting (IEDM)*, pages 205–208, Washington, DC, USA, December 1997.
- [35] R. E. Jacobson, S. F. Ray, G. G. Attridge, and N. R. Axford. *The Manual of Photography: Photographic and Digital Imaging*. Focal Press, Waltham, MA, USA, 2000.

- [36] J.R. Janesick. *Scientific Charge-Coupled Devices*. SPIE, Bellingham, WA, USA, 2001.
- [37] L. G. Shapiro and G. C. Stockman. *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, USA, 2001.
- [38] B. W. Keelan. *Handbook of Image Quality: Characterization and Prediction*. Marcel Dekker Inc., New York, NY, USA, 2002.
- [39] F. Rooms, W. Philips, and P. Van Oostveldt. Integrated approach for estimation and restoration of photon-limited images based on steerable pyramids. In *Proc. IEEE EURASIP Conference Focussed on Video/Image Processing and Multimedia Communications (EC-VIP-MC)*, pages 131–136, Zagreb, Croatia, July 2003.
- [40] R. Ramanath, W. E. Snyder, Y. Yoo, and M. S. Drew. Color image processing pipeline. *IEEE Signal Processing Magazine*, 22(1):34–43, January 2005.
- [41] J. Nakamura. *Image Sensors and Signal Processing for Digital Still Cameras*. CRC Press, Boca Raton, FL, USA, 2005.
- [42] W. C. Kao, S. H. Wang, L. Y. Chen, and S. Y. Lin. Design considerations of color image processing pipeline for digital cameras. *IEEE Transactions on Consumer Electronics*, 52(4):1144–1152, November 2006.
- [43] J. G. Pellegrino, J. Zeibel, R. G. Driggers, and P. Perconti. *Infrared Camera Characterization*. CRC Press, Boca Raton, FL, USA, 2006.
- [44] L. MacDonald. *Digital Heritage: Applying Digital Imaging to Cultural Heritage*. Butterworth-Heinemann, Oxford, UK, 2006.
- [45] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, USA, 2007.
- [46] M. A. Covington. *Digital SLR Astrophotography*. Cambridge University Press, Cambridge, UK, 2007.
- [47] J. Ohta. *Smart CMOS Image Sensors and Applications*. CRC Press, Boca Raton, FL, USA, 2008.

- [48] C. Boncelet. Image noise models. In A. Bovik, editor, *The Essential Guide to Image Processing*, pages pp. 143–167, Academic Press, Boston, USA, July 2009.
- [49] K. Dabov. *Image and Video Restoration with Nonlocal Transform-Domain Filtering*. PhD thesis, Department of Information Technology, Tampere University of Technology, Finland, 2010.
- [50] J. Frank. *Three-dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in their Native State*. Oxford University Press, NY, USA, second edition, 2006.
- [51] J. Frank. *Electron Tomography: Methods for Three-dimensional Visualisation for Structures in the Cell*. Springer, New York, second edition, 2006.
- [52] D. A. Evans, P. P. Allport, G. Casse, A. R. Faruqi, B. Gallop, R. Henderson, M. Prydderch, R. Turchetta, M. Tyndel, J. Velthuis, G. Villani, and N. Waltham. CMOS active pixel sensors for ionising radiation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 546(1-2):281–285, July 2005.
- [53] A. R. Faruqi and D. M. Cattermole. Pixel detectors for high-resolution cryo-electron microscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 549(1-3):192–198, September 2005.
- [54] A.R. Faruqi, D. M. Cattermole, and C. Raeburn. Applications of pixel detectors to electron microscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 512(1-2):310–317, October 2003.
- [55] A.R. Faruqi, D. M. Cattermole, and C. Raeburn. Direct electron detection methods in electron microscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 513(1-2):317–321, November 2003.
- [56] A. R. Faruqi, D. M. Cattermole, R. Henderson, B. Mikulec, and C. Raeburn. Evaluation of a hybrid pixel detector for electron microscopy. *Ultramicroscopy*, 94(3-4):263–276, April 2003.

- [57] A. R. Faruqi, R. Henderson, M. Pryddetch, P. Allport, and A. Evans. Direct single electron detection with a CMOS detector for electron microscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 546(1-2):170–175, July 2005.
- [58] Y. Mäkinen, L. Azzari, and A. Foi. Exact transform-domain noise variance for collaborative filtering of stationary correlated noise. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 185–189, Taipei, Taiwan, September 2019.
- [59] Y. Mäkinen, L. Azzari, and A. Foi. Collaborative filtering of correlated noise: Exact transform-domain variance for improved shrinkage and patch matching. *IEEE Transactions on Image Processing*, 29:8339–8354, August 2020.
- [60] M. Mäkitalo and A. Foi. Optimal inversion of the Anscombe transformation in low-count Poisson image denoising. *IEEE Transactions on Image Processing*, 20(1):99–109, July 2010.
- [61] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, October 2008.
- [62] R. A. Boie and I. J. Cox. An analysis of camera noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(6):671–674, June 1992.
- [63] M. Colom, M. Lebrun, A. Buades, and J. M. Morel. Nonparametric multiscale blind estimation of intensity-frequency-dependent noise. *IEEE Transactions on Image Processing*, 24(10):3162–3175, October 2015.
- [64] M. Colom, M. Lebrun, A. Buades, and J. M. Morel. A non-parametric approach for the estimation of intensity-frequency dependent noise. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 4261–4265, Paris, France, October 2014.

- [65] M. Colom. *Multiscale Noise Estimation and Removal for Digital Images*. PhD thesis, Dpt. de Matemàtiques i Informàtica, Universitat de les Illes Balears, Spain, 2014.
- [66] L. Azzari, L.R. Borges, and A. Foi. Modeling and estimation of signal-dependent and correlated noise. In M. Bertalmío, editor, *Denoising of Photographic Images and Video: Fundamentals, Open Challenges and New Trends*, Advances in Computer Vision and Pattern Recognition, pages 1–36, Switzerland, 2018. Springer.
- [67] L. Azzari. *Models and Methods for Estimation and Filtering of Signal-Dependent Noise in Imaging*. PhD thesis, Department of Information Technology, Tampere University of Technology, Finland, 2016.
- [68] C. C. Enz and G. C. Temes. Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization. *Proceedings of the IEEE*, 84(11):1584–1614, November 1996.
- [69] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):299–314, February 2008.
- [70] S. W. Hasinoff, F. Durand, and W. T. Freeman. Noise-optimal capture for high dynamic range photography. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 553–560, San Francisco, CA, USA, June 2010.
- [71] H. Tian, B. Fowler, and A. E. Gamal. Analysis of temporal noise in CMOS photodiode active pixel sensor. *IEEE Journal of Solid-State Circuits*, 36(1):92–101, January 2001.
- [72] A. Foi, S. Alenius, V. Katkovnik, and K. Egiazarian. Noise measurement for raw-data of digital imaging sensors by automatic segmentation of nonuniform targets. *IEEE Sensors Journal*, 7(10):1456–1461, October 2007.
- [73] J. Boulanger, J. B. Sibarita, C. Kervrann, and P. Bouthemy. Non-parametric regression for patch-based fluorescence microscopy image sequence denoising. In *Proc. IEEE International Symposium on*

- Biomedical Imaging: From Nano to Macro (ISBI)*, pages 748–751, Paris, France, May 2008.
- [74] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J. B. Sibarita, and J. Salamero. Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. *IEEE Transactions on Medical Imaging*, 29(2):442–454, February 2010.
- [75] F. Luisier, C. Vonesch, T. Blu, and M. Unser. Fast Haar-wavelet denoising of multidimensional fluorescence microscopy data. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 310–313, Boston, MA, USA, July 2009.
- [76] A. Foi. Practical denoising of clipped or overexposed noisy images. In *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, pages 1–5, Lausanne, Switzerland, August 2008.
- [77] A. Foi. Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, 89(12):2609–2629, December 2009.
- [78] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.
- [79] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey. Complex wavelet structural similarity: A new image similarity index. *IEEE Transactions on Image Processing*, 18(11):2385–2401, November 2009.
- [80] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In *Proc. IEEE Asilomar Conference on Signals, Systems and Computers (ACSSC)*, volume 2, pages 1398–1402, Pacific Grove, CA, USA, November 2003.
- [81] K. Okarma. Colour image quality assessment using structural similarity index and singular value decomposition. In L. Bolc, J. L. Kulikowski, and K. Wojciechowski, editors, *Computer Vision and Graphics*, volume 5337 of *Lecture Notes in Computer Science*, pages 55–65, Berlin, Heidelberg, 2008. Springer.

- [82] A. Kolaman and O. Y. Pecht. Quaternion structural similarity: A new quality index for color images. *IEEE Transactions on Image Processing*, 21(4):1526–1536, April 2012.
- [83] J. Zujovic, T. N. Pappas, and D. L. Neuhoff. Structural similarity metrics for texture analysis and retrieval. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 2225–2228, Cairo, Egypt, November 2009.
- [84] D. Brunet, E. R. Vrscay, and Z. Wang. On the mathematical properties of the structural similarity index. *IEEE Transactions on Image Processing*, 21(4):1488–1499, April 2012.
- [85] R. Dosselmann and X. D. Yang. A comprehensive assessment of the structural similarity index. *Signal, Image and Video Processing*, 5(1):81–91, November 2009.
- [86] H. R. Sheikh and A. C. Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2):430–444, February 2006.
- [87] W. O. Saxton and W. Baumeister. The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of Microscopy*, 127(2):127–138, August 1982.
- [88] P. A. Penczek. Resolution measures in molecular electron microscopy. *Methods in Enzymology*, 482:73–100, 2010.
- [89] K. Bodduna and R. Siddavatam. A novel algorithm for detection and removal of random valued impulse noise using cardinal splines. In *Proc. IEEE India Conference (INDICON)*, pages 1003–1008, Kochi, India, December 2012.
- [90] S. Jayasree, K. Bodduna, P. K. Pattnaik, and R. Siddavatam. An expeditious cum efficient algorithm for salt-and-pepper noise removal and edge-detail preservation using cardinal spline interpolation. *Journal of Visual Communication and Image Representation*, 25(6):1349–1365, August 2014.

- [91] T. Iijima. Basic theory on normalization of pattern (in case of typical one-dimensional pattern). *Bulletin of the Electrotechnical Laboratory*, 26:368–388, 1962. In Japanese.
- [92] E. A. Nadaraya. On estimating regression. *SIAM Journal on Theory of Probability and its Applications*, 9(1):141–142, 1964.
- [93] G. S. Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 64(4):359–372, December 1964.
- [94] C. J. Stone. Consistent nonparametric regression. *The Annals of Statistics*, 5(4):595–620, July 1977.
- [95] W. S. Cleveland and S. J. Devlin. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610, 1988.
- [96] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.
- [97] J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London, UK, 1996.
- [98] J. Astola, V. Katkovnik, and K. Egiazarian. *Local Approximation Techniques in Signal and Image Processing*, volume PM157. SPIE, Bellingham, WA, USA, 2006.
- [99] A. Foi. *Anisotropic Nonparametric Image Processing: Theory, Algorithms and Applications*. PhD thesis, Dipartimento Di Matematica, Politecnico di Milano, Italy, 2005.
- [100] P. Hall, S. J. Sheather, M. C. Jones, and J. S. Marron. On optimal data-based bandwidth selection in kernel density estimation. *Biometrika*, 78(2):263–269, June 1991.
- [101] J. Fan and I. Gijbels. Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(2):371–394, July 1995.

- [102] M. C. Jones, J. S. Marron, and S. J. Sheather. Progress in Data-based Bandwidth Selection for Kernel Density Estimation. Technical report, Dept. of Statistics, North Carolina State University, 1992.
- [103] S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3):683–690, 1991.
- [104] D. Ruppert, S. J. Sheather, and M. P. Wand. An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, 90(432):1257–1270, 1995.
- [105] M. C. Jones, J. S. Marron, and S. J. Sheather. A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, 91(433):401–407, 1996.
- [106] A. Goldenshluger and A. Nemirovski. On spatially adaptive estimation of nonparametric regression. *Mathematical Methods of Statistics*, 6(2):135–170, 1997.
- [107] V. Katkovnik. A new method for varying adaptive bandwidth selection. *IEEE Transactions on Image Processing*, 47(9):2567–2571, September 1999.
- [108] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola. Directional varying scale approximations for anisotropic signal processing. In *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, pages 101–104, Vienna, Austria, September 2004.
- [109] H. Takeda, S. Farsiu, and P. Milanfar. Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2):349–366, February 2007.
- [110] L. P. Yaroslavsky. *Digital Picture Processing: An Introduction*. Springer Verlag, 1985.
- [111] J. S. Lee. Digital image smoothing and the sigma filter. *Computer Vision, Graphics and Image Processing*, 24(2):255–269, November 1983.
- [112] S. M. Smith and J. M. Brady. SUSAN - A new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78, May 1997.

- [113] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, page 839–846, Bombay, India, January 1998.
- [114] M. Elad. On the origin of the bilateral filter and ways to improve it. *IEEE Transactions on Image Processing*, 11(10):1141–1151, October 2002.
- [115] J. Wang, D. Xu, C. Lang, and B. Li. An adaptive tone mapping method for displaying high dynamic range images. *Journal of Information Science and Engineering*, 26(3):977–990, May 2010.
- [116] B. Zhang and J. P. Allebach. Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE Transactions on Image Processing*, 17(5):664–678, May 2008.
- [117] M. Zhang and B. K. Gunturk. Compression artifact reduction with adaptive bilateral filtering. In M. Rabbani and R. L. Stevenson, editors, *Visual Communications and Image Processing*, volume 7257, pages 413 – 423, San Jose, CA, USA, 2009. International Society for Optics and Photonics, SPIE.
- [118] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, Cambridge, UK, 2001.
- [119] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, New York, 2002.
- [120] T. F. Chan and J. Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. SIAM, Philadelphia, 2005.
- [121] F. Guichard, L. Moisan, and J. M. Morel. A review of PDE models in image processing and image analysis. *Journal de Physique IV*, 12(1):137–154, March 2002.
- [122] N. Sochen, R. Kimmel, and A. M. Bruckstein. Diffusions and confusions in signal and image processing. *Journal of Mathematical Imaging and Vision*, 14(3):195–210, May 2001.

- [123] D. Barash. A fundamental relationship between bilateral filtering, adaptive smoothing and the nonlinear diffusion equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6):844–847, June 2002.
- [124] A. P. Witkin. Scale-space filtering. In *Proc. International Joint Conference on Artificial Intelligence*, volume 2, pages 945–951, Karlsruhe, West Germany, August 1983.
- [125] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50(5):363–370, August 1984.
- [126] T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3):234–254, March 1990.
- [127] L. Florack. *Image Structure*, volume 10 of *Computational Imaging and Vision*. Kluwer, Dordrecht, 1997.
- [128] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990.
- [129] F. Catté, P. L. Lions, J. M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, March 1991.
- [130] R. T. Whitaker and S. M. Pizer. A multi-scale approach to nonuniform diffusion. *CVGIP: Image Understanding*, 57(1):99–110, January 1993.
- [131] G.-H. Cottet and L. Germain. Image processing through reaction combined with nonlinear diffusion. *Mathematics of Computation*, 61(204):659–673, October 1993.
- [132] J. Weickert. Theoretical foundations of anisotropic diffusion in image processing. In W. Kropatsch, R. Klette, F. Solina, and R. Albrecht, editors, *Theoretical Foundations of Computer Vision*, volume 11 of *Computing Supplement*, pages 221–236. Springer, Vienna, 1996.
- [133] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart, 1998.

- [134] G. Steidl and T. Teuber. Diffusion tensors for processing sheared and rotated rectangles. *IEEE Transactions on Image Processing*, 18(12):2640–2648, December 2009.
- [135] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, November 1992.
- [136] P. Blomgren, T. F. Chan, P. Mulet, and C. K. Wong. Total variation image restoration: Numerical methods and extensions. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 384–387, Santa Barbara, CA, USA, October 1997.
- [137] P. Blomgren and T. F. Chan. Color TV: Total variation methods for restoration of vector valued images. *IEEE Transactions on Image Processing*, 7(3):304–309, March 1998.
- [138] R. R. Coifman and D. L. Donoho. Translation invariant denoising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets in Statistics*, volume 103 of *Lecture Notes in Statistics*, pages 125–150. Springer, New York, 1995.
- [139] C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238, March 1995.
- [140] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, volume 22 of *University Lecture Series*. AMS, Providence, 2001.
- [141] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, April 1997.
- [142] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling and Simulation*, 4(2):460–489, 2005.
- [143] E. Tadmor, S. Nezzar, and L. Vese. A multiscale image representation using hierarchical (BV-L2) decompositions. *Multiscale Modeling and Simulation*, 2(4):554–579, 2004.

- [144] G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In X.C. Tai, K. Morken, M. Lysaker, and K.A. Lie, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 5567 of *Lecture Notes in Computer Science*, pages 477–489. Springer, Berlin, Heidelberg, 2009.
- [145] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. Wiley, New York, 1948.
- [146] F. Luisier. *The SURE-LET Approach to Image Denoising*. PhD thesis, 'Ecole polytechnique f'ed'erale de Lausanne, Switzerland, 2010.
- [147] J. S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2):165–168, March 1980.
- [148] D. C. Knill and W. Richards. *Perception as Bayesian Inference*. Cambridge University Press, Cambridge, UK, 1996.
- [149] J. Bernardo and A. F. M. Smith. *Bayesian Theory*. Wiley Series in Probability and Statistics. Wiley, Chichester, 2000.
- [150] A. Hyvärinen, J. Hurri, and P. O. Hoyer. *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision*, volume 31 of *Computational Imaging and Vision*. Springer, London, 2009.
- [151] P. M. Lee. *Bayesian Statistics: An Introduction*. Wiley, Chichester, fourth edition, 2012.
- [152] W.H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America A*, 62(1):55–69, January 1972.
- [153] J. Shen. Inpainting and the fundamental problem of image processing. *SIAM news*, 36(5):1–4, June 2003.
- [154] S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 860–867, San Diego, CA, USA, June 2005.

- [155] C. S. Won and R. M. Gray. *Stochastic Image Processing*. Information Technology: Transmission, Processing and Storage. Kluwer, Norwell, MA, USA, 2004.
- [156] Y. Weiss and W. T. Freeman. What makes a good model of natural images? In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, Minneapolis, MN, USA, June 2007.
- [157] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1033–1038, Kerkyra, Greece, September 1999.
- [158] J. S. De Bonet. Noise Reduction Through Detection of Signal Redundancy. Technical report, Rethinking Artificial Intelligence, MIT AI Lab, 1997.
- [159] N. Lu. *Fractal Imaging*. Academic Press, London, 1997.
- [160] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 60–65, San Diego, CA, USA, June 2005.
- [161] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, 4(2):490–530, 2005.
- [162] C. Kervrann and J. Boulanger. Optimal spatial adaptation for patch-based image denoising. *IEEE Transactions on Image Processing*, 15(10):2866–2878, October 2006.
- [163] C. Kervrann and J. Boulanger. Local adaptivity to variable smoothness for exemplar-based image regularization and representation. *International Journal of Computer Vision*, 79(1):45–69, August 2008.
- [164] C. A. Deledalle, V. Duval, and J. Salmon. Non-local methods with shape-adaptive patches (NLM-SAP). *Journal of Mathematical Imaging and Vision*, 43(2):103–120, June 2012.
- [165] D. Van De Ville and M. Kocher. SURE-based non-local means. *IEEE Signal Processing Letters*, 16(11):973–976, November 2009.

- [166] V. Doré and M. Cheriet. Robust NL-means filter with optimal pixelwise smoothing parameter for statistical image denoising. *IEEE Transactions on Signal Processing*, 57(5):1703–1716, May 2009.
- [167] J. Froment. Parameter-free fast pixelwise non-local means denoising. *Image Processing On Line*, 4:300–326, November 2014.
- [168] S. P. Awate and R. T. Whitaker. Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 44–51, San Diego, CA, USA, June 2005.
- [169] S. Kindermann, S. Osher, and P. M. Jones. Deblurring and denoising of images by nonlocal functionals. *Multiscale Modeling and Simulation*, 4(4):1091–1115, 2005.
- [170] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling and Simulation*, 7(3):1005–1028, November 2008.
- [171] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. *Multiscale Modeling and Simulation*, 6(2):595–630, July 2007.
- [172] G. Peyré, S. Bogleux, and L. Cohen. Non-local regularization of inverse problems. In D. Forsyth, P. Torr, and A. Zisserman, editors, *Computer Vision (ECCV)*, Lecture Notes in Computer Science, pages 57–68, Berlin, Heidelberg, 2008. Springer.
- [173] A. Elmoataz, O. Lezoray, and S. Bogleux. Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing. *IEEE Transactions on Image Processing*, 17(7):1047–1060, July 2008.
- [174] P. Chatterjee and P. Milanfar. Patch-based near-optimal image denoising. *IEEE Transactions on Image Processing*, 21(4):1635–1649, April 2012.
- [175] A. Levin and B. Nadler. Natural image denoising: Optimality and inherent bounds. In *Proc. IEEE Conference on Computer Vision and*

- Pattern Recognition (CVPR)*, pages 2833–2840, Colorado Springs, CO, USA, June 2011.
- [176] S. Lyu and E. P. Simoncelli. Modeling multiscale subbands of photographic images with fields of Gaussian scale mixtures. *IEEE Transactions on Image Processing*, 31(4):693–706, April 2009.
- [177] G. Yu, G. Sapiro, and S. Mallat. Image modeling and enhancement via structured sparse model selection. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1641–1644, Hong Kong, China, September 2010.
- [178] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity. *IEEE Transactions on Image Processing*, 21(5):2481–2499, May 2012.
- [179] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 479–486, Barcelona, Spain, November 2011.
- [180] A. De Decker, J. A. Lee, and M. Verleysen. Performance assessment of patch-based bilateral denoising. In *Proc. Medical Image Analysis and Description for Diagnosis Systems (MIAD)*, pages 52–61, Porto, Portugal, January 2009. Insticc Press.
- [181] A. de Decker, J. A. Lee, and M. Verlysen. Variance stabilizing transformations in patch-based bilateral filters for Poisson noise image denoising. In *Proc. IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3673–3676, Minneapolis, MN, USA, September 2009.
- [182] A. Buades, J. Duran, and J. Navarro. Motion-compensated spatio-temporal filtering for multi-image and multimodal super-resolution. *International Journal of Computer Vision*, 127(10):1474–1500, July 2019.
- [183] M. Mignotte. A non-local regularization strategy for image deconvolution. *Pattern Recognition Letters*, 29(16):2206–2212, December 2008.
- [184] O. Lezoray, V. T. Ta, and A. Elmoataz. Nonlocal graph regularization for image colorization. In *Proc. IEEE International Conference on*

- Pattern Recognition (ICPR)*, pages 1–4, Tampa, FL, USA, December 2008.
- [185] Y. Chen, J. Ma, Q. Feng, L. Luo, P. Shi, and W. Chen. Nonlocal prior Bayesian tomographic reconstruction. *Journal of Mathematical Imaging and Vision*, 30(2):133–146, November 2007.
- [186] A. Buades. *Image and Film Denoising by Non-local Means*. PhD thesis, Dpt. de Matemàtiques i Informàtica, Universitat de les Illes Balears, Spain, 2006.
- [187] F. Guichard and J. M. Morel. A note on two classical enhancement filters and their associated PDEs. *International Journal of Computer Vision*, 52(2-3):153–160, May 2003.
- [188] J. G. M. Schavemaker, M. J. T. Reinders, J. J. Gerbrands, and E. Backer. Image sharpening by morphological filtering. *Pattern Recognition*, 33(6):997–1012, June 2000.
- [189] A. Buades, B. Coll, and J. M. Morel. Neighborhood filters and PDEs. *Numerische Mathematik*, 105(1):1–34, October 2006.
- [190] A. Buades, B. Coll, and J. M. Morel. The staircasing effect in neighborhood filters and its solution. *IEEE Transactions on Image Processing*, 15(6):1499–1505, June 2006.
- [191] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In *Proc. ACM Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 257–266, San Antonio, Texas, USA, July 2002.
- [192] S. Paris, P. Kornprobst, and J. Tumblin. Bilateral filtering: Theory and applications. *Foundations and Trends in Computer Graphics and Vision*, 4(1):1–73, August 2009.
- [193] T. Brox, M. Welk, G. Steidl, and J. Weickert. Equivalence results for TV diffusion and TV regularisation. In L. D. Griffin and M. Lillholm, editors, *Scale-Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 86–100. Springer, Berlin, Heidelberg, 2003.

- [194] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. *SIAM Journal on Numerical Analysis*, 42(2):686–713, 2004.
- [195] P. Peter, J. Weickert, A. Munk, T. Krivobokova, and H. Li. Justifying tensor-driven diffusion from structure-adaptive statistics of natural images. In X.-C. Tai, E. Bae, T. F. Chan, and M. Lysaker, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, volume 8932 of *Lecture Notes in Computer Science*, pages 263–277. Springer, Berlin, 2015.
- [196] M. Lebrun. *From Theory to Practice, a Tour of Image Denoising*. PhD thesis, Centre de Mathématiques et de Leurs Applications, École Normale Supérieure Paris-Saclay, France, 2014. tel-01114299.
- [197] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.
- [198] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7):909–996, October 1988.
- [199] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, second edition, 1999.
- [200] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.
- [201] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Communications on Pure and Applied Mathematics*, 44(2):141–183, March 1991.
- [202] R. Coifman and Y. Meyer. *Ondelettes Et Opérateurs III*. Opérateurs multilinéaires, Actualités Mathématiques. Hermann, Paris, 1991.
- [203] J. P. Antoine, P. Vandergheynst, and R. Murenzi. Two-dimensional directional wavelets in image processing. *International Journal of Imaging Systems and Technology*, 7(3):152–165, September 1996.

- [204] D. Labate, W. Q. Lim, G. Kutyniok, and G. Weiss. Sparse multidimensional representation using shearlets. In M. Papadakis, A. F. Laine, and M. A. Unser, editors, *Wavelets XI*, volume 5914, pages 254 – 262, San Diego, CA, USA, September 2005. International Society for Optics and Photonics, SPIE.
- [205] F. Malgouyres. Noise selection approach to image restoration. In A. F. Laine, M. A. Unser, and A. Aldroubi, editors, *Wavelets: Applications in Signal and Image Processing IX*, volume 4478, pages 34 – 41, San Diego, CA, USA, December 2001. International Society for Optics and Photonics, SPIE.
- [206] L. P. Yaroslavsky. Local adaptive image restoration and enhancement with the use of DFT and DCT in a running window. In M. A. Unser, A. Aldroubi, and A. F. Laine, editors, *Wavelet Applications in Signal and Image Processing IV*, volume 2825, pages 2–13, Denver, CO, USA, October 1996. International Society for Optics and Photonics, SPIE.
- [207] L. P. Yaroslavsky and M. Eden. *Fundamentals of Digital Optics: Digital Signal Processing in Optics and Holography*. Birkhaeuser, Boston, 1996.
- [208] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, September 1994.
- [209] D. L. Donoho. Denoising by soft thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, May 1995.
- [210] R. Oktem, L. Yaroslavsky, and K. Egiazarian. Signal and image denoising in transform domain and wavelet shrinkage: A comparative study. In *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, pages 1–4, Rhodes, Greece, September 1998.
- [211] L. P. Yaroslavsky, , K. O. Egiazarian, and J. T. Astola. Transform domain image restoration methods: Review, comparison, and interpretation. In E. R. Dougherty and J. T. Astola, editors, *Nonlinear Image Processing and Pattern Analysis XII*, volume 4304, pages 155–169, San Jose, CA, USA, May 2001. International Society for Optics and Photonics, SPIE.

- [212] D. L. Donoho and I. M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432):1200–1224, 1995.
- [213] S. Durand and M. Nikolova. Restoration of wavelet coefficients by minimizing a specially designed objective function. In O. Faugeras and N. Paragios, editors, *Proc. IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision (VLSM)*, Nice, France, October 2003. INRIA.
- [214] T. F. Chan and H. M. Zhou. Total variation improved wavelet thresholding in image compression. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 391–394, Vancouver, BC, Canada, September 2000.
- [215] S. Durand and J. Froment. Reconstruction of wavelet coefficients using total-variation minimization. *SIAM Journal on Scientific Computing*, 24(5):1754–1767, 2003.
- [216] S. Lintner and F. Malgouyres. Solving a variational image restoration model which involves L constraints. *Inverse Problems*, 20(3):815–831, March 2004.
- [217] A. Chambolle, R. A. De Vore, N. Y. Lee, and B. J. Lucier. Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7(3):319–335, March 1998.
- [218] E. P. Simoncelli and E. H. Adelson. Noise removal via Bayesian wavelet coring. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 1, pages 379–382, Lausanne, Switzerland, September 1996.
- [219] L. Sendur and I. W. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on Image Processing*, 50(11):2744–2756, November 2002.
- [220] A. Pizurica and W. Philips. Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising. *IEEE Transactions on Image Processing*, 15(3):654–665, March 2006.

- [221] T. Blu and F. Luisier. The SURE-LET approach to image denoising. *IEEE Transactions on Image Processing*, 16(11):2778–2786, November 2007.
- [222] F. Yan, L. Cheng, and S. Peng. A new interscale and intrascale orthonormal wavelet thresholding for SURE-based image denoising. *IEEE Signal Processing Letters*, 15:139–142, January 2008.
- [223] H. Yu, L. Zhao, and H. Wang. Image denoising using trivariate shrinkage filter in the wavelet domain and joint bilateral filter in the spatial domain. *IEEE Transactions on Image Processing*, 18(10):2364–2369, October 2009.
- [224] D. D. Muresan and T. W. Parks. Adaptive principal components and image denoising. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 101–104, Barcelona, Spain, September 2003.
- [225] P. O. Hoyer. Independent component analysis in image denoising. Master’s thesis, Department of Computer Science and Engineering, Helsinki University of Technology, Finland, 1999.
- [226] A. Foi, K. Dabov, V. Katkovnik, and K. Egiazarian. Shape-adaptive DCT for denoising and image reconstruction. In N. M. Nasrabadi, S. A. Rizvi, E. R. Dougherty, J. T. Astola, and K. O. Egiazarian, editors, *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, volume 6064, pages 203 – 214, San Jose, CA, USA, February 2006. International Society for Optics and Photonics, SPIE.
- [227] A. Foi. *Pointwise Shape-adaptive DCT Image Filtering and Signal-dependent Noise Estimation*. PhD thesis, Department of Information Technology, Tampere University of Technology, Finland, 2007.
- [228] P. Chatterjee and P. Milanfar. Clustering-based denoising with locally learned dictionaries. *IEEE Transactions on Image Processing*, 18(7):1438–1451, July 2009.
- [229] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, December 2006.

- [230] J. Mairal, G. Sapiro, and M. Elad. Learning multiscale sparse representations for image and video restoration. *Multiscale Modeling and Simulation*, 7(1):214–241, December 2007.
- [231] W. Dong, G. Shi, and X. Li. Nonlocal image restoration with bilateral variance estimation: A low-rank approach. *IEEE Transactions on Image Processing*, 22(2):700–711, February 2013.
- [232] C. Knaus. *Dual-Domain Image Denoising*. PhD thesis, Department of Computer Science, University of Bern, Switzerland, 2016.
- [233] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. A nonlocal and shape-adaptive transform-domain collaborative filtering. In *Proc. International Workshop on Local and Non-Local Approximations in Image Processing (LNLA)*, Lausanne, Switzerland, August 2008.
- [234] S. Mallat. Geometrical grouplets. *Applied and Computational Harmonic Analysis*, 26(2):161–180, March 2009.
- [235] G. Peyré. Image processing with nonlocal spectral bases. *Multiscale Modeling and Simulation*, 7(2):703–730, July 2008.
- [236] G. Steidl and J. Weickert. Relations between soft wavelet shrinkage and total variation denoising. In L. Van Gool, editor, *Pattern Recognition (DAGM)*, volume 2449 of *Lecture Notes in Computer Science*, pages 198–205. Springer, Berlin, 2002.
- [237] P. Mrázek, J. Weickert, and G. Steidl. Correspondences between wavelet shrinkage and nonlinear diffusion. In L. D. Griffin and M. Lillholm, editors, *Scale-Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 101–116. Springer, Berlin, 2003.
- [238] B. Dong, Q. Jiang, and Z. Shen. Image restoration: Wavelet frame shrinkage, nonlinear evolution PDEs, and beyond. *Multiscale Modeling and Simulation*, 15(1):606–660, January 2017.
- [239] N. Pierazzo, M. Lebrun, M.E. Rais, J. M. Morel, and G. Facciolo. Non-local dual image denoising. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 813–817, Paris, France, October 2014.

- [240] N. Pierazzo. *Advances in Image Denoising*. PhD thesis, Centre de Mathématiques et de Leurs Applications, École Normale Supérieure Paris-Saclay, France, 2016. NNT:2016SACLN036. tel-01504686.
- [241] C. Knaus and M. Zwicker. Dual-domain image denoising. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 440–444, Melbourne, VIC, Australia, September 2013.
- [242] C. Knaus and M. Zwicker. Progressive image denoising. *IEEE Transactions on Image Processing*, 23(7):3114–3125, July 2014.
- [243] C. Knaus and M. Zwicker. Dual-domain filtering. *SIAM Journal on Imaging Sciences*, 8(3):1396–1420, April 2015.
- [244] H. C. Burger and S. Harmeling. Improving denoising algorithms via a multi-scale meta-procedure. In R. Mester and M. Felsberg, editors, *Pattern Recognition (DAGM)*, volume 6835 of *Lecture Notes in Computer Science*, pages 206–215, Berlin, Heidelberg, 2011. Springer.
- [245] F. Estrada, D. Fleet, and A. Jepson. Stochastic image denoising. In A. Cavallaro, S. Prince, and D. Alexander, editors, *Proc. British Machine Vision Conference (BMVC)*, pages 117.1–117.11, London, UK, September 2009. BMVA Press.
- [246] D. Gnanadurai and V. Sadasivam. Image de-noising using double density wavelet transform based adaptive thresholding technique. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(1):141–152, March 2005.
- [247] H. Q. Li, S. Q. Wang, and C. Z. Deng. New image denoising method based wavelet and curvelet transform. In *Proc. IEEE International Conference on Information Engineering (ICIE)*, pages 136–139, Taiyuan, Chanxi, China, July 2009.
- [248] J. L. Starck, E. J. Candès, and D. L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6):670–684, June 2002.
- [249] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 2272–2279, Kyoto, Japan, October 2009.

- [250] J. Sulam, B. Ophir, and M. Elad. Image denoising through multi-scale learnt dictionaries. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 808–812, Paris, France, October 2014.
- [251] G. Facciolo, N. Pierazzo, and J. M. Morel. Conservative scale recombination for multiscale denoising (the devil is in the high frequency detail). *SIAM Journal on Imaging Sciences*, 10(3):1603–1626, July 2017.
- [252] N. Pierazzo, J. M. Morel, and G. Facciolo. Multi-Scale DCT Denoising. *Image Processing On Line*, 7:288–308, October 2017.
- [253] H. Burger, S. Christopher, and S. Harmeling. Learning how to combine internal and external denoising methods. In J. Weickert, M. Hein, and B. Schiele, editors, *Pattern Recognition (GCPR)*, volume 8142 of *Lecture Notes in Computer Science*, pages 121–130, Berlin, Heidelberg, 2013. Springer.
- [254] H. Huanjing, X. Sun, J. Yang, and F. Wu. CID: Combined image denoising in spatial and frequency domains using web images. In *Proc. IEEE International Conference of Computer Vision and Pattern Recognition (CVPR)*, pages 2933–2940, Columbus, OH, USA, June 2014.
- [255] C. Lee, C. Lee, and C. S. Kim. An MMSE approach to nonlocal image denoising: Theory and practical implementation. *Journal of Visual Communication and Image Representation*, 23(3):476–490, April 2012.
- [256] J. Hays and A. A. Efros. Scene completion using millions of photographs. *ACM Transactions on Graphics*, 26(3):4–11, July 2007.
- [257] A. Torralba, R. Fergus, and W. T. Freeman. 80 million tiny images: A large data set for nonparametric object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11):1958–1970, November 2008.
- [258] A. Levin, B. Nadler, F. Durand, and W. T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, *Computer Vision (ECCV)*, volume 7576 of *Lecture Notes in Computer Science*, pages 73–86, Berlin, Heidelberg, 2012. Springer.

- [259] N. Pierazzo and M. Rais. Boosting “shotgun denoising” by patch normalization. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1115–1119, Melbourne, VIC, Australia, September 2013.
- [260] S. Pyatykh and J. Hesser. MMSE estimation for Poisson noise removal in images. *arXiv:1512.00717*, December 2015.
- [261] A. K. Oh, T. Z. Harmany, and R. M. Willett. To e or not to e in Poisson image reconstruction. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 2829–2833, Paris, France, October 2014.
- [262] K. Timmermann and R. D. Nowak. Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging. *IEEE Transactions on Information Theory*, 45(3):846–842, April 1999.
- [263] E. D. Kolaczyk. Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds. *Statistica Sinica*, 9(1):119–135, January 1999.
- [264] R. D. Nowak. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing*, 8(10):1408–1419, October 1999.
- [265] A. Antoniadis and T. Sapatinas. Wavelet shrinkage for natural exponential families with quadratic variance functions. *Biometrika*, 88(3):805–820, October 2001.
- [266] H. Lu, Y. Kim, and J. M. M. Anderson. Improved Poisson intensity estimation: Denoising application using Poisson data. *IEEE Transactions on Image Processing*, 13(8):1128–1135, August 2004.
- [267] E. D. Kolaczyk and R. D. Nowak. Multiscale likelihood analysis and complexity penalized estimation. *The Annals of Statistics*, 32(2):500–527, April 2004.
- [268] P. N. Patil and A. T. A. Wood. Counting process intensity estimation by orthogonal wavelet methods. *Bernoulli*, 10(1):1–24, February 2004.

- [269] S. Sardy, A. Antoniadis, and P. Tseng. Automatic smoothing with wavelets for a wide class of distributions. *Journal of Computational and Graphical Statistics*, 13(2):399–421, 2004.
- [270] R. M. Willett and R. D. Nowak. Multiscale Poisson intensity and density estimation. *IEEE Transactions on Information Theory*, 53(9):3171–3187, September 2007.
- [271] S. Lefkimmiatis, P. Maragos, and G. Papandreou. Bayesian inference on multiscale models for Poisson intensity estimation: Applications to photon-limited image denoising. *IEEE Transactions on Image Processing*, 18(8):1724–1741, August 2009.
- [272] F. Luisier, C. Vonesch, T. Blu, and M. Unser. Fast interscale wavelet denoising of Poisson-corrupted images. *Signal Processing*, 90(2):415–427, February 2010.
- [273] R. M. Willett and R. D. Nowak. Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging. *IEEE Transactions on Medical Imaging*, 22(3):332–350, March 2003.
- [274] M. Collins, S. Dasgupta, and R. E. Schapire. A generalization of principal components analysis to the exponential family. In S. Becker, S. Thrun, and K. Obermayer, editors, *Advances in Neural Information Processing Systems (NIPS)*, pages 617–624, Vancouver, British Columbia, Canada, December 2001.
- [275] J. Salmon, Z. Harmany, C. A. Deledalle, and R. Willett. Poisson noise reduction with non-local PCA. *Journal of Mathematical Imaging and Vision*, 48(2):279–294, February 2014.
- [276] R. Giryes and M. Elad. Sparsity-based Poisson denoising with dictionary learning. *IEEE Transactions on Image Processing*, 23(12):5057–5069, October 2014.
- [277] F. J. Anscombe. The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, 35(3/4):246–254, December 1948.
- [278] M. Mäkitalo and A. Foi. A closed-form approximation of the exact unbiased inverse of the Anscombe variance-stabilizing transformation.

- IEEE Transactions on Image Processing*, 20(9):2697–2698, September 2011.
- [279] L. Azzari and A. Foi. Variance stabilization for noisy+estimate combination in iterative Poisson denoising. *IEEE Signal Processing Letters*, 23(8):1086–1090, August 2016.
- [280] M. Mäkitalo and A. Foi. On the inversion of the Anscombe transformation in low-count Poisson image denoising. In *Proc. IEEE International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*, pages 26–32, Tuusula, Finland, August 2009.
- [281] M. Mäkitalo and A. Foi. Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. *IEEE Transactions on Image Processing*, 22(1):91–103, January 2013.
- [282] M. Mäkitalo and A. Foi. Noise parameter mismatch in variance stabilization, with an application to Poisson-Gaussian noise estimation. *IEEE Transactions on Image Processing*, 23(12):5348–5359, December 2014.
- [283] P. Fryzlewicz and G. P. Nason. A Haar-Fisz algorithm for Poisson intensity estimation. *Journal of Computational and Graphical Statistics*, 13(3):621–638, 2004.
- [284] B. Zhang, J. M. Fadili, and J. L. Starck. Wavelets, ridgelets, and curvelets for Poisson noise removal. *IEEE Transactions on Image Processing*, 17(7):1093–1108, July 2008.
- [285] P. Besbeas, I. De Feis, and T. Sapatinas. A comparative simulation study of wavelet shrinkage estimators for Poisson counts. *International Statistical Review*, 72(2):209–237, August 2004.
- [286] M. Jansen. Multiscale Poisson data smoothing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):27–48, February 2006.
- [287] M. Mäkitalo. *Exact Unbiased Inverse of the Anscombe Transformation and its Poisson-Gaussian Generalization*. PhD thesis, Department of Information Technology, Tampere University of Technology, Finland, 2013.

- [288] S. V. Venkatakrisnan, C. A. Bouman, and B. Wohlberg. Plug-and-play priors for model based reconstruction. In *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 945–948, Austin, TX, USA, December 2013.
- [289] A. Rond, R. Giryes, and M. Elad. Poisson inverse problems by the plug-and-play scheme. *Journal of Visual Communication and Image Representation*, 41:96–108, November 2016.
- [290] M. Lebrun, A. Buades, and J. M. Morel. Implementation of the non-local Bayes (NL-Bayes) image denoising algorithm. *Image Processing On Line*, 3:1–42, June 2013.
- [291] P. Chatterjee and P. Milanfar. Is denoising dead? *IEEE Transactions on Image Processing*, 19(4):895–911, April 2010.
- [292] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, July 2017.
- [293] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila. Noise2Noise: Learning image restoration without clean data. In J. Dy and A. Krause, editors, *Proc. International Conference on Machine Learning (ICML)*, volume 80 of *Proceedings of Machine Learning Research*, pages 2965–2974, Stockholm, Sweden, July 2018.
- [294] D. Valsesia, G. Fracastoro, and E. Magli. Image denoising with graph-convolutional neural networks. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 2399–2403, Taipei, Taiwan, September 2019.
- [295] C. Cruz, A. Foi, V. Katkovnik, and K. Egiazarian. Nonlocality-reinforced convolutional neural networks for image denoising. *IEEE Signal Processing Letters*, 25(8):1216–1220, August 2018.
- [296] Y. Lu, A. Zhong, Q. Li, and B. Dong. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. In J. Dy and A. Krause, editors, *Proc. International Conference on Machine Learning (ICML)*, volume 80 of *Proceedings of Machine Learning Research*, pages 3276–3285, Stockholm, Sweden, July 2018.

- [297] T. Alt and J. Weickert. Translating diffusion, wavelets, and regularization into residual networks. *arXiv:2002.02753*, February 2020.
- [298] T. Alt, P. Peter, J. Weickert, and K. Schrader. Translating numerical concepts for PDEs into neural architectures. In A. Elmoataz, J. Fadili, Y. Quéau, J. Rabin, and L. Simon, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 12679 of *Lecture Notes in Computer Science*, pages 294–306. Springer, Cham, 2021.
- [299] Y. Chen and T. Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1256–1272, June 2017.
- [300] T. Alt and J. Weickert. Learning a generic adaptive wavelet shrinkage function for denoising. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2018–2022, Barcelona, Spain, May 2020.
- [301] L. Ruthotto and E. Haber. Deep neural networks motivated by partial differential equations. *Journal of Mathematical Imaging and Vision*, 62:352–364, April 2020.
- [302] T. Alt and J. Weickert. Learning integrodifferential models for image denoising. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2045–2049, Toronto, ON, Canada, June 2021.
- [303] V. Aurich and J. Weule. Non-linear Gaussian filters performing edge preserving diffusion. In G. Sagerer, S. Posch, and F. Kummert, editors, *Mustererkennung 1995*, pages 538–545. Springer, Berlin, 1995.
- [304] P. Irrera, I. Bloch, and M. Delplanque. A flexible patch based approach for combined denoising and contrast enhancement of digital X-ray images. *Medical Image Analysis*, 28:33–45, February 2016.
- [305] M. H. Alkinani and M. R. El-Sakka. Patch-based models and algorithms for image denoising: A comparative review between patch-based images denoising methods for additive noise reduction. *EURASIP Journal on Image and Video Processing*, 2017(1):1–27, August 2017.

- [306] M. Niknejad, J. M. Bioucas-Dias, and M. A. T. Figueiredo. Class-specific Poisson denoising by patch-based importance sampling. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1247–1251, Beijing, China, September 2017.
- [307] M. A. T. Figueiredo. Synthesis versus analysis in patch-based image priors. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1338–1342, New Orleans, LA, USA, March 2017.
- [308] R. Ali, P. Yunfeng, and R. U. Amin. A novel Bayesian patch-based approach for image denoising. *IEEE Access*, 8:38985–38994, February 2020.
- [309] C. Coartazar, M. Elgueta, J. D. Rossi, and N. Wolanski. How to approximate the heat equation with Neumann boundary conditions by nonlocal diffusion problems. *Archive for Rational Mechanics and Analysis*, 187(1):137–156, January 2008.
- [310] J. Lellmann, K. Papafitsoros, C. Schönlieb, and D. Spector. Analysis and application of a nonlocal Hessian. *SIAM Journal on Imaging Sciences*, 8(4):2161–2202, July 2015.
- [311] T. Mengesha and D. Spector. Localization of nonlocal gradients in various topologies. *Calculus of Variations and Partial Differential Equations*, 52:253–279, January 2015.
- [312] Y. Hafiene, J. M. Fadili, and A. Elmoataz. Continuum limits of nonlocal p-Laplacian variational problems on graphs. *SIAM Journal on Imaging Sciences*, 12(4):1772–1807, August 2019.
- [313] C. Ren, X. He, Y. Pu, and T. Q. Nguyen. Enhanced non-local total variation model and multi-directional feature prediction prior for single image super resolution. *IEEE Transactions on Image Processing*, 28(4):3778–3793, August 2019.
- [314] J. M. Mazón, J. D. Rossi, and J. Toledo. Nonlocal perimeter, curvature and minimal surfaces for measurable sets. *Journal d’Analyse Mathématique*, 138(1):235–279, July 2019.

- [315] Z. Shi, S. Osher, and W. Zhu. Generalization of the weighted nonlocal Laplacian in low dimensional manifold model. *Journal of Scientific Computing*, 75(2):638–656, May 2018.
- [316] L. Wang, S. Zhou, X. Lin, T. Qi, X. Yin, and Y. Yang. A novel adaptive image zooming method based on nonlocal Cahn–Hilliard equation. *Knowledge-Based Systems*, 166:118–131, February 2020.
- [317] I. Cohen and G. Gilboa. Introducing the p-Laplacian spectra. *Signal Processing*, 167:107281, February 2020.
- [318] M. Cárdenas. *Nonlocal Evolutions in Image Processing*. PhD thesis, Department of Mathematics and Computer Science, Saarland University, Germany, July 2018.
- [319] M. Lebrun. An analysis and implementation of the BM3D image denoising method. *Image Processing On Line*, 2:175–213, August 2012.
- [320] Y. Hou, J. Xu, M. Liu, G. Liu, L. Liu, F. Zhu, and L. Shao. NLH: A blind pixel-level non-local method for real-world image denoising. *IEEE Transactions on Image Processing*, 29:5121–5135, March 2020.
- [321] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. IEEE Sixth International Conference on Computer Vision (ICCV)*, pages 839–846, Bombay, India, January 1998.
- [322] A. Buades, B. Coll, and J. M. Morel. Non-local means denoising. *Image Processing On Line*, 1:208–212, September 2011.
- [323] Eastman Kodak Company. Kodak true color image suite. <http://r0k.us/graphics/kodak/>, 1999. Online.
- [324] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: Can plain neural networks compete with BM3D? In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2392–2399, Providence, RI, USA, June 2012.
- [325] P. Chatterjee and P. Milanfar. Patch-based near-optimal image denoising. *IEEE Transactions on Image Processing*, 21(4):1635–1649, April 2012.

- [326] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, November 2003.
- [327] M. Mäkitalo and A. Foi. Poisson-Gaussian denoising using the exact unbiased inverse of the generalized Anscombe transformation. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1081–1084, Kyoto, Japan, March 2012.
- [328] J. Froment. Parameter-free fast pixelwise non-local means denoising. *Image Processing On Line*, 4:300–326, November 2014.
- [329] S. Laine, T. Karras, J. Lehtinen, and T. Aila. High-quality self-supervised deep image denoising. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems (NIPS)*, pages 6970–6980, Vancouver, Canada, December 2019.
- [330] A. Davy and T. Ehret. GPU acceleration of NL-means, BM3D and VBM3D. *Journal of Real-Time Image Processing*, 18:57–74, February 2020.
- [331] A. I. El-Fallah and G. E. Ford. The evolution of mean curvature in image filtering. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 1, pages 298–302, Austin, TX, USA, November 1994.
- [332] M. Nitzberg and T. Shiota. Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):826–833, August 1992.
- [333] K.O. Reidel. Corner-preserving anisotropic diffusion and junction detection using the structure tensor. In W. Förstner, J. M. Buhmann, A. Faber, and P. Faber, editors, *Mustererkennung*, pages 164–171, Berlin, Heidelberg, 1999. Springer.
- [334] H. Schar. Diffusion-like reconstruction schemes from linear data models. In K. Franke, K.-R. Müller, B. Nickolay, and R. Schäfer, editors, *Pattern Recognition (DAGM)*, volume 4174 of *Lecture Notes in Computer Science*, pages 51–60. Springer, Berlin, 2006.

- [335] E. Franken and R. Duits. Crossing-preserving coherence-enhancing diffusion on invertible orientation scores. *International Journal of Computer Vision*, 85(253-278), December 2009.
- [336] S. Zimmer, S. Didas, and J. Weickert. A rotationally invariant block matching strategy improving image denoising with non-local means. In *Proc. International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*, Lausanne, Switzerland, August 2008.
- [337] J. Weickert, M. Welk, and M. Wickert. L2-stable nonstandard finite differences for anisotropic diffusion. In A. Kuijper, T. Pock, K. Bredies, and H. Bischof, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 7893 of *Lecture Notes in Computer Science*, pages 380–391. Springer, Berlin, 2013.
- [338] S. Grewenig, S. Zimmer, and J. Weickert. Rotationally invariant similarity measures for nonlocal image denoising. *Journal of Visual Communication and Image Representation*, 22(2):117–130, February 2011.
- [339] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser. Multiframe SURE-LET denoising of timelapse fluorescence microscopy images. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 149–152, Paris, France, May 2008.
- [340] A. M. Hasan, A. Melli, K. A. Wahid, and P. Babyn. Denoising low-dose CT images using multi-frame blind source separation and block matching filter. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2(4):279–287, July 2018.
- [341] W. Dong, G. Li, G. Shi, X. Li, and Y. Ma. Low-rank tensor approximation with Laplacian scale mixture modeling for multiframe image denoising. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 442–449, Santiago, Chile, December 2015.
- [342] W. Dong, T. Huang, G. Shi, Y. Ma, and X. Li. Robust tensor approximation with Laplacian scale mixture modeling for multiframe image and video denoising. *IEEE Journal of Selected Topics in Signal Processing*, 12(6):1435–1448, December 2018.

- [343] R. Hao and Z. Su. A patch-based low-rank tensor approximation model for multiframe image denoising. *Journal of Computational and Applied Mathematics*, 329:125–133, February 2018.
- [344] L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu. Sparsity based denoising of spectral domain optical coherence tomography images. *Biomedical Optics Express*, 3(5):927–942, 2012.
- [345] E. Gil-Rodrigo, J. Portilla, D. Miraut, and R. Suarez-Mesa. Efficient joint Poisson-Gauss restoration using multi-frame L2-relaxed-L0 analysis-based sparsity. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1385–1388, Brussels, Belgium, September 2011.
- [346] L. Zhang, S. Vaddadi, H. Jin, and S. K. Nayar. Multiple view image denoising. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1542–1549, Miami, FL, USA, June 2009.
- [347] A. Buades, B. Coll, and J. M. Morel. Denoising image sequences does not require motion estimation. In *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 70–74, Como, Italy, September 2005.
- [348] M. Tico. Multi-frame image denoising and stabilization. In *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, pages 1–4, Lausanne, Switzerland, August 2008.
- [349] K. Dabov, A. Foi, and K. O. Egiazarian. Video denoising by sparse 3d transform-domain collaborative filtering. In *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, pages 145–149, Poznan, Poland, September 2007.
- [350] T. Buades, Y. Lou, J. M. Morel, and Z. Tang. A note on multi-image denoising. In *Proc. IEEE International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*, pages 1–15, Tuusula, Finland, August 2009.
- [351] A. Buades, Y. Lou, J. M. Morel, and Z. Tang. Multi Image Noise Estimation and Denoising. *HAL Archives*, hal-00510866, version 1, August 2010.

- [352] A. Buades, J. L. Lisani, and M. Miladinović. Patch-based video denoising with optical flow estimation. *IEEE Transactions on Image Processing*, 25(6):2573–2586, June 2016.
- [353] A. Davy, T. Ehret, J. M. Morel, P. Arias, and G. Facciolo. A non-local CNN for video denoising. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 2409–2413, Taipei, Taiwan, September 2019.
- [354] A. Davy, T. Ehret, J. M. Morel, P. Arias, and G. Facciolo. Video denoising by combining patch search and CNNs. *Journal of Mathematical Imaging and Vision*, 63:73–88, October 2020.
- [355] C. Godard, K. Matzen, and M. Uyttendaele. Deep burst denoising. In V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, editors, *Computer Vision (ECCV)*, volume 11219 of *Lecture Notes in Computer Science*, pages 560–577. Springer, Berlin, 2018.
- [356] B. Mildenhall, J. T. Barron, J. Chen, D. Sharlet, R. Ng, and R. Carroll. Burst denoising with kernel prediction networks. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2502–2510, Salt Lake City, UT, USA, June 2018.
- [357] L. Alvarez, M. Esclarin, M. Lefebure, and J. Sanchez. A PDE model for computing the optical flow. In *Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicaciones*, pages 1349–1356, Las Palmas de Gran Canaria, Spain, September 1999.
- [358] N. Monzon, A. Salgado, and J. Sanchez. Regularization strategies for discontinuity-preserving optical flow methods. *IEEE Transactions on Image Processing*, 25(4):1580–1591, April 2016.
- [359] N. Monzon, A. Salgado, and J. Sanchez. Robust discontinuity preserving optical flow methods. *Image Processing On Line*, 6:165–182, November 2016.
- [360] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *Computer Vision (ECCV - Part IV)*, volume 3024 of *Lecture Notes in Computer Science*, pages 25–36. Springer, Berlin, 2004.

- [361] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime TV-L1 optical flow. In F. A. Hamprecht, C. Schnörr, and B. Jähne, editors, *Pattern Recognition (DAGM)*, volume 4713 of *Lecture Notes in Computer Science*, pages 214–223, Berlin, 2007. Springer.
- [362] S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, November 2010.
- [363] B. Ummenhofer and T. Brox. Dense 3D reconstruction with a hand-held camera. In A. Pinz, T. Pock, H. Bischof, and F. Leberl, editors, *Pattern Recognition (DAGM/OAGM)*, volume 7476 of *Lecture Notes in Computer Science*, pages 103–112, Berlin, Heidelberg, 2012. Springer.
- [364] Y. Hou, C. Zhao, D. Yang, and Y. Cheng. Comments on “Image denoising by sparse 3-D transform-domain collaborative filtering”. *IEEE Transactions on Image Processing*, 20(1):268–270, January 2011.
- [365] R. Molina, M. Vega, J. Mateos, and A. K. Katsaggelos. Variational posterior distribution approximation in Bayesian super resolution reconstruction of multispectral images. *Applied and Computational Harmonic Analysis*, 24(2):251–267, March 2008.
- [366] Q. Yuan, L. Zhang, and H. Shen. Multiframe super-resolution employing a spatially weighted total variation model. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(3):379–392, March 2012.
- [367] Y. Zhang, G. Wu, P. T. Yap, Q. Feng, J. Lian, W. Chen, and D. Shen. Reconstruction of super-resolution lung 4D-CT using patch-based sparse representation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 925–931, Providence, RI, USA, June 2012.
- [368] P. Krämer, J. Benois-Pineau, and J. P. Domenger. Local object-based super-resolution mosaicing from low-resolution video. *Signal Processing*, 91(8):1771–1780, August 2011.
- [369] K. Choi, C. Kim, M. H. Kang, and J. B. Ra. Resolution improvement of infrared images using visible image information. *IEEE Signal Processing Letters*, 18(10):611–614, October 2011.

- [370] W. W. Zou and P. C. Yuen. Very low resolution face recognition problem. *IEEE Transactions on Image Processing*, 21(1):327–340, January 2012.
- [371] X. Zhang, M. Tang, and R. Tong. Robust super resolution of compressed video. *The Visual Computer*, 28(12):1167–1180, January 2012.
- [372] F. Zhou, W. Yang, and Q. Liao. A coarse-to-fine subpixel registration method to recover local perspective deformation in the application of image super-resolution. *IEEE Transactions on Image Processing*, 21(1):53–66, January 2012.
- [373] I. Begin and F. P. Ferrie. PSF recovery from examples for blind super-resolution. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume 5, pages 421–424, San Antonio, TX, USA, October 2007.
- [374] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 349–356, Kyoto, Japan, September 2009.
- [375] A. Marquina and S. Osher. Image super-resolution by TV-regularization and Bregman iteration. *Journal of Scientific Computing*, 37(3):367–382, July 2008.
- [376] J. Yang, J. Wright, T.S. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2681–2873, November 2010.
- [377] A. Tatem, H. Lewis, P. Atkinson, and M. Nixon. Super-resolution target identification from remotely sensed images using a Hopfield neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 39(4):781–796, April 2001.
- [378] A. Tatem, H. Lewis, P. Atkinson, and M. Nixon. Super-resolution land cover pattern prediction using a Hopfield neural network. *Remote Sensing of Environment*, 79(1):1–14, January 2002.
- [379] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multi-frame super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, October 2004.

- [380] F. Lin, C. Fookes, V. Chandran, and S. Sridharan. Investigation into optical flow super-resolution for surveillance applications. In B. C. Lovell and A. J. Maeder, editors, *APRS Workshop on Digital Image Computing: Pattern Recognition and Imaging for Medical Applications*, pages 73–78, Brisbane, February 2005.
- [381] T. Pham, L. Vliet, and K. Schutte. Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Advances in Signal Processing*, 2006(083268), December 2006.
- [382] L. Zhang, H. Zhang, H. Shen, and P. Li. A super-resolution reconstruction algorithm for surveillance images. *Signal Processing*, 90(3):848–859, March 2010.
- [383] F. Knoll, K. Bredies, T. Pock, and R. Stollberger. Second-order total generalized variation (TGV) for MRI. *Magnetic Resonance in Medicine*, 65(2):480–491, February 2011.
- [384] F. Stanco, S. Battiato, and G. Gallo. *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*. CRC Press, Boca Raton, 2011.
- [385] D. Kosmopoulos, N. Doulamis, and A. Voulodimos. Bayesian filter based behavior recognition in workflows allowing for user feedback. *Computer Vision and Image Understanding*, 116(3):422–434, March 2012.
- [386] Q. Yuan, L. Zhang, and H. Shen. Multiframe super-resolution employing a spatially weighted total variation model. *IEEE Transactions on Circuits, Systems and Video Technology*, 22(3):379–392, March 2012.
- [387] X. Li, P. Mooney, S. Zheng, C. Booth, M. Braunfeld, S. Gubbens, D. Agard, and Y. Cheng. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods*, 10(6):584–590, May 2013.
- [388] K. Nasrollahi and T. B. Moeslund. Super-resolution: A comprehensive survey. *Machine Vision and Applications*, 25(6):1423–1468, June 2014.
- [389] A. Laghrib, A. Hakim, and S. Raghay. A combined total variation and bilateral filter approach for image robust super resolution. *EURASIP*

Journal on Advances in Signal Processing, June 2015. Article Number - 19.

- [390] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4959–4962, Milan, Italy, July 2015.
- [391] A. Doulamis, N. Doulamis, C. Ioannidis, C. Chrysouli, N. Grammalidis, K. Dimitropoulos, C. Potsiou, E. Stathopoulou, and M. Ioannides. 5D modelling: An efficient approach for creating spatiotemporal predictive 3D maps of large-scale cultural resources. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(5):61–68, August 2015.
- [392] R.Y. Tsai and T.S. Huang. Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing*, 1(2):317–339, November 1984.
- [393] M. Elad and Y. Hel-Or. A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193, August 2001.
- [394] N. Nguyen, P. Milanfar, and G. Golub. Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *IEEE Transactions on Image Processing*, 10(9):1299–1308, September 2001.
- [395] S. Chaudhuri. *Super-resolution Imaging*. Springer, New York, 2000.
- [396] G. Rochefort, F. Champagnat, G. Le Besnerais, and J. F. Giovannelli. An improved observation model for super-resolution under affine motion. *IEEE Transactions on Image Processing*, 15(11):3325–3337, November 2006.
- [397] H. He and L. P. Kondi. An image super-resolution algorithm for different error levels per frame. *IEEE Transactions on Image Processing*, 15(3):592–603, March 2006.

- [398] H. Shen, L. Zhang, B. Huang, and P. Li. A MAP approach for joint motion estimation, segmentation, and super resolution. *IEEE Transactions on Image Processing*, 16(2):479–490, February 2007.
- [399] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang. A total variation regularization based super-resolution reconstruction algorithm for digital video. *EURASIP Journal on Advances in Signal Processing*, December 2007. Article number - 074585.
- [400] D. Mitzel, T. Pock, T. Schoenemann, and D. Cremers. Video super resolution using duality based TV-L1 optical flow. In J. Denzler, G. Notni, and H. Süße, editors, *Pattern Recognition (DAGM)*, volume 5748 of *Lecture Notes in Computer Science*, pages 432–441. Springer, Berlin, 2009.
- [401] P. Milanfar. *Super-resolution Imaging*. CRC Press, Boca Raton, 2011.
- [402] S. P. Belekos, N. P. Galatsanos, and A. K. Katsaggelos. Maximum a posteriori video super-resolution using a new multichannel image prior. *IEEE Transactions on Image Processing*, 19(6):1451–1464, June 2010.
- [403] X. Li, Y. Hu, X. Gao, D. Tao, and B. Ning. A multi-frame image super-resolution method. *Signal Processing*, 90(2):405–414, February 2010.
- [404] J. Tian and K. K. Ma. A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342, September 2011.
- [405] S. D. Babacan, R. Molina, and A. K. Katsaggelos. Variational Bayesian super resolution. *IEEE Transactions on Image Processing*, 20(4):984–999, April 2011.
- [406] X. Zhang, J. Jiang, and S. Peng. Commutability of blur and affine warping in super-resolution with application to joint estimation of triple-coupled variables. *IEEE Transactions on Image Processing*, 21(4):1796–1808, April 2012.
- [407] Z. Wang and F. Qi. On ambiguities in super-resolution modeling. *IEEE Signal Processing Letters*, 11(8):678–681, August 2004.

- [408] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, May 2003.
- [409] J. Weickert, S. Grewenig, C. Schroers, and A. Bruhn. Cyclic schemes for PDE-based image analysis. *International Journal of Computer Vision*, 118(3):275–299, July 2016.
- [410] M. J. Black and P. Anandan. A framework for robust estimation of optical flow. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 231–236, Berlin, Germany, May 1993.
- [411] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987.
- [412] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer Vision, Graphics and Image Processing*, 37(3):362–385, March 1987.
- [413] A. R. Rao and B. G. Schunck. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53(2):157–185, March 1991.
- [414] E. Franken, R. Duits, and B. ter Haar Romeny. Nonlinear diffusion on the 2D Euclidean motion group. In F. Sgallari, F. Murli, and N. Paragios, editors, *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 4485 of *Lecture Notes in Computer Science*, pages 461–472. Springer, Berlin, 2007.
- [415] M. Mühlich and T. Aach. Analysis of multiple orientations. *IEEE Transactions on Image Processing*, 18(7):1424–1437, July 2009.
- [416] L. R. Williams and D. W. Jacobs. Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, 9(4):837–858, May 1997.
- [417] E. J. Bekkers, M. Loog, B. M. ter Haar Romeny, and R. Duits. Template matching via densities on the roto-translation group. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(2):452–466, February 2018.

- [418] K. Krissian, G. Malandain, and N. Ayache. Directional anisotropic diffusion applied to segmentation of vessels in 3D images. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 345–348. Springer, Berlin, 1997.
- [419] E. Payot, R. Guillemaud, Y. Troussset, and F. Preteux. An adaptive and constrained model for 3D X-ray vascular reconstruction. In P. Grangeat and J.-L. Amans, editors, *Three-Dimensional Image Reconstruction in Radiation and Nuclear Medicine*, volume 4 of *Computational Imaging and Vision*, pages 47–57. Springer, Dordrecht, 1996.
- [420] L. Alvarez, P. L. Lions, and J. M. Morel. Image selective smoothing and edge detection by nonlinear diffusion II. *SIAM Journal on Numerical Analysis*, 29(3):845–866, May 1991.
- [421] R. Carmona and S. Zhong. Adaptive smoothing respecting feature directions. *IEEE Transactions on Image Processing*, 7(3):353–358, March 1998.
- [422] H. M. Gach, C. Tanase, and F. Boada. 2D & 3D Shepp-Logan phantom standards for MRI. In *Proc. IEEE International Conference on Systems Engineering*, pages 521–526, Las Vegas, NV, USA, August 2008.
- [423] L. A. Shepp and B. F. Logan. The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science*, 21(3):21–43, June 1974.

Chapter 13

Own Publications

Published

1. K. Bodduna and J. Weickert. Evaluating data terms for variational multi-frame super-resolution. In F. Lauze, Y. Dong, A. B. Dahl, editors, *Scale Space and Variational Methods (SSVM)*, volume 10302 of *Lecture Notes in Computer Science*, pages 590-601, Springer, Cham, 2017.
2. K. Bodduna, J. Weickert, and A. S. Frangakis. Hough-based evolutions for enhancing structures in 3D electron microscopy. In M. Vento, G. Percannella, editors, *Computer Analysis of Images and Patterns (CAIP)*, volume 11678 of *Lecture Notes in Computer Science*, pages 102-112, Springer, Cham, 2019.
3. K. Bodduna and J. Weickert. Enhancing patch-based methods with inter-frame connectivity for denoising multi-frame images. In *Proc. IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 2414-2418.
4. K. Bodduna and J. Weickert. Poisson Noise Removal Using Multi-Frame 3D Block Matching. In *Proc. 8th IEEE European Workshop on Visual Information Processing (EUVIP)*, Rome, Italy, Oct. 2019, pp. 58-63.
5. K. Bodduna, J. Weickert, and M. Cárdenas. Multi-frame Super-resolution from Noisy Data. In A. Elmoataz, J. Fadili, Y. Quéau, J. Rabin, L. Simon, editors, *Scale Space and Variational Methods in Computer Vision*

(SSVM), volume 12679 of *Lecture Notes in Computer Science*, pages 565-577, Springer, Cham, 2021.

6. K. Bodduna and J. Weickert. Removing Multi-frame Gaussian Noise by Combining Patch-based Filters with Optical Flow. *Journal of Electronic Imaging*, 30(3): 033031, June 2021.

Technical Reports

7. K. Bodduna and J. Weickert. Image Denoising with Less Artefacts: Novel Non-linear Filtering on Fast Patch Reorderings. *arXiv:2002.00638*, February 2020.

In Progress

8. M. Cárdenas, K. Bodduna, and J. Weickert. A Non-linear Scale-space Framework for Continuous Non-local Evolutions. Part I: Symmetric Case. *Contains content from disc diffusion chapter*.
9. M. Cárdenas, K. Bodduna, and J. Weickert. A Non-linear Scale-space Framework for Continuous Non-local Evolutions. Part II: Non-symmetric Case. *Contains content from sector diffusion chapter*.
10. K. Bodduna and J. Weickert. Importance of Filter Shape in Image Denoising: A New Stripe-based Diffusion Technique.
11. K. Bodduna and J. Weickert. Robust Image Denoising with Less Artefacts: Non-linear Filtering on Fast Patch Reorderings. *A journal extension of technical report 7*.

Chapter 14

Abbreviations

| | |
|----------------|--|
| AF | average then filter |
| AWGN | additive white Gaussian noise |
| BM3D | 3D block matching |
| CCD | charge-coupled device |
| CED | coherence-enhancing diffusion |
| CF | combined filtering |
| CMOS | complementary metal oxide semiconductor |
| cryo-EM | cryo-electron microscopy |
| CTF | contrast transfer function |
| DCT | discrete cosine transform |
| DD | disc diffusion |
| DDID | dual-domain image denoising |
| DnCNN | denoising convolutional neural network |
| EED | edge-enhancing diffusion |
| FA | filter then average |
| FRC | Fourier ring correlation |
| GTF | ground truth flow |
| HE | Hough-based evolution |
| HR | high-resolution |
| FPN | fixed pattern noise |
| LFSP | linear filtering on smooth patch orderings |
| LPA | local polynomial approximation |
| LR | low-resolution |

| | |
|---------------|--|
| MF | multiple reference frame <i>filtering</i> |
| MSE | mean squared error |
| NFPR | non-linear <i>filtering</i> on fast patch reordderings |
| NLDD | non-local <i>dual-domain denoising</i> |
| NLB | non-local Bayes |
| NLM | non-local means |
| NSD | non-symmetric stripe diffusion |
| PDE | partial differential equation |
| PSNR | peak signal-to-noise ration |
| SD | sector diffusion |
| SF | single reference frame <i>filtering</i> |
| SOF | sub-optimal flow |
| SR | super-resolution |
| SSD | symmetric stripe diffusion |
| SSIM | structural similarity index measure |
| VNLNET | video non-local network |
| VST | variance stabilising transformation |
| V-BM4D | video - 4D block matching |
| V-NLB | video - non-local Bayes |

Chapter 15

List of Symbols

| | |
|-------------------------------|---|
| a_i, b_i | normalisation constants in NFPR |
| \mathbf{A} | discrete approximation of the continuous gradient operator |
| \mathbf{A}_{HD} | homogenous diffusion operator |
| \mathbf{A}_{SD} | sector diffusion operator |
| \mathbf{A}_{EED} | edge enhancing diffusion operator |
| $B_{i,\rho}$ | disc B or radius ρ around pixel i |
| B_{search} | search area in NFPR |
| $\mathbf{B}, \mathbf{B}^\top$ | blur operator that models point spread function |
| c | normalisation constant in sector diffusion |
| \mathbf{D} | 2×2 diffusion tensor and downsampling factor in SR |
| \mathbf{D}^\top | upsampling operator |
| e | standard mathematical constant |
| \mathbf{e}_i | error vector that depicts noise in low resolution images |
| E | variational energy functional |
| E_{data} | data term of energy functional |
| E_{smooth} | smoothness term of energy functional |
| f | raw input image |
| \mathbf{f}^{nr} | non-registered multi-frame data |
| \mathbf{f} | registered multi-frame data |
| g | diffusivity function |
| h | weighting function for inter-patch distances |
| h_1, h_2 | Gaussian smoothing functions for sector diffusion |
| H_1, H_2 | dimensions of high resolution image |
| $H(\hat{\mathbf{x}})$ | point spread function in Fourier space for cryo-EM |
| i | index for central pixel in diffusion process |
| $I(\hat{\mathbf{x}})$ | image in Fourier space for cryo-EM |

| | |
|---|---|
| j | index for neighbouring pixel in diffusion process |
| J | spatial weighting function for diffusion processes |
| k | iteration number for image evolution process |
| k_{\max} | maximum iterations for image evolution process |
| L | total number of frames for multi-frame denoising |
| L_1, L_2 | dimensions of low resolution image |
| m | minimiser or maximiser of functional |
| M | total number of pixels in the discrete image grid and also total number of sectors/stripes in a diffusion model |
| M_i | total number of pixels in \mathbf{P}_i and $\mathbf{P}_i^{\text{add}}$ together |
| n | noise distribution |
| $n_{\mathcal{G}}$ | Gaussian noise distribution |
| $n_{\mathcal{P}}$ | Poissonian noise distribution |
| N | total number of pixels in set \mathbf{P}_i |
| N_H | size of high resolution image |
| N_L | size of low resolution image |
| $O(\hat{\mathbf{x}})$ | object in Fourier space for cryo-EM |
| p_{post} | posterior probability function |
| p_{prior} | prior of posterior probability function |
| p_{like} | likelihood of posterior probability function |
| P_j | set of sectors in which pixel j is present |
| $\mathbf{P}_i, \mathbf{P}_i^{\text{add}}$ | nearest neighbours to pixel i acc. to inter-patch distances |
| P_ℓ | reference patch in BM3D/NLB from frame ℓ |
| $\mathcal{P}(P_\ell)$ | set of most similar patches to reference patch P_ℓ |
| $q_{i,j}$ | entries of matrix \mathbf{Q} |
| \mathbf{Q} | system matrix for diffusion-based evolutions |
| \mathbb{R} | real number line |
| \mathbb{R}^+ | positive real number line |
| r_i | ring number i in Fourier Ring Correlation measure |
| S_ℓ | set of pixels within sector/stripe ℓ |
| T | threshold in voting process of Hough transform |
| u | processed image |
| u_σ | Gaussian smoothing of u in Cartesian domain in general or the domain defined by inter-patch distances for NFPR |
| \hat{u} | Fourier transformation of image u |
| $\mathbf{u}^{\text{initial}}$ | denoised image after first step in BM3D/NLB |
| $\mathbf{u}^{\text{final}}$ | final denoised image in BM3D/NLB |
| v | ground truth image |
| $\mathbf{v}_1, \mathbf{v}_2$ | eigenvectors of diffusion tensor |
| w_s | spatial weight for image denoising |
| w_t | tonal weight for image denoising |

| | |
|----------------------------------|--|
| w_{nbd} | neighbourhood weight for image denoising |
| \mathbf{w} | motion vector in optical flow computation |
| \mathbf{W}_i | warping operator for each low resolution image |
| \mathbf{W}^\top | backward registration operator |
| x, y, z | positions within image domain |
| $\hat{\mathbf{x}}$ | frequency vector in Fourier space |
| α_{SR} | smoothness term weight in SR variational model |
| $\alpha, \alpha_{\text{OF}}$ | gradient term weight in optical flow |
| χ | multilication factor that controls amount of Poissonian noise and also indicator function in multi-frame BM3D/NLB |
| $\partial_t u$ | derivative of u in temporal dimension t |
| $\partial_\theta u$ | derivative of u in direction θ |
| $\partial_\theta^+ u$ | one-sided derivative of u in orientation θ |
| $\partial_{\eta\eta} u$ | double derivative of u in direction η |
| ϵ | ensures strict convexity of optical flow functional |
| η | downsampling factor of warping scheme in optical flow |
| η_1, η_2 | inner and outer iterations in optical flow scheme |
| $\gamma, \gamma_{\text{OF}}$ | smoothness term weight in optical flow |
| λ | parameter within diffusivity function and also within regularisation function of optical flow |
| $\mu_{\mathcal{G}}$ | mean of Gaussian noise distribution |
| ∇u | gradient of u - either spatial or spatio-temporal |
| ω | successive over-relaxation parameter in optical flow scheme |
| Ω | image domain |
| $ \Omega $ | Cardinality of set Ω |
| $\Omega(\mathbf{x}, \mathbf{h})$ | neighbourhood around \mathbf{x} whose size is determined by \mathbf{h} |
| π | standard mathematical constant |
| Φ | regularisation function in optical flow |
| Ψ | function responsible for dealing with outliers in optical flow energy functional |
| ρ | radius of disc for DD, SD, NSD and SSD |
| ρ_{search} | radius of search area in NFPR |
| ρ_{sim} | disc radius in NFPR for patch-matching |
| ρ_t | half of the thickness of each stripe in NSD and SSD |
| ρ_1 | radius of sphere in HE method |
| ρ_2 | half length of line segment in HE method |
| $\sigma_{\mathcal{G}}$ | standard deviation of Gaussian noise distribution |
| σ_{noise} | standard deviation of mixed Poissonian-Gaussian noise |
| τ | time step size of image evolution process |
| \perp | perpendicular to |