Synthetic Image Data for the Development, Analysis and Evaluation of Image Processing Methods

Niklas Rottmayer

Vom Fachbereich Mathematik
der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau
zur Verleihung des akademischen Grades

Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.)

genehmigte Dissertation

Erstgutachterin	Prof. Dr. Claudia Redenbach
Zweitgutachter	Dr. Jesùs Angulo Lopez (HDR)
Datum der Disputation	27.02.2026

D 386

Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

Acknowledgements

This work was supported by the Franco-German University and German Federal Ministry of Education and Research (BMBF), poSt: Synthetic Data for Machine Learning Segmentation of Highly Porous Structures from FIB-SEM Nano-tomographic Data, Funding Number: 01IS21054A.

First, I want to express my deepest gratitude to my supervisor Prof. Dr. Claudia Redenbach for giving me the opportunity to do my PhD under her guidance and providing me with excellent supervision. She was always available, offered valuable guidance and insights, and put me on the right path with her advice. Next, I want to thank all my colleagues and coworkers at the Image Processing department at ITWM and the Statistics Group at the RPTU who created a wonderful work environment for me. I am grateful for all their support, help and advice that they offered over the years, especially, in our numerous coffee break chats. Lastly, I would like to thank my family for their constant love and aid in my journey. Without you, this would not been possible. I am deeply grateful for all the advice and input from my brother Jan when discussing scientific and non-scientific matters, the assistance and support from my parents throughout my life, and Lara whose unconditional love constantly inspires me to grow.

Abstract

The rise of artificial intelligence (AI) as a general solution in image processing problems, such as segmentation and detection tasks, has sparked growing interest in generating synthetic image data. A significant factor is that real data is often not available in sufficient quantity or quality for the thorough development, evaluation, and analysis of solutions. The reasons for this are diverse: image acquisition can be costly, labeling may require unfeasible and error-prone manual annotation, or the data may have low variability that does not allow for generalization. With the increase in computational power over the last few decades, synthetic data has become an essential component in the development of solutions as it addresses the shortcomings of real data. Furthermore, it provides almost unlimited diversity, is easily acquired, and comes with perfect annotation.
This thesis consists of three parts that highlight the significance of synthetic data in developing, validating, and analyzing methods and results in image processing.
In the first part, we present a convex optimization method for removing stripe artifacts. These elongated and parallel corruptions appear frequently with various imaging techniques including light-sheet fluorescence microscopy (LSFM), focused ion beam scanning electron microscopy (FIB-SEM), and remote sensing. Our approach offers intuitive parametrization and is highly flexible to different scenarios of image structures and stripes. We demonstrate the effectiveness and advantages of our approach over existing solutions by comparing them across real images from LSFM, FIB-SEM, and remote sensing through visual inspection. Based on synthetic LSFM data obtained by simulating physical light propagation we enrich our analysis by comparing the processed images to ground truth data and quantitatively confirming the performance observed on real data.
In the second part, we discuss the assessment of quality for results in binary image segmentation tasks by comparing a large variety of established traditional metrics and distance-based approaches. This includes a distance-based metric that we propose, which captures the spatial distribution of errors while offering desirable properties such as normalization and interpretability. Using predominantly synthetic data and some real segmentation results, we perform a thorough analysis of the segmentation metrics across diverse conditions. This demonstrates the robustness and effectiveness of our metric in distinguishing errors near the surface from those farther away across different structural contexts. We illustrate its inclusion to real-world segmentation tasks by extending a previous study on segmenting cracks in CT images.
In the third part, we conduct a systematic study of morphologically diverse geometric structures with the goal to characterize the morphology of 3D spatial structures in terms of their anisotropy, scale and angularity based on the measurement of common morphological features. Geometries are entirely generated using stochastic models whose parameters intuitively translate to the studied concepts and yield morphologically diverse structures. Using classical machine learning approaches for dimensionality reduction our study finds a series of measures for anisotropy, scale and angularity based on linear combination of the morphological features that distinguish between synthetic structures accordingly. We exemplify and discuss their use on real data for modeling of general spatial structures.

Zusammenfassung

Mit dem Aufstieg künstlicher Intelligenz (KI) als Universallösung für Probleme der Bildverarbeitung wie Segmentierung oder Erkennung wuchs der Bedarf an synthetisch hergestellten Bilddaten. Realdaten stehen oftmals nicht in ausreichender Menge oder Qualität zur Verfügung, um gründliche Auswertungen durchzuführen und robuste Lösungen zu entwickeln. Das Bildverfahren kann kostenintensiv sein, die Erstellung von Segmentierungen kann fehleranfällig und unpraktikabel sein oder die Daten sind schlichtweg nicht vielfältig genug, um generalisieren zu können. Mit zunehmender Rechenleistung in den letzten Jahrzehnten wurde die Erzeugung und Verwendung synthetischer Bilder zu einem essenziellen Schritt in der Entwicklung von Methoden, da sie Defizite der Realdaten ausgleichen oder diese sogar ganz ersetzen können. Dabei bieten sie eine fast unbeschränkte Vielfalt, sind vergleichsweise einfach zu erzeugen und kommen mit perfekter Annotation.
Diese Arbeit ist unterteilt in drei Teile, welche die vielfältigen Anwendungsmöglichkeiten und Vorteile synthetischer Daten für die Entwicklung, Validierung und Analyse in der Bildverarbeitung vorstellen.
Im ersten Teil präsentieren wir ein konvexes Optimierungsverfahren zur Entfernung von Streifenartefakten. Die oftmals langen und parallelen Artefakte treten häufig bei den verschiedensten Aufnahmeverfahren wie Lichtscheiben-Fluoreszenzmikroskopie (LSFM), Rasterelektronenmikroskopie mit fokussiertem Ionenstrahl (FIB-REM) oder Fernerkundung auf. Unser Ansatz stellt eine intuitive Parametrisierung bereit und kann flexibel auf eine Vielfalt an Bildstrukturen und Streifenartefakten angewendet werden. Wir demonstrieren die Fähigkeiten und Vorteile unseres Verfahrens gegenüber existierenden Lösungen, indem wir diese auf Realbildern von LSFM, FIB-REM und Fernerkundung visuell vergleichen. Synthetische LSFM Daten, welche wir durch Simulation von Lichtausbreitung erhalten, ermöglichen es, die beobachteten Ergebnisse auch quantitativ zu bestätigen.
Im zweiten Teil diskutieren wir die qualitative Bewertung von Bildsegmentierungen. Wir vergleichen eine Vielzahl etablierter klassischer und Distanz-basierter Metriken. Letztere beinhalten eine von uns vorgestellte Metrik, welche die räumliche Verteilung von Fehlern berücksichtigt und Eigenschaften wie Normalisierung und Interpretierbarkeit bietet. Mit großteils synthetischen, aber auch realen Segmentierungsergebnissen analysieren wir die Metriken in verschiedenen Umständen. Wir zeigen, dass unsere Metrik robust und effektiv oberflächennahe von distanzierten Fehler unterscheiden kann und nutzen sie um eine bereits durchgeführte Studie eines realen Segmentierungsproblems zu vertiefen.
Im dritten Teil führen wir eine systematische Studie morphologisch vielfältiger Geometrien durch, mit dem Ziel die Anisotropie, Größe und Kantigkeit allgemeiner 3D Strukturen anhand gängiger morphologischer Merkmale zu charakterisieren. Zur Erzeugung synthetischer Geometrien nutzen wir stochastische Modelle, deren Parameter intuitiv die Morphologie beeinflussen und entsprechende Vielfältigkeit liefern. Mit klassischen Dimensionsreduktionsmethoden leiten wir eine Reihe von Maßen für Anisotropie, Größe und Kantigkeit her, welche Linearkombinationen der berechneten Merkmale sind und die synthetischen Daten voneinander unterscheiden können. Wir zeigen und diskutieren deren Übertragbarkeit und Anwendung auf reale Daten für die Modellierung von Strukturen.

Contents

Symbols

$\mathbb{N}$	Natural numbers
$\mathbb{Z}$	Integers
$\mathbb{R}$	Real numbers
$\bar{\mathbb{R}}$	Extended real numbers
$\mathbb{C}$	Complex numbers
$\operatorname{conv}(X)$	Convex hull of a set X
$\operatorname{aff}(X)$	Affine hull of a set X
$\operatorname{ri}(X)$	Relative interior of a set X
B_r	d-dimensional ball of radius r
N_K	Normal cone of a convex set K
P_K	Orthogonal projection onto a convex set K
$\operatorname{dom} f$	Effective domain of function f
ι_X	Indicator function of a set X
$\Gamma_0(\mathbb{R}^d)$	Space of proper, convex, lsc functions on $\mathbb{R}^d$
$\partial f(x)$	Subdifferential of a function f at x
$\operatorname{prox}_{\tau f}$	Proximal mapping of a function f
f_τ	Moreau envelope of a function f
S_τ	Soft-shrinkage function
H_τ	Huber function
C_τ	Coupled soft-shrinkage function
f^*	Fenchel conjugate of a function f
s_X	Support function of a set X
$\oplus$	Minkowski addition
$\ominus$	Minkowski subtraction
$\circ$	Morphological opening
$\bullet$	Morphological closure
$\check{X}$	Reflection of a set X
$\mathcal{B}$	Borel σ-algebra
$\mathcal{F}$	System of closed sets
$\mathcal{G}$	System of open sets
$\mathcal{C}$	System of compact sets
$\mathbb{P}$	Probability measure
T_Ξ	Capacity functional of a random closed set Ξ
$\mathrm{SO}(d)$	Special orthogonal group of dimension d
Λ_Ξ	Intensity measure of a point process Ξ
$\mathbb{E}$	Expectation
V_d	d-dimensional Lebesgue measure
$\mathbb{Q}$	Mark distribution
$\mathcal{K}$	System of convex bodies

$\mathcal{R}$	Convex ring
$\mathcal{S}$	Extended convex ring
$V_k(X)$	k-th intrinsic volume of a set X
Γ	Gamma function
$\chi(X)$	Euler characteristic of a set X
$V_{V,k}(\Xi)$	k-th intrinsic volume density of a random closed set Ξ
$\mathcal{L}^k$	System of k-dimensional linear subspaces
$L^\perp$	Orthogonal complement of a linear subspace L
$p_k(X, L)$	k-th total projection of a set X onto a linear subspace L
$p_{V,k}(\Xi, L)$	k-th total projection density of a random closed set Ξ onto a linear subspace L
$\ell(\Xi, L)$	Mean chord lengths of a random closed set Ξ in the direction defined by L
∂X	Surface of a set X
S^{d-1}	d-dimensional unit sphere
$\mathrm{Nor}(K)$	Generalized normal bundle of the convex set K
$\Lambda_k(K, \cdot)$	k-th generalized curvature measures of convex set K
$M_r(K, A)$	Local parallel of a convex set K at distance r determined by A
$\mathbb{T}^p$	Space of symmetric p-tensors on $\mathbb{R}^d$
$\otimes$	Symmetric tensor product
$\mathrm{O}(d)$	Orthogonal group of dimension d
$\Phi_k^{n,m}(X)$	Minkowski tensor of the convex set X
$\Phi_{V,k}^{0,m}(\Xi)$	Minkowski tensor density of a random closed set Ξ
$\mathbb{L}^d$	d-dimensional homogeneous lattice
$\mathcal{F}^k(F)$	Set of k-dimensional faces of a convex polytope F
$X \sqcap \mathbb{F}$	Gauss digitization of a set X with respect to an adjacency system $\mathbb{F}$
X^C	Complement of a set X
$\mathbb{F}_{\mathrm{loc}}$	Local adjacency system
$\mathbb{F}$	Adjacency system
$\mathbb{L}^k$	k-dimensional section lattice of a lattice $\mathbb{L}^d$
${}^\perp\mathbb{L}^k$	Translative complement of section lattice $\mathbb{L}^k$
$\hat{V}$	Estimate of V
T^d	d-Torus of length 1
$*$	(Cyclic) convolution
$\oslash$	Element-wise division
$\odot$	Element-wise product
$\mathscr{F}$	Discrete Fourier transform

List of Abbreviations

AI	Artificial intelligence
LSFM	Light-sheet fluorescence microscopy
FIB-SEM	Focused ion beam scanning electron microscopy
lsc	lower semi-continuous
PDHGMp	Primal dual hybrid gradient method with extrapolation of dual variable
DFT	Discrete Fourier transform
IDFT	Inverse discrete Fourier transform
FFT	Fast Fourier transform
GSR	General stripe remover
VSNR	Variational stationary noise remover
MDSR	Multi-directional stripe remover
NSCT	Non-subsampled contourlet transformation
PSNR	Peak signal-to-noise ratio
MS-SSIM	Multi-scale structural similarity index measure
DSC	Dice similarity coefficient
TP	True positives
TN	True negatives
FP	False positives
FN	False negatives
MCC	Matthews correlation coefficient
SCC	Surface consistency coefficient
dAHD	Directed average Hausdorff distance
ACC	Accuracy
TNR	True negative rate
PPV	Positive predictive value
AHD	Average Hausdorff distance
ASSD	Average symmetric surface distance
EDT	Euclidean distance transformation
PCA	Principal component analysis
LDA	Linear discriminant analysis
t-SNE	t-distributed stochastic neighbor embedding
UMAP	Uniform manifold approximation and projection

List of Figures

List of Tables

1 Introduction

"A picture is worth a thousand words." This famous adage underscores the importance of vision in human perception of the world. No wonder image processing has become one of the most prominent mathematical fields ubiquitous in our everyday lives. From printed images in newspapers and digital images on social media to moving images such as movies, the merits of image processing surround us and have become an inseparable component of modern society.

The steady increase in computational power of electronic hardware and improvements in optical systems, e.g., microscopes and cameras, over the past century have enabled the transfer and processing of increasingly complex image data. With it came the growing list of image-based tasks such as autonomous driving and visual inspection which require extracting necessary information from images. Artificial intelligence (AI) solutions have emerged as 'the' general solution to all such tasks for several reasons. They are easy to apply and transfer between similar problems, and their performance typically increases with more data. However, this comes at a cost. It requires exceedingly large amounts of data to train AI solutions and validate that they perform their dedicated tasks correctly. This has led to a massive demand for suitable images which oftentimes cannot be met by regular means. The reasons for this are diverse: image acquisition can be costly, labeling may require unfeasible and error-prone manual annotation, or real data may exhibit low variability that does not allow for generalization.

The improvements in computing capabilities also laid the groundwork for the generation of increasingly complex *synthetic data* produced by algorithms or simulations, e.g., of physical light transport. Synthetic data offers numerous advantages which increase its appeal for AI applications. Compared to real images, it is incredibly easy to obtain and cheap to produce in abundance. At the same time, the diversity of the synthetic data can be controlled by input parameters and is limited only by the underlying models used to generate it. Most notably, edge cases that are rarely found in real images can be constructed. One of the major benefits for image processing in particular, is that synthetic data comes with perfect annotations making it ideal for image classification or segmentation tasks where labeling is typically tedious or entirely unfeasible. Hence, training AI models often relies entirely on synthetic data while real images are used only for validation. Synthetic data is not limited to AI models and is used in the same ways to improve the development, evaluation and analysis of traditional approaches.

This thesis consists of three parts that highlight the significance of synthetic data in developing, validating, and analyzing methods and results in image processing.

In the first part, we present an approach based on convex optimization for removing stripe artifacts. These elongated and parallel corruptions are frequently observed across various imaging techniques including light-sheet fluorescence microscopy (LSFM), focused ion beam scanning electron microscopy (FIB-SEM), and remote sensing. Their removal is required not only to improve visual quality but also to enable further analysis and image processing of the data. Our approach, which offers an intuitive parametrization, is highly flexible to different scenarios of image structures and stripes. We demonstrate

its capabilities and advantages over existing solutions through an extensive parameter study and a rigorous comparison with established methods across images originating from LSFM, FIB-SEM, and remote sensing using visual inspection. Our analysis is enriched with synthetic LSFM data obtained by simulating physical light propagation on simple stochastic models. The ground truth data provided by this procedure allows for comparison with the processed images and the calculation of selected image quality metrics which quantitatively confirms and validates the performance we observe on real data.
In the second part, we turn to the topic of quality assessment for image segmentation tasks. Despite years of research and progress, traditional quality metrics that rely solely on a simple contingency table have remained the gold standard for assessing segmentation quality in the literature, primarily because of tradition. By analyzing and comparing a wide variety of established traditional metrics with modern approaches that include geometric information about the relative positions of pixels we demonstrate clear benefits that distance-based metrics provide. In particular, we propose and include our own metric which captures the spatial distribution of incorrectly labeled pixels. It offers desirable properties such as normalization and interpretability that are typically found only in traditional metrics. Our analysis and comparison build on a synthetic dataset that combines stochastic processes with common morphological operations in image processing to yield highly controlled and diverse geometric settings. The study demonstrates the robustness and effectiveness of our metric in distinguishing between errors found near the surface of a structure and those farther away across different geometric contexts. We illustrate its inclusion into real-world segmentation tasks by extending a previous study on segmenting cracks in CT images.
In the third part, we examine geometric structures focusing on morphological properties such as anisotropy, scale or angularity. These characteristics are studied in various scientific fields including materials science and biology where they help improve our understanding of materials and cells. However, quantifying these abstract concepts is a challenging task that typically imposes strong restrictions on the structures under investigation. Using a morphologically diverse set of synthetic geometries we conduct an exploratory study aiming to characterize the morphology for a broad set of structures by calculating common morphological features. Geometries are entirely generated using stochastic models whose parameters intuitively translate to the studied concepts and yield morphologically diverse structures. Using classical machine learning approaches for dimensionality reduction we not only demonstrate that geometries can be distinguished but also derive corresponding measures that quantify anisotropy, scale and angularity of the structures. We demonstrate that these measures are interpretable and generalize beyond the synthetic data considered by showing their use in deriving simple models of real structures.

2 Mathematical Preliminaries

This chapter presents the mathematical theory and notation required throughout this thesis. We begin by visiting the theory of convex functions and their properties that make them desirable for establishing optimization problems. We continue by discussing methods for finding the solution to convex optimization problems which are frequently encountered in image processing. Afterwards, we dedicate a section to the theory of stochastic and integral geometry. This versatile field covers several aspects of classical image processing, ranging from simple operations on images and stochastic models for random sets to the characterization of geometries with specialized features. Another aspect covered by this topic is the discretization of sets into a pixel image and its influence on the calculation of geometric features. As a final topic, we discuss the famous Fourier series and discrete Fourier transform which are regularly applied in signal and image processing to decompose information into different periodic oscillations.

2.1 Convex Optimization

In this section, we introduce convex functions and their optimization which lays the foundation for numerous methods in image processing for tasks such as inpainting, active contours, segmentation or denoising. This part mainly follows traditional sources such as the books by Ekeland and Temam [36] and Rockafellar [103, 104] but also takes inspiration from more recent contributions by Chambolle and Pock [16] and Burger et al. [12].

2.1.1 Convexity

Definition 2.1. A set $K \subseteq \mathbb{R}^d$ is *convex* if

$$tx + (1-t)y \in K \quad \text{for all } x, y \in K \text{ and } t \in [0,1].$$

In other words, the convexity of a set can be understood as the property that any line segment connecting a pair of elements must be entirely contained in the set itself. This concept extends towards any finite set of elements $x_1, \dots, x_n \in K$ via the *convex combinations*

$$x = \sum_{j=1}^{n} t_j x_j \quad \text{for } t_j \geq 0, j = 1, \dots, n \text{ and } \sum_{j=1}^{n} t_j = 1. \tag{2.1}$$

Definition 2.1 is equivalent to all convex combinations of points in K to be contained in K. For an arbitrary set $X \subseteq \mathbb{R}^d$ we call the set of all convex combinations x of points in X the *convex hull* $\operatorname{conv}(X)$. It is the smallest convex set that contains X, in particular, $\operatorname{conv}(X) = X$ if and only if X is convex. The condition $t_j \geq 0$ in (2.1) ensures that a convex combination always lies within the polyhedron defined by the corners $x_1, \dots, x_n$.

The relaxation of this condition yields an *affine combination*

$$\sum_{j=1}^{n} t_j x_j \quad \text{with } x_1, \dots, x_n \in X \text{ and } \sum_{j=1}^{n} t_j = 1.$$

We denote the *affine hull* $\operatorname{aff}(X)$ as the set of all affine combinations of points $x_1, \dots, x_n \in X$. The interior of X when regarded as a subset of its affine hull is called the *relative interior* $\operatorname{ri}(X)$, i.e.,

$$\operatorname{ri}(X) = \{x \in \operatorname{aff}(X) \,:\, \exists \epsilon > 0, (x + B_\epsilon) \cap \operatorname{aff}(X) \subseteq X\}$$

with the ball of radius $\epsilon > 0$ centered in the origin $B_\epsilon = B_\epsilon(0)$. For a convex set $K \subseteq \mathbb{R}^d$ we denote the *normal cone* of C at a point $x_0 \in K$ as

$$N_K(x_0) \coloneqq \{p \in \mathbb{R}^d \,:\, \langle p, x - x_0 \rangle = 0 \text{ for all } x \in K\}$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^d$.

Proposition 2.2. Let $K \subseteq \mathbb{R}^d$ be a non-empty, closed, convex set. Then for any $x_0 \in \mathbb{R}^d$ there exists a unique element $\hat{x}_0 = P_K(x_0) \in K$ called the *orthogonal projection* of x_0 onto K so that

$$\|\hat{x} - x_0\|_2 = \inf_{x \in K} \|x - x_0\|_2 .$$

The element is uniquely characterized by the relation

$$\langle x_0 - \hat{x}_0, y - \hat{x}_0 \rangle \leq 0 \quad \text{for all } y \in K.$$

It is not difficult to see, that the orthogonal projection for any $x \notin K$ fulfills the relation

$$x - P_K(x) \in N_K(P_K(x)).$$

The concept of convexity for sets can be extended towards functions

$$f : \mathbb{R}^d \to \bar{\mathbb{R}} \coloneqq \mathbb{R} \cup \{-\infty, \infty\}$$

which map into the *extended real numbers* $\bar{\mathbb{R}}$ which includes the values $+\infty$ and $-\infty$. We call the set $\operatorname{dom}(f) \coloneqq \{x \in \mathbb{R}^d \,:\, f(x) < \infty\}$ where f does not attain infinity the *effective domain*.

Definition 2.3. A function $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ is

(i) *convex* if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \tag{2.2}$$

for all $x, y \in \operatorname{dom}(f)$ and $t \in [0, 1]$.

(ii) *strictly convex* if the inequality (2.2) is strict for all $t \in (0, 1)$ and $x \neq y$.

(iii) *strongly convex* with modulus $\mu > 0$ if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\mu}{2} t(1-t) \|y - x\|_2^2$$

for all $x, y \in \operatorname{dom}(f)$ and $t \in [0, 1]$.

The definition of strong convexity for a function f is equivalent to $f - \frac{\mu}{2}\|\cdot\|_2^2$ being convex. This implies that there exists a function "below" f that is still convex. Clearly, strong convexity implies strict convexity.

Example 2.4. For a set $X \subseteq \mathbb{R}^d$ the *indicator function* ι_X is defined as

$$\iota_X(x) = \begin{cases} 0 & \text{for } x \in X, \\ +\infty & \text{for } x \notin X. \end{cases}$$

It is convex if and only if X is convex. ι_X is often hidden in constrained optimization problems since

$$\min_{x \in X} f(x) = \left\{ \min_{x \in \mathbb{R}^d} f(x) \text{ s.t. } x \in X \right\} = \min_{x \in \mathbb{R}^d} \{f(x) + \iota_X(x)\}.$$

Definition 2.5. A function $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ is

(i) *proper* if $f(x) > -\infty$ for all $x \in \mathbb{R}^d$ and $f \neq +\infty$.

(ii) *lower semi-continuous* (lsc) if

$$f(x) \leq \liminf_{y \to x} f(y) = \sup_{\epsilon > 0} \inf_{y \in B_\epsilon(x)} f(y) \quad \text{for all } x \in \mathbb{R}^d.$$

(iii) *coercive* if

$$\lim_{\|x\|_2 \to \infty} f(x) = \infty.$$

Properties such as convexity and lower semi-continuity are preserved under standard transformations. In particular, the sums of convex functions are convex, if well-defined, and the point-wise supremum of a family of convex functions is also convex. The same is true for lsc functions. The provided definition of convexity considers only two points and their connecting line segment, as with convex sets. The extension towards convex combinations can be made as before and is known as *Jensen's inequality*

$$f\left(\sum_{j=1}^n t_j x_j\right) \leq \sum_{j=1}^n t_j f(x_j)$$

for all $x_j \in \mathbb{R}^d, t_j \geq 0, j = 1, \ldots, n$ with $\sum_{j=1}^n t_j = 1$. We denote the set of all proper, convex, lsc functions by $\Gamma_0(\mathbb{R}^d)$.
An important consideration for optimization are existence, uniqueness and globality of an optimum which are necessary to form a well-defined optimization problem. Under certain conditions, convex functions provide all of them.

Theorem 2.6. Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ be a proper function and $S := \operatorname{argmin}_{x \in \mathbb{R}^d} f(x)$.

i) If f is lsc and coercive, then $\inf_{x \in \mathbb{R}^d} f(x)$ is attained and $S \neq \emptyset$ is compact.

ii) If f is convex, then local minimizers are global minimizers and S is convex.

iii) If f is strictly convex, then $|S| \in \{0, 1\}$, i.e., if a minimizer exists, it is unique.

This result represents the foundation of convex optimization since none of these restrictions are particularly limiting. Most importantly, we get that for any coercive and strictly convex $f \in \Gamma_0(\mathbb{R}^d)$ there exists a unique minimizer $\hat{x} \in \mathbb{R}^d$ which fulfills

$$f(\hat{x}) = \min_{x \in \mathbb{R}^d} f(x).$$

By relaxing the strict convexity we only loose the uniqueness of the minimizer but not its existence or globality.

2.1.2 Subdifferentials

Differentiability is one of the core concepts employed to optimize functions, especially if they are non-linear. It allows to characterize critical points by a vanishing derivative and establish iterative solvers such as gradient descent. For convex functions, differentiability cannot be guaranteed on the entire domain and is a strong condition to set in general. For example, the 1-norm $\|\cdot\|_1$ is not differentiable in 0 but a popular choice as cost function.

Definition 2.7. A function $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ is *subdifferentiable* at $x_0 \in \mathbb{R}^d$ if $f(x_0) \in \mathbb{R}$ and there exists a point $p \in \mathbb{R}^d$ so that

$$f(x) \geq f(x_0) + \langle p, x - x_0 \rangle \quad \text{for all } x \in \mathbb{R}^d.$$

The element p is called *subgradient* of f at x_0. The set of all subgradients at x_0 is called *subdifferential* of f at x_0 and denoted by $\partial f(x_0)$.

This definition generalizes the concept of classical differentiability and coincides with it if and only if a unique subgradient exists. Moreover, it can be shown that for proper, convex functions the subdifferential is non-empty, at least on the relative interior of its domain, and bounded on the interior. The subdifferential behaves analogously to the classical differential under transformation with one key difference. It is only subadditive, that is, $\partial(f+g)(x_0) \subseteq (\partial f + \partial g)(x_0)$ for proper, convex functions f and g. Equality holds if the respective domains overlap sufficiently, i.e., $\mathrm{ri}(\mathrm{dom}(f)) \cap \mathrm{ri}(\mathrm{dom}(g)) \neq \emptyset$ [36]. The subdifferential allow us to generalize the condition of critical points.

Theorem 2.8 (*Fermat's rule*)**.** Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ be a proper function. Then $\hat{x} \in \mathbb{R}^d$ is a global minimizer of f if and only if $0 \in \partial f(\hat{x})$.

This essential theorem follows directly from the definition of the subdifferential since $0 \in \partial f(\hat{x})$ implies $f(x) \geq f(\hat{x})$ for all $x \in \mathbb{R}^d$ and vice versa.

2.1.3 Proximal Mapping

We have seen that convex functions $f \in \Gamma_0(\mathbb{R}^d)$ can provide desirable properties for minimization problems under additional constraints. However, we have yet to discuss, how an optimum can be obtained. In the case, where f is differentiable a common strategy to find a minimum is the gradient descent iteration

$$x^{(k+1)} = x^{(k)} - \tau \nabla f(x^{(k)})$$

with constant step size $\tau > 0$. It converges under strict conditions such as Lipschitz continuity of f and small enough $\tau > 0$. A relaxation of conditions can be achieved using

implicit gradient descent

$$x^{(k+1)} = x^{(k)} - \tau \nabla f(x^{(k+1)}) \tag{2.3}$$

but it is unclear if and how such a step can be computed.

Definition 2.9. Let $f \in \Gamma_0(\mathbb{R}^d)$ and $\tau > 0$. The *proximal mapping* of f is defined as

$$\operatorname{prox}_{\tau f}(x) \coloneqq \operatorname*{argmin}_{y \in \mathbb{R}^d} \left\{ \frac{1}{2\tau} \|x - y\|_2^2 + f(y) \right\}.$$

The corresponding *Moreau envelope* or *Moreau-Yoshida regularization* is

$$f_\tau(x) \coloneqq \min_{y \in \mathbb{R}^d} \left\{ \frac{1}{2\tau} \|x - y\|_2^2 + f(y) \right\}.$$

Example 2.10.

(i) For the indicator function $f = \iota_K$ of a non-empty, closed, convex set $K \subseteq \mathbb{R}^d$, the proximal mapping is the orthogonal projection onto K, i.e.,

$$\operatorname{prox}_{\tau \iota_K}(x) = P_K(x) = \operatorname*{argmin}_{y \in K} \|x - y\|_2^2 .$$

and the Moreau envelope relates to the squared distance from K via

$$f_\tau(x) = \frac{1}{2\tau} \min_{y \in K} \|x - y\|_2^2 .$$

(ii) The proximal mapping of the 1-norm $g = \|\cdot\|_1$

$$S_\tau(x) \coloneqq \operatorname{prox}_{\tau \|\cdot\|_1}(x) = (S_\tau(x)_j)_{j=1}^d \quad \text{with} \quad S_\tau(x)_j = \begin{cases} x_j - \tau & \text{for } x_j > \tau, \\ 0 & \text{for } x_j \in [-\tau, \tau], \\ x_j + \tau & \text{for } x_j < -\tau, \end{cases}$$

is called the (component-wise) *soft-shrinkage function* with threshold τ. The corresponding Moreau envelope

$$H_\tau(x) \coloneqq g_\tau(x) = (H_\tau(x)_j)_{j=1}^d \quad \text{with} \quad H_\tau(x)_j = \begin{cases} \frac{1}{2\tau} x_j^2 & \text{for } x_j \in [-\tau, \tau], \\ |x_j| - \frac{\tau}{2} & \text{otherwise}, \end{cases}$$

is known as *Huber function*.

(iii) Let $v = (x, y, z)$ with $x, y, z \in \mathbb{R}^d$ and

$$\|v\|_{2,1} = \sum_{i=1}^d \sqrt{x_i^2 + y_i^2 + z_i^2}.$$

The proximal mapping

$$C_\tau(v) = \operatorname{prox}_{\tau \|\cdot\|_{2,1}}(x) = (C_\tau(v)_j)_{j=1}^d \quad \text{with} \quad C_\tau(v)_j = v_j \left(1 - \frac{\tau}{\sqrt{x_j^2 + y_j^2 + z_j^2}} \right)$$

is a *coupled soft-shrinkage function*.

The proximal mapping is well-defined, i.e., there exists a unique minimizer which can be deduced by exploiting strong convexity and Theorem 2.6. We can rewrite the implicit gradient descent (2.3) into the equation

$$0 = \frac{x^{(k+1)} - x^{(k)}}{\tau} + \nabla f(x^{(k+1)}).$$

Using basic subdifferential calculus and Fermat's rule, $x^{(k+1)}$ can be identified as a minimizer of the function

$$F(x) = \frac{1}{2\tau} \left\| x^{(k)} - x \right\|_2^2 + f(x)$$

which corresponds to the previously introduced proximal mapping, i.e.,

$$x^{(k+1)} = \operatorname{prox}_{\tau f}(x^{(k)}). \tag{2.4}$$

This is known as the *proximal point algorithm* and since $\operatorname{prox}_{\tau f}$ is well-defined for any $f \in \Gamma_0(\mathbb{R}^d)$, we do not require f to be differentiable for this update. While f may not be differentiable, it can be shown that the Moreau envelope f_τ is differentiable with $\nabla f_\tau(x) = \frac{1}{\tau}(x - \operatorname{prox}_{\tau f}(x))$ and fulfills a Lipschitz condition with constant $1/\tau$. By inserting this relation into (2.4) we get

$$x^{(k+1)} = \operatorname{prox}_{\tau f}(x^{(k)}) = x^{(k)} - \tau \nabla f_\tau(x^{(k)})$$

which is again an explicit gradient descent step. This shows that the implicit gradient descent for f is identical to the explicit gradient descent for f_τ and indicates that f and f_τ share the same set of minimizers which fulfill the fixpoint equation $\hat{x} = \operatorname{prox}_{\tau f}(\hat{x})$. We refer to [16] for additional details on the convergence behavior.

2.1.4 Conjugate Functions

Definition 2.11. Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$. Then the *(Fenchel) conjugate* of f is defined as

$$f^*(p) \coloneqq \sup_{x \in \mathbb{R}^d} \left\{ \langle p, x \rangle - f(x) \right\}.$$

The function $f^{**} \coloneqq (f^*)^*$ is called the *biconjugate* of f.

Example 2.12. For the indicator function ι_X of a set $X \subseteq \mathbb{R}^d$ the conjugate is given by

$$s_X(x) = \iota_X^*(x) = \sup_{x \in \mathbb{R}^d} \left\{ \langle p, x \rangle - \iota_X(x) \right\} = \sup_{x \in X} \langle p, x \rangle$$

which is called the *support function* of X.

The conjugate is by definition convex and lsc since it is the point-wise supremum of affine functions $g_x(p) \coloneqq \langle p, x \rangle - f(x)$ with $x \in \operatorname{dom}(f)$, which themselves are convex and lsc. Therefore, the biconjugate is also convex and lsc and represents the closest convex, lsc function to approximate f. For a proper function, we have $f = f^{**}$ if and only if $f \in \Gamma_0(\mathbb{R}^d)$. The conjugate and the subdifferential are closely linked by the following proposition.

Proposition 2.13 (*Legendre-Fenchel equality*)**.** If $f \in \Gamma_0(\mathbb{R}^d)$ and $x \in \mathrm{dom}(f)$, it holds that

$$p \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(p) \quad \Leftrightarrow \quad \langle p, x \rangle = f(x) + f^*(p).$$

When combined with Fermat's rule, this result provides an alternative to finding a global minimizer $\hat{x}$ of a function f by searching for $\hat{x} \in \partial f^*(0)$. It also enables us to derive the *Moreau decomposition* [9, Theorem 6.44 and 6.45]

$$\mathrm{prox}_{\tau f}(x) + \tau \, \mathrm{prox}_{\frac{1}{\tau} f^*}\left(\frac{x}{\tau}\right) = x \quad \text{for all } x \in \mathbb{R}^d, \tau > 0$$

which connects the proximal mappings of a function and its conjugate. Most importantly, once we have calculated one of the proximal mappings, we immediately obtain the other through this relation. In imaging applications it is common that either the proximal mapping of the function itself or its conjugate are easy to compute so that both are available without complex calculations.

2.1.5 Minimization of Convex Functions

In the following, we consider the convex optimization problem

$$\inf_{x \in \mathbb{R}^d} \{g(x) + h(Ax)\} \tag{2.5}$$

with functions $g \in \Gamma_0(\mathbb{R}^d)$, $h \in \Gamma_0(\mathbb{R}^m)$ and $A \in \mathbb{R}^{m \times d}$. Problems of this form are often encountered in imaging applications and can be rewritten into a min-max formulation using the Fenchel conjugate. Recall, that for $h \in \Gamma_0(\mathbb{R}^d)$ it holds that $h^{**} = h$ such that (2.5) is equivalent to

$$(P) \quad \inf_{x \in \mathbb{R}^d} \sup_{y \in \mathbb{R}^m} \{g(x) - h^*(y) + \langle Ax, y \rangle\}$$

which is called the *primal problem*. We refer to the function $L(x, y) = g(x) - h^*(y) + \langle Ax, y \rangle$ as the *Lagrangian* of this problem. Under very weak conditions it holds that the optimal value of (P) is finite and that the ordering of inf and sup can be exchanged, e.g., it is already sufficient if $h(0) < \infty$ and g is continuous at 0 [11, Satz 6.68]. This establishes the *dual problem*

$$(D) \quad \sup_{y \in \mathbb{R}^m} \inf_{x \in \mathbb{R}^d} \{g(x) - h^*(y) + \langle Ax, y \rangle\} = \sup_{y \in \mathbb{R}^m} \{-h^*(y) - g^*(-A^* y)\}.$$

Under the same assumptions as above, it can be shown that at least one solution $\hat{y}$ to the dual problem exists [36]. In this case, any solution $\hat{x}$ of (P) corresponds to a saddle point of the Lagrangian, i.e.,

$$L(x, \hat{y}) \leq L(\hat{x}, \hat{y}) \leq L(\hat{x}, y) \quad \text{for all } x \in \mathbb{R}^d, y \in \mathbb{R}^m.$$

Saddle points of the Lagrangian are equivalently characterized by the *Kuhn-Tucker* conditions [104]

$$0 \in \partial g(\hat{x}) + A^* \hat{y} \quad \text{and} \quad 0 \in \partial h^*(\hat{y}) - A\hat{x}.$$

Based on this, one approach to obtain a saddle point of L is given by the strategy of alternating between optimizing the primal and the dual problem via

$$x^{(k+1)} = \operatorname{prox}_{\tau g}(x^{(k)} - \tau A^* y^{(k)})$$
$$y^{(k+1)} = \operatorname{prox}_{\sigma h^*}(y^{(k)} + \sigma A x^{(k+1)})$$

with $\tau, \sigma > 0$ where the same ideas as in (2.4) were employed. Such methods are known as *primal-dual algorithms*. Over the years several modifications and variants based on this idea were proposed that improve the conditions for convergence and its speed. The pseudo-code of one such method is given in Algorithm 1.

Algorithm 1: Primal dual hybrid gradient method with extrapolation of dual variable (PDHGMp)[12]

Initialization: $x^{(0)} \in \mathbb{R}^d, b^{(0)} = \bar{b}^{(0)} \in \mathbb{R}^m, \tau, \sigma > 0$ with $\tau\sigma < \frac{1}{\|A\|_2^2}, \theta \in (0, 1]$

for $k = 0, 1, \ldots$ **do**

$$x^{(k+1)} = \operatorname{prox}_{\tau g}\left(x^{(k)} - \tau\sigma A^* \bar{b}^{(k)}\right)$$
$$y^{(k+1)} = \operatorname{prox}_{\frac{1}{\sigma}h}\left(b^{(k)} + A x^{(k+1)}\right)$$
$$b^{(k+1)} = b^{(k)} + A x^{(k+1)} - y^{(k+1)}$$
$$\bar{b}^{(k+1)} = b^{(k+1)} + \theta(b^{(k+1)} - b^{(k)})$$

end

Theorem 2.14. Let $g \in \Gamma_0(\mathbb{R}^d)$, $h \in \Gamma_0(\mathbb{R}^m)$, $A \in \mathbb{R}^{m \times d}$, $\theta \in (0, 1]$ and $\tau, \sigma > 0$ with $\tau\sigma < 1/\|A\|_2^2$. Suppose that the Lagrangian $L(x, y) \coloneqq g(x) - h^*(y) + \langle Ax, y\rangle$ has a saddle point. Then the sequence $\{(x^{(k)}, \sigma b^{(k)})\}_k$ generated by the PDHGMp converges to a saddle point of L.

We want to highlight that no additional restrictions on g and h are necessary for the convergence to a solution. We refer to [12] for further details on different optimization strategies for convex optimization problems and their corresponding convergence criteria.

2.2 Stochastic and Integral Geometry

This section introduces the theory of several mathematical concepts in image processing including mathematical morphology, the formulation of stochastic models to generate random sets, the derivation of representative object features for convex bodies and their estimation based on discretizations. In general, we follow the books by Ohser and Schladitz [91], Schneider and Weil [113] and Chiu et al. [27]. However, the object features are also inspired by the work of Hug and colleagues [56, 59], the book by Jensen and Kiderlen [62] and the book from Ohser and Mücklich [90].

2.2.1 Morphological Operations

Definition 2.15. Let $X, Y \subseteq \mathbb{R}^d$. The *Minkowski sum* and *difference* of X and Y are defined by

$$X \oplus Y \coloneqq \{x + y \,:\, x \in X, y \in Y\}, \text{ and } X \ominus Y \coloneqq (X^C \oplus Y)^C$$

where X^C denotes the complement of X. The operator $\oplus$ is called *Minkowski addition* and $\ominus$ *Minkowski subtraction*.

These operations can be equivalently written as

$$X \oplus Y = \bigcup_{x \in X} (Y + x) \quad \text{and} \quad X \ominus Y = \bigcap_{y \in Y} (X + y).$$

We remark that the Minkowski subtraction is not the inverse of the Minkowski addition, i.e. $(X \oplus Y) \ominus Y \neq X$ in general. In image processing, it is more common to use the notion as *(morphological) dilation* $X \oplus \check{Y}$ and *erosion* $X \ominus \check{Y}$ with *structuring element* Y using the *reflection* $\check{Y} = -Y$. In practice, we typically assume that Y is centered such that $0 \in Y$ but this is not necessary. The dilation can be understood as the set of all points covered when attaching $\check{Y}$ to each point in X. Similarly, the erosion is the set of all points which are not covered when attaching $\check{Y}$ to each point in the complement X^C.

Definition 2.16. Let $X, Y \subseteq \mathbb{R}^d$. The *(morphological) opening* and *closure* of the set X with structuring element Y are defined by

$$X \circ Y = (X \ominus \check{Y}) \oplus Y, \text{ and } X \bullet Y = (X \oplus \check{Y}) \ominus Y.$$

It can be shown that $X \circ Y \subseteq X \subseteq X \bullet Y$ holds true. A geometric interpretation for these relations can be given as follows. The initial erosion in the opening can remove smaller parts of the set which the subsequent Minkowski addition cannot restore. In case of the closure, the dilation fills up small holes in the set which the subsequent Minkowski subtraction cannot add again. These interpretations present the main application of the opening and closure, that is, to clean up small structures or holes which are not relevant but hinder further processing of the data. The set X is said to be *morphologically closed w.r.t* Y if $X \bullet Y = X$. Similarly, it is called *morphologically open w.r.t.* Y if $X \circ Y = X$. It is *morphologically regular w.r.t.* Y if it is morphologically open and closed with respect to Y. We call a set X morphologically closed/open/regular if there exists an $\epsilon > 0$ such that X is morphologically closed/open/regular w.r.t. the ball B_ϵ.

2.2.2 Random Closed Sets

Let E be a locally compact space with countable base and denote its Borel σ-algebra by $\mathcal{B} = \mathcal{B}(E)$. We denote the systems of closed, open, and compact subsets of E by $\mathcal{F}, \mathcal{G}$ and $\mathcal{C}$, respectively. All three systems contain the empty set.

Definition 2.17. The topology generated by

$$\{\mathcal{F}^C : C \in \mathcal{C}\} \cup \{\mathcal{F}_G : G \in \mathcal{G}\},$$

where $\mathcal{F}^C := \{F \in \mathcal{F} : F \cap C = \emptyset\}$ and $\mathcal{F}_G := \{F \in \mathcal{F} : F \cap G \neq \emptyset\}$ is called the *topology of closed convergence* on $\mathcal{F}$.

In the following, $\mathcal{F}$ will always be equipped with the topology of closed convergence. The σ-algebra $\mathcal{B}(\mathcal{F})$ of Borel sets on $\mathcal{F}$ is generated by each of the systems $\{\mathcal{F}^C : C \in \mathcal{C}\}$, $\{\mathcal{F}_C : C \in \mathcal{C}\}$, $\{\mathcal{F}^G : G \in \mathcal{G}\}$ or $\{\mathcal{F}_G : G \in \mathcal{G}\}$.

Definition 2.18. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space. A *random closed set* is a $(\Sigma, \mathcal{B}(\mathcal{F}))$-measurable map $\Xi : \Omega \to \mathcal{F}$. The *distribution* of Ξ is the image measure $\mathbb{P}_\Xi := \Xi(\mathbb{P})$ of $\mathbb{P}$ under Ξ given by

$$\mathbb{P}_\Xi(F) = \mathbb{P}(\Xi \in F) = \mathbb{P}(\{\omega \in \Omega : \Xi(\omega) \in F\}), \quad F \in \mathcal{B}(\mathcal{F}).$$

Random closed sets can be understood as set valued random variables such that the concepts of independence, joint distributions, etc. known from real valued random variables apply as well. We call two random closed sets Ξ_1 and Ξ_2 *stochastically equivalent* if they have the same distribution which we denote by $\Xi_1 \overset{d}{=} \Xi_2$. This allows to define two properties which are frequently desired. Let $\mathrm{SO}(d)$ denote the special orthogonal group in dimension d.

Definition 2.19. A random closed set Ξ in $\mathbb{R}^d$ is

(i) *stationary* if $\Xi + x \overset{d}{=} \Xi$ for all $x \in \mathbb{R}^d$.

(ii) *isotropic* if $\theta\Xi \overset{d}{=} \Xi$ for all rotations $\theta \in \mathrm{SO}(d)$.

In other words, stationarity and isotropy of a random closed set refer to the shift and rotation invariance of its distribution. Especially stationarity is an important assumption in practice since it enables to estimate properties of an entire model from a local observation window. A random closed set is often associated with its *capacity functional*

$$T_\Xi(C) = \mathbb{P}_\Xi(\mathcal{F}_C) = \mathbb{P}(\Xi \cap C \neq \emptyset), \quad C \in \mathcal{C}$$

which has comparable properties to the distribution function of real-valued random variables. Most importantly, it uniquely characterizes the distribution of the random closed set. This result is known as *Choquet's theorem* [113, Theorem 2.2.1]. Stationarity and isotropy are equivalent to the capacity functional fulfilling $T_\Xi(C) = T_\Xi(C - z)$ for all $z \in \mathbb{R}^d$ and $T_\Xi(C) = T_\Xi(\theta C)$ for all $\theta \in \mathrm{SO}(d)$, respectively.

Definition 2.20. A random closed set Ξ is a *point process* if its values are almost surely locally finite, i.e., the number of points $\#(\Xi \cap C) < \infty$ for any $C \in \mathcal{C}$. The associated measure $\Lambda_\Xi(X) = \mathbb{E}(\#(\Xi \cap X))$ is called *intensity measure* of Ξ. If $\#(\Xi \cap \{\omega\}) \leq 1$ for all $\omega \in \Omega$ the process is *simple*. We can write a simple point process as $\Xi = \{x_1, x_2, \dots\} \subseteq E$.

Example 2.21. A special case of point processes are *Poisson (point) processes* which fulfill the two conditions:

(i) For each $X \in \mathcal{B}$ with $\Lambda_\Xi(X) < \infty$ the random variable $\#(\Xi \cap X)$ is Poisson distributed with parameter $\Lambda_\Xi(X)$.

(ii) For pairwise disjoint sets $X_1, \dots, X_n \in \mathcal{B}$ the random variables $\#(\Xi \cap X_1), \dots, \#(\Xi \cap X_n)$ are independent.

If additionally $\Lambda_\Xi(X) = \lambda V_d(X)$ for some $\lambda > 0$ where V_d is the d-dimensional Lebesgue measure, the process is called *homogeneous Poisson process* with *intensity* λ.

Point processes are commonly used in the modeling of positions of objects and can incorporate interactions such as repulsion, attraction or independence. A natural extension of

a point process Ξ_0 in $\mathbb{R}^d$ is to attach a *mark* in the form of a compact set $C_j \in \mathcal{C}$ to each point x_j. These marks are typically centered in some sense, e.g., $0 \in C_j, j = 1, 2, \dots$, and drawn from a common *mark distribution* $\mathbb{Q}$. The set $\{(x_j, C_j)_j\}$ is called a *marked point process* and defines a random closed set via $\Xi = \bigcup_j (x_j + C_j)$ which is also called a *germ grain model*. In general, the mark distribution and point process need not be independent. However, independence is regularly assumed which enables writing the intensity measure Λ_Ξ of Ξ as the product measure

$$\Lambda_\Xi = \Lambda_{\Xi,0} \times \mathbb{Q},$$

where $\Lambda_{\Xi,0}$ is the intensity measure of Ξ_0. In this case, $\mathbb{Q}$ is referred to as the *distribution of the typical particle*.

Definition 2.22. A random closed set Ξ in $\mathbb{R}^d$ is a *(stationary) Boolean model* if there exists a homogeneous Poisson process Ξ_0 with intensity $\lambda > 0$ and a mark distribution $\mathbb{Q}$ such that

$$\Xi = \bigcup_j (x_j + C_j)$$

where $\Xi_0 = \{x_1, x_2, \dots\}$ and C_j are drawn iid from $\mathbb{Q}$ independently of the locations x_j.

2.2.3 Intrinsic Volumes

In the following let $\mathcal{K}$ denote the system of *convex bodies*, i.e., non-empty, compact and convex sets in $\mathbb{R}^d$. By $\mathcal{R}$ we denote the *convex ring* whose elements are the finite union of convex bodies, also called *polyconvex sets*. The *extended convex ring* $\mathcal{S}$ is the system of all sets which are *locally polyconvex*, that is,

$$\mathcal{S} := \{F \in \mathcal{F} : F \cap K \in \mathcal{R} \text{ for all } K \in \mathcal{K}\}.$$

Definition 2.23. Let $K \in \mathcal{K}$. The *intrinsic volumes* $V_0, \dots, V_d$ of K are defined through the *Steiner formula*

$$V_d(K \oplus B_r) = \sum_{k=0}^{d} r^{d-k} \kappa_{d-k} V_k(K), \quad r \geq 0 \tag{2.6}$$

where V_d is the d-dimensional Lebesgue measure and $\kappa_k = \pi^{k/2}/\Gamma(1 + k/2)$ is the volume of the k-dimensional unit ball with the gamma function Γ. $K \oplus B_r$ is called *parallel body* of K at distance $r \geq 0$.

The intrinsic volumes are equivalent to the *Minkowski functionals* or *quermassintegrals* found in the literature and differ only in their normalization constants and indexing order.

Definition 2.24. A function $f : \mathcal{K} \to \mathbb{R}$ is

(i) *invariant under rigid motions* if

$$f(\theta K + x) = f(K) \quad \text{for all } x \in \mathbb{R}^d \text{ and } \theta \in \mathrm{SO}(d).$$

(ii) *additive* if

$$f(K_1 \cup K_2) + f(K_1 \cap K_2) = f(K_1) + f(K_2)$$

for all $K_1, K_2 \in \mathcal{K}$ with $K_1 \cup K_2 \in \mathcal{K}$.

(iii) *continuous* if

$$f(K_n) \overset{n\to\infty}{\longrightarrow} f(K)$$

for all $K_n, K \in \mathcal{K}$ such that $K_n \overset{n\to\infty}{\longrightarrow} K$ with respect to the Hausdorff distance.

The Lebesgue measure is invariant under rigid motions, additive and continuous. From the Steiner formula (2.6) we obtain immediately that the intrinsic volumes inherit these properties. We remark, that the restriction towards convex bodies does not suffice when modeling geometric structures in practice. An important result to bridge the gap is *Groemer's extension theorem* [113, Theorem 14.4.2] which states that additive and continuous functions on $\mathcal{K}$ can be extended to additive but not necessarily continuous functions on $\mathcal{R}$. Most importantly, the intrinsic volumes can be extended to the convex ring via the inclusion-exclusion formula

$$V_k(X) = \sum_{j=1}^{n} V_k(K_i) - \sum_{j=1}^{n-1} \sum_{l=j+1}^{n} V_k(K_j \cap K_l) + \ldots + (-1)^{n+1} V_k\left(\bigcap_{j=1}^{n} K_j\right), \tag{2.7}$$

for $k = 0, \ldots, d$ where $X = \bigcup_{j=1}^{n} K_j$ and $K_j \in \mathcal{K}$ for $j = 1, \ldots, n$. For the cases $d = 2$ and $d = 3$ the intrinsic volumes have well-known interpretations on the convex ring, namely:

$d=2$		$d=3$	
V_2	area A	V_3	volume
$2V_1$	boundary length / perimeter	$2V_2$	surface area
V_0	Euler characteristic χ	πV_1	integral of mean curvature
		$\frac{1}{2}V_1$	mean width (for convex bodies only)
		V_0	Euler characteristic χ

Example 2.25. The *Euler(-Poincaré) characteristic* or *Euler number* $\chi = V_0$ on $\mathcal{R}$ is the additive extension of the constant function $\chi(K) = 1$ for all $K \in \mathcal{K}$. In particular it holds that for pair-wise disjoint $K_1, \ldots, K_n \in \mathcal{K}$ that

$$\chi\left(\bigcup_{j=1}^{n} K_j\right) = n.$$

Theorem 2.26 (*Hadwiger's Characterization Theorem*)**.** Let $f : \mathcal{K} \to \mathbb{R}$ be a function which is invariant under rigid motions, additive and continuous. Then there exist constants $c_0, \ldots, c_d \in \mathbb{R}$ such that

$$f(K) = \sum_{k=0}^{d} c_k V_k(K) \quad \text{for all } K \in \mathcal{K}.$$

This result is incredibly important in practice as it means that typical object features can be represented using the intrinsic volumes, at least those, which are independent of position and orientation, additive and continuous. In particular, $d + 2$ or more additive object features always carry redundant information about K. The concept of intrinsic volumes can be extended to random closed sets which are usually not finite. The idea is to consider the random closed set Ξ in $\mathbb{R}^d$ observed in a compact observation window $W \in \mathcal{G}$

and calculate the *intrinsic volume densities*

$$V_{V,k}(\Xi) \coloneqq \lim_{r\to\infty} \frac{\mathbb{E}(V_k(\Xi \cap rW))}{V_d(rW)}, \quad k = 0, \dots, d-1 \tag{2.8}$$

$$V_{V,d}(\Xi) \coloneqq \frac{\mathbb{E}(V_d(\Xi \cap rW))}{V_d(rW)}. \tag{2.9}$$

The densities are well-defined under weak conditions on the random closed set Ξ [113, Theorem 9.2.1], that is,

(i) Ξ is stationary.

(ii) The realizations of Ξ almost surely belong to $\mathcal{S}$.

(iii) Ξ fulfills the integrability condition

$$\mathbb{E}(2^{\#(\Xi\cap[0,1]^d)}) < \infty$$

where $\#(\Xi \cap [0,1]^d)$ is the smallest number m such that there exist $K_1, \dots, K_m \in \mathcal{K}$ such that $\Xi = \bigcup_{j=1}^m K_j$.

We remark that condition (i) ensures that we can generalize from the observation window to the entire structure since the placement of W does not matter and local properties are expected to be the same everywhere. Condition (ii) guarantees that the intersection $\Xi \cap rW$ is almost surely a polyconvex set so that the intrinsic volumes are well-defined. Lastly, condition (iii) presents a weak constraint on the complexity of the structure to guarantee that the intrinsic volumes are finite.

Theorem 2.27. Let Ξ be a Boolean model in $\mathbb{R}^d$ with intensity $\lambda > 0$ and distribution of the typical mark $\mathbb{Q}$ with convex grains. Furthermore, let $\bar{V}_k(\Xi)$ denote the mean intrinsic volumes of the typical mark. Then, the intrinsic volume densities $V_{V,k}(\Xi)$ fulfill the equations

$$V_{V,d}(\Xi) = 1 - e^{-\lambda \bar{V}_d(\Xi)}$$

and

$$V_{V,k}(\Xi) = e^{-\lambda \bar{V}_d(\Xi)} \left(\lambda \bar{V}_k(\Xi) - \sum_{j=2}^{d-k} \frac{(-1)^j}{j!} \sum_{\substack{m_1,\dots,m_j=k+1 \\ m_1+\dots+m_j=(j-1)d+k}}^{d-1} \prod_{i=1}^{j} c_{k,m_i} \lambda \bar{V}_{m_i}(\Xi) \right)$$

for $k = 0, \dots, d-1$ with

$$c_{k,m} = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)}.$$

These formulae highlight that the intrinsic volume densities of a Boolean model are determined entirely by the intensity of its point process and the expected intrinsic volumes of its typical mark. Conversely, this shows that no information beyond this can be obtained from measurements of the densities. For $d = 3$ the formulae are known as Miles' formulae [91] and read

$$V_{V,3}(\Xi) = 1 - e^{-\lambda \bar{V}_3(\Xi)},$$

$$
\begin{aligned}
V_{V,2}(\Xi) &= e^{-\lambda \bar{V}_3(\Xi)} \lambda \bar{V}_2(\Xi), \\
V_{V,1}(\Xi) &= e^{-\lambda \bar{V}_3(\Xi)} \left(\lambda \bar{V}_1(\Xi) - \frac{\lambda^2 \pi^2}{32} \bar{V}_2(\Xi)^2 \right), \\
V_{V,0}(\Xi) &= e^{-\lambda \bar{V}_3(\Xi)} \left(\lambda - \frac{\lambda^2}{4\pi} \bar{V}_1(\Xi) \bar{V}_2(\Xi) + \frac{\lambda^3 \pi}{384} \bar{V}_2(\Xi)^3 \right).
\end{aligned}
$$

Proposition 2.28 (*Principal kinematic formulae*)**.** Let μ be the rotation invariant probability measure of $\mathrm{SO}(d)$ and $X, Y \in \mathcal{R}$. It holds, that

$$
\int_{\mathrm{SO}(d)} \int_{\mathbb{R}^d} V_j(X \cap (\theta Y + x)) V_d(\mathrm{d}x) \mu(\mathrm{d}\theta) = \sum_{k=j}^{d} \alpha_{d,j,k} V_k(X) V_{d+j-k}(Y) \tag{2.10}
$$

for $j = 0, \ldots, d$ with

$$
\alpha_{d,j,k} = \frac{\Gamma\left(\frac{k+1}{2}\right) \Gamma\left(\frac{d+j-k+1}{2}\right)}{\Gamma\left(\frac{j+1}{2}\right) \Gamma\left(\frac{d+1}{2}\right)}.
$$

This formula is often considered in a special case. Let $\mathcal{L}^k$ denote the system of k-dimensional linear subspaces of $\mathbb{R}^d$, $L_k^\perp$ be the orthogonal complement to $L_k \in \mathcal{L}^k$ and $V_{L_k^\perp}$ be the Lebesgue measure on $L_k^\perp$. Then by replacing Y with a k-dimensional linear subspace L_k in (2.10) yields *Crofton's intersection formulae*

$$
\int_{\mathcal{L}^k} \int_{L_k^\perp} V_j(X \cap (L_k + x)) V_{L_k^\perp}(\mathrm{d}x) \mu(\mathrm{d}L_k) = \alpha_{d,j,k} V_{d+j-k}(X), \tag{2.11}
$$

see [114]. An important application of this formula is the estimation of the intrinsic volumes based on a discrete image which will be discussed in Section 2.2.6.

2.2.4 Basic Directional Characteristics

In the following, we expand on Crofton's intersection formulae for $j = 0$. Let the k-th *total projection*

$$
p_k(X, L_k) = \int_{L_k^\perp} \chi(X \cap (L_k + x)) V_{L_k^\perp}(\mathrm{d}x)
$$

denote the inner integral of (2.11). In the case $d = 3$ we refer to $p_1(X, L_1)$ as the *area of total projection* of X on the plane $L_1^\perp$ which w.l.o.g. contains the origin. Analogously, $p_2(X, L_2)$ is known as the *length of total projection* of X on the line $L_2^\perp$ containing the center. We can extend the total projections to random closed sets by defining the corresponding *total projection densities* as

$$
p_{V,k}(\Xi, L_k) := \lim_{r \to \infty} \frac{\mathbb{E}(p_k(\Xi \cap rW, L_k))}{V_d(rW)}, \quad k = 1, \ldots, d-1.
$$

For $k = 1, d-1$ and under a sufficient smoothness assumptions [90, 91] these densities are linked to the directional distributions of the surface normals and boundary line lengths of Ξ known as *rose of normal directions* and *rose of intersections*. They also allow us to

define the *mean chord length*

$$\ell(\Xi, L_1) \coloneqq \frac{V_{V,d}(\Xi)}{p_{V,1}(\Xi, L_1)}$$

of Ξ in the direction defined by L_1 which intuitively measures the average length of a line segment found in $\Xi \cap L_1$.

2.2.5 Minkowski Tensors

The intrinsic volumes have proven useful in the analysis of geometric structures since they relate to several quantities of interest. Generalizations towards vector and tensor-valued functions known as *Minkowski tensors* exist and are based on a local formulation of the Steiner formula and a relaxation of conditions.
Denote the product space $\Sigma^d \coloneqq \mathbb{R}^d \times S^{d-1}$ with the unit sphere S^{d-1} in $\mathbb{R}^d$. Let ∂K be the surface of $K \in \mathcal{K}$. We call a pair $(x, u) \in \Sigma^d$ a *support element* of K if $x \in \partial K$ and u is an outer unit normal of K at x, i.e. $u \in N_K(x)$ with $\|u\|_2 = 1$. The set of all support elements of K is called the *generalized normal bundle* and denoted by $\mathrm{Nor}(K)$. With the orthogonal projection from Proposition 2.2 it holds that

$$u_K(x) = \left(P_K(x), \frac{x - P_K(x)}{d(K, x)}\right) \in \mathrm{Nor}(K) \text{ for all } x \in \mathbb{R}^d \backslash K.$$

Definition 2.29. Let $K \in \mathcal{K}$. The *generalized curvature measures* or *support measures* $\Lambda_0(K, \cdot), \ldots, \Lambda_{d-1}(K, \cdot)$ of K are defined through the *local Steiner formula*

$$V_d(M_r(K, A)) = \sum_{k=0}^{d-1} r^{d-k} \kappa_{d-k} \Lambda_k(K, A),$$

where $M_r(K, A) \coloneqq \{x \in (K \oplus B_r) \backslash K : (P_K(x), u_K(x)) \in A\}$ is called *local parallel set* of K at distance $r \geq 0$ determined by $A \in \mathcal{B}(\Sigma^d)$.

We note that $M_r(K, \Sigma^d) = (K \oplus B_r) \backslash K$ and $V_k(K) = \Lambda_k(K, \Sigma^d)$. It can be shown that the generalized curvature measures have similar properties to the intrinsic volumes, that is, they fulfill properties corresponding to rigid motion invariance, additivity and continuity. For $p \in \mathbb{N}_0$, a *p-tensor*, or *tensor of rank p* on $\mathbb{R}^d$ is defined as a p-linear mapping from $(\mathbb{R}^d)^p$ to $\mathbb{R}$. It is *symmetric* if it is invariant under permutation of its arguments. Let $\mathbb{T}^p$ be the space of symmetric p-tensors on $\mathbb{R}^d$. We denote the symmetric tensor product of two tensors a, b as $ab = a \otimes b$ and write $x^p = x \otimes \ldots \otimes x$ (with p factors) for the p-fold symmetric tensor product of $x \in \mathbb{R}^d$. Let $\mathrm{O}(d)$ denote the orthogonal group in dimension d.

Definition 2.30. A function $f : \mathcal{K} \to \mathbb{T}^p$

(i) is *rotation covariant* if

$$f(\theta K) = \theta f(K) \quad \text{for all } \theta \in \mathrm{O}(d).$$

(ii) has *polynomial translation behavior* if there exist tensors $f_{p-k}(K) \in \mathbb{T}^{p-k}$ for $k =$

$0, \dots, p$ so that

$$f(K+x) = \sum_{k=0}^{p} \frac{1}{k!} f_{p-k}(K) x^k \quad \text{for all } x \in \mathbb{R}^d.$$

(iii) is *isometry covariant* if it is rotation covariant and has polynomial translation behavior.

Additivity and continuity are defined analogous to Definition 2.24. By Groemer's extension theorem [113, Theorem 14.4.2] we again obtain additive extensions to $\mathcal{R}$ for any arbitrary continuous and additive function f on $\mathcal{K}$.

Example 2.31. Let $Q = Q(x, y) = \langle x, y \rangle$ for $x, y \in \mathbb{R}^d$. The constant function

$$f : \mathcal{K} \to \mathbb{T}^2, K \mapsto Q$$

is additive, continuous and isometry covariant.

Definition 2.32. Let $K \in \mathcal{K}$ and $n, m \geq 0$. Then the *Minkowski tensors* are defined as

$$\Phi_k^{n,m}(K) \coloneqq \frac{1}{n!m!} \frac{\omega_{d-k}}{\omega_{d-k+m}} \int_{\Sigma^d} x^n u^m \Lambda_k(K, \mathrm{d}(x,u)), \quad k = 0, \dots, d-1,$$
$$\Phi_d^{n,0}(K) \coloneqq \frac{1}{n!} \int_{\mathbb{R}^d} x^n \mathrm{d}x$$

where $\omega_k = 2\pi^{k/2}/\Gamma(k/2)$ is the surface area of S^{k-1}.

For convenience we set $\Phi_k^{n,m} = 0$ if $k \notin \{0, \dots, d\}$, $n \notin \mathbb{N}_0$, $m \notin \mathbb{N}_0$ or $k = d$ and $m \neq 0$. The Minkowski tensors fulfill a Steiner-type formula

$$\Phi_d^{n,0}(K \oplus B_r) = \sum_{k=0}^{d+n} r^{d+n-k} \kappa_{d-k} V_k^{(n)}(K), \quad r \geq 0$$

where $V_k^{(n)}(K) = \sum_{m \in \mathbb{N}_0} \Phi_{k-n+m}^{n-m,m}(K)$ [62].

Theorem 2.33. Let $p \in \mathbb{N}_0$ and $f : \mathcal{K} \to \mathbb{T}^p$ be a function which is isometry covariant, additive and continuous. Then there exist real coefficients so that f is a linear combination of $Q^l \Phi_k^{n,m}$ where l, n, m are so that $2l + n + m = p$ and $k \in \{0, \dots, d\}$.

This result can be interpreted as an extension of Theorem 2.26 with one key difference. For $p \geq 2$ the functions $Q^l \Phi_k^{n,m}$ are no longer linearly independent and satisfy a series of linear relations

$$Q \sum_{m \in \mathbb{N}_0} \Phi_{k-n+m}^{n-m,m-2}(K) = 2\pi \sum_{m \in \mathbb{N}_0} m \Phi_{k-n+m}^{n-m,m}(K) \quad \text{for all } K \in \mathcal{K}$$

known as *McMullen relations* [84]. In [59, 62] the construction of several bases for $p \leq 3$ is discussed and proven. The concept of the Minkowski tensors can be extended for the case $n = 0$ to random closed sets which are usually not finite. The idea is to consider

the random closed set Ξ observed through a compact observation window $W \in \mathcal{G}$ and calculate the *Minkowski tensor densities* defined by

$$\Phi_{V,k}^{0,m}(\Xi) \coloneqq \lim_{r\to\infty} \frac{\mathbb{E}(\Phi_k^{0,m}(\Xi \cap rW))}{V_d(rW)}, \quad k = 0, \ldots, d-1.$$

The densities are well-defined under weak conditions on the random closed set Ξ and fulfill similar equations to the ones presented in Theorem 2.27 [57].

2.2.6 Lattices and Digitization

In practice, geometric structures are not imaged continuously but discretely in the form of a d-dimensional array of data points known as *pixels* or *voxels* on a *homogeneous (point) lattice*

$$\mathbb{L}^d \coloneqq \left\{ x \in \mathbb{R}^d \,:\, x = \sum_{j=1}^{d} \lambda_j u_j, \lambda_j \in \mathbb{Z} \right\}$$

where $u_1, \ldots, u_d \in \mathbb{R}^d$ form a basis of $\mathbb{R}^d$ and $U = (u_1, \ldots, u_d) \in \mathbb{R}^{d\times d}$ is the matrix of column vectors. The special case $\mathbb{L}^d = a\mathbb{Z}^d$ with $a > 0$ is called *cubic lattice*. The *closed unit cell* associated with a lattice is the Minkowski sum $C = [0, u_1] \oplus \ldots \oplus [0, u_d]$ of line segments connecting the origin to the lattice points u_j. Its volume is given by $\mathrm{vol}(C) = |\det(U)|$ and we call $1/\mathrm{vol}(C)$ the *pixel density* or *sampling rate* of $\mathbb{L}^d$. The set of all lattice cells $\{C + x \,:\, x \in \mathbb{L}^d\}$ covers $\mathbb{R}^d$, i.e., $\mathbb{R}^d = \bigcup_{x\in\mathbb{L}^d}(C + x)$. The basis of a lattice is not uniquely determined but in the context of image processing it is preferential to choose the basis with minimal norms $\|u_1\|_2, \ldots, \|u_d\|_2$.

Definition 2.34. Let $\mathbb{L}^d$ be a homogeneous lattice with unit cell C and let $X \subseteq \mathbb{R}^d$. The set

$$\bigcup_{x\in X\cap\mathbb{L}^d} (C_0 + x)$$

with centered unit cell $C_0 = \frac{1}{2}(C + \check{C})$ is called the *Gauss digitization* of the set X with respect to $\mathbb{L}^d$. The set $X \cap \mathbb{L}^d$ is called the $\mathbb{L}^d$-sampling of X or *set of foreground pixels* while $X^C \cap \mathbb{L}^d$ is the *set of background pixels*.

Clearly, the Gauss digitization and $\mathbb{L}^d$-sampling of a set X carry the same information about X. The latter can be locally described by *pixel configurations* $\xi_l \subseteq \mathcal{F}^0(C)$ where the index $l \in \{0, \ldots, \nu_d\}$ with $\nu_d = 2^{2^d} - 1$ is constructed as follows. Let $\mathcal{F}^0(C) = \{x_0, \ldots, x_{2^d-1}\}$ denote the set of vertices of the unit cell where

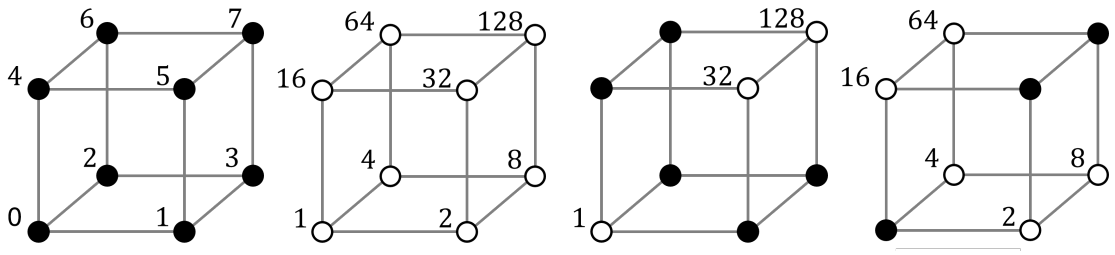
$$x_m = \sum_{j=1}^{d} \lambda_j u_j \quad \text{with} \quad m = \sum_{j=1}^{d} 2^{j-1}\lambda_j \text{ and } \lambda_j \in \{0, 1\}.$$

Then each pixel configuration $\xi_l \subseteq \mathcal{F}^0(C)$ can be uniquely identified by the index

$$l = \sum_{j=0}^{2^d-1} 2^j \mathbb{1}_{\xi_l}(x_j), \quad l \in \{0, \ldots, \nu_d\}$$

with characteristic function $\mathbb{1}_{\xi_l}(x) = 1$ if $x \in \xi_l$ and zero otherwise, see Figure 2.1. By

construction it holds that $\xi_{\nu_d} = \mathcal{F}^0(C)$, $\xi_0 = \emptyset$ and $\xi_{\nu_d - l} = \xi_{\nu_d} \backslash \xi_l$. For each pixel configuration ξ_l we denote its convex hull by $F_l = \mathrm{conv}(\xi_l) \subseteq C$. In addition we write $\mathcal{F}^j(F)$ for the set of all j-dimensional faces of the convex polytope F and $\mathcal{F}^j(\mathbb{F}) = \{\mathcal{F}^j(F) : F \in \mathbb{F}\}$ for a set of convex polytopes $\mathbb{F}$.



(a) Vertex labels m (b) $l = 255$ (c) $l = 161$ (d) $l = 94$

Figure 2.1: Visualization of the numbering for pixel configurations based on the vertex labels (a). (b-d) display configurations and corresponding vertex weights.

Definition 2.35. A set $\mathbb{F}_{\mathrm{loc}} \subseteq \{F_0, \ldots, F_{\nu_d}\}$ is a *local adjacency system* if the following conditions are fulfilled:

(i) $\emptyset, C \in \mathbb{F}_{\mathrm{loc}}$,

(ii) if $F \in \mathbb{F}_{\mathrm{loc}}$ then $\mathcal{F}^j(F) \subset \mathbb{F}_{\mathrm{loc}}$ for $j = 0, \ldots, \dim F$,

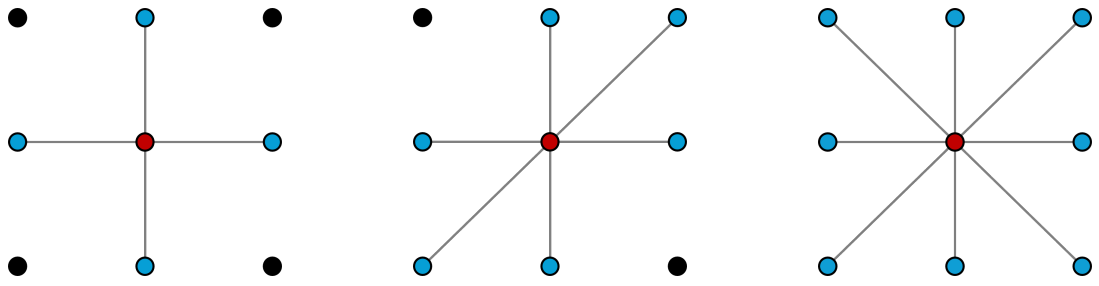
(iii) if $F_j, F_k \in \mathbb{F}_{\mathrm{loc}}$ and $\mathrm{conv}(F_j \cup F_k) \notin \mathbb{F}_{\mathrm{loc}}$ then $F_j \cap F_k, \overline{F_j \backslash F_k}, \overline{F_k \backslash F_j} \in \mathbb{F}_{\mathrm{loc}}$,

(iv) if $F_{j_1}, \ldots, F_{j_m} \in \mathbb{F}_{\mathrm{loc}}$ and $F = \bigcup_{k=1}^m F_{j_k}$ is convex then $F \in \mathbb{F}_{\mathrm{loc}}$ for $m = 2, \ldots, \#\mathbb{F}_{\mathrm{loc}}$.

The set $\mathbb{F} = \bigcup_{x \in \mathbb{L}^d} (\mathbb{F}_{\mathrm{loc}} + x)$ is called an *adjacency system* of the lattice $\mathbb{L}^d$.

The pair $\Gamma(\mathbb{F}) = \{\mathcal{F}^0(\mathbb{F}), \mathcal{F}^1(\mathbb{F})\}$ is called the *neighborhood graph* for an adjacency system $\mathbb{F}$. All vertices of Γ are of the same order which is called the *connectivity* of $\mathbb{F}$.

Example 2.36. For $d = 2$ there exist only 3 unique adjacency systems $\mathbb{F}_4$, $\mathbb{F}_6$ and $\mathbb{F}_8$ on a cubic lattice. They are characterized by their number of neighbors and can be distinguished by their neighborhood graphs, see Figure 2.2.



(a) $\mathbb{F}_4$ (b) $\mathbb{F}_6$ (c) $\mathbb{F}_8$

Figure 2.2: Neighborhood graphs of adjacency systems in 2D.

Definition 2.37. Let $X \subseteq \mathbb{R}^d$ and $\mathbb{F}$ be an adjacency system. The set

$$X \sqcap \mathbb{F} := \bigcup_{F \in \mathbb{F} : \mathcal{F}^0(F) \subseteq X} F$$

is the *discretization* of X with respect to $\mathbb{F}$.

The discretization $X \sqcap \mathbb{F}$ can be understood as a reconstruction of X from its $\mathbb{L}^d$-sampling using polyhedra derived from the adjacency system. It therefore reflects a more refined approximation of X compared to the Gauss digitization which does not consider relations between neighboring pixels.

2.2.7 Estimation of Intrinsic Volumes

In the following, let $\mathbb{L}^d = a\mathbb{Z}^d$ with $a > 0$ be a cubic lattice and P be a d-dimensional polytope. We denote the number of its j-dimensional faces by $\#\mathcal{F}^j(P)$. The Euler characteristic of P, see Example 2.25, can be computed by the *Euler-Poincaré formula* [113]

$$\chi(P) = \sum_{j=0}^{d} (-1)^j \#\mathcal{F}^j(P),$$

which can be derived from (2.7). This relation can be exploited when considering the discretization $X \sqcap \mathbb{F}$ of a compact set X with respect to the adjacency system $\mathbb{F}$ such that the calculation of the Euler characteristic reduces to

$$\chi(X \sqcap \mathbb{F}) = \sum_{l=0}^{\nu_d} \chi_0^d(\xi_l \sqcap \mathbb{F}) h_l \tag{2.12}$$

where χ_0^d are edge-corrected localizations of χ [91] for the pixel configuration ξ_l and h_l for $l = 0, \ldots, \nu_d$ are the number their occurrences in $X \cap \mathbb{L}^d$. Most importantly, $\chi_0^d(\xi_l \sqcap \mathbb{F})$ are weights that depend entirely on the adjacency system but not on X and h_l depend only on the sampling $X \cap \mathbb{L}^d$ but are independent of the adjacency system.
When choosing an adjacency system $\mathbb{F}$ for the discretization of X, an appropriate adjacency system $\mathbb{F}_c$ for the discretization of X^C needs to be selected.

Definition 2.38. The pair $(\mathbb{F}, \mathbb{F}_c)$ is called a *pair of complementary adjacency systems* if $(X \sqcap \mathbb{F}) \cap (X^C \sqcap \mathbb{F}_c) = \emptyset$ and

$$\chi(X \sqcap \mathbb{F}) = (-1)^{d+1} \chi(X^C \sqcap \mathbb{F}_c)$$

for all compact $X \subseteq \mathbb{R}^d$. An adjacency system is called *self-complementary* if $(\mathbb{F}, \mathbb{F})$ is a pair of complementary adjacency systems.

The set X^C is neither compact nor polyconvex. However, the Euler characteristic of X^C and $X^C \sqcap \mathbb{F}_c$ can be introduced via *Hadwiger's recursive formula* [91]. We note, that a complementary adjacency system $\mathbb{F}_c$ for a given $\mathbb{F}$ need not necessarily exist. It suffices to check the definition of complementary adjacency systems for all the pixel configurations ξ_l instead of all compact $X \subseteq \mathbb{R}^d$.

Proposition 2.39. Let $\mathbb{F}$ be an adjacency system on $\mathbb{L}^d$. Then for any compact and morphologically regular $X \subseteq \mathbb{R}^d$ it holds that

$$\lim_{a \searrow 0} \chi(X \sqcap a\mathbb{F}) = \chi(X).$$

This means that the Euler characteristic is *multi-grid convergent* for sets that are morphologically regular. The estimation of the remaining intrinsic volumes can be derived from the Crofton's intersection formula (2.11) with $j = 0$ which yields

$$V_{d-k}(X) = \frac{1}{\alpha_{d,0,k}} \int_{\mathcal{L}^k} \underbrace{\int_{L_k^\perp} \chi(X \cap (L_k + x)) V_{L_k^\perp}(\mathrm{d}x)}_{p_k(X, L_k)} \, \mu(\mathrm{d}L_k), \quad k = 1, \ldots, d-1.$$

Hence, the intrinsic volumes can be calculated through the Euler characteristic measured on lower-dimensional subspaces.

Definition 2.40. A pair $(\mathbb{L}^k, {}^\perp\mathbb{L}^{d-k})$ with $k \in \{1, \ldots, d-1\}$ is called a *k-dimensional section lattice* $\mathbb{L}^k$ of $\mathbb{L}^d$ with translative complement ${}^\perp\mathbb{L}^{d-k}$, if there exists a basis $v_1, \ldots, v_d$ of $\mathbb{L}^d$ with

(i) $\mathbb{L}^k = (v_1, \ldots, v_k)\mathbb{Z}^k$

(ii) ${}^\perp\mathbb{L}^{d-k} = (v_{k+1}, \ldots, v_d)\mathbb{Z}^{d-k}$

(iii) there is an $x \in \mathcal{F}^0(\check{C})$ with $\{v_1, \ldots, v_k\} \subseteq \mathcal{F}^0(C + x)$ where C is the unit cell of $\mathbb{L}^d$.

Analogous to our understanding of the lattice $\mathbb{L}^d$ as a discretization of $\mathbb{R}^d$, a k-dimensional section lattice $\mathbb{L}^k$ and its translative complement ${}^\perp\mathbb{L}^{d-k}$ are discretizations of the k-dimensional linear subspace $L_k = \operatorname{span}(\mathbb{L}^k)$ and its orthogonal complement. In particular, we have that the lattice is the union of shifts of the section lattice over its complement, i.e.,

$$\mathbb{L}^d = \bigcup_{x \in {}^\perp\mathbb{L}^{d-k}} \mathbb{L}^k + x.$$

In comparison to $\mathbb{R}^d$ where linear subspaces can be arbitrarily oriented, section lattices are bound to the lattice $\mathbb{L}^d$ which results in only finitely many k-dimensional section lattices.

Example 2.41. For $d = 3$ and $\mathbb{L}^3 = \mathbb{Z}^3$ there exists a total of 13 section lattices $\mathbb{L}^k_j$ for $k = 1$ or $k = 2$ which can be constructed as shown in Table 2.1

j	Basis of $\mathbb{L}_j^1$	Basis of ${}^{\perp}\mathbb{L}_j^2$	j	$\mathbb{L}_j^2$	Basis of ${}^{\perp}\mathbb{L}_j^1$
1	$\{u_1\}$	$\{u_2, u_3\}$	1	$\{u_1, u_2\}$	$\{u_3\}$
2	$\{u_2\}$	$\{u_1, u_3\}$	2	$\{u_1, u_3\}$	$\{u_2\}$
3	$\{u_3\}$	$\{u_1, u_2\}$	3	$\{u_2, u_3\}$	$\{u_1\}$
4	$\{u_1 + u_2\}$	$\{u_1, u_3\}$	4	$\{u_1, u_2 + u_3\}$	$\{u_3\}$
5	$\{-u_1 + u_2\}$	$\{u_1, u_3\}$	5	$\{u_1, -u_2 + u_3\}$	$\{u_3\}$
6	$\{u_1 + u_3\}$	$\{u_1, u_2\}$	6	$\{u_2, u_1 + u_3\}$	$\{u_3\}$
7	$\{-u_1 + u_3\}$	$\{u_1, u_2\}$	7	$\{u_2, -u_1 + u_3\}$	$\{u_3\}$
8	$\{u_2 + u_3\}$	$\{u_1, u_3\}$	8	$\{u_3, u_1 + u_2\}$	$\{u_1\}$
9	$\{-u_2 + u_3\}$	$\{u_1, u_3\}$	9	$\{u_3, -u_1 + u_2\}$	$\{u_1\}$
10	$\{u_1 + u_2 + u_3\}$	$\{u_1, u_2\}$	10	$\{u_1 + u_3, u_2 + u_3\}$	$\{u_3\}$
11	$\{-u_1 + u_2 + u_3\}$	$\{u_1, u_2\}$	11	$\{-u_1 + u_3, u_2 + u_3\}$	$\{u_3\}$
12	$\{u_1 - u_2 + u_3\}$	$\{u_1, u_2\}$	12	$\{-u_1 + u_3, -u_2 + u_3\}$	$\{u_3\}$
13	$\{-u_1 - u_2 + u_3\}$	$\{u_1, u_2\}$	13	$\{u_1 + u_3, -u_2 + u_3\}$	$\{u_3\}$

Table 2.1: Bases of the 13 section lattices $\mathbb{L}_j^k$ of $\mathbb{Z}^3$ for $k = 1$ and $k = 2$.

The definition of section lattices allows us to discretize the integrals in Crofton's intersection formula (2.11) as follows. Consider the k-dimensional subspace $L_k = \operatorname{span}(\mathbb{L}^k)$. The inner integral $p_k(X, L_k)$ consist only of Euler characteristics $\chi(X \cap (L_k + x))$ calculated on the lower-dimensional subspace shifted across its orthogonal complement. We obtain the Euler characteristics analogous to (2.12) and estimate

$$p_k(X, L_k) \approx \hat{p}_k(X, L_k) = \frac{\operatorname{vol}(C)}{\operatorname{vol}(C^k)} \sum_{l=0}^{\nu_k} \chi_0^k(\xi_l \sqcap \mathbb{F}^k) h_l^{(k)}$$

where ν_k, $\chi_0^k(\xi_l \sqcap \mathbb{F}^k)$ and $h_l^{(k)}$ are the respective quantities from before but with respect to the section lattice $\mathbb{L}^k$ with adjacency system $\mathbb{F}^k$. The factor $\operatorname{vol}(C)/\operatorname{vol}(C^k)$ is a correction factor that arises by projecting the unit cell of the translative complement onto the orthogonal complement $L_k^{\perp} = \operatorname{span}(\mathbb{L}^k)^{\perp}$. We remark that $h_l^{(k)}$ can be deduced from the pixel configurations on $\mathbb{L}^d$ via a linear transformation yielding

$$\hat{p}_k(X, L_k) = a^{d-k} \sum_{l=0}^{\nu_d} w_l^{(k)} h_l$$

with weights $w_l^{(k)}$ that are independent of X. For the outer integral of (2.11) we apply a simple quadrature rule over the finitely many subspaces $L_k^j = \operatorname{span}(\mathbb{L}_j^k)$, $j = 1, \ldots, m_k$ yielding

$$\int_{\mathcal{L}^k} p_k(X, L_k) \mu(\mathrm{d}L_k) \approx \sum_{j=1}^{m_k} \gamma_j^{(k)} \hat{p}_k(X, L_k^j)$$

with suitable weights $\gamma_j^{(k)}$. In higher dimension there exists no natural choice of weights as the k-dimensional subspaces in $\mathcal{L}^k$ are not uniformly scattered. However, for $d = 2$ and $d = 3$ intuitive choices are possible [91, pp. 156f]. By combining the discretizations of

integrals we can estimate the intrinsic volumes as

$$\hat{V}_{d-k}(X) = \frac{a^{d-k}}{\alpha_{d,0,k}} \sum_{l=0}^{\nu_d} v_l^{(k)} h_l,$$

$k = 1, \ldots, d$ where the weights $v_l^{(k)} = \sum_{j=1}^{m_k} \gamma_j^{(k)} w_l^{(j,k)}$ are independent of X and h_l are the number of occurrences of the pixel configurations ξ_l in $X \cap \mathbb{L}^d$. Table 5.2 in [91] presents different weights that were proposed for the calculation of the surface area and integral of mean curvature in the case $d = 3$ on a cubic lattice. We complete the set of estimators by that of the volume V_d and Euler number $\chi = V_0$

$$\hat{V}_d(X) = a^d \sum_{l=0}^{\nu_d} v_l^{(0)} h_l \quad \text{and} \quad \hat{V}_0(X) = \sum_{l=0}^{\nu_d} v_l^{(d)} h_l$$

with coefficients $v_l^{(0)} = \mathrm{vol}(C)/a^d$ for odd l and $v_l^{(0)} = 0$ otherwise, and $v_l^{(d)} = \chi_0^d(\xi_l \sqcap \mathbb{F})$.

2.2.8 Estimation of Intrinsic Volume Densities for Stationary Random Sets

The ideas of estimating the intrinsic volumes extends directly to the intrinsic volume densities of a stationary random closed set Ξ which is observed through a compact window $W \subseteq \mathbb{R}^d$ on $\mathbb{L}^d$. Following the definitions (2.8) and (2.9) we obtain estimators via

$$\hat{V}_{V,d-k}(\Xi) = \frac{\sum_{l=0}^{\nu_d} v_l^{(k)} \tilde{h}_l}{a^k \alpha_{d,0,k} \sum_{l=0}^{\nu_d} \tilde{h}_l}, \quad k = 0, \ldots, d$$

where $\tilde{h}_l$ is number of occurrences of the pixel configurations l in $\Xi \cap (W \ominus \check{C}) \cap \mathbb{L}^d$. The estimation based on the reduced window $W \ominus \check{C}$ is due to edge effects and not necessary for estimating $V_{V,d}(\Xi)$ [91].

2.3 Fourier Transform

The Fourier transformation is an essential component in modern signal and image processing. It can be understood as a decomposition of the signal or image into constituent frequencies, i.e., a transition from the classical spatial domain to the frequency domain which is often referred to as *Fourier domain*. In the following, we introduce the theory of the Fourier series and the discrete Fourier transform which applies to discrete and finite signals and images. This section is largely based on [52] but also utilized parts of [13, 47, 98, 129].

2.3.1 Fourier Series

In the following we consider 1-periodic functions $f : \mathbb{R}^d \to \mathbb{C}$ with $f(x + k) = f(x)$ for all $k \in \mathbb{Z}^d$. Then f can be uniquely identified with a function on the *d-torus* T^d of length 1, which is usually considered as the quotient space $\mathbb{R}^d/\mathbb{Z}^d$ or a representative interval $[0, 1)^d$. We consider the Banach space of continuous functions on T^d as $C(T^d)$ with norm

$$\|f\|_{C(T^d)} = \max_{x \in T^d} |f(x)|.$$

Further, $L^2(T^d)$ is the Hilbert space of L^2-integrable functions on T^d with the scalar product and norm

$$\langle f, g\rangle_{L^2(T^d)} \coloneqq \int_{[0,1)^d} f(x)\overline{g(x)}\mathrm{d}x, \quad \|f\|_{L^2(T^d)} \coloneqq \left(\int_{[0,1)^d} |f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}},$$

i.e., $f \in L^2(T^d)$ if $\|f\|_{L^2(T^d)} < \infty$. Similarly, the Banach space $L^1(T^d)$ of L^1-integrable functions is defined through the norm

$$\|f\|_{L^1(T^d)} \coloneqq \int_{[0,1)^d} |f(x)|\mathrm{d}x < \infty.$$

It can be shown that $L^1(T^d) \supset L^2(T^d) \supset C(T^d)$ holds true. We denote the *(cyclic) convolution* on T^d as

$$(f * g)(x) \coloneqq \int_{[0,1)^d} f(y)g(x-y)\mathrm{d}y.$$

Proposition 2.42. The system of functions

$$\left\{e^{2\pi i\langle k,x\rangle} \,:\, k \in \mathbb{Z}^d\right\}$$

forms an orthonormal basis of $L^2(T^d)$.

From this result we can deduce that for any function $f \in L^2(T^d)$, there exists a unique representation

$$f = \sum_{k\in\mathbb{Z}^d} c_k(f)e^{2\pi i\langle k,\cdot\rangle} \quad \text{with} \quad c_k(f) = \langle f, e^{2\pi i\langle k,\cdot\rangle}\rangle_{L^2(T^d)}, \tag{2.13}$$

which we call the *Fourier series* of f. Its coefficients $c_k(f)$ are referred to as the *Fourier coefficients* of f and the set $\{|c_k(f)|\}_{k\in\mathbb{Z}^d}$ as the *spectrum* of f. It can be shown that the mapping $f \mapsto \{c_k(f)\}_{k\in\mathbb{Z}^d}$ is an isometry from $L^2(T^d)$ onto the space of square-summable sequences ℓ^2. The definition of the Fourier coefficients extends towards functions $f \in L^1(T^d)$. However, in this case the convergence of the corresponding Fourier series is not guaranteed. A sufficient condition for convergence almost everywhere is that $\sum_{k\in\mathbb{Z}^d} |c_k(f)| < \infty$.

Proposition 2.43. Let $f, g \in L^1(T^d)$, $x_0 \in T^d$ and $k_0 \in Z^d$. Then it holds

(i) $c_k(\lambda_1 f + \lambda_2 g) = \lambda_1 c_k(f) + \lambda_2 c_k(g)$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$,

(ii) $c_{k+k_0}(f) = c_k(e^{-2\pi i\langle k_0,\cdot\rangle}f)$ and $c_k(f(\cdot - x_0)) = e^{-2\pi i\langle k,x_0\rangle}c_k(f)$,

(iii) $c_k(f * g) = c_k(f)c_k(g)$,

(iv) $c_k(\partial^\alpha f) = (2\pi ik)^\alpha c_k(f)$ for all $\alpha \in \mathbb{N}_0^d$ if $f \in C^\alpha(T^d)$.

These properties show that the Fourier coefficients behave well under classical manipulations. Especially (iii) and (iv) stand out since they show that computationally expensive operations such as the convolution or differentiation become easily computable through the Fourier series via simple multiplications.

Example 2.44. Consider the trigonometric polynomial

$$D_n^d(x) = \sum_{|k|\leq n} e^{2\pi i\langle k,x\rangle} \text{ with } n \in \mathbb{N}$$

which is known as *(circular) Dirichlet kernel*. Its Fourier coefficients are $c_k(D_n) = 1$ if $|k| \leq n$ and 0 otherwise. Hence, for any $f \in L^1(T^d)$ it holds that

$$S_n(x) = (f * D_n^d)(x) = \sum_{|k|\leq n} c_k(f)e^{-2\pi i\langle k,x\rangle}$$

by which is called the n-th *Fourier partial sum* of f. For the special case $d = 1$ it holds that

$$D_n(x) = D_n^1(x) = \frac{\sin((2n+1)\pi x)}{\sin(\pi x)} \text{ for } x \in [0, 1).$$

The Dirichlet kernel is a special case of a *low-pass filter*, that is, a function f whose Fourier coefficients fulfill $|\hat{f}(k)| \approx 0$ for large $|k|$. Analogously, a *high-pass filter* is characterized as a function with Fourier coefficients $|\hat{f}(k)| \approx 0$ for small $|k|$.

Definition 2.45. A sequence of $L^1(T^d)$ functions $\{\varphi_n\}_{n\in\mathbb{N}}$ is an *approximate identity* on T^d if

(i) there exists a constant $c > 0$ such that $\|\varphi_n\|_{L^1(T^d)} \leq c$ for all $n \in \mathbb{N}$.

(ii) $\int_{T^d} \varphi_n(x)\mathrm{d}x = 1$ for all $n \in \mathbb{N}$.

(iii) $\lim_{n\to\infty} \int_{\substack{x\in T^d \\ \|x\|_2\geq\delta}} |\varphi_n(x)|\mathrm{d}x = 0$ for all $\delta > 0$.

Approximate identities can be interpreted as a sequence of functions which approximate the Dirac delta function

$$\delta_0(x) = \begin{cases} \infty & \text{if } x = 0, \\ 0 & \text{otherwise} \end{cases}$$

as $n \to \infty$. Since δ_0 is the identity element of the convolution, i.e., $f = f * \delta_0$, a natural approach to approximate a function $f \in L^1(T^d)$ is to calculate the convolutions $f * \varphi_n$ and let n tend to ∞. Indeed, for approximate identities it can be shown that for any $f \in C(T^d)$ uniform convergence is obtained, i.e.,

$$\lim_{n\to\infty} \|\varphi_n * f - f\|_{C(T^d)} = 0,$$

see [52, Theorem 1.2.19]. Sadly, the Dirichlet kernels $\{D_n\}_{n\in\mathbb{N}}$ are not an approximate identity since they violate property (i) and we need to look for an alternative. Consider the *Fejer kernels*

$$F_n^d(x) = \prod_{j=1}^{d} \left(\frac{1}{n+1} \sum_{k=0}^{n} D_k(x_j) \right) = \frac{1}{(n+1)^d} \prod_{j=1}^{d} \left(\frac{\sin((n+1)\pi x_j)}{\sin(\pi x_j)} \right)^2.$$

Proposition 2.46. The Fejer kernels $\{F_n^d\}_{n\in\mathbb{N}}$ form an approximate identity on T^d.

This implies that for any $f \in C(T^d)$ the sequence $f * F_n^d$ converges uniformly to f which is known as *Weierstraß approximation theorem* [52]. We remark that the trigonometric functions $e^{-2\pi i\langle k,\cdot\rangle}$ in the Fourier series always have an absolute value of 1. Therefore, it is not surprising that the Fourier coefficients $c_k(f)$ must decay as $|k|$ grows larger, i.e.,

$$\lim_{|k|\to\infty} |c_k(f)| \to 0$$

for any $f \in L^1(T^d)$. This is known as the *lemma of Riemann-Lebesgue* [52] and indicates that higher frequencies are typically dominated by lower frequencies.

2.3.2 Discrete and Fast Fourier Transform (FFT)

Let $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$ and $N = n_1 \cdot \ldots \cdot n_d$. We denote the homogeneous lattice $\mathbb{L}^d = \{x \in \mathbb{R}^d \,:\, x = k \oslash n, k \in \mathbb{Z}^d\}$ with element-wise division $\oslash$. In the following we consider periodic functions $f : \mathbb{L}^d \to \mathbb{C}$ with $f(x+k) = f(x)$ for all $k \in \mathbb{Z}^d$. Similar to the previous section, f can be uniquely identified with a function on the *discretized d-torus* $\hat{T}^d$ which can be considered as the quotient space $\mathbb{L}^d/\mathbb{Z}^d$ or a representative set of sampling points $S = \{j \oslash n \,:\, j \in J\}$ with the set of indices

$$J = \{(j_1, \ldots, j_d) \in \mathbb{Z}^d \,:\, 0 \le j_m \le n_m - 1 \text{ for all } m = 1, \ldots, d\}.$$

The latter representation holds because the number of sampling points $|S| = N$ is finite. In particular, we write $f = (f_j)_{j\in J} := (f(x))_{x\in S} \in \mathbb{C}^n$ with $\mathbb{C}^n := \mathbb{C}^{n_1\times\ldots\times n_d}$. By applying the trapezoidal rule we approximate the integrals of the Fourier coefficients (2.13) by

$$c_k(f) \approx \frac{1}{N}\sum_{x\in S} f(x)e^{-2\pi i\langle k,x\rangle} = \frac{1}{N}\sum_{j\in J} f_j e^{-2\pi i\langle k, j\oslash n\rangle} =: \frac{1}{N}\hat{f}_k. \tag{2.14}$$

We remark that the approximations are periodic with $\hat{f}_k = \hat{f}_{k+j}$ for every $j \in n \odot \mathbb{Z}^d$ with element-wise multiplication $\odot$. In particular, only N unique approximations exist. It can be shown that the *aliasing formula*

$$\frac{1}{N}\hat{f}_k = \sum_{j\in n\odot\mathbb{Z}^d} c_{k+j}(f)$$

holds true which reveals how the infinite series of Fourier coefficient relates to the finite set of approximations. An intuitive interpretation of the aliasing formula is that high frequencies become indistinguishable from lower ones due to the discrete sampling.

Proposition 2.47. The system of tensors

$$\left\{(e^{2\pi i\langle k, j\oslash n\rangle})_{j\in J} \,:\, k \in J\right\}$$

forms an orthonormal basis of $\mathbb{C}^n$.

This result is the equivalent to Proposition 2.42 and implies that the representation of the estimates in (2.14) is unique. Further, it can be shown that the mapping $f \mapsto \frac{1}{\sqrt{N}}(\hat{f}_k)_{k\in J}$ is unitary. Most importantly, an inverse mapping exists.

Definition 2.48. Let $n \in \mathbb{N}^d$. The linear map

$$\mathscr{F} : \mathbb{C}^n \to \mathbb{C}^n, \quad f \mapsto \mathscr{F}(f) = \left(\sum_{j \in J} f_j e^{-2\pi i \langle k, j \oslash n \rangle} \right)_{k \in J} \tag{2.15}$$

is the *discrete Fourier transform* (DFT) on $\mathbb{C}^n$. The inverse is given by

$$\mathscr{F}^{-1} : \mathbb{C}^n \to \mathbb{C}^n, \quad \hat{f} \mapsto \mathscr{F}^{-1}(\hat{f}) = \left(\frac{1}{N} \sum_{k \in J} \hat{f}_k e^{2\pi i \langle k, j \oslash n \rangle} \right)_{j \in J}$$

and called the *inverse discrete Fourier transform* (IDFT) on $\mathbb{C}^n$.

Despite considering only approximations, we still refer to $\hat{f}_k$ as the Fourier coefficients of f and $\{|\hat{f}_k|\}_{k \in J}$ as the spectrum.

Example 2.49.

(i) In the case $d = 1$ where $f \in \mathbb{C}^n$ is a vector, the DFT can be written in matrix form as $\mathscr{F}(f) = \mathscr{F}_n f$ with the *Fourier matrix* $\mathscr{F}_n := \left(e^{-2\pi i k j / n}\right)_{j,k=0}^{n-1}$.

(ii) For $d = 2$ with $f \in \mathbb{C}^{n_1 \times n_2}$ we can express the DFT via a matrix product

$$\mathscr{F}(f) = \mathscr{F}_{n_1} f \mathscr{F}_{n_2}.$$

We can expand and rewrite the entries in the DFT (2.15) as

$$\begin{aligned} \sum_{j \in J} f_j e^{-2\pi i \langle k, j \oslash n \rangle} &= \sum_{j_1=0}^{n_1-1} \cdots \sum_{j_d=0}^{n_d-1} f_{j_1,\ldots,j_d} e^{-2\pi i \langle (k_1,\ldots,k_d), \left(\frac{j_1}{n_1},\ldots,\frac{j_d}{n_d}\right) \rangle} \\ &= \sum_{j_1=0}^{n_1-1} e^{-2\pi i k_1 j_1 / n_1} \left(\cdots \left(\underbrace{\sum_{j_d=0}^{n_d-1} e^{-2\pi i k_d j_d / n_d} f_{j_1,\ldots,j_d}}_{=\hat{f}^{(d)}} \right) \right) \end{aligned}$$

and notice that $\hat{f}^{(d)}$ can be interpreted as an entry of the 1-dimensional DFT of

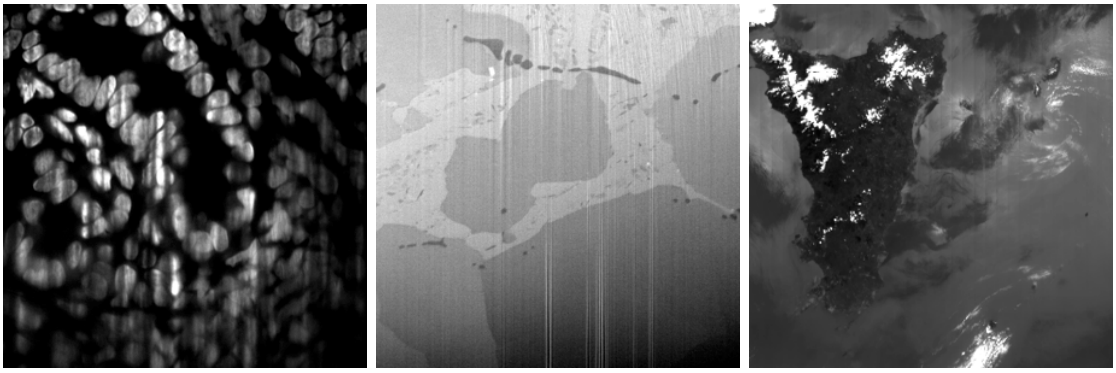
$$f^{(d)} = f\left(\frac{j_1}{n_1}, \ldots, \frac{j_{d-1}}{n_{d-1}}, \cdot \right).$$

More precisely, we can calculate the DFT by sequentially applying 1-dimensional DFTs along each individual dimension, i.e., the DFT is separable. Computing the DFT directly as given by (2.15) has complexity $\mathcal{O}(N^2)$ in arithmetic operations. However, computations can be optimized using a *divide and conquer* strategy yielding the *fast Fourier transform* (FFT) with complexity $\mathcal{O}(N \log(N))$. For more details on this approach in $d = 1$ we refer to [98]. Higher dimensions follow immediately by the separability.

3 Stripe Removal

3.1 Introduction

Stripe artifacts are a common problem across several imaging techniques including light-sheet fluorescence microscopy (LSFM), focused ion beam scanning electron microscopy (FIB-SEM) and remote sensing. In general, these artifacts are characterized by highly elongated shapes of low width which point in a common direction. Their removal also known as *destriping* is required not only to improve visual quality but also to enable further analysis and image processing of the data. The examples shown in Figure 3.1 highlight the diversity of image structures and stripes encountered in practice. While stripes are periodic, thin, long and of low visibility in (c), much more severe artifacts of varying width, smaller length and higher intensity can be seen in (a).



(a) LSFM (b) FIB-SEM (c) Remote sensing

Figure 3.1: Stripe corruptions in LSFM image of cultivated mouse intestine cells (a), FIB-SEM image of a tin bronze cast alloy (b) and remote sensing data from a Terra MODIS satellite (c).

There exists a large body of literature dealing with stripe removal including a recently published article by Ricci and colleagues [101] which provides a comprehensive overview of previous research on stripe removal in LSFM. One approach is the prevention of stripe artifacts, e.g., using digital scanned light beams [39, 68, 69], multi-directional illumination [60, 111] or polarization filters [35]. Alternatively, post-processing methods can be employed to remove the stripes after acquisition such as filtering of Fourier coefficients [19, 72, 89] or variational methods [17, 43]. Lastly, hybrid solutions combining elements from both can be considered, for example, by combining multiple views [96, 122]. Preventing the formation of artifacts would be preferable but typically requires specialized hardware equipment or can limit the performance, mainly in terms of speed [40–42]. In contrast, post-processing is cheap and requires only computational power. Furthermore, corrupted images may already exist and specimens cannot be re-imaged, e.g., due to degradation

and aging. For these reasons, it is desirable to have powerful post-processing tools for stripe removal available.
In this chapter, we propose, explain and validate a general solution for removing stripe artifacts named general stripe remover (GSR). We quantitatively compare it to previous approaches across images from LSFM, FIB-SEM and remote sensing. Our solution generalizes exceptionally well on all data tested. Thanks to intuitively adjustable parameters, it is highly flexible and customizable for different appearances of structures and stripes. The source code is shared on GitHub [108] together with comprehensive guidance for application without further knowledge. The method belongs to the class of variational methods and is compared against selected state-of-the-art solutions from the same category and Fourier filtering methods. The performance is quantitatively evaluated with the help of synthetic LSFM data obtained by physically correct simulation of light transport with the Python package *biobeam* [135]. We include simulation data since it provides ground truth images that are unavailable for experimentally acquired images. This enables an objective assessment of performance using common quality metrics such as peak signal-to-noise ratio or the multi-scale structural similarity index measure. The evaluation is supplemented by comparing line profiles and using the curtaining metric, a measure of stripe corruptions proposed by Roldán [107].
We remark that parts of this chapter have already been published by the author of this thesis in [110].

3.2 Methods

Most methods proposed for stripe removal fall into one of two categories, *Fourier filtering* and *variational methods*. We focus on these categories, as they generalize well to several imaging methods, variations in image structures and stripe appearances. Other approaches include average filtering [3, 10], histogram matching [14, 97], spline interpolation [126] and, recently, neural networks [76, 94, 131, 134]. However, these methods are usually tailored for a specific appearance of images and stripes such that they are harder to transfer to other scenarios.

3.2.1 Variational Methods

Variational methods are applied to tasks such as denoising, segmentation, active contours [127] or destriping [46] by minimizing a task-specific convex objective function that is constructed by penalizing unwanted features of the uncorrupted image and stripes. Therefore, a minimizer of the function has desirable properties such as being free of stripe artifacts. Under mild assertions on the function, a unique minimizer exists and can be approximated using well-studied optimization algorithms, see Section 2.1.
Several variational methods for removing stripes were proposed in the past. The contributions [38, 43] represent stripes using elementary stripe patterns. Other propositions assume sparsity and employ a low-rank assumption on the stripe image [18, 20, 138, 143]. Alternatively, [70] includes additional information from the Fourier domain. In the following, we introduce, motivate and explain our objective function which builds on previous work by Fitschen et al. [46]. A similar objective function to ours was also explored by Liu et al. [75] but only for remote sensing images. In contrast to previous implementations, our objective function aims to effectively address stripes, retain image structures, and

provide a reasonable intensity profile simultaneously.
Let $n = (n_x, n_y, n_z) \in \mathbb{N}^3$ contain the image sizes. We consider a corrupted image as $u_0 \in [0,1]^n$ where $[0,1]^n = [0,1]^{n_x \times n_y \times n_z}$ and assume that it can be decomposed as $u_0 = u + s$ into a clean image u and stripes s. Let ∇_x, ∇_y and ∇_z denote the directional difference operators for the three coordinate directions and assume that stripes point in y-direction. In general, variational methods are special convex optimization problems

$$\operatorname*{argmin}_{u+s=u_0} D(u) + N(s)$$

whose objective functions are typically composed of a data term $D(u)$ and noise term $N(s)$, which penalize undesired properties of the clean image and stripes, respectively.

3.2.1.1 General Stripe Remover (GSR)

We propose the optimization problem

$$\operatorname*{argmin}_{u+s=u_0} \underbrace{\mu_1 \left\|\nabla^* u\right\|_{2,1} + \iota_{[0,1]^n}(u)}_{D(u)} + \underbrace{\left\|\nabla_y s\right\|_1 + \mu_2 \left\|s\right\|_1}_{N(s)}, \quad \mu_1, \mu_2 > 0 \tag{3.1}$$

where $\nabla^* = (\nabla_x, \nabla_y, \rho_z \nabla_z)$ and

$$\begin{aligned} \left\|\nabla^* u\right\|_{2,1} &= \sum_{i,j,k} \sqrt{(\nabla_x u)_{i,j,k}^2 + (\nabla_y u)_{i,j,k}^2 + (\rho_z \nabla_z u)_{i,j,k}^2}, \quad \rho_z \in [0,1], \\ \left\|\nabla_y s\right\|_1 &= \sum_{i,j,k} |(\nabla_y s)_{i,j,k}| \quad \text{and} \quad \left\|s\right\|_1 = \sum_{i,j,k} |s_{i,j,k}|. \end{aligned} \tag{3.2}$$

The individual terms making up the objective function can be motivated and interpreted as follows:

(i) $\left\|\nabla^* u\right\|_{2,1}$ encodes that the clean image contains only few strong edges, i.e., it is piece-wise smooth. This term is known as *total variation* [4] for $\rho_z = 1$.

(ii) $\iota_{[0,1]^n}(u)$ ensures that the input and clean image share the same value range. In particular, intensities exceeding the logical bounds, e.g., negative values, are prohibited which prevents the need for brightness adjustments after processing.

(iii) $\left\|\nabla_y s\right\|_1$ emphasizes the property that stripes have large differences orthogonal and small differences parallel to the stripe direction.

(iv) $\left\|s\right\|_1$ promotes sparsity in the stripe image and reflects that typically only a small part of the image is directly affected by artifacts.

The parameter ρ_z in (3.2) can be used to incorporate that information along the z-axis may be less reliable and coherent, e.g., due to lower resolution which is common for LSFM and FIB-SEM images. A natural choice for ρ_z is the ratio of the voxel sizes along the x and y-axes compared to the z-axis. Alternatively, $\rho_z = 0$ can be considered to process images in 2D, i.e., slice-by-slice. μ_1 and μ_2 are free parameters that control the stripe removal, regularization and smoothness of the outcome. They enable us to adapt GSR to different conditions ranging from severe corruptions to weak obscurities or ambiguities.

For a detailed analysis and discussion about the choice of μ_1 and μ_2 we refer to Section 3.2.1.1.3. A basic intuition is the following: The parameter μ_1 adjusts the general strength of stripe removal. Larger values result in a stronger removal at the potential cost of affecting stripe-like structures and introducing smoothing artifacts. μ_2 counteracts μ_1 and controls how strongly the properties of stripes must be present for a structure to be removed. Increasing the value of μ_2 impairs the removal by enforcing more sparsity in the stripe image. This especially affects elements which do not possess stripe-like features.

3.2.1.1.1 Optimization

We optimize problem (3.1) using the PDHGMp which was introduced as Algorithm 1 in Section 2.1.5. For this we identify the problem with

$$\underset{u+s=u_0}{\operatorname{argmin}} \underbrace{\iota_{[0,1]^n}(u)}_{=g(x)} + \underbrace{\mu_1 \left\|\nabla^* u\right\|_{2,1} + \left\|\nabla_y s\right\|_1 + \mu_2 \left\|s\right\|_1}_{=h(Ax)}, \quad \mu_1, \mu_2 > 0, \tag{3.3}$$

where $x = (u, s)$ and $Ax \coloneqq (\nabla^* u, \nabla_y s, s)$ to apply the algorithm. We obtain the iterative optimization procedure

Algorithm 2: PDHGMp applied to (3.3)

Initialization: $u^{(0)} = u_0, s^{(0)} = 0, b_j^{(0)} = \bar{b}_j^{(0)} = 0, j = 1,2,3, \theta = 1, \tau = \sigma = 0.28$

for $k = 0, 1, \ldots$ **do**

$$\begin{aligned}
(u^{(k+1)}, s^{(k+1)}) &= P_{[0,1]^n}\left((u^{(k)} - \tau\sigma\left(\nabla_x, \rho_z\nabla_z\right)^T \bar{b}_1^{(k)}, s^{(k)} - \tau\sigma\nabla_y^T \bar{b}_2^{(k)} - \tau\sigma\bar{b}_3^{(k)})\right) \\
y_1^{(k+1)} &= C_{\frac{\mu_1}{\sigma}}\left(b_1^{(k)} + \left(\nabla_x, \rho_z\nabla_z\right) u^{(k+1)}\right) \\
y_2^{(k+1)} &= S_{\frac{1}{\sigma}\|\cdot\|_1}\left(b_2^{(k)} + \nabla_y s^{(k+1)}\right) \\
y_3^{(k+1)} &= S_{\frac{\mu_2}{\sigma}\|\cdot\|_1}\left(b_3^{(k)} + s^{(k+1)}\right) \\
b_1^{(k+1)} &= b_1^{(k)} + \left(\nabla_x, \rho_z\nabla_z\right) u^{(k+1)} - y_1^{(k+1)} \\
b_2^{(k+1)} &= b_2^{(k)} + \nabla_y s^{(k+1)} - y_2^{(k+1)} \\
b_3^{(k+1)} &= b_3^{(k)} + s^{(k+1)} - y_3^{(k+1)} \\
\bar{b}_j^{(k+1)} &= b_j^{(k+1)} + \theta(b_j^{(k+1)} - b_i^{(k)}) \quad \text{for } j = 1,2,3.
\end{aligned}$$

end

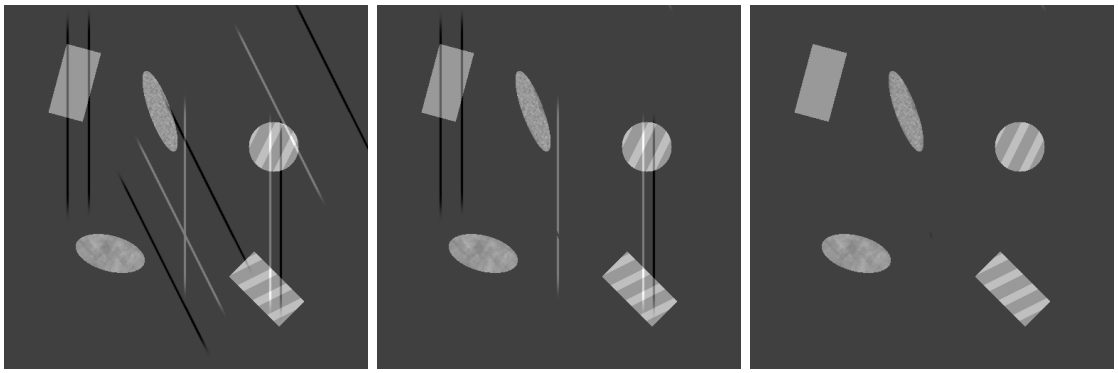
that utilizes the orthogonal projection, coupled soft-shrinkage and soft-shrinkage functions found in Example 2.10. The algorithm converges to a solution according to Theorem 2.14. We discuss the necessary number of iteration steps in practice in Section 3.2.1.1.3.

3.2.1.1.2 Arbitrary Stripe Directions

In this work, we assume that stripe artifacts are parallel to the y-axis which is usually encountered in practice. We note that stripe directions parallel to any of the other coordinate axes can easily be rotated to this case. Incorporating arbitrary stripe directions requires the alteration of the optimization problem to

$$\underset{u+s=u_0}{\operatorname{argmin}} \mu_1 \left\|\nabla^* u\right\|_{2,1} + \iota_{[0,1]^n}(u) + \left\|\nabla_\theta s\right\|_1 + \mu_2 \left\|s\right\|_1, \quad \mu_1, \mu_2 > 0 \tag{3.4}$$

where ∇_θ is an adequate directional difference operator for the stripe direction θ. In particular, we only replace the penalization term $\|\nabla_y s\|_1$ with $\|\nabla_\theta s\|_1$ in (3.1). Liu and colleagues [75] present a series of oblique difference operators for directions in 2D. This extension towards arbitrary stripe directions enables the removal of multi-directional stripe artifacts, i.e., stripes in multiple known directions $\theta_j, j = 1, \ldots, m$ for $m \in \mathbb{N}$. This can be achieved by successively solving (3.4) for each θ_j which is exemplified in Figure 3.2 for a synthetic test image.



(a) Input (b) Oblique stripe removal (c) Vertical stripe removal

Figure 3.2: Stripe removal results for multi-directional stripes by successively processing the two stripe directions. Corrupted input (a), intermediate result after removing oblique stripes (b) and final result after vertical stripe removal (c). For both removals $\mu_1 = 7/30$, $\mu_2 = 1/300$ was used.

3.2.1.1.3 Parameter Selection

We want to give advice on the selection and tuning of the parameters. Figure 3.3 displays the results of a synthetic test image for different parameter pairs of μ_1 and μ_2. The influence of μ_1 and μ_2 become most apparent when considering the extreme cases, e.g., the columns $\mu_2 = 1/600$ and $\mu_2 = 2/75$ showcase that μ_1 controls the general strength of reduction. The larger the value of μ_1, the stronger the reduction of stripes but also the greater the risk of affecting and removing image structures. Conversely, the rows $\mu_1 = 1/10$ and $\mu_1 = 1/2$ show that increasing μ_2 results in a reduced influence on image structures but at the risk of ignoring stripes. Scaling the parameter pair by a comparable factor larger than one we can address less ideal stripes which for example deviate in direction or have larger widths while maintaining the impact on non-stripe structures. The ability to increase the parameters is limited by the impact on image structures. In combination, the parameters allow to adjust the strength of stripe removal, reduce smoothing artifacts and prevent modification of image structures for different imaging methods and samples. Depending on the severity of stripe corruptions and their appearance, different choices of μ_1 and μ_2 are advised and optimal. In general, $\mu_1 = 1/3$ and $\mu_2 = 1/300$ is a good starting point and the intervals $\mu_1 \in [0.1, 0.5]$ and $\mu_2 \in [0.0016, 0.017]$ were never exceeded. If stripe artifacts are less severe, $\mu_1 = 1/6$ or $\mu_1 = 7/30$ with $\mu_2 = 1/300$ is sufficient. Persistent stripes which are for example short and similar to structures can be addressed using a stronger removal via $\mu_1 = 1/2, \mu_2 = 1/60$ or $\mu_1 = 2/5, \mu_2 = 1/150$. We note that the values of μ_1 and μ_2 are unusual fractions due to rescaling that was applied in the original implementation.

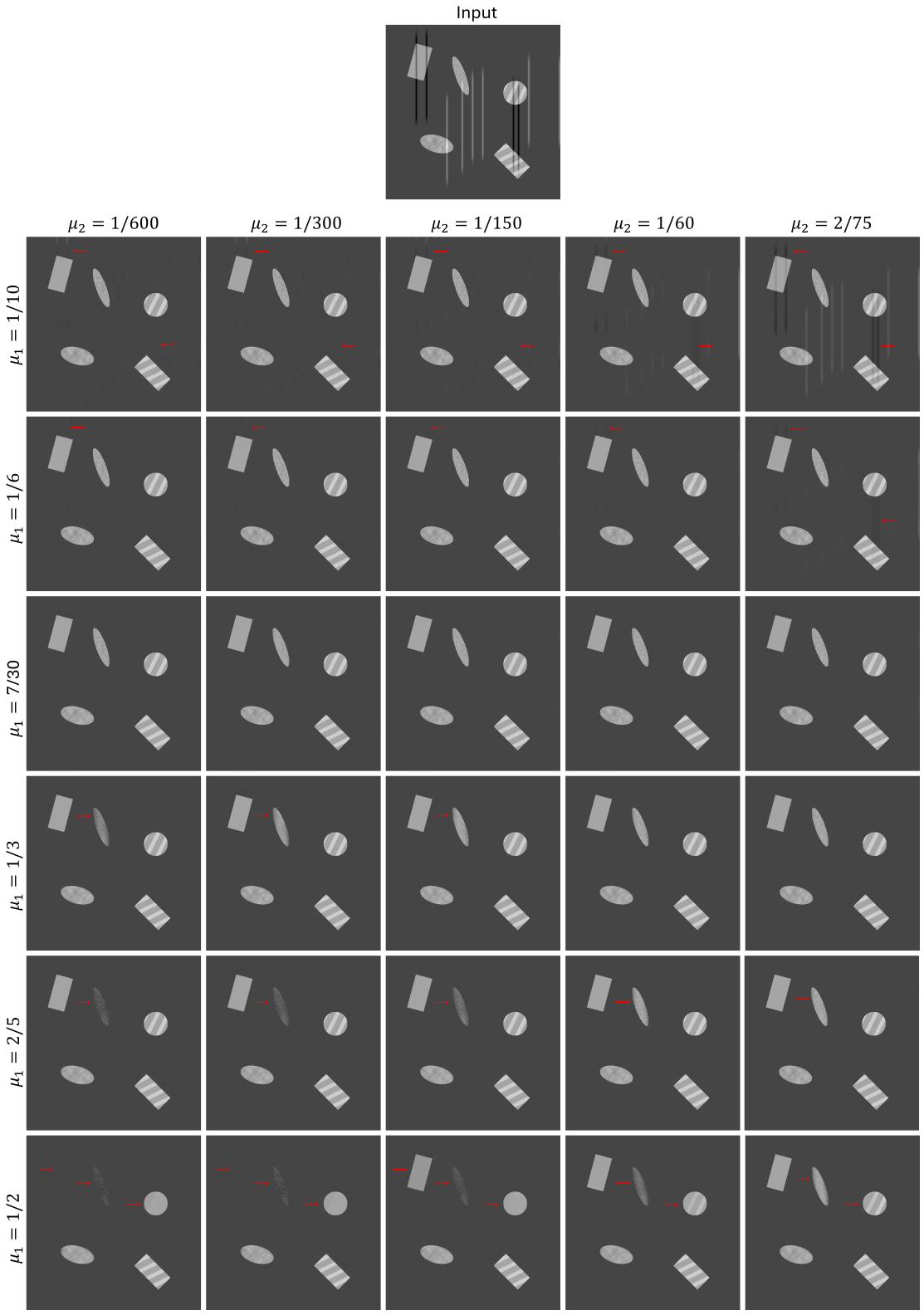
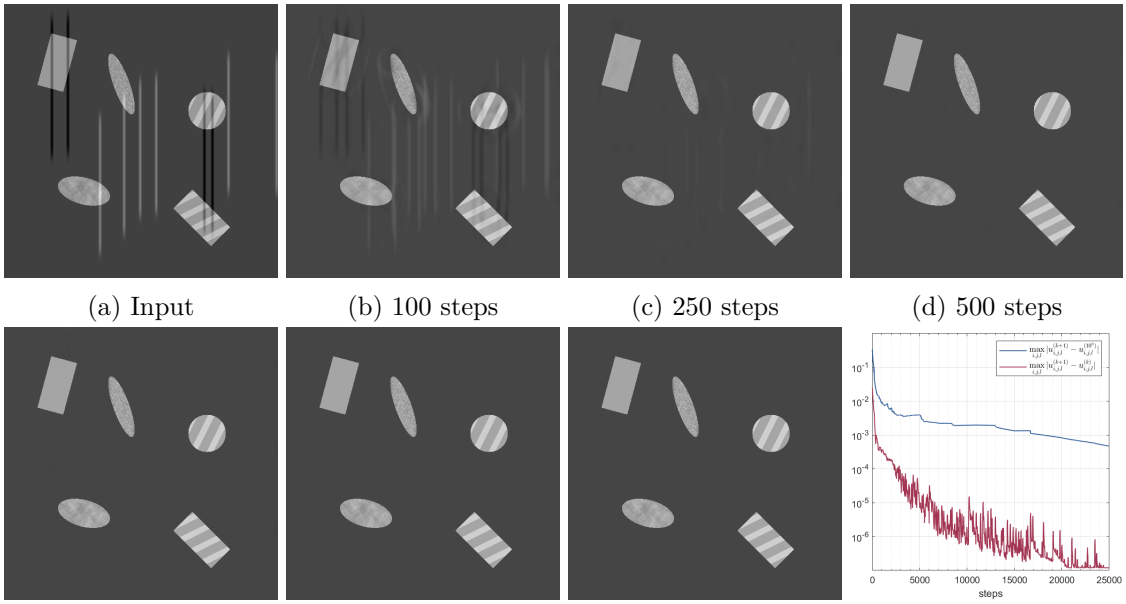


Figure 3.3: Results of GSR for different combination of parameters. The red arrows indicate image artifacts.

In contrast to the adjustable parameters μ_1 and μ_2, the number of iteration steps can be fixed beforehand and does not require tuning dependent on the image structures or severity of stripe corruptions. Based on the results shown in Figure 3.4 we recommend 1000 steps since the outcome is already visually indistinguishable from subsequent results, compare (e-g). The maximum absolute pixel-wise difference when compared to the result after 10^6 steps is below 0.01, i.e., we are reasonably well converged already. We note that after approximately 18000 steps this difference stays below 0.001 if higher precision is desired.



(a) Input (b) 100 steps (c) 250 steps (d) 500 steps

(e) 1000 steps (f) 5000 steps (g) 10000 steps (h) Convergence plot

Figure 3.4: Intermediate results with $\mu_1 = 7/30$ and $\mu_2 = 1/300$ for corrupted image (a) at different iteration steps (b-g) and corresponding quantification of the convergence behavior (h).

3.2.1.1.4 Objective Scaling

It is not uncommon for optimization algorithms to be influenced by the scale of values. The objective function (3.1) operates under the assumption that the corrupted input and cleaned output share the same range of values between 0 and 1. This is not restrictive at all since any image can be transformed accordingly using its minimal and maximum intensity. We consider the rescaled problem

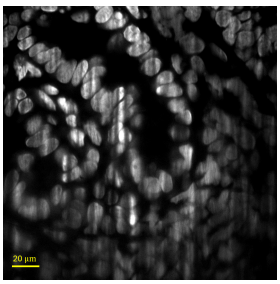
$$\underset{u+s=u_0}{\operatorname{argmin}} \lambda \left(\mu_1 \left\| \nabla^* u \right\|_{2,1} + \iota_{[0,1]^n}(u) + \left\| \nabla_y s \right\|_1 + \mu_2 \left\| s \right\|_1 \right), \quad \mu_1, \mu_2, \lambda > 0 \tag{3.5}$$

which is equivalent to (3.1), i.e., it shares the same properties and optimizers. Application of Algorithm 1 results in a modified version of Algorithm 2 where only the updating steps

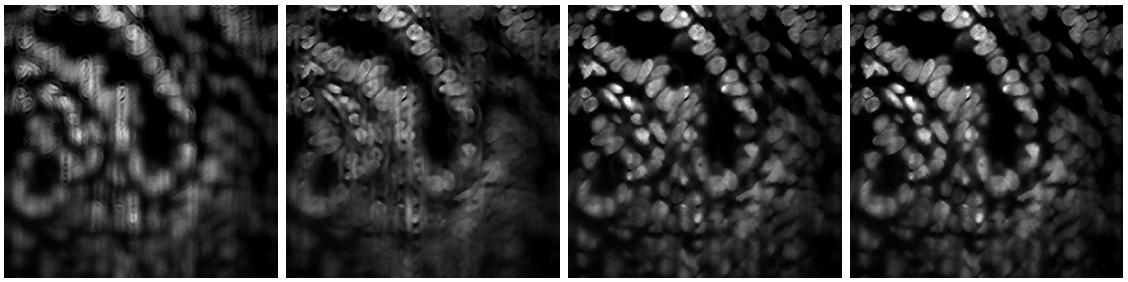
$$\begin{aligned} y_1^{(k+1)} &= C_{\frac{\lambda\mu_1}{\sigma}} \left(b_1^{(k)} + (\nabla_x, \rho_z \nabla_z)\, u^{(k+1)} \right), \\ y_2^{(k+1)} &= S_{\frac{\lambda}{\sigma}\|\cdot\|_1} \left(b_2^{(k)} + \nabla_y s^{(k+1)} \right), \end{aligned}$$

$$y_3^{(k+1)} = S_{\frac{\lambda\mu_2}{\sigma}\|\cdot\|_1}\left(b_3^{(k)} + s^{(k+1)}\right)$$

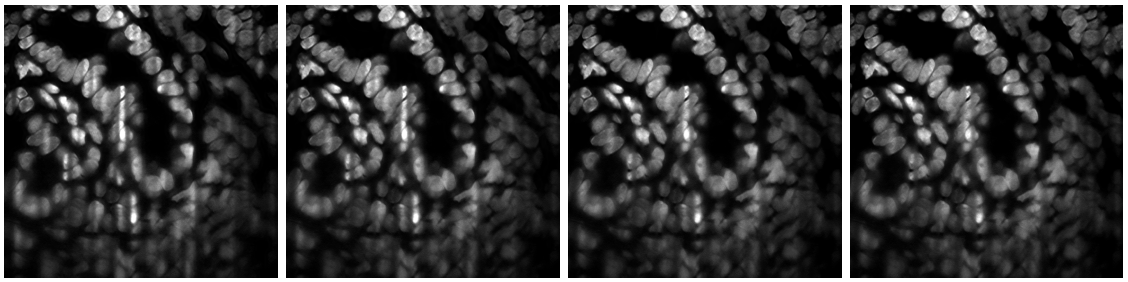
differ, i.e., the threshold of the shrinkage functions are scaled by λ. The results displayed in Figure 3.5 indicate that by adapting λ the convergence behavior changes significantly. We notice that $\lambda = 1$ results in an initial blurring which is slowly resolved with further steps. In contrast, $\lambda = 0.1$ yields no such artifacting and slowly removes stripe artifacts without additional effects. We argue that this difference in behavior is caused by the adapted soft-shrinkage thresholds that define the range of values which are annihilated and shrunk in the updating steps. Therefore, smaller λ result in smaller changes and vice versa. The described influence of λ goes beyond the shown example and can be observed in general. However, the visual improvement does not necessarily translate towards a faster convergence. In fact, there exists no universally optimal value for λ due to unpredictable dependencies on the choice of μ_1, μ_2, and the image itself. Nonetheless, scaling by $\lambda = 0.1$ or comparable values enables the implementation of early stopping, e.g., for real-time processing, since any intermediate result presents a partially destriped image. Since our work focuses on the comparison of fully destriped images, we consider $\lambda = 1$.



(a) Input



(b) 50 steps (c) 100 steps (d) 250 steps (e) 500 steps



(f) 50 steps (g) 100 steps (h) 250 steps (i) 500 steps

Figure 3.5: Intermediate results with $\mu_1 = 2/5$ and $\mu_2 = 1/150$ for corrupted image (a) at different iteration steps for $\lambda = 1$ (unscaled) (b-e) and $\lambda = 0.1$ (f-i).

3.2.1.2 Variational Stationary Noise Remover (VSNR)

For comparison we consider the variational stationary noise remover (VSNR) proposed by Fehrenbach et al. [43] due to its performance, notoriety and availability. The core principle of the method is pattern matching with elementary noise patterns $\psi_1, \dots, \psi_m \in \mathbb{R}^n$ with $m \in \mathbb{N}$ and corresponding location maps $\lambda_1, \dots, \lambda_m \in \mathbb{R}^n$. For stripe removal, suitable patterns can for example be line segments or anisotropic Gaussians. Then, the stripe image is decomposed into $s = \sum_{j=1}^{m} \lambda_j * \psi_j$ where $*$ marks the discrete convolution. The authors propose a general framework designed for removing arbitrary noise patterns, i.e., not necessarily stripes. For destriping, the following optimization problem was derived

$$\underset{u+s=u_0}{\operatorname{argmin}} \underbrace{\|\nabla u\|_{1,\epsilon}}_{D(u)} + \underbrace{\sum_{j=1}^{m} \alpha_j \|\lambda_j\|_1}_{N(s)} \quad \text{s.t.} \quad s = \sum_{j=1}^{m} \lambda_j * \psi_j \tag{3.6}$$

where

$$\|\nabla u\|_{1,\epsilon} = \sum_{i,j,k} H_\epsilon \left(\sqrt{(\nabla_x u)_{i,j,k}^2 + (\nabla_y u)_{i,j,k}^2 + (\nabla_z u)_{i,j,k}^2} \right) \tag{3.7}$$

utilizes the Huber function H_ϵ introduced in Example (2.10). (3.7) can be interpreted as a differentiable approximation of the total variation (3.2) for small ϵ. Problem (3.6) is optimized by a primal dual algorithm which can be found in [15, Algorithm 2]. The motivation and interpretation of the individual terms is the following:

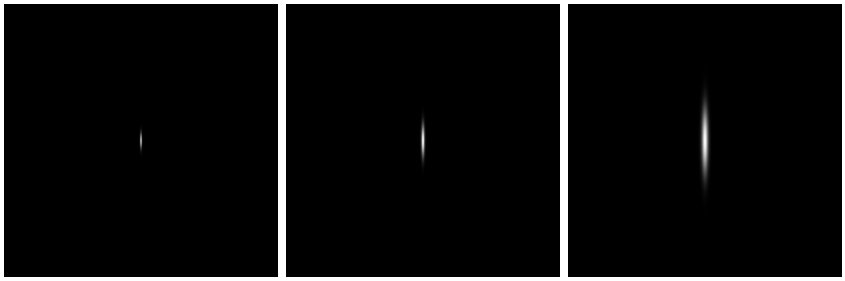
(i) $\|\nabla u\|_{1,\epsilon}$ encodes that the clean image contains only few strong edges, i.e., it is piecewise smooth.

(ii) $\sum_{j=1}^{m} \alpha_j \|\lambda_j\|_1$ emphasizes sparsity in the location maps. It reflects that typically only few stripes affect the image directly.

The parameters α_j for $j = 1, \dots, m$ are free parameters that control the strength of stripe removal and regularization. They enable us to adapt the method to different conditions including severe or minor corruptions. In general, smaller α_j result in a stronger removal of the pattern ψ_j but may affect similarly shaped image structures. In this chapter, we consider the VSNR in its 2D version, i.e., without the variation in z-direction in (3.7), and a selection of $m = 3$ stripe patterns. The patterns are derived from Gabor filters

$$g(x, y) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2} \right) \cos\left(\frac{2\pi x}{\lambda} \right)$$

with parameters $\gamma = 0.1, \lambda = 0.1$ and $\sigma \in \{0.5, 1, 2\}$, see Figure 3.6, to obtain differently sized stripe patterns. This makes the VSNR applicable to a variety of stripe artifacts since patterns can be combined in a superposition to remove corruptions which do not directly relate to an individual pattern.

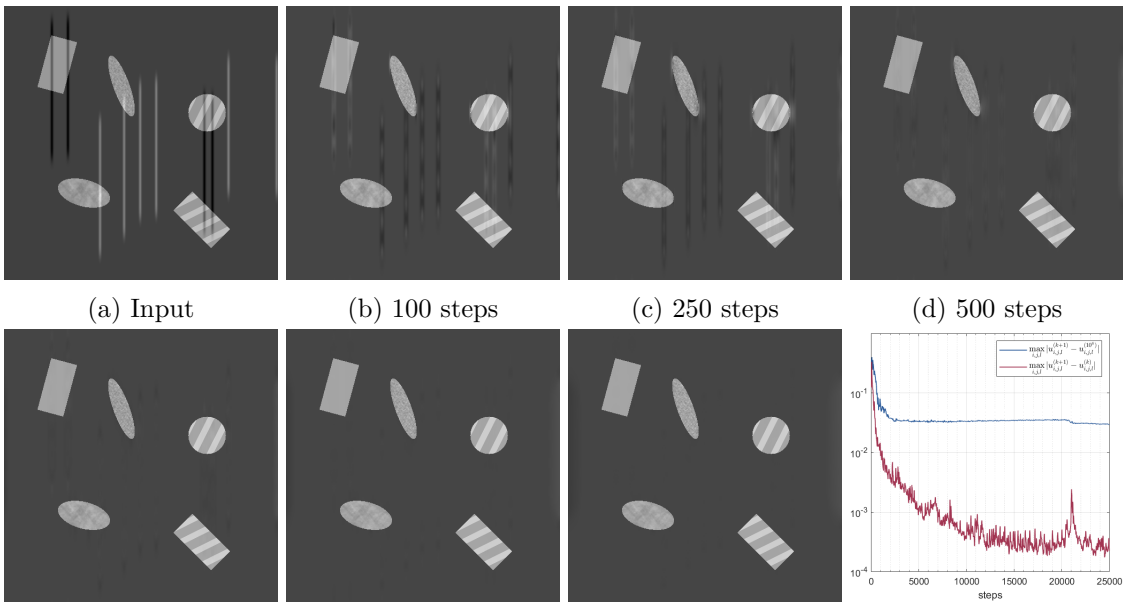


(a) $\sigma = 0.5$ (b) $\sigma = 1$ (c) $\sigma = 2$

Figure 3.6: Selected stripe patterns derived from Gabor filters.

3.2.1.2.1 Parameter Selection

In comparison to the GSR, no general recommendation for choosing the parameters can be given. The parameters α_1, α_2 and α_3 control the penalty for removing the respective patterns from the image with smaller values corresponding to stronger removal. In our experiments the values $\alpha_j \in [3, 13], j = 1, 2, 3$ were never exceeded since smaller values produced artifacts and larger values yielded negligible removal. However, beyond restricting their range it is not intuitive to choose good initial parameters and even more difficult to determine how to tune them. This is caused by the complex interactions of patterns which are combined to remove stripes of different forms and sizes.



(a) Input (b) 100 steps (c) 250 steps (d) 500 steps

(e) 1000 steps (f) 5000 steps (g) 10000 steps (h) Convergence plot

Figure 3.7: Intermediate results with $\alpha_1 = \alpha_2 = \alpha_3 = 3$ for corrupted image (a) at different iteration steps (b-g) and corresponding quantification of the convergence behavior (h).

In contrast to the pattern-related parameters α_1, α_2 and α_3, the number of iteration steps can be fixed beforehand and does not require tuning dependent on the image structures or severity of stripe corruptions. Following the results shown in Figure 3.7 we recommend at

least 3000-5000 steps for a sufficiently converged solution but we can still observe minor visual improvements between 5000 and 10000 steps.

3.2.2 Fourier Filtering

The Fourier transform is a mathematical decomposition of an image into constituent frequencies. If the fundamental assumptions on stripes being parallel and thin are fulfilled, stripe information is primarily encoded in frequency coefficients in a small band orthogonal to the stripe direction. In most cases, natural structures are dominated by low frequencies and concentrate near central coefficients. This decisive difference is exploited by Fourier filtering approaches using masked filters [74, 119, 137], a decision-based algorithm [19] or a prior structural decomposition [72, 89] to restrict tampering to stripe-related coefficients. The latter approach reduces the influence on image structures the most. While wavelet-Fourier filtering [89] is commonly referenced, the multi-directional stripe remover (MDSR) by Liang et al. [72] uses the non-subsampled contourlet transform (NSCT) [30] which improves the separation of stripes and structures and results in less interference during removal. We have further enhanced the original method through slight modification in the filtering step which we abbreviate by MDSR$^+$ and explain in the following.

3.2.2.1 Multi-directional Stripe Remover (MDSR$^+$)

The MDSR [72] is a Fourier filtering-based stripe removal method. It closely follows the ideas proposed by Münch et al. [89] of combining a structural decomposition with selected filtering of stripe-related subimages in the Fourier domain, see Figure 3.8.

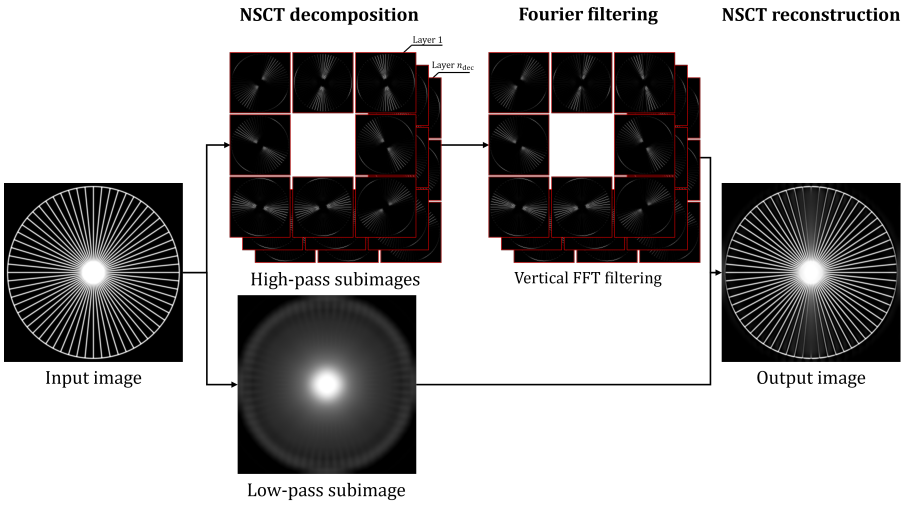


Figure 3.8: Illustration of the MDSR workflow.

The initial image decomposition of the MDSR is the non-subsampled contourlet transformation (NSCT) [31] which is composed of two parts. First, the image u_0 is decomposed by scale into a series of images using high-pass and low-pass filters from a filter pyramid with decomposition depth n_{dec}. Afterward, the high-pass images of each layer are divided into n_{dir} different directional wedges using a directional filter bank. The flowchart visualizing the filtering procedure is shown in Figure 3.9. We refer to [31, 142] for more details.

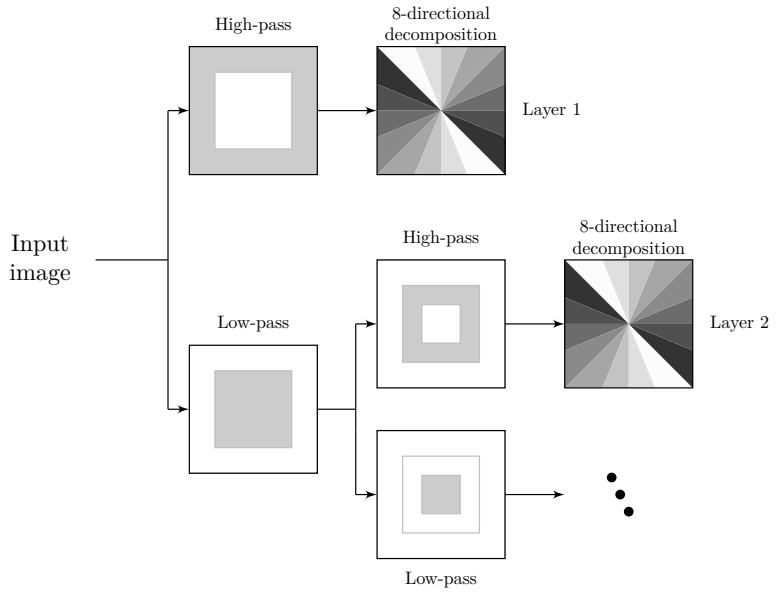
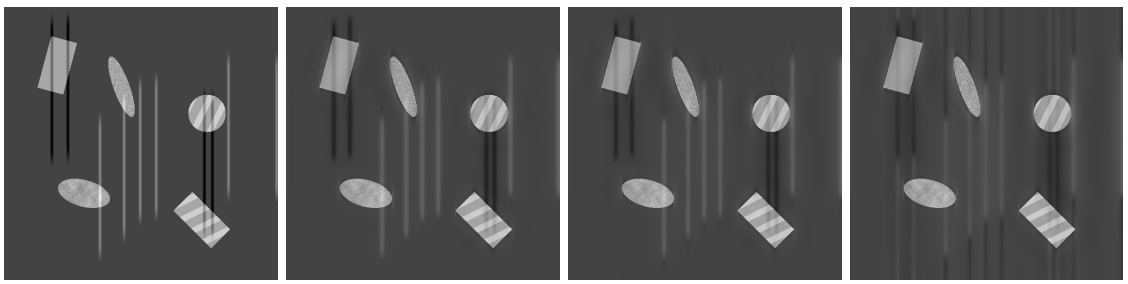


Figure 3.9: NSCT flowchart for the example with an 8-directional decomposition showing the idealized filter responses when using the z-transform [92].

Following the decomposition, subimages $u_0^{(j)}$ attributed to directions close to the stripe direction are further processed with a typical Fourier filtering procedure:

$$\hat{u}_0^{(j)} = \mathscr{F}^{-1}\left[\left(1 - \exp\left(-\frac{y^2}{2\sigma_j^2}\right)\right) \odot \mathscr{F}\left(u_0^{(j)}\right)\right], \quad \sigma_j = \sigma \cdot \exp -\frac{\theta_j^2}{2\sigma_a^2}$$

where $\odot$ is the element-wise multiplication, $\mathscr{F}$ the discrete Fourier transform, y the vertical distance to the image center, θ_j the angle of the subimage to the stripe direction, σ the Gaussian damping parameter and σ_a a decay factor which reduces the strength of filtering for subimages the further attributed directions deviate from the stripe direction. Following the results shown in 3.10, we further restrict filtering only to subimages of directions with an angle $\theta \in [-\pi/4, \pi/4]$ to the stripe direction which significantly reduces artifacting. This method is referred to as MDSR$^+$ to differentiate it from the original proposition [72] which filtered all subimages.



(a) Input (b) $|\theta_i| \leq \pi/8$ (c) $|\theta_i| \leq \pi/4$ (d) $|\theta_i| \leq \pi/2$

Figure 3.10: MDSR$^+$ results when filtering restricted directions using $\sigma = 10$ and $\sigma_a = 0.3$. (d) is the result of the original MDSR.

The MDSR$^+$ depends on a total of four parameters: $n_{\text{dir}}, n_{\text{dec}}, \sigma$ and σ_a. However, $n_{\text{dir}} = 8$ yields a sufficient directional decomposition and n_{dec} must only be large enough such that the scale of stripes is captured by the NSCT. In our experiments, $n_{\text{dec}} \in \{4, 5, 6\}$ sufficed depending on the image size. Both parameters require no additional adjustment. The main parameter of this method is σ controlling the strength of damping Fourier coefficients. Larger values correspond to a stronger stripe removal. The parameter σ_a determines the influence on directions which deviate from the stripe direction. Larger values enable the reduction of artifacts which are shorter or slightly oblique.

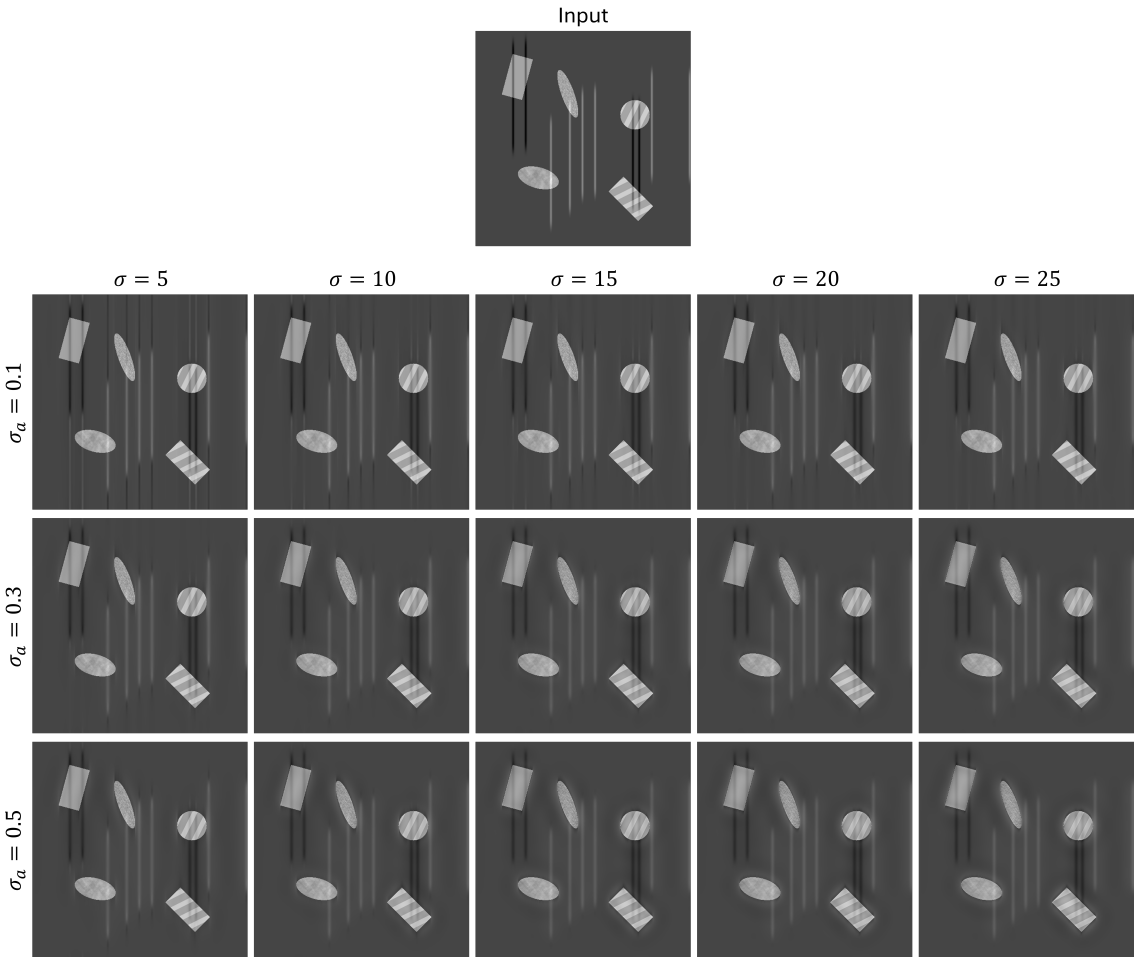


Figure 3.11: Results of MDSR$^+$ for different combination of parameters.

3.2.2.1.1 Parameter Selection

In the following, we provide guidance on selecting and tuning parameters using the results shown in Figure 3.11. In general, we find that $\sigma \in [10, 15]$ and $\sigma_a = 0.3$ effectively reduces stripes with minimal artifacting. Increasing σ beyond this can improve stripe removal but oftentimes introduces significant smoothing artifacts in image structures. Additionally, we recommend $\sigma_a = 0.3$ as smaller values result in the extension of stripes along their direction and larger values introduce glowing artifacts around the structures' edges. However, we note that $\sigma_a = 0.5$ can be beneficial in cases where stripes are not ideally vertical.

3.2.3 Simulation of LSFM Images

Creating synthetic LSFM images requires the simulation of light propagation through media to accurately capture effects such as scattering and absorption. We achieve this using the Python package *biobeam* [135] and by modeling geometric structures via a refractive index distribution (rid) and fluorescence distribution (fld). The rid contains complex-valued refractive indices. The real part corresponds to the classical refractive index which describes the scattering behavior. The imaginary part is the absorption coefficient and yields effects such as light attenuation. We refer to [73] for typical refractive indices for different parts of a biological sample. Ground truth images were generated by setting the rid to the constant value $n_0 = 1.33$ which is approximately the refractive index of water, the main constituent of biological cells. The fld $f \in \mathbb{R}^n_{\geq 0}$ contains non-negative factors which represent the proportion of re-emitted light by fluorescence when subjected to illumination. This indirectly models the real process of staining a sample with a fluorescent marker which is a specialized protein that emits light when illuminated at specific wavelengths. Besides geometrical distributions, we select the wavelength of illumination light λ, the numerical apertures of illumination and detection optics $\mathrm{NA}_{\mathrm{ill}}$ and $\mathrm{NA}_{\mathrm{det}}$ and the voxel size. Furthermore, we set the illumination direction along the y-axis. Afterwards, an image can be simulated by the following steps:

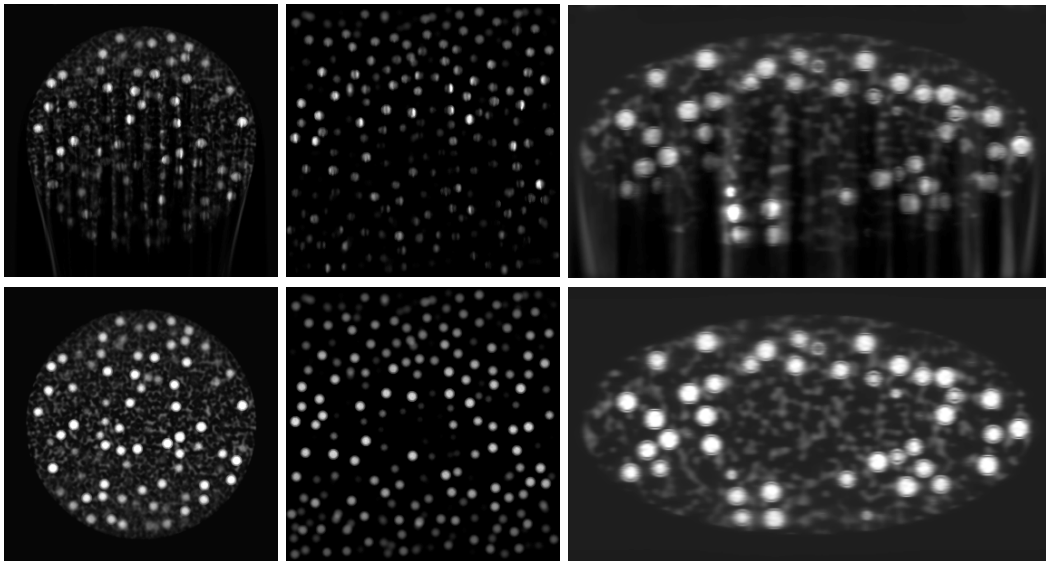
(i) **Illumination:** A complex valued x-z cross-section of a cylindrical light-sheet is generated from $\mathrm{NA}_{\mathrm{ill}}$ and λ. With wave-optical propagation based on the rid as described in [105], the cross-section is propagated through the sample. This yields the illuminated volume I_{ill}.

(ii) **Fluorescence:** The amount of light re-emitted through fluorescence is obtained via $I_{\mathrm{fl}} = I_{\mathrm{ill}} \odot f$ with element-wise multiplication $\odot$. It represents the "fluorescence response" to the illumination.

(iii) **Acquisition:** The 2D image of the section in focus is acquired by convolution with the detection point-spread function h_{det} determined by $\mathrm{NA}_{\mathrm{det}}$ and slice extraction, i.e.,

$$I_{\mathrm{det}} = (I_{\mathrm{fl}} * h_{\mathrm{det}})_{z=\bar{z}}.$$

(iv) **3D Imaging:** A 3D image stack is acquired by repeating steps (i)-(iii) for different focal planes.



(a) Cell cluster (b) Dispersed cells (c) Embryo

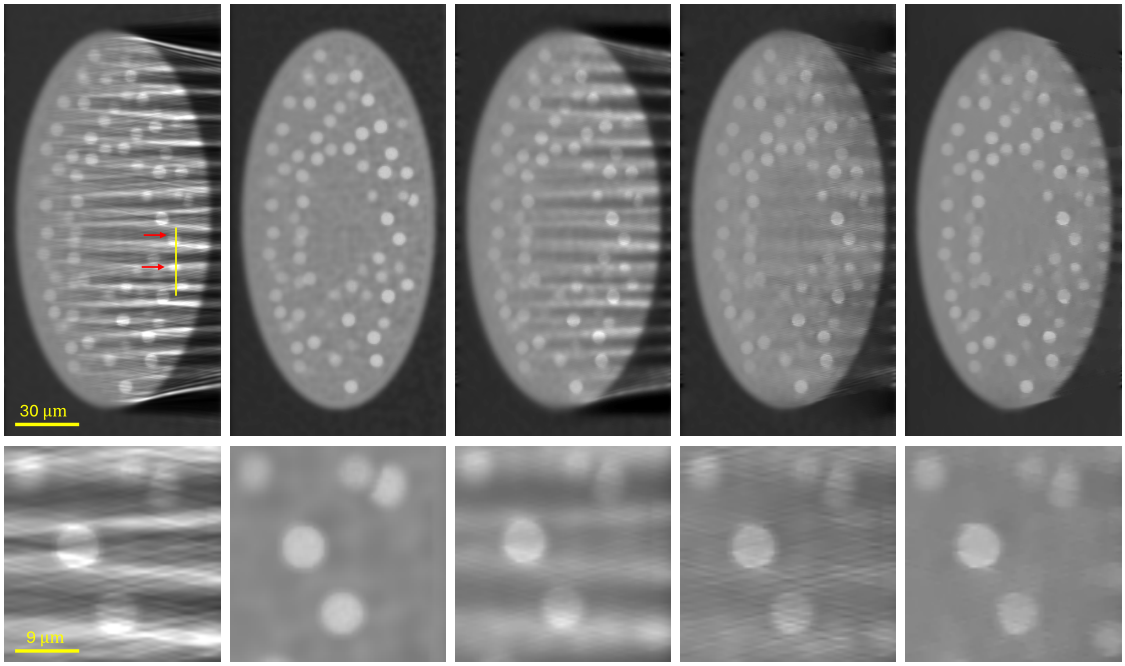
Figure 3.12: Simulated images (top) and ground truths (bottom) for different simulation examples.

For model generation in this thesis, we sample non-overlapping spheres as foreground and combine the structures with 3D Perlin noise. Refractive indices were chosen based on reference values given in [73]. The values for the fld were chosen to produce visually appealing images that look similar to real LSFM images.

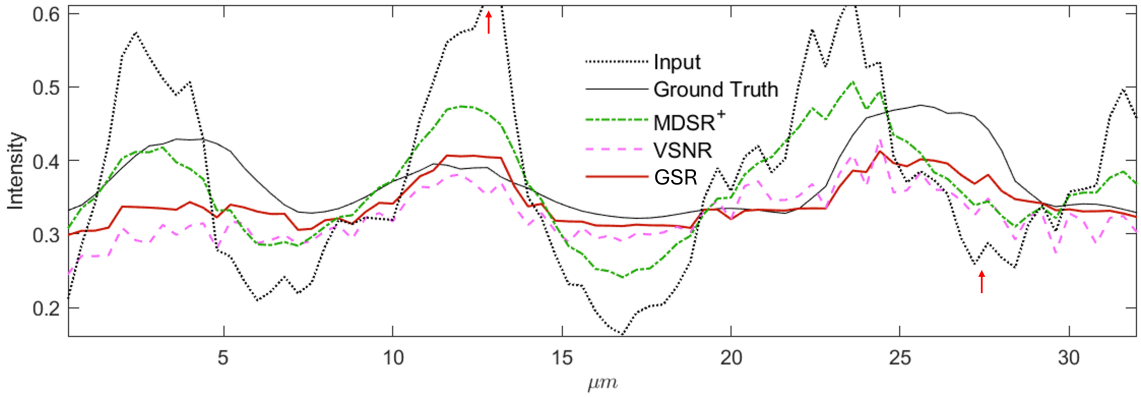
3.3 Results

In this section, we evaluate, compare and analyze GSR against previously published solutions in the categories of Fourier filtering and variational methods. In particular, we consider the MDSR$^+$, see Section 3.2.2, and the VSNR [43], see Section 3.2.1.2. Both methods are well-known, validated for different imaging methods, and freely available. For the MDSR$^+$ we used the NSCT with 8 directions. The maximum possible depth, which depends on the image size, was set between 4 and 6. The VSNR was used as discussed in Section 3.2.1.2 with three differently sized stripe patterns created from Gabor filters to ensure general applicability. The GSR was used with $\rho_z = 0$ as MDSR$^+$ and VSNR were only available for 2D processing. The number of optimization steps for GSR and VSNR was set to 25000 to obtain a sufficiently accurate approximation of the optimum. The remaining parameters of the three stated methods were optimized via grid search with numerical and visual assessment. For the numerical evaluation of synthetic data with ground truth we utilize common performance metrics: peak signal-to-noise ratio (PSNR) and the multi-scale structural similarity index measure (MS-SSIM) [132]. Furthermore, we compare the line profiles of the result and ground truth where a line is placed along a selected segment in both images and the Euclidean norm of their difference is calculated. Additionally, we use the curtaining metric proposed by Roldán [107]. In contrast to the other metrics, it exclusively measures stripe corruptions based on the ideas which motivate Fourier filtering methods, see Section 3.2.2. We report the absolute difference between the curtaining metrics of the processed image and the ground truth

as ∂Curtaining. This choice accounts for naturally occurring stripe-like patterns in the image structures and treats over- and under-performance alike.



(a) Input (b) Ground truth (c) MDSR$^+$ $\sigma = 25$ $\sigma_a = 0.3$ (d) VSNR $\alpha = [3, 5, 10]$ (e) GSR $\mu_1 = 1/3$ $\mu_2 = 1/300$ $\rho_z = 0$



(f) Line profiles

Figure 3.13: Destriping results of simulated LSFM data of a synthetic Drosophila-like embryo model.

3.3.1 Synthetic Data

Figure 3.13 shows results on a synthetic Drosophila-like embryo model inspired by [135] and obtained by physically accurate modeling of light propagation in LSFM imaging, see Section 3.2.3 and [105]. Significant differences between the methods can be spotted. While MDSR$^+$ reduces stripes significantly, larger artifacts and thin oblique stripes remain visible in the outcome. The latter become only visible at closer inspection. In comparison,

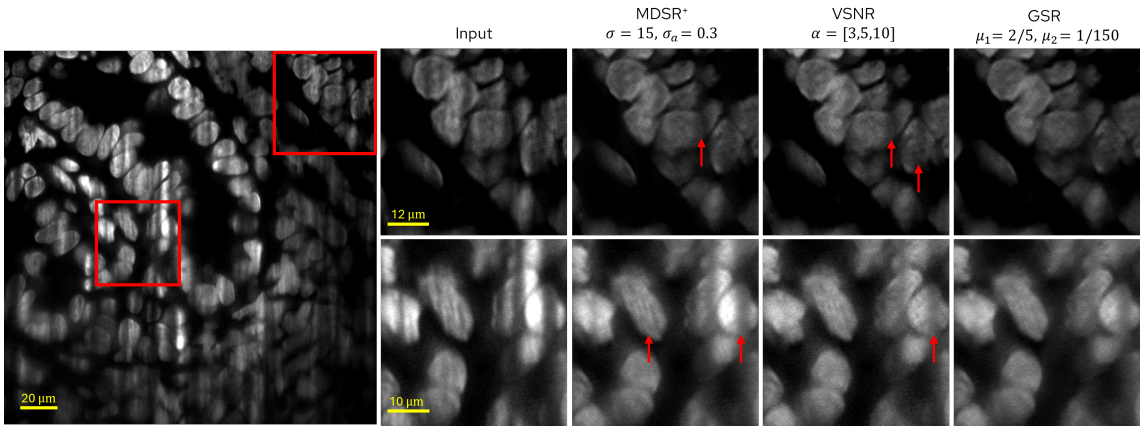
VSNR can reduce wider artifacts more effectively leaving only weak remnants. However, fine oblique artifacts are entirely unaffected by the removal. Only GSR removes all stripes, including the trails, with only faint artifacts remaining, particularly at the right edge of the body. The line plot supports our findings. MDSR$^+$ shrinks the profile but is unable to adjust wider artifacts correctly. On the other hand, the line profile of VSNR has fewer extremes but contains small perturbations reflecting the struggle with fine artifacts. GSR is the only one that achieves both a significant reduction in stripes and a flat profile. The numerical values in Table 3.1 further confirm prior observations with GSR outperforming the others across all metrics.

Metric	**Input**	**Ground Truth**	**MDSR$^+$**	**VSNR**	**GSR**
PSNR↑	24.25	∞	$29.02^{+20\%}$	$31.06^{+28\%}$	$35.65^{+47\%}$
MS-SSIM↑	0.901	1	$0.940^{+4\%}$	$0.919^{+2\%}$	$0.969^{+7.5\%}$
∂Curtaining↓	0.184	0	$0.056^{-70\%}$	$0.074^{-60\%}$	$0.036^{-80\%}$
Line difference↓	1.055	0	$0.523^{-50\%}$	$0.589^{-44\%}$	$0.389^{-63\%}$

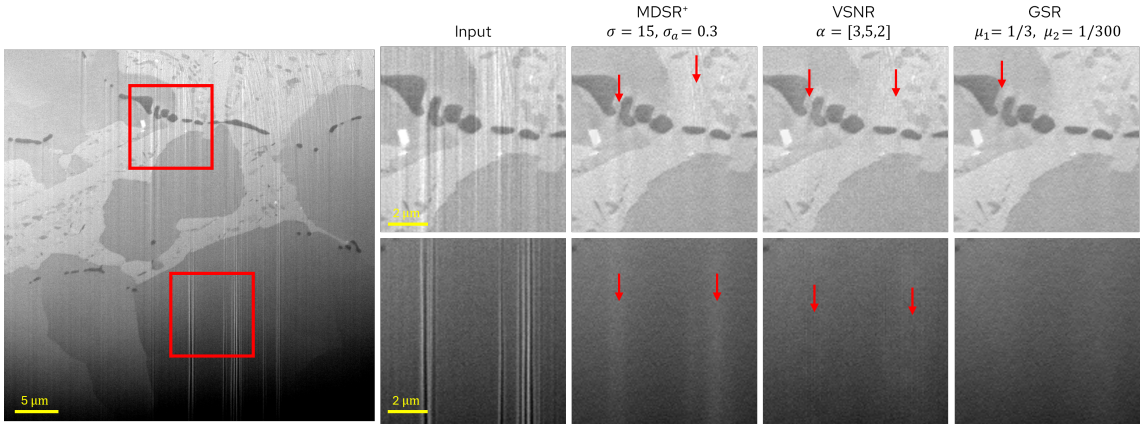
Table 3.1: Quantitative evaluation of results on synthetic data from Figure 3.13. Arrows indicate better performance. Percentage improvements from the input are given in green. As reference, the curtaining metric of the ground truth is 0.90.

3.3.2 Real Data

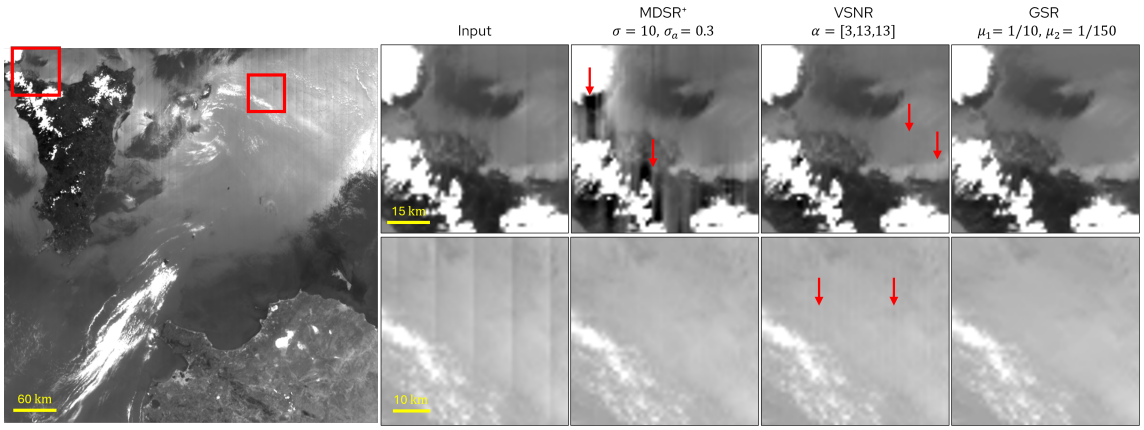
In Figure 3.14 we show results obtained by processing real data from LSFM, FIB-SEM and remote sensing. We observe similar behavior to the synthetic data with VSNR leaving very thin remnants and MDSR$^+$ having insufficient reduction. The LSFM image highlights the capabilities of variational methods in general. While MDSR$^+$ only reduces stripes, both VSNR and GSR show near-perfect on-par performance. However, the FIB-SEM and remote sensing images reveal shortcomings of VSNR which leaves very thin stripes in the result. MDSR$^+$ even introduces severe image artifacts in areas of high contrast in the remote sensing image despite performing well in low contrast regions. On the other hand, GSR delivers visually perfect results for FIB-SEM and remote sensing in the sense that stripes are entirely removed and image structures remain intact.



(a) LSFM



(b) FIB-SEM



(c) Remote sensing

Figure 3.14: Stripe removal results on real data acquired on different systems. The red arrows highlight areas of particular interest.

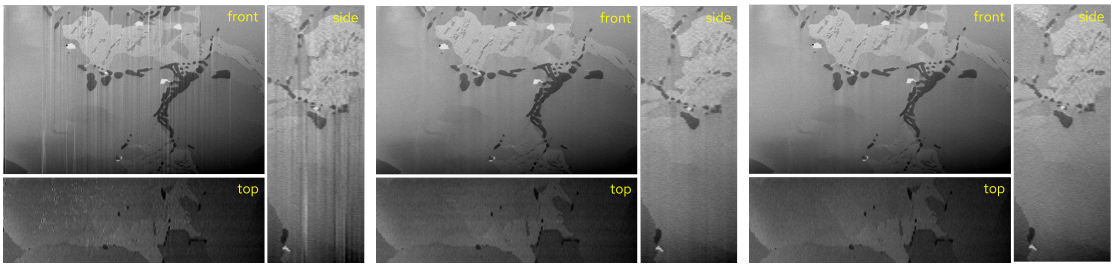
3.3.3 Extended Results for GSR

In the following, we provide a series of results for the GSR to establish its full capabilities and effectiveness in removing stripe artifacts. First, we show the benefits of processing a

corrupted image in 3D ($\rho_z \neq 0$) compared to processing it 2D ($\rho_z = 0$), i.e., slice-by-slice. Afterwards, we display results from a diverse set of LSFM and FIB-SEM images that underscore the universal applicability of GSR. Lastly, we address the limitations of the GSR and discuss mitigation strategies for these scenarios.

3.3.3.1 Advantages of 3D Processing

Figure 3.15 compares the results of removing stripe artifacts in 3D against 2D. The stripes visible in the side view usually correspond to the stripe artifacts that are visible in the front view. However, some are caused by variations in the brightness levels between the slices and are a result of the imaging procedure. These brightness variations are visible in the top view as horizontal stripes and are not affected when processing in 2D. In comparison, the result obtained by processing in 3D shows greater consistency in brightness in the side and top views where both the stripe artifacts and variations in brightness between slices are resolved. The performance in the front view remains indistinguishable. The removal of inconsistent brightness levels is a side product of using the total variation term in the objective function (3.1) since it penalizes brightness changes in all directions, including the slice direction. In particular, variations between individual slices that are not consistent with image structures are removed.

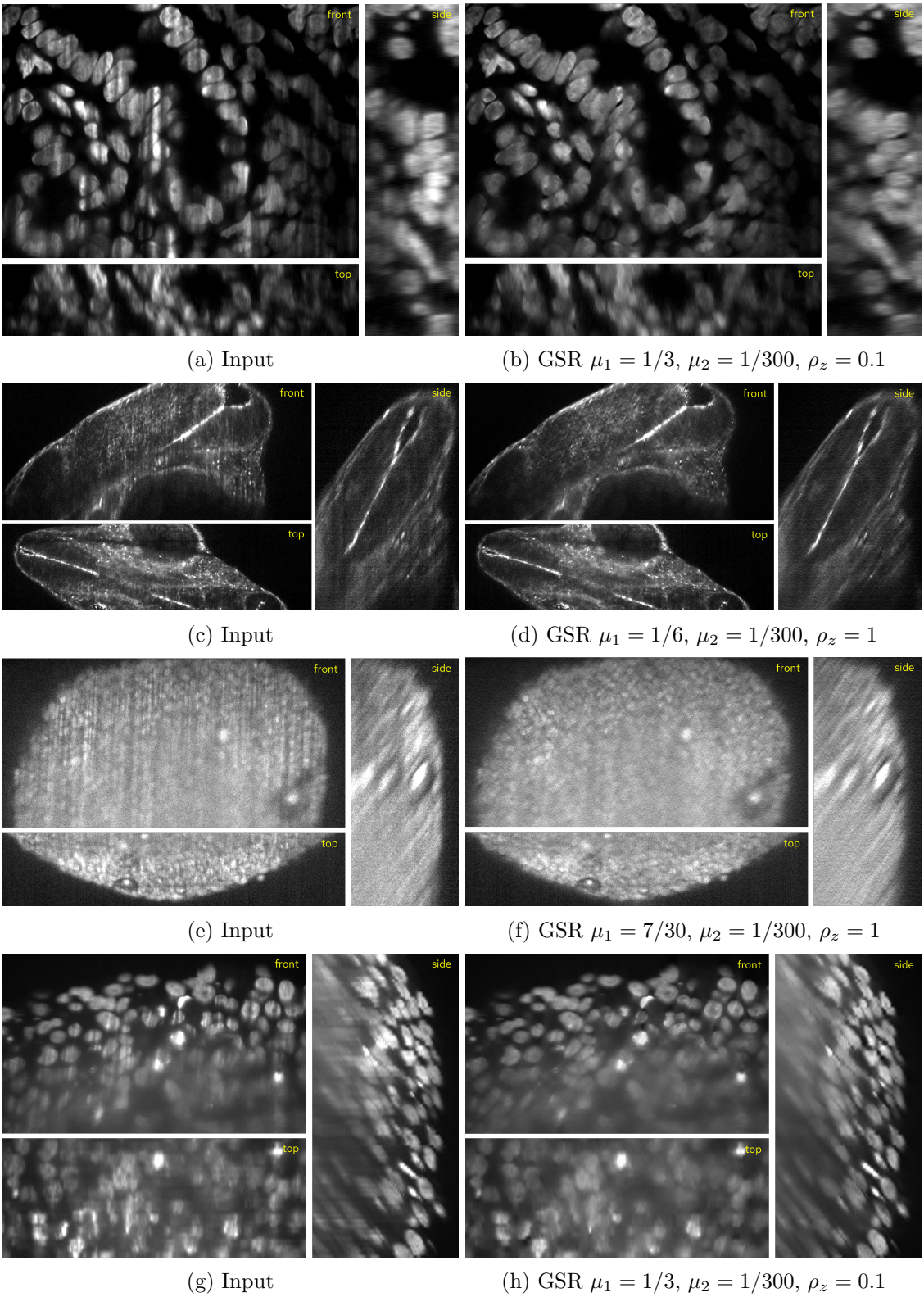


(a) Input (b) GSR, $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 0$ (c) GSR, $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 0.5$

Figure 3.15: Comparison of 2D and 3D destriping using GSR on a FIB-SEM image of a tin bronze cast alloy.

3.3.3.2 General Applicability

The results displayed in the Figures 3.16 and 3.17 reveal that GSR can remove stripe artifacts effectively in all tested scenarios. This includes corruptions that strongly impair the perceived visual quality and stripes that deviate from the ideal assumption of being thin and perfectly vertical. In most instances, the stripe removal is performed with near perfection and only rarely do we observe more minor issues.



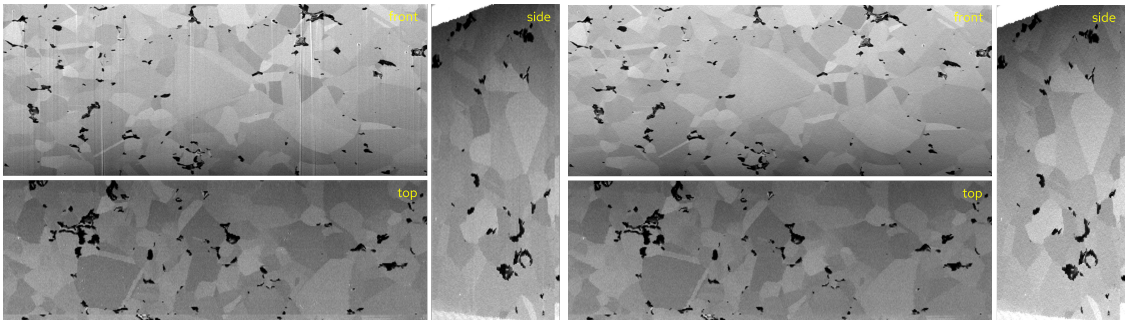
(a) Input (b) GSR $\mu_1 = 1/3$, $\mu_2 = 1/300$, $\rho_z = 0.1$

(c) Input (d) GSR $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 1$

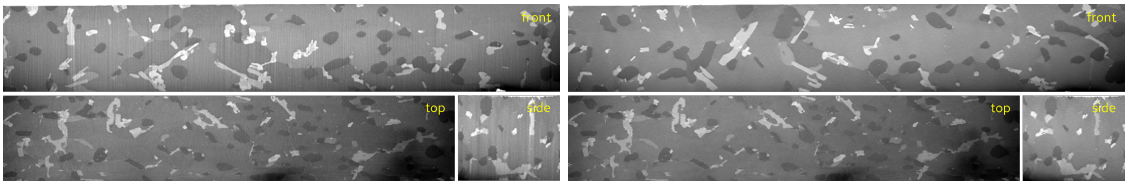
(e) Input (f) GSR $\mu_1 = 7/30$, $\mu_2 = 1/300$, $\rho_z = 1$

(g) Input (h) GSR $\mu_1 = 1/3$, $\mu_2 = 1/300$, $\rho_z = 0.1$

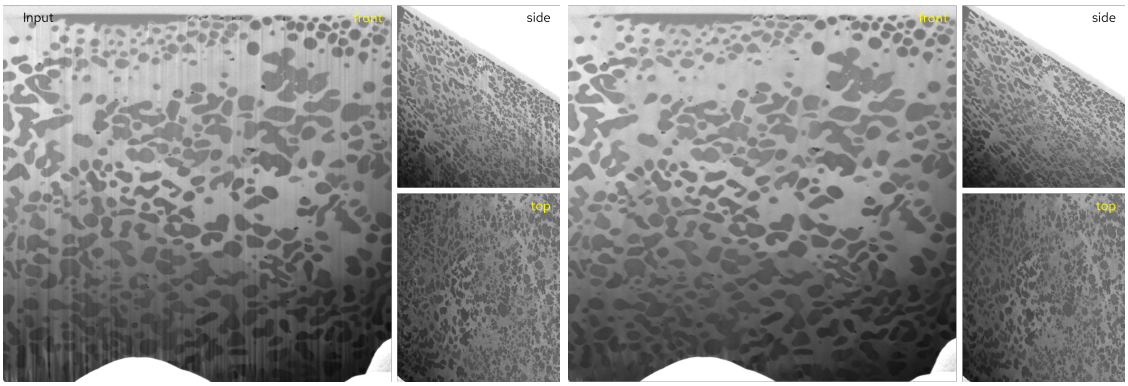
Figure 3.16: Stripe removal results of GSR on LSFM images of cultivated mouse intestine cells (a-b), embryo of a zebra-fish larva (c-d), cluster of HeLa cells (e-h).



(a) Input (b) GSR $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 0.5$



(c) Input (d) GSR $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 0.5$



(e) Input (f) GSR $\mu_1 = 1/6$, $\mu_2 = 1/300$, $\rho_z = 0.25$

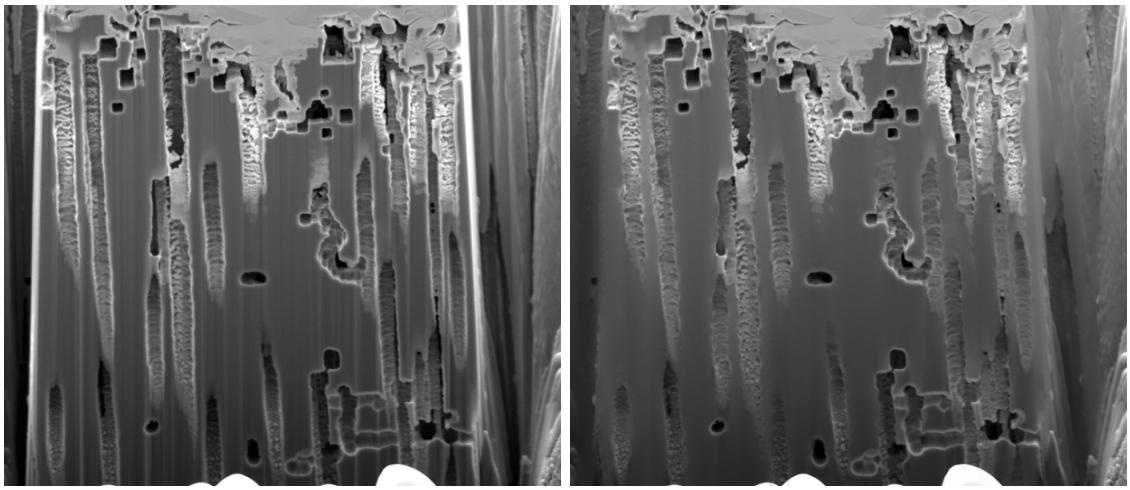
Figure 3.17: Stripe removal results of GSR on FIB-SEM images of carbon nanotube reinforced metal matrix composite (a-b), AlSi13 (c-d), CuCr (e-f).

3.3.4 Limitations

While GSR produces universally great results, there exist some limitations that should not be overlooked including incomplete stripe removal and smoothing of image structures in challenging conditions. In the following, we provide examples for limitations and comment on their potential mitigation.

When the image structures are similar to the corrupting stripes, insufficient stripe removal or artifacts can be observed depending on the choice of parameters. The similarity of structures and stripes can occur in two ways. First, elongated image structures share the same direction as the stripes displayed in Figure 3.18. This typically results in a reduction of these image structures. Second, stripes and structures share similar scales such as the larger cells shown in Figure 3.16 (a-b) and (g-h). Depending on the parameter choice, the outcome displays either insufficient reduction of these stripes or smoothing artifacts beyond the edges of affected structures. This limitation can often be overcome by adapting the parameters, e.g., by reducing the strength of removal when structures are affected.

Another limitation is displayed in Figure 3.16 (a-b) where severe corruptions are contained that suppress image information entirely. In these areas, only the level of brightness is corrected but previously lost information such as texture cannot be reconstructed. This problem can only be resolved by preventing severe artifacts from forming, see [101] for more details on prevention methods.



(a) Input (b) GSR $\mu_1 = 1/10$, $\mu_2 = 1/600$, $\rho_z = 1$

Figure 3.18: FIB-SEM image of porous aluminum containing vertical structures which are affected by stripe removal.

3.3.5 Computational Cost and Complexity

The presented methods differ significantly in their effort for application to a corrupted image. In the following, we compare the methods against each other based on this complexity and explain the process to achieve good or optimal results. We note that all the methods were compared in Matlab relying on [30] for MDSR$^+$ and the implementation of VSNR provided by the author Pierre Weiss [43]. The latter was made faster using a GPU. The implementation of GSR also utilizes the GPU and could be further improved on with parallelization. The processing times presented in Table 3.2 were collected on a laptop (AMD R7 5800H, 16GB RAM, RTX 3050 Mobile) for an image of size 940 x 940 pixels.

Method	**Recommended**	**Time per Trial**	**#Trials**	**Post-Processing**
MDSR$^+$	$\sigma_a = 0.3$	429s	≤ 3	required
VSNR	3000 steps	185s	> 5	required
GSR	1000 steps	2.3s	≤ 3	none

Table 3.2: Computational complexity for applying the methods to an image of size 940 x 940 pixels. The number of trials for finding good parameters is a conservative estimate based on our experience.

For the application of GSR adequate initial parameters need to be selected. This selection and possible adjustments are intuitive and discussed in Section 3.2.1.1.3. While expertise can help in initializing the parameters so that no further adjustments are needed, it is

not required. Good results were achieved after at most three attempts following our recommendations. The result of GSR does not require further post-processing such as the adjustment of brightness by construction. As stated in Section 3.2.1.1.3 we recommend using 1000 optimization steps for a well-converged solution. This yields a processing time of only 2.3s. However, depending on the complexity of image structures fewer steps sufficed.
Using MDSR$^+$ requires the specification of σ and σ_a. $\sigma_a = 0.3$ is generally good and may only be increased to 0.5 if the stripe directions deviate from being vertical. Finding a good σ is similar to the parameter tuning of GSR. With some expertise it can be specified directly but on average, fewer than three different values are necessary for good results. The processing time was 429s. Post-processing the intensity profile is almost certainly required, e.g., using clipping or rescaling.
For VSNR, we consider the pre-processing step of choosing adequate stripe patterns. Finding good patterns is a trial-and-error procedure. In combination with long processing times we deem it unfeasible for general use and have resorted to a set of well-performing patterns that cover different sizes of stripes. Finding good hyperparameters α_i is essential but no standard choice that consistently performed well could be found. The initial choice and subsequent adjustments are difficult due to their unexpected interplay. Typically, more than five different sets of parameters are needed to achieve good results. Based on Section 3.2.1.2.1 we recommend using at least 3000 optimization steps for a well-converged solution. This results in the processing time of 185s. The outcome needs the same post-processing as MDSR$^+$ of adjusting the intensity profile.

3.4 Discussion

The results demonstrate that GSR performs consistently well across different imaging methods in terms of stripe removal quality and retention of image details. Besides quality, throughput and usability are important aspects. We found GSR to perform around 100 times faster than MDSR$^+$ and VSNR as shown in Table 3.2. As explained in Section 3.2.1.1.3, GSR is comparatively easy to use because it relies on only two parameters (VSNR, for example, uses three). Moreover, the parameters μ_1 and μ_2 correspond to the strength of removal and precaution towards retaining details, respectively, which can be intuitively understood. Good results can already be achieved for default values $\mu_1 = 1/30$ and $\mu = 1/300$. Guidance for fine-tuning is provided in Section 3.2.1.1.3.
The consistency in quality of GSR across all shown settings can be attributed to the non-restrictive formulation of its objective function with only rough assumptions on properties of the image and stripes. Therefore, deviations in stripe direction and thickness remain adequately penalized such that a general application for stripe removal is provided. However, limitations exist. In particular, the alignment of image structures with the stripe direction should be avoided as both will be affected by the removal, see Section 3.3.4 for more details.
The insufficient removal by MDSR$^+$ for more sophisticated stripes is due to a considerable overlap in area of stripe and image-related frequency coefficients when structures and stripes live on similar scales or stripe directions deviate from the assumption. Depending on the choice of parameters, either insufficient stripe removal or severe modification of image structures are the result. We deem the prior to be preferential since the latter renders the outcome unusable for further analysis or visualization which is why we only

showed these results. The introduction of artifacts in the remote sensing image is surprising as the setting reflects an ideal scenario for this approach with thin, long and periodic stripes. However, the high contrast in some areas produces unforeseen problems by filling separating spaces with smooth stripe-like artifacts.
As demonstrated by the results of Figure 3.14 (a) VSNR can efficiently remove stripes. However, this performance comes at a higher effort than with GSR. Finding adequate patterns is a trial-and-error procedure. In combination with long processing times we deem it unfeasible for general use. Hyperparameters strongly influence the quality of the result. In our hands, VSNR required more iterations of parameters to yield good results. Contrary to GSR, no standard choice that always performed well could be found for VSNR. We note that VSNR shows limitations in removing thin or oblique stripes. More image-specific patterns may improve the results. However, each pattern introduces an additional hyperparameter, further increasing the complexity.
For our analysis, we supplement real data with synthetic LSFM data. We used a highly accurate simulation method that yields realistic, irregular stripe artifacts. It is important to note that realistic models of the refractive index distribution within the sample are crucial. Therefore, considerable effort and rigor were applied to create a realistic model following [135]. While simulations have been employed for the assessment of LSFM performance such as [105], so far stripe removal algorithms have not been evaluated based on the synthetic data despite the clear benefits: the objective assessment by comparison to the ground truth obtained by "turning off" the cause of the stripe artifacts yielding unperturbed images for comparison. This approach enabled the use of quality metrics such as PSNR and thus a more accurate assessment of quality compared to the previously employed approach of superimposing a clean image with independent stripes [72] or purely relying on visual assessment.
Roldán's curtaining metric [107] complements previously established metrics well for assessing quality on the synthetic data. Since it is specifically designed to measure the amount of stripe corruptions in an image it is more robust against other differences between the image and its ground truth, e.g., different levels of brightness. We did not apply the metric directly. Instead, we used the absolute difference with the ground truth. We have found that a poor choice of parameters that leads to overly aggressive stripe removal often results in a seemingly better curtaining value, although the methods perform poorly across other relevant aspects, such as structure preservation. By adjusting for naturally occurring imperfections in the structures using the ground truth, the metric captures improvements in image quality more accurately and aligns with visual inspection and other metrics.
The results shown in this chapter were processed as 2D images. This does not reduce the significance of our findings but only simplifies visualization and comparison with imaging methods that deliver 2D images such as remote sensing and with algorithms that operate only in 2D. However, we emphasize that GSR can process 3D images as well, as demonstrated by Figure 3.16 and 3.17. Importantly, image stacks are not processed slice-by-slice but entirely in 3D. This improves consistency between slices and prevents errors in image structures that result from ignoring relations along the stack direction as shown in Figure 3.15. However, in some cases, the impact is limited due to inferior axial resolution which impairs the separation of image structures and stripes on this axis.

3.5 Conclusion

In this chapter, we have demonstrated the potential of an improved variational method using a specially adapted objective function as a general solution for removing stripe artifacts. Results for a variety of image structures and different imaging methods together with quantitative assessment demonstrated the method's potential. An important aspect for the adoption of the method, which we made publicly available on GitHub [108], is the correspondence of the parameters to intuitive image properties enabling non-specialist users to readily achieve optimized destriping and preservation of image structures. Furthermore, we substantiate that it is highly effective in removing stripes and features improved speed and performance compared to previously published solutions. The generality of the approach is backed by using images from three different imaging methods: LSFM, FIB-SEM and remote sensing. A quantitative comparison on synthetic LSFM data using several quality metrics helped to better understand the capabilities and limitations of the stripe removal methods. It confirmed that our solution provides the best destriping while retaining image structures which coincides with the visual assessment of real images.

4 Segmentation Quality

4.1 Introduction

Segmenting images into meaningful regions plays a vital role in a wide range of applications including medical imaging, autonomous driving and materials science. Accurate segmentation is often a prerequisite for downstream processing such as quantitative analysis or classification. In the supervised setting, quality assessment relies on comparing the segmentation result with a ground truth annotation. This assessment is essential when developing, improving, and comparing methods on different datasets, e.g., when training machine learning models.

Quality metrics such as the Dice similarity coefficient (DSC) or the Hausdorff distance [123] are frequently used in segmentation tasks and have proven useful across many domains. However, these metrics are not without limitations. For instance, DSC summarizes performance solely using ratios of correctly and incorrectly annotated pixels regardless of their position in the image. Like many other metrics, it is sensitive to class imbalances. In contrast, the Hausdorff distance and its descendants collect spatial information from errors to assess quality. However, they yield unnormalized values because they scale with the distance and number of errors. This limits the interpretability of individual values and complicates comparisons between different datasets and segmentations. In general, all quality metrics have certain limitations since they reduce an entire image to a scalar [100]. Therefore, a precise assessment of performance requires the use of multiple quality metrics at once.

The literature related to the analysis, comparison and application of quality metrics in segmentation tasks is extensive, with several studies and surveys comparing multiple metrics against each other on various datasets. Wang et al. [133] provide an overview of available quality metrics for supervised and unsupervised tasks and discuss the corresponding literature. Reinke et al. [100] offer a detailed list of quality metrics, their limitations and pitfalls that have to be considered in applications. Other articles such as [34, 78, 95, 123] analyze and compare selected quality metrics across various types of image data in detailed studies. They primarily cover traditional metrics that can be expressed through the entries of a contingency table, the basic cardinalities: true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN). Most importantly, they do not consider the positions of pixels and their geometric relations. Chicco and coauthors published multiple articles showing that the Matthews correlation coefficient (MCC) is superior to other traditional metrics for general quality assessment [22–25].

Distance-based metrics often fall short in discussions of quality and contain only a handful of metrics that are derived by weighing incorrectly annotated pixels by their distance to the structures' surfaces. This results in unnormalized values that highly depend on the geometry and are therefore difficult to interpret without additional information. Several attempts were made to improve the robustness, reliability and interpretability of the Hausdorff distance such as averaging distances, using quantiles [123], adjusting the nor-

malization [5] or restricting the comparison only to the boundaries [50]. However, none of the approaches resolved all the issues at once. Nevertheless, distance-based quality assessment provides fundamental insight into the spatial distribution of errors and can benefit quality assessment. In particular, it allows us to determine whether predictions match the general shape of the ground truth and differ only in regions close to or far from its surface. The importance of spatial information for segmentation tasks was also demonstrated by Fend et al. [44] who saw the performance of AI segmentation models significantly increase when using a distance-based loss function.
In this chapter, we introduce and analyze the quality metric surface consistency coefficient (SCC) which evaluates the severity of errors based on their spatial location, distinguishing between errors within a specified range and those that extend beyond it. The metric addresses issues associated with previous distance-based metrics. It is normalized, quantifies performance based on spatial information and is easy to interpret and compare across different datasets. In addition, SCC remains independent of the rate of errors. Instead of trying to assess quality alone, SCC aims to be combined with a traditional metric such as accuracy to evaluate both the quantity of errors and their geometric distribution independently. This reduces the use of redundant information and enables a more precise assessment of quality and comparison of segmentations. To achieve this, SCC summarizes the geometric distribution of errors into a scalar value which differentiates between clustering of errors near or far from the geometric boundary. We rigorously analyze and validate SCC across a diverse set of geometries combined with synthetic segmentation results generated by introducing systematic errors. By comparing with established quality metrics such as DSC, MCC and the directed average Hausdorff distance (dAHD) we demonstrate the utility of SCC for assessing quality and comparing available segmentations. The integration of SCC into real-world segmentation tasks is exemplified by extending a study performed by Barisin et al. [7] who compared the performance of various segmentation methods on a semi-synthetic dataset of cracks in concrete using traditional metrics. The inclusion of SCC reveals previously undetected information about the performance of the individual segmentation methods.
We remark that parts of this chapter have already been published by the author of this thesis in [109].

4.2 Methods

We consider a binary segmentation or annotation of an image with pixels $X \subset \mathbb{Z}^d$ by the partition $S = \{S^0, S^1\}$, i.e., $S^0 \cap S^1 = \emptyset$ and $S^0 \cup S^1 = X$. The sets S^0 and S^1 are the back- and foreground, respectively. Equivalently, we say that a pixel $x \in S^0$ has value 0 and $x \in S^1$ has value 1. Let S_{gt} denote the ground truth and S_{pr} be a predicted annotation. We denote the set of incorrectly labeled pixels by $E := (S^0_{\mathrm{pr}} \cap S^1_{\mathrm{gt}}) \cup (S^0_{\mathrm{gt}} \cap S^1_{\mathrm{pr}})$. In the following, we will give a brief overview of established metrics and provide their formulae under a unified notation for the case of binary segmentation. This section is inspired by [100]. By $\uparrow$ and $\downarrow$ we indicate whether larger or smaller values are considered better.

4.2.1 Traditional Metrics

The set of all pixels can be divided into four groups, the *basic cardinalities*

$$\mathrm{TP} = |S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|, \quad \mathrm{FN} = |S^1_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|, \quad \mathrm{FP} = |S^0_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|, \quad \mathrm{TN} = |S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|$$

which fulfill $n := |X| = \mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN}$. They are often shown as a contingency table and suffice to express a large quantity of established metrics used for image segmentation tasks.

4.2.1.1 Accuracy

The *accuracy*

$$\mathrm{ACC} = \frac{|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}| + |S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|}{|X|} = \frac{\mathrm{TP} + \mathrm{TN}}{n} \in [0,1]^{\uparrow}$$

is the ratio of correctly predicted pixels over all pixels. It is the most intuitive measurement of quality and gives equal weight to each pixel.

4.2.1.2 True Positive Rate

The *true positive rate*, also known as *recall* or *sensitivity*,

$$\mathrm{TPR} = \frac{|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}|} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \in [0,1]^{\uparrow}$$

is defined as the ratio of correctly predicted foreground pixels. It neglects the prediction of the background entirely such that the trivial prediction of only foreground yields a perfect score.

4.2.1.3 True Negative Rate

The *true negative rate*, *specificity* or *selectivity*

$$\mathrm{TNR} = \frac{|S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|}{|S^0_{\mathrm{gt}}|} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \in [0,1]^{\uparrow}$$

is the ratio of correctly predicted background pixels and the counterpart to TPR. Since foreground pixels are ignored, the trivial prediction of only background produces a perfect score.

4.2.1.4 Positive Predictive Value

The *positive predictive value* or *precision* is defined as

$$\mathrm{PPV} = \frac{|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{pr}}|} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \in [0,1]^{\uparrow}$$

and represents the probability that a pixel predicted as foreground is correctly annotated. It is almost unaffected by FN errors.

4.2.1.5 Negative Predictive Value

The *negative predictive value* is given by

$$\mathrm{NPV} = \frac{|S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|}{|S^0_{\mathrm{pr}}|} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FN}} \in [0,1]^{\uparrow}$$

and reflects the probability that a pixel predicted as background is correctly annotated. It is the counterpart to PPV and barely influenced by FP errors.

4.2.1.6 Balanced Accuracy

The *balanced accuracy*

$$\mathrm{BAC} = \frac{1}{2}\left(\mathrm{TPR} + \mathrm{TNR}\right) = \frac{1}{2}\left(\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} + \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}\right) \in [0,1]^{\uparrow}$$

is the arithmetic mean of the predictive values TPR and TNR. Since fore- and background are weighted equally, the influence of an error depends on their class affiliation and the class distributions.

4.2.1.7 F_β-Score

The F_β-score [26]

$$F_\beta = \frac{(\beta^2 + 1) \cdot \mathrm{PPV} \cdot \mathrm{TPR}}{\beta^2 \cdot \mathrm{PPV} + \mathrm{TPR}} \in [0,1]^{\uparrow}, \quad \beta > 0$$

adaptively weighs PPV and TPR according to the parameter β. The harmonic mean, i.e., $\beta = 1$, is known as *Dice coefficient* and reduces to

$$\mathrm{DSC} = F_1 = \frac{2|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}| + |S^1_{\mathrm{pr}}|} = \frac{2\,\mathrm{TP}}{2\,\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \in [0,1]^{\uparrow}.$$

The F_β score prioritizes the correct prediction of the foreground over the background.

4.2.1.8 Intersection over Union

The *intersection over union* or *Jaccard index* [61]

$$\mathrm{IoU} = \frac{|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}} \cup S^1_{\mathrm{pr}}|} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} = \frac{\mathrm{DSC}}{2 - \mathrm{DSC}} \in [0,1]^{\uparrow}$$

is defined as the ratio of correctly predicted foreground pixels to all pixels that are either predicted or belong to the foreground. Greater emphasis is given to the correct prediction of the foreground than the background. The last equality shows that IoU and DSC are related and produce equivalent rankings.

4.2.1.9 Matthews Correlation Coefficient

The *Matthews correlation coefficient* [81] or *phi-coefficient*

$$\mathrm{MCC} = \frac{\mathrm{TP} \cdot \mathrm{TN} - \mathrm{FP} \cdot \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}} \in [-1, 1]^{\uparrow}$$

is an estimate of the Pearson correlation coefficient between the ground truth and predicted annotation, i.e.,

$$\mathrm{MCC} = \frac{\mathrm{Cov}(S_{\mathrm{gt}}, S_{\mathrm{pr}})}{\sigma_{S_{\mathrm{gt}}} \sigma_{S_{\mathrm{pr}}}}$$

with covariance estimate $\mathrm{Cov}(S_{\mathrm{gt}}, S_{\mathrm{pr}})$ and estimated standard deviations $\sigma_{S_{\mathrm{gt}}}$ and $\sigma_{S_{\mathrm{pr}}}$. Large values can be obtained only if predictions of fore- and background are simultaneously good. A value of 0 implies that the prediction is equivalent to random guessing. In general, the interpretation of its value is difficult [25, 100].

4.2.1.10 Cohen Kappa

The *Cohen kappa* [28] is given by

$$\kappa = \frac{2(\mathrm{TP} \cdot \mathrm{TN} - \mathrm{FP} \cdot \mathrm{FN})}{(\mathrm{TP} + \mathrm{TN})(\mathrm{FP} + \mathrm{FN}) + 2\,\mathrm{TP} \cdot \mathrm{TN} + \mathrm{FP}^2 + \mathrm{FN}^2} \in [-1, 1]^{\uparrow}$$

and measures the normalized agreement between the two annotations in the form

$$\kappa = \frac{P_a - P_c}{1 - P_c}$$

where P_a is the estimated probability of agreement and P_c is the probability to agree by chance. This can be interpreted as a chance-corrected version of accuracy or the Dice coefficient [25]. κ shares various properties with MCC, e.g., a value of 0 reflects that the prediction is similar to random guessing. Interpretation of its value can be difficult in general [100].

4.2.1.11 Consistency Errors

The *consistency errors* are a family of measures based on the *local refinement errors*

$$E_{S_{\mathrm{gt}}, S_{\mathrm{pr}}}(x) = \begin{cases} \frac{|S^1_{\mathrm{gt}} \setminus S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}|} = \frac{\mathrm{FN}}{\mathrm{TP+FN}} & \text{if } x \in S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}} \\ \frac{|S^1_{\mathrm{gt}} \setminus S^0_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}|} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} & \text{if } x \in S^1_{\mathrm{gt}} \cap S^0_{\mathrm{pr}} \\ \frac{|S^0_{\mathrm{gt}} \setminus S^1_{\mathrm{pr}}|}{|S^0_{\mathrm{gt}}|} = \frac{\mathrm{TN}}{\mathrm{TN+FN}} & \text{if } x \in S^0_{\mathrm{gt}} \cap S^1_{\mathrm{pr}} \\ \frac{|S^0_{\mathrm{gt}} \setminus S^0_{\mathrm{pr}}|}{|S^0_{\mathrm{gt}}|} = \frac{\mathrm{FP}}{\mathrm{TN+FP}} & \text{if } x \in S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}} \end{cases}$$

and $E_{S_{\mathrm{pr}}, S_{\mathrm{gt}}}(x)$ accordingly. The measures are defined in [79, 80] as

$$
\begin{aligned}
\mathrm{LCE} &= \frac{1}{n}\sum_{x\in X}\min\left(E_{S_{\mathrm{gt}},S_{\mathrm{pr}}}(x), E_{S_{\mathrm{pr}},S_{\mathrm{gt}}}(x)\right) \in [0,1]^{\downarrow} \\
&= \frac{\mathrm{TP}}{n}\cdot\min\left(\frac{\mathrm{FN}}{\mathrm{TP}+\mathrm{FN}}, \frac{\mathrm{FP}}{\mathrm{TP}+\mathrm{FP}}\right) + \frac{\mathrm{TN}}{n}\cdot\min\left(\frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}, \frac{\mathrm{FN}}{\mathrm{TN}+\mathrm{FN}}\right) \\
&+ \frac{\mathrm{FP}}{n}\cdot\min\left(\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}, \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\right) + \frac{\mathrm{FN}}{n}\cdot\min\left(\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}\right), \\
\mathrm{GCE} &= \frac{1}{n}\min\left(\sum_{x\in X}E_{S_{\mathrm{gt}},S_{\mathrm{pr}}}(x), \sum_{x\in X}E_{S_{\mathrm{pr}},S_{\mathrm{gt}}}(x)\right) \in [0,1]^{\downarrow} \\
&= \frac{2}{n}\min\left\{\frac{\mathrm{TP}\cdot\mathrm{FN}}{\mathrm{TP}+\mathrm{FN}} + \frac{\mathrm{TN}\cdot\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}, \frac{\mathrm{TP}\cdot\mathrm{FP}}{\mathrm{TP}+\mathrm{FP}} + \frac{\mathrm{TN}\cdot\mathrm{FN}}{\mathrm{TN}+\mathrm{FN}}\right\} \\
\mathrm{BCE} &= \frac{1}{n}\sum_{x\in X}\max\left(E_{S_{\mathrm{gt}},S_{\mathrm{pr}}}(x), E_{S_{\mathrm{pr}},S_{\mathrm{gt}}}(x)\right) \in [0,1]^{\downarrow} \\
&= \frac{\mathrm{TP}}{n}\cdot\max\left(\frac{\mathrm{FN}}{\mathrm{TP}+\mathrm{FN}}, \frac{\mathrm{FP}}{\mathrm{TP}+\mathrm{FP}}\right) + \frac{\mathrm{TN}}{n}\cdot\max\left(\frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}, \frac{\mathrm{FN}}{\mathrm{TN}+\mathrm{FN}}\right) \\
&+ \frac{\mathrm{FP}}{n}\cdot\max\left(\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}, \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\right) + \frac{\mathrm{FN}}{n}\cdot\max\left(\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}\right)
\end{aligned}
$$

and called *local*, *global* and *bidirectional consistency error*, respectively. It holds that BCE $\geq$ GCE $\geq$ LCE. By definition, these metrics match the predicted labels to the ground truth labels since they were originally designed for semantic segmentation. This means that the metric may compare the foreground of the prediction with the background of the ground truth and vice versa if the prediction is poor. In particular, inverting the annotation yields the same value. Interpreting the metrics is difficult and GCE was even discarded by the authors themselves.

4.2.1.12 Volumetric Similarity

The *volumetric similarity*

$$
\mathrm{VS} = 1 - \frac{||S^1_{\mathrm{gt}}| - |S^1_{\mathrm{pr}}||}{|S^1_{\mathrm{gt}}| + |S^1_{\mathrm{pr}}|} = 1 - \frac{|\mathrm{FN}-\mathrm{FP}|}{2\,\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \in [0,1]^{\uparrow}
$$

is defined as the ratio of the difference in the number of foreground pixels to the sum of foreground pixels in both annotations. Hence, it does not consider whether the pixels are predicted correctly. For instance, a perfect score is achieved if the number of foreground pixels in the prediction equals those in the ground truth regardless of their location or correctness.

4.2.1.13 Segmentation Covering

The *segmentation covering* [2]

$$
\begin{aligned}
\mathrm{SC} &= \frac{1}{n}\left(|S^0_{\mathrm{gt}}|\frac{|S^0_{\mathrm{gt}}\cap S^0_{\mathrm{pr}}|}{|S^0_{\mathrm{gt}}\cup S^0_{\mathrm{pr}}|} + |S^1_{\mathrm{gt}}|\frac{|S^1_{\mathrm{gt}}\cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}\cup S^1_{\mathrm{pr}}|}\right) \\
&= \frac{1}{n}\left(\frac{\mathrm{TP}(\mathrm{TP}+\mathrm{FN})}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} + \frac{\mathrm{TN}(\mathrm{TN}+\mathrm{FP})}{\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}\right) \in [0,1]^{\uparrow}
\end{aligned}
$$

is a weighted average of the IoU and its equivalent when switching fore- and background with weights proportional to the respective frequency of each class.

4.2.1.14 Larsen Index

The *Larsen index* [71]

$$\begin{aligned}\mathcal{L} &= \frac{1}{n}\left(|S^0_{\mathrm{gt}}|\frac{2|S^0_{\mathrm{gt}} \cap S^0_{\mathrm{pr}}|}{|S^0_{\mathrm{gt}}| + |S^0_{\mathrm{pr}}|} + |S^1_{\mathrm{gt}}|\frac{2|S^1_{\mathrm{gt}} \cap S^1_{\mathrm{pr}}|}{|S^1_{\mathrm{gt}}| + |S^1_{\mathrm{pr}}|}\right) \\ &= \frac{1}{n}\left(\frac{2(\mathrm{TP}+\mathrm{FN})\mathrm{TP}}{2\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} + \frac{2(\mathrm{TN}+\mathrm{FP})\mathrm{TN}}{2\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}\right) \in [0,1]^{\uparrow}\end{aligned}$$

is the weighted average of DSC and its equivalent when switching fore- and background with weights proportional to the respective frequency of each class. In [85, 86] the Larsen index is introduced with a different formula as the geometric mean of DSC and its counterpart.

4.2.1.15 Intraclass Correlation

The *intraclass correlation* [120]

$$\mathrm{ICC} = \frac{\mathrm{MS}_b - \mathrm{MS}_w}{\mathrm{MS}_b + \mathrm{MS}_w} \in [-1,1]^{\uparrow}$$

with

$$\begin{aligned}\mathrm{MS}_b &= \frac{2}{n-1}\left(\mathrm{TP}(1-\mu)^2 + \mathrm{TN}\mu^2 + (\mathrm{FP}+\mathrm{FN})\left(\frac{1}{2}-\mu\right)^2\right) \\ \mathrm{MS}_w &= \frac{\mathrm{FP}+\mathrm{FN}}{2n}, \qquad \mu = \frac{2\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}{2n}\end{aligned}$$

measures the conformity between the ground truth and prediction. It is a consistent but biased estimator [120] of the variance ratio $\rho = \frac{\sigma_B^2}{\sigma_B^2+\sigma_W^2}$ where σ_B^2 denotes the variance caused by differences between the segmentations and σ_W^2 denotes the variance caused by differences between the points within the segmentations, see [123]. The estimate is calculated using the mean squares between the segmentations MS_b and the mean squares within the segmentations MS_w. The ICC was originally defined as part of a family of correlation coefficients, see [120] for more details. It is sometimes mistakenly called "interclass" correlation [123, 133] which is a term that is rather associated with the Pearson correlation coefficient [82], i.e., MCC.

4.2.2 Traditional Metrics: Pair-Counting

Another approach to measuring segmentation quality is provided by so-called pair-counting. The fundamental idea is to disregard the labeling as fore- and background and instead focus on consistency and inconsistency in the affiliations of pixel pairs between the prediction and ground truth. The corresponding *pair-counting cardinalities* are

$$P_{11} = |\{(x,y) \in X \times X \,|\, \exists i,j \in \{0,1\} \,:\, x,y \in S^i_{\mathrm{gt}}, x,y \in S^j_{\mathrm{pr}}\}|$$

$$
\begin{aligned}
&= \frac{1}{2}\left(\mathrm{TP}(\mathrm{TP}-1) + \mathrm{FP}(\mathrm{FP}-1) + \mathrm{FN}(\mathrm{FN}-1) + \mathrm{TN}(\mathrm{TN}-1)\right) \\
P_{10} &= |\{(x,y) \in X \times X \,|\, \exists i,j \in \{0,1\} \,:\, x,y \in S^i_{\mathrm{gt}}, x \in S^j_{\mathrm{pr}}, y \notin S^j_{\mathrm{pr}}\}| \\
&= \frac{1}{2}\left((\mathrm{TP}+\mathrm{FN})^2 + (\mathrm{TN}+\mathrm{FP})^2 - (\mathrm{TP}^2 + \mathrm{TN}^2 + \mathrm{FP}^2 + \mathrm{FN}^2)\right) \\
P_{01} &= |\{(x,y) \in X \times X \,|\, \exists i,j \in \{0,1\} \,:\, x \in S^i_{\mathrm{gt}}, y \notin S^i_{\mathrm{gt}}, x,y \in S^j_{\mathrm{pr}}\}| \\
&= \frac{1}{2}\left((\mathrm{TP}+\mathrm{FP})^2 + (\mathrm{TN}+\mathrm{FN})^2 - (\mathrm{TP}^2 + \mathrm{TN}^2 + \mathrm{FP}^2 + \mathrm{FN}^2)\right) \\
P_{00} &= |\{(x,y) \in X \times X \,|\, \exists i,j \in \{0,1\} \,:\, x \in S^i_{\mathrm{gt}}, y \notin S^i_{\mathrm{gt}}, x \in S^j_{\mathrm{pr}}, y \notin S^j_{\mathrm{pr}}\}| \\
&= \frac{n(n-1)}{2} - P_{11} - P_{10} - P_{01}.
\end{aligned}
$$

The metrics derived from these can clearly be expressed by the basic cardinalities.

4.2.2.1 Rand Index

The *Rand index* [99] is defined as

$$\mathrm{RI} = \frac{P_{11} + P_{00}}{P_{11} + P_{10} + P_{01} + P_{00}} \in [0,1]^{\uparrow}$$

and presents the pair-counting equivalent to the accuracy. The rescaled version

$$\mathcal{M} = n(n-1)(1-\mathrm{RI})$$

is known as *Mirkin metric* [86].

4.2.2.2 Adjusted Rand Index

The *adjusted Rand index* [58]

$$\mathrm{ARI} = \frac{2(P_{11}P_{00} - P_{10}P_{01})}{P_{10}^2 + P_{01}^2 + 2P_{11}P_{00} + (P_{11}+P_{00})(P_{10}+P_{01})} \in [-1,1]^{\uparrow}$$

is the pair-counting equivalent to the Cohen kappa and measures the agreement between the pixel pairs of two samples.

4.2.2.3 Fowlkes-Mallow Index

The *Fowlkes-Mallow index* [48] is defined by

$$\mathrm{FMI} = \sqrt{\frac{P_{11}}{P_{11}+P_{10}} \cdot \frac{P_{11}}{P_{11}+P_{01}}} \in [0,1]^{\uparrow}.$$

It can be connected to the product of the pair-counting equivalents of PPV and TPR which can be interpreted as the conditional probabilities that a pair of pixels which lie in the same region in the ground truth also lie in the same region in the prediction and vice versa.

4.2.2.4 Pair-counting Jaccard Index

The *pair-counting Jaccard index* is defined as

$$\mathrm{JAC}_{\mathrm{pc}} = \frac{P_{11}}{P_{11} + P_{10} + P_{01}} \in [0,1]^{\uparrow}$$

and the pair-counting equivalent to the IoU.

4.2.3 Traditional Metrics: Information Theory

Segmentation quality can also be assessed through the means of information theory based on the *entropies* that are defined as

$$H(S) = -\mathbb{P}(S^0)\log(\mathbb{P}(S^0)) - \mathbb{P}(S^1)\log(\mathbb{P}(S^1)), \quad S \in \{S_{\mathrm{gt}}, S_{\mathrm{pr}}\}$$
$$H(S_{\mathrm{gt}}, S_{\mathrm{pr}}) = -\sum_{i,j=0}^{1} \mathbb{P}(S_{\mathrm{gt}}^i, S_{\mathrm{pr}}^j)\log(\mathbb{P}(S_{\mathrm{gt}}^i, S_{\mathrm{pr}}^j))$$

with the estimated probabilities

$$\mathbb{P}(S_{\mathrm{gt}}^1) = \frac{|S_{\mathrm{gt}}^1|}{|X|} = \frac{TP+FN}{n}, \qquad \mathbb{P}(S_{\mathrm{gt}}^0) = \frac{|S_{\mathrm{gt}}^0|}{|X|} = \frac{TN+FP}{n},$$
$$\mathbb{P}(S_{\mathrm{pr}}^1) = \frac{|S_{\mathrm{pr}}^1|}{|X|} = \frac{TP+FP}{n}, \qquad \mathbb{P}(S_{\mathrm{pr}}^0) = \frac{|S_{\mathrm{pr}}^0|}{|X|} = \frac{TN+FN}{n},$$
$$\mathbb{P}(S_{\mathrm{gt}}^1, S_{\mathrm{pr}}^1) = \frac{|S_{\mathrm{gt}}^1 \cap S_{\mathrm{pr}}^1|}{|X|} = \frac{TP}{n}, \qquad \mathbb{P}(S_{\mathrm{gt}}^1, S_{\mathrm{pr}}^0) = \frac{|S_{\mathrm{gt}}^1 \cap S_{\mathrm{pr}}^0|}{|X|} = \frac{FN}{n},$$
$$\mathbb{P}(S_{\mathrm{gt}}^0, S_{\mathrm{pr}}^1) = \frac{|S_{\mathrm{gt}}^0 \cap S_{\mathrm{pr}}^1|}{|X|} = \frac{FP}{n}, \qquad \mathbb{P}(S_{\mathrm{gt}}^0, S_{\mathrm{pr}}^0) = \frac{|S_{\mathrm{gt}}^0 \cap S_{\mathrm{pr}}^0|}{|X|} = \frac{TN}{n}.$$

The entropies can be interpreted to capture the uncertainty of information.

4.2.3.1 Mutual Information

The *mutual information* [128]

$$\mathrm{MI} = H(S_{\mathrm{gt}}) + H(S_{\mathrm{pr}}) - H(S_{\mathrm{gt}}, S_{\mathrm{pr}}) \in [0, \log(4)]^{\uparrow}$$

can be interpreted as the amount of information one segmentation has about the other. [63] suggest the normalized equivalent $\overline{\mathrm{MI}} = \mathrm{MI}/\log(4)$.

4.2.3.2 Variation of Information

The *variation of information* [85] is defined as

$$\mathrm{VOI} = 2H(S_{\mathrm{gt}}, S_{\mathrm{pr}}) - H(S_{\mathrm{gt}}) - H(S_{\mathrm{pr}}) \in [0, \log(4)]^{\downarrow}$$

and quantifies the loss or gain in information between the prediction and ground truth. Normalization via $\overline{\mathrm{VOI}} = \mathrm{VOI}/\log(4)$ is suggested in [85].

4.2.4 Distance-based Metrics

We can also measure the quality of a segmentation based on the geometric information of the ground truth and prediction. More precisely, we can take into account how far away incorrectly predicted pixels are from the structure. For this we denote by

$$d_{\partial \mathrm{gt}}(x) = \begin{cases} \min_{y \in S^1_{\mathrm{gt}}} \|x - y\|_2 & \text{if } x \in S^0_{\mathrm{gt}} \\ \min_{y \in S^0_{\mathrm{gt}}} \|x - y\|_2 & \text{if } x \in S^1_{\mathrm{gt}}, \end{cases} \quad d_{\partial \mathrm{pr}}(x) = \begin{cases} \min_{y \in S^1_{\mathrm{pr}}} \|x - y\|_2 & \text{if } x \in S^0_{\mathrm{pr}} \\ \min_{y \in S^0_{\mathrm{pr}}} \|x - y\|_2 & \text{if } x \in S^1_{\mathrm{pr}} \end{cases}$$

the Euclidean distances of a pixel $x \in X$ to the surface of the ground truth and prediction, respectively.

4.2.4.1 Average Hausdorff Distance

The *directed average Hausdorff distance* [123] is defined as

$$\mathrm{dAHD}(S_{\mathrm{pr}}, S_{\mathrm{gt}}) = \frac{1}{|X|} \sum_{x \in E} d_{\partial \mathrm{gt}}(x) \geq 0 \downarrow$$

where incorrectly predicted pixels are weighted by their distance to the surface of the ground truth. The symmetric version of this metric,

$$\mathrm{AHD} = \max(\mathrm{dAHD}(S_{\mathrm{pr}}, S_{\mathrm{gt}}), \mathrm{dAHD}(S_{\mathrm{gt}}, S_{\mathrm{pr}}))$$

is known as *average Hausdorff distance*.

4.2.4.2 Average Symmetric Surface Distance

The *average symmetric surface distance* [100] is given by

$$\mathrm{ASSD}(S_{\mathrm{pr}}, S_{\mathrm{gt}}) = \frac{1}{|\partial S^1_{\mathrm{gt}}| + |\partial S^1_{\mathrm{pr}}|} \left(\sum_{x \in \partial S^1_{\mathrm{pr}}} d_{\partial \mathrm{gt}}(x) + \sum_{x \in \partial S^1_{\mathrm{gt}}} d_{\partial \mathrm{pr}}(x) \right) \geq 0 \downarrow$$

where $\partial S^1_{\mathrm{gt}}$ and $\partial S^1_{\mathrm{pr}}$ are the boundaries of the foreground of the ground truth and prediction. The metrics can be understood as the average distance between the surfaces of the prediction and the ground truth.

4.2.4.3 Surface Consistency Coefficient

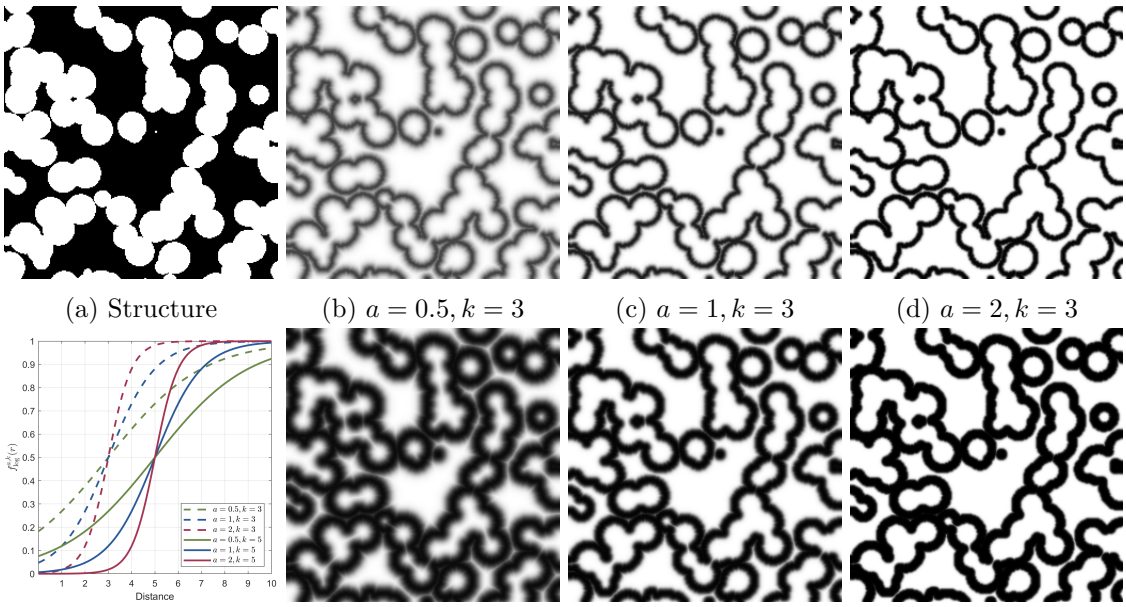
Suppose that there exist incorrectly labeled pixels, i.e., $E \neq \emptyset$. We propose the *surface consistency coefficient* [109]

$$\mathrm{SCC}^{a,k}(S_{\mathrm{pr}}, S_{\mathrm{gt}}) \coloneqq \frac{1}{|E|} \sum_{x \in E} f^{a,k}_{\log}(d_{\partial \mathrm{gt}}(x)) \in [0, 1] \tag{4.1}$$

with a logistic weighting function

$$f^{a,k}_{\log}(r) = \frac{1}{1 + \exp\left(-a(r - k)\right)} \in [0, 1], \quad a > 0 \text{ and } k \geq 0.$$

Compared to previous distance-based metrics, SCC attains values in the interval $[0, 1]$ since it is normalized by the number of errors $|E|$ and the weights of predicted pixels are bounded by 1. This also means that SCC does not scale directly with the quantity of errors. The scaling parameter a and shift parameter k provided by the logistic weighting function are adjustable and can be used to adapt the metric to a desired understanding of proximity depending on its application, see Figure 4.1.



(a) Structure (b) $a = 0.5, k = 3$ (c) $a = 1, k = 3$ (d) $a = 2, k = 3$

(e) Weight functions (f) $a = 0.5, k = 5$ (g) $a = 1, k = 5$ (h) $a = 2, k = 5$

Figure 4.1: Section image of a 3D volume of overlapping spheres (a). Weight maps (b-d, f-h) using different parameter combinations for $f_{\log}^{a,k}$ (e). $a = 1$, $k = 5$ defines a reasonable proximity range and transition speed for this geometry.

The parameter k is called the *proximity range* and marks the distance from the surface at which pixels transition from being considered still close to rather distant, indicated by $f_{\log}^{a,k}(k) = 0.5$. The parameter a controls the speed of this transition and can be interpreted as the uncertainty in the proximity range. We suggest choosing $a \approx 4/(k_{\max} - k)$ where $k_{\max} > k$ is the smallest distance from the surface at which a pixel is without a doubt far away. Then, the quantity $2(k_{\max} - k)$ defines the *transition width* over which the function $f_{\log}^{a,k}$ rises from approximately 0 to 1. In other words, it reflects the range over which we become increasingly confident that a pixel is distant. This concept is illustrated in Figure 4.2. In the limit $a \to \infty$, $f_{\log}^{a,k}$ converges to a step function, i.e., a direct transition from proximal to distant. In this case, the value of SCC becomes the fraction of errors found outside the proximity range. In general, small values indicate that errors are predominantly found near the surface, while large values imply that errors appear mostly outside the proximity range.

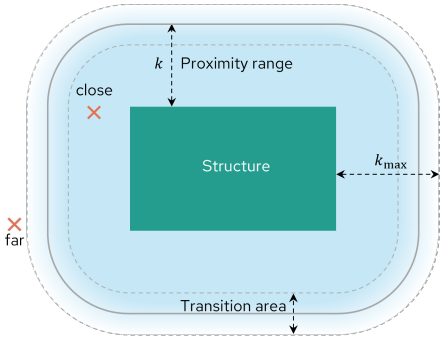


Figure 4.2: Illustration of the proximity range, related quantities and interpretations based on the foreground.

4.2.5 Synthetic Data with Geometric Context

For assessing segmentation quality we consider a large dataset consisting of 1860 segmentation images generated in two steps. First, we sample a series of diverse 3D geometries of size 512^3 pixels from stochastic processes. The structures are obtained by placing a series of particles which are either spheres, cubes, cylinders, cuboids or ellipsoids uniformly random. We make the shapes comparable by choosing their parametrizations, see Table 4.1, so that their ratios of surface area to volume are the same. Each particle is rotated uniformly with respect to its center of mass so that the resulting stochastic model is isotropic. For each particle shape we generate a geometry with (volume) density of approximately 10%, 30%, 50% and 70% resulting in a total of 20 different geometric structures which are visualized in Figure 4.3. For a density of 10%, particles were placed via rejection sampling to ensure they do not overlap. In all other cases, realizations were drawn from Boolean models. The variation in shapes and densities ensure a broad coverage of geometric diversity including porous and dense, angular and rounded structures.

Shape	**Parameters** [px]
Sphere	$radius = 15$
Cube	$edge\ length = 30$
Cylinder	$radius = 10.5$, $height = 210$
Cuboid	$edge\ lengths = (14.33, 43, 143.33)$
Ellipsoid	$semiaxes = (8.46, 25.39, 84.63)$

Table 4.1: Parameters of each particle shape used for generating 3D geometries. The ratio of surface area to volume for each shape is approximately 0.2.

Afterward, we obtain different segmentation results for each geometry by systematically altering the structures, i.e., by omitting pixels from the foreground or adding others to it. The modifications to the geometry are repeated for different error percentages ranging from 1% to 15% of the total pixel count, if possible. We employ three primary modification strategies: First, we use morphological operations to introduce errors near to the surface. Second, we combine the Euclidean distance transformation (EDT) and thresholding to set incorrect labels furthest from the surface. Third, we introduce errors randomly according to selected distributions. In the following, we provide a brief explanation of the seven synthetic segmentations that we derive using these strategies, see Figure 4.4 for examples.

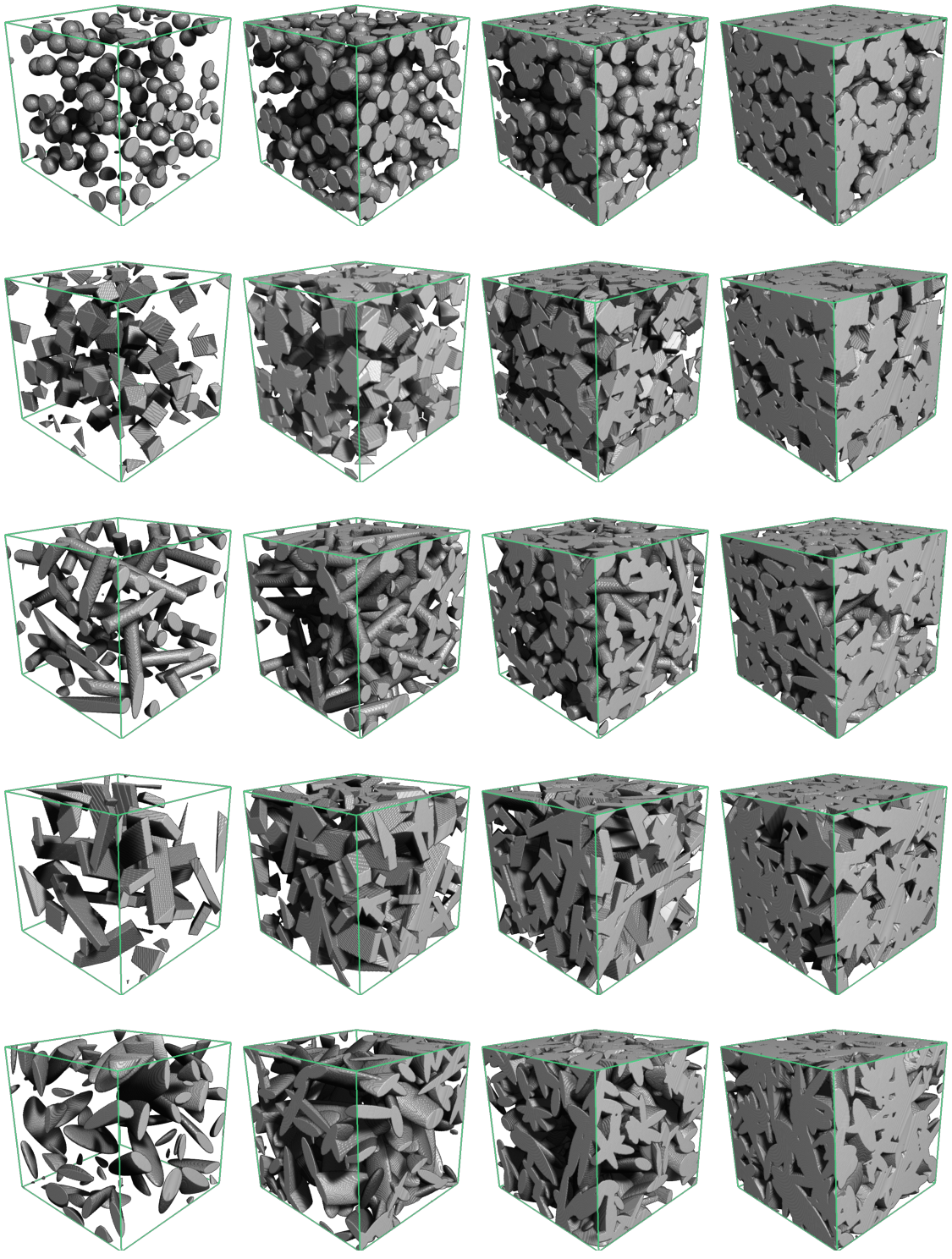
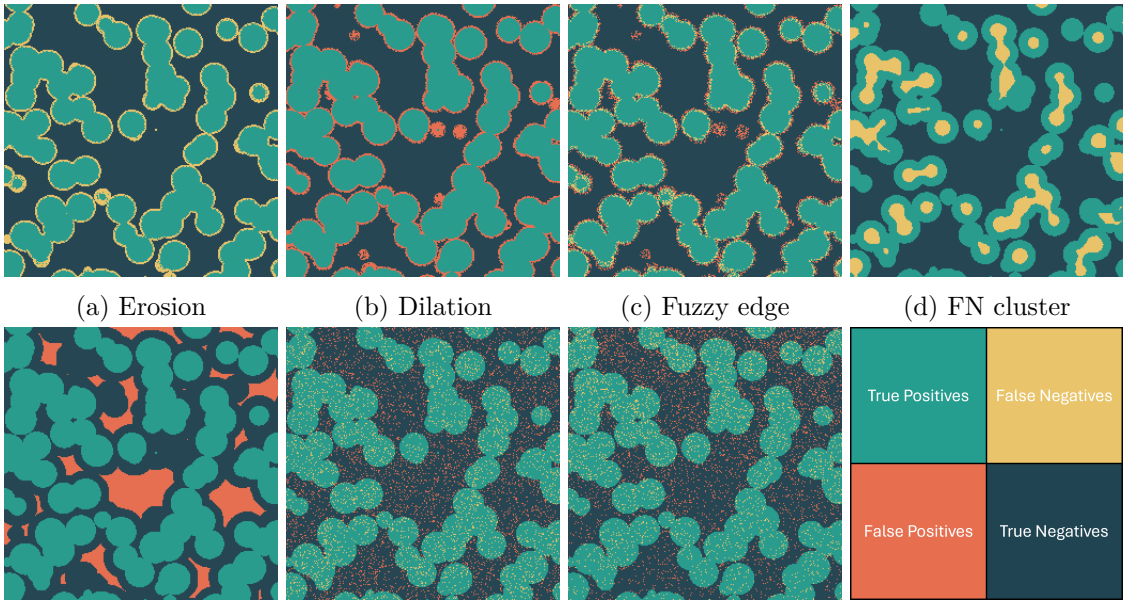


Figure 4.3: Renderings of the 3D structures considered to generate synthetic segmentations with densities 10%, 30%, 50% and 70% (left to right) and particle shapes spheres, cubes, cylinders, cuboids and ellipsoids (top to bottom)



(a) Erosion (b) Dilation (c) Fuzzy edge (d) FN cluster

(e) FP cluster (f) Uniform (g) Nonuniform (h) Color legend

Figure 4.4: Synthetic segmentations created by introducing systematic errors with 15% error rate visualized on a 256^2 section image from a 3D volume of overlapping spheres.

4.2.5.1 Proximate Errors

- **Erosion.** Application of the morphological erosion with a ball as structuring element. The ball's size depends on the targeted error rate. Precise error rates are achieved by uniformly sampling pixels at the edge of the eroded structure until the desired rate is reached.
- **Dilation.** Application of the morphological dilation with a ball as structuring element. The ball's size depends on the targeted error rate. Precise error rates are achieved by uniformly sampling pixels at the edge of the dilated structure until the desired rate is reached.
- **Fuzzy edge.** Consider the two structures obtained by erosion and dilation with the respective error rate. Pixels in the band defined by the difference between them are uniformly sampled so that half the errors are false positives and half are false negatives.

4.2.5.2 Distant Errors

- **FN cluster.** Calculation of the EDT on the foreground followed by thresholding to determine pixels deepest inside the geometry. The threshold depends on the targeted error rate. Precise error rates are achieved by uniformly sampling pixels with a distance close to the threshold until the desired rate is reached.
- **FP cluster.** Calculation of the EDT on the background followed by thresholding to determine the pixels furthest from the geometry. The threshold depends on the

targeted error rate. Precise error rates are achieved by uniformly selecting pixels with a distance close to the threshold until the desired rate is reached.

4.2.5.3 Random Errors

- **Uniform.** Pixels are sampled uniformly with an incorrect label yielding salt and pepper noise.
- **Nonuniform.** Pixels are sampled from a Poisson point process with inhomogeneous intensity that decreases linearly along the y-axis. The intensity depends on the targeted error rate.

4.3 Results and Discussion

In this section, we evaluate and compare the previously introduced quality metrics for image segmentation using primarily the synthetic dataset established in Section 4.2.5. On the synthetic data, we considered $SCC^{1,5}$ since the corresponding proximity range is suitable for the structures' sizes in our opinion. By conducting a correlation analysis we first identify relationships between the metrics and determine which metrics carry redundant information. Afterward, we investigate the behavior of selected metrics including SCC in more detail under the controlled circumstances provided by the data that vary in the quantity and spatial distribution of errors. Lastly, we exemplify the integration of SCC into real segmentation tasks by extending a study by Barisin et al. [7] who analyzed and compared the performance of different segmentation methods for crack segmentation in CT images.

4.3.1 Correlation Analysis

We evaluated all metrics introduced in Section 4.2 on the entire dataset of synthetic segmentation data presented in Section 4.2.5 which incorporates variations in error quantity and their spatial distributions across different underlying geometries. Based on this, pairwise Pearson correlation coefficients between the metrics were calculated and are visualized in Figure 4.6. With few exceptions, we observe moderate to strong correlations among the traditional metrics which are primarily due to their shared design principle. They rely entirely on the basic cardinalities to establish a value indicative of quality that improves with lower error rates and worsens with higher error rates. Nonetheless, some exceptions such as TNR and DSC show no correlation because they measure independent aspects of quality. TNR emphasizes only the accurate prediction of the background while DSC prioritizes the foreground. The generally high correlations among the traditional metrics suggest that only a few, or even one metric suffice for evaluating quality based on the error quantity since they carry mostly redundant information. For this, we suggest using easily interpretable metrics such as ACC rather than more complex metrics whose interpretation is less intuitive. Alternatively, MCC could be considered since it has advantages over other traditional metrics as demonstrated by Chicco and colleagues [22–25]. In instances where the priority of assessing quality lies in the fore- or background, respectively, the use of different traditional metrics such as DSC or TNR is also feasible.

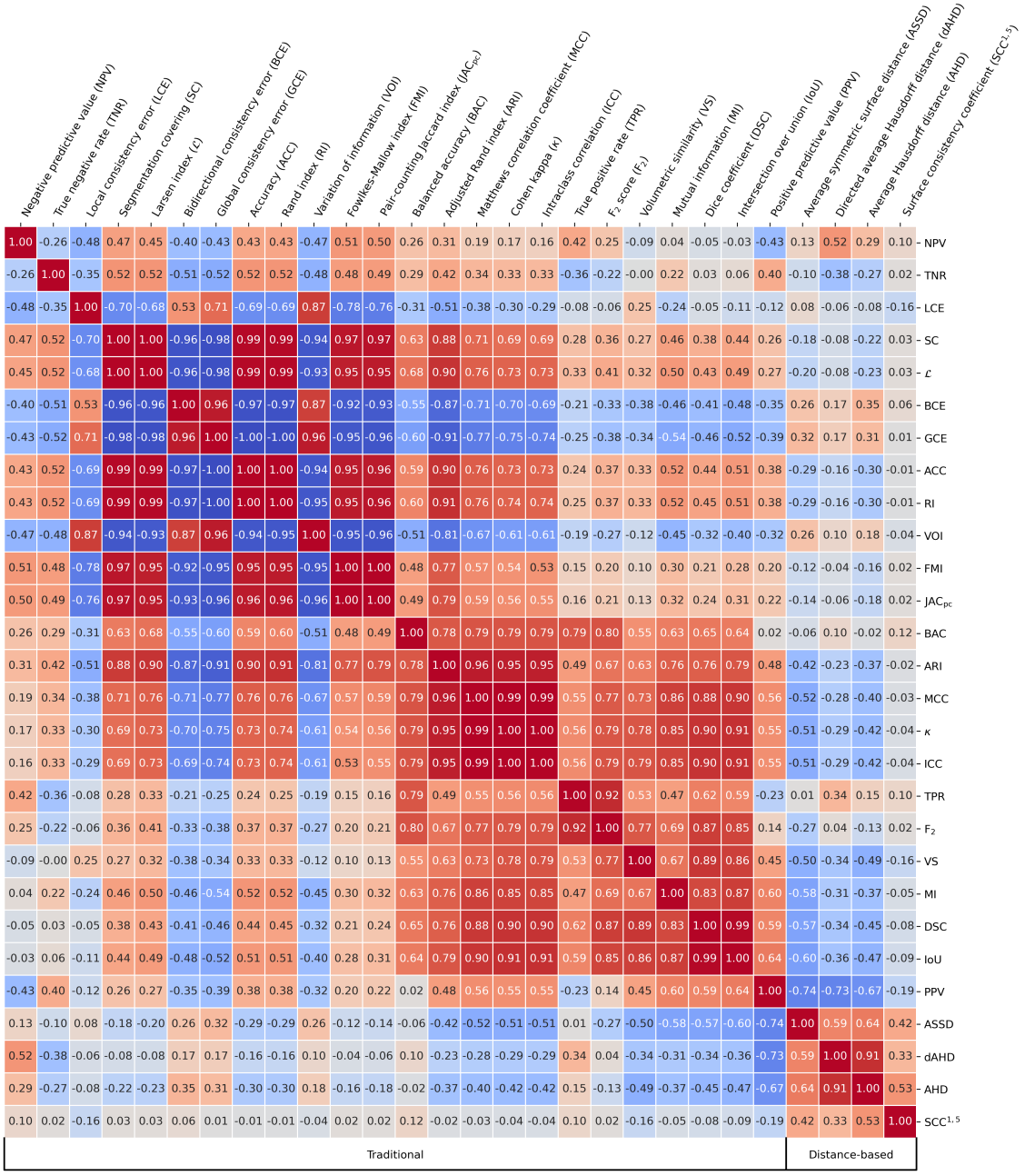


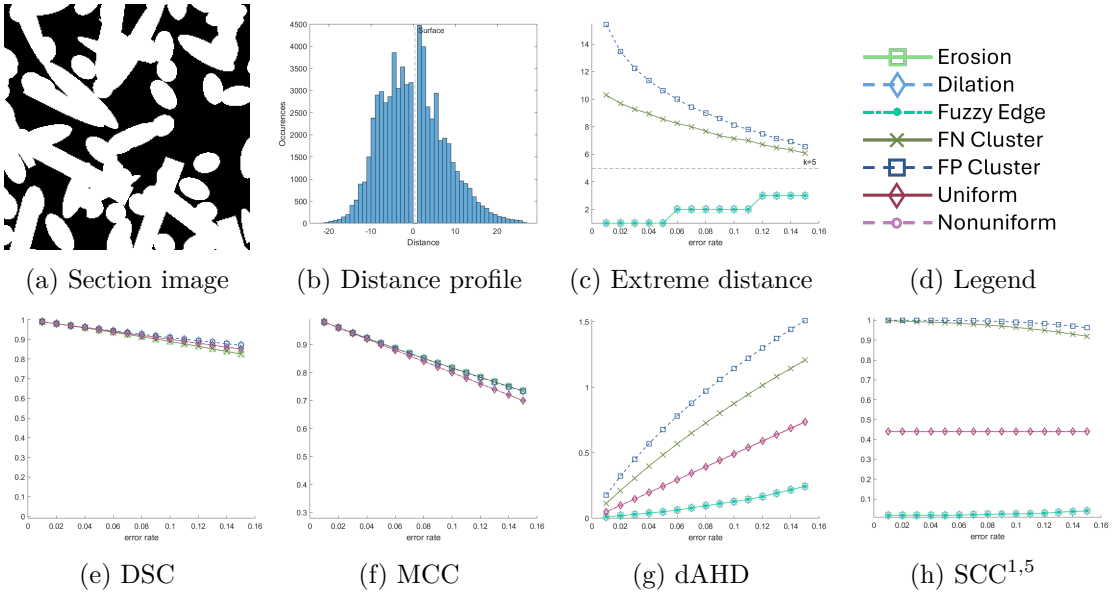
Figure 4.5: Pair-wise Pearson correlation coefficients between the different quality metrics calculated from the entire set of synthetic segmentations.

Distance-based metrics and traditional metrics exhibit significantly weaker correlations with one another. This discrepancy arises because distance-based metrics incorporate spatial information which is entirely absent in traditional metrics. Nonetheless, some weak to moderate correlation coefficients can be found, e.g., for dAHD, AHD, and ASSD in relation with PPV. While influenced by distances to the structures' surfaces, these distance-based metrics also scale with error quantity which is already captured by traditional metrics. As a result, they are not independent of traditional metrics.

In contrast, SCC stands out for showing no correlation with traditional metrics and only moderate correlation with other distance-based metrics. This indicates that SCC provides distinct insights into the spatial distribution of errors without the influence of error quantity. In particular, this property unique to SCC enables the independent assessment of quality based on error quantity and spatial distribution when combined with a traditional metric such as ACC.

4.3.2 Study on Selected Geometries

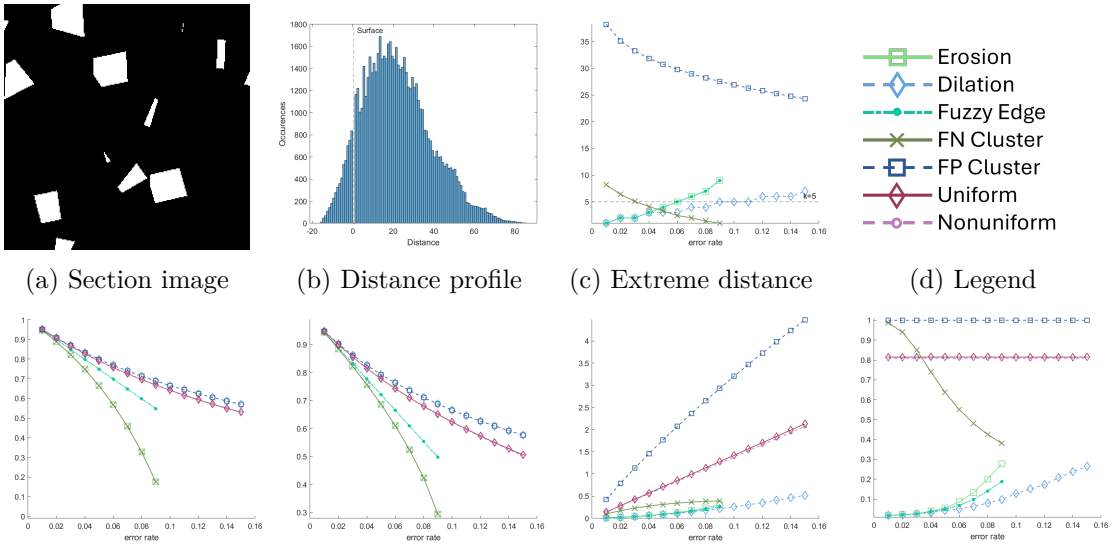
In the following, we perform a rigorous analysis of SCC and other selected metrics on a subset of the synthetic segmentation data established in Section 4.2.5. More precisely, we evaluate and compare the metrics on all segmentations derived from selected geometries. We validate the use of SCC for quality assessment of segmentation results and highlight several key properties that set it apart from other quality metrics. We compare SCC only with DSC, MCC and dAHD following the results in Section 4.3.1 which showcased that all remaining metrics have a moderate to high Pearson correlation coefficient with at least one of the selected metrics. We also incorporate ACC through the error rate since ACC $= 1 -$ *error rate*. We remark that for a density of 10% we can only produce segmentations for erosion and FN cluster up to an error rate of 10%, see Figures 4.7 and 4.9 (a).



(a) Section image (b) Distance profile (c) Extreme distance (d) Legend

(e) DSC (f) MCC (g) dAHD (h) $SCC^{1,5}$

Figure 4.6: Comparison of quality metrics for systematic errors on an image of overlapping cylinders of radius 10.5 pixels, height 210 pixels and volume density 50%. Sectional image of the structure (a). Histogram of distances from the surface split into foreground (negative) and background (positive) pixels (b). Largest distance of an error to the surface for proximate errors and smallest distance to the surface for distant errors (c). Legend (d). Metrics for varying error rate (e-h).



(a) Section image (b) Distance profile (c) Extreme distance (d) Legend

(e) DSC (f) MCC (g) dAHD (h) $SCC^{1,5}$

Figure 4.7: Comparison of quality metrics for systematic errors on an image of non-overlapping cubes with side length 30 pixels and volume density 10%. Sectional image of the structure (a). Histogram of distances from the surface split into foreground (negative) and background (positive) pixels (b). Largest distance of an error to the surface for proximate errors and smallest distance to the surface for distant errors (c). Legend (d). Metrics for varying error rate (e-h).

4.3.2.1 Traditional Metrics

The results presented in Figures 4.6 and 4.7 summarize the evaluation of the metrics for all segmentations derived from a Boolean cylinder model with 50% density and non-overlapping cubes with 10% density, respectively. They confirm that DSC and MCC are only influenced by the quantity and type of errors, but are independent of their spatial distribution. In particular, the systematic errors that introduce the same amounts of FPs and FNs for a given structure produce the same value of DSC and MCC, e.g., dilation and FP cluster. This unsurprisingly persists across all geometries and follows immediately from their respective definitions. In class-balanced settings, see Figure 4.6, where fore- and background appear equally, both traditional metrics relate linearly to the error rate. Hence, they do not provide additional characterization of quality. In class-imbalanced cases, see Figure 4.7, MCC and DSC prioritize the correct labeling of the less frequent foreground as shown by the steeper descent for systematic errors that affect the foreground, e.g., erosion and FN cluster. This is often desirable and presents their main advantage over using the error rate or ACC which are easier to interpret.

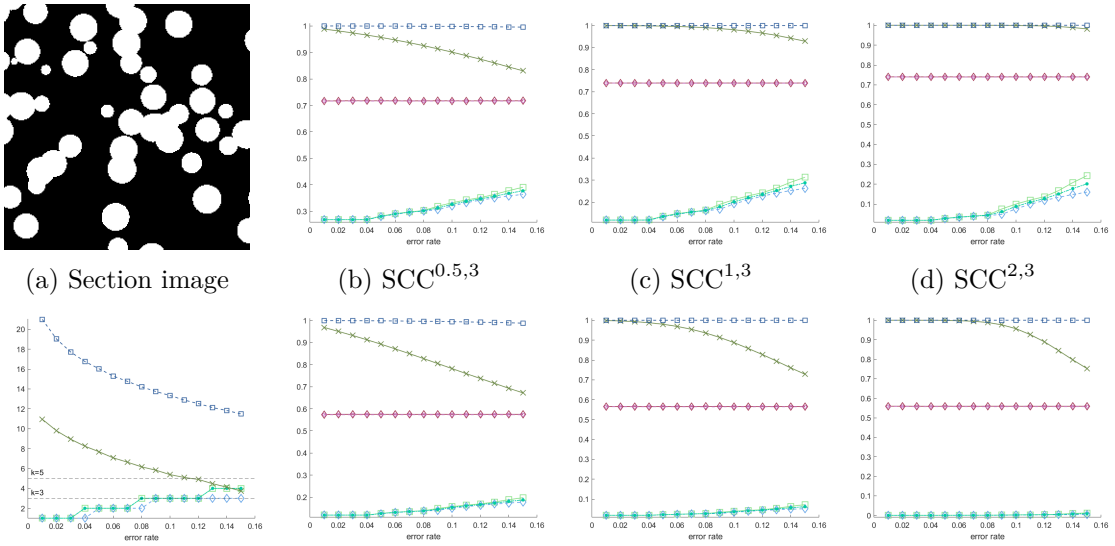
4.3.2.2 Distance-based Metrics

Compared to traditional metrics, the distance-based metrics dAHD and SCC behave very differently across the individual systematic errors. They can separate proximate from distant errors and are not influenced by the error type, i.e., FN or FP. dAHD achieves this by incorporating both the quantity of errors and their distance from the surface showing

an almost linear scaling with the error rate in Figures 4.6 and 4.7 (g). The scaling for distant errors is distinctively larger than for proximate errors. However, individual values of dAHD are difficult to interpret and compare. This has two reasons. First, the same value of dAHD can be obtained through different combinations of error quantity and their positions, e.g., few distant or many proximal errors. Second, there exists no universal supremum across geometries such that a deduction of quality is not possible from a specific value alone, compare the scales in Figures 4.6 and 4.7 (g). Nevertheless, a combination of dAHD with traditional metrics would still allow for comparative statements about quality differences between segmentations based on the spatial distributions of errors. However, this requires considerable expertise and effort since the influence of various factors must be recognized.

The problems identified with dAHD are completely resolved by SCC. Most notably, it is easily interpretable and uniquely characterizes error positioning independent of error type or quantity, at least in the class-balanced setting. Proximal errors are consistently mapped to zero and distant errors to one, see Figure 4.6 (h). Depending on the application and the user's preference, either extreme may reflect better quality. As previously mentioned, the value of SCC relates to the fraction of errors found outside the proximity range. For the considered random segmentations which place errors independent of the structures, this translates to the fraction of volume outside the proximity range. The respective fractions of 43% for the structure in Figure 4.6 and 82% in Figure 4.7 align perfectly with our measurements of SCC.

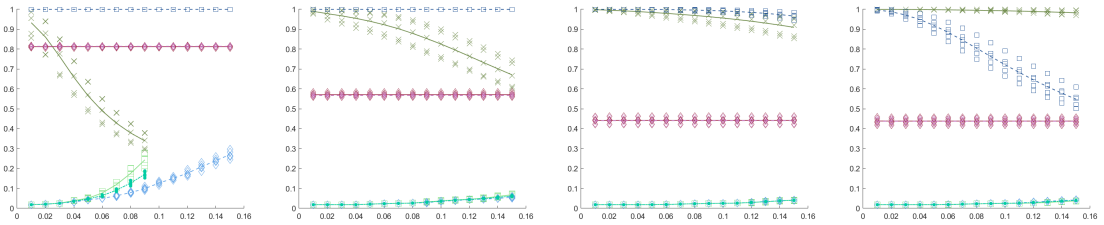
The interpretability of SCC is further substantiated by the results shown in Figure 4.7 (h). Due to the low-density geometry larger error rates are achieved by introducing errors near or beyond the proximity range with the exception of dilation, see (c). This violation of the proximity range directly translates into changes in SCC for the affected segmentations. In other words, the proximate errors introduce incorrect labels that are not considered proximate, and similarly, FN cluster introduces incorrect labels that are not distant anymore. Both are correctly reflected by SCC. The same argument applies when we consider different parameter combinations of a and k as shown in Figure 4.8. Here, the structure and distance distribution remain the same but the definition of proximity changes. Deviations from the anticipated values for proximate and distant errors are observed for larger error rates and again relate directly to violations of the respective proximity ranges. This demonstrates that SCC can comprehensively quantify the spatial distribution of errors regardless of the error rate and type. In addition, SCC yields consistent values across comparable geometries as shown in Figure 4.9 since they provide comparable distance profiles and therefore spatial error distributions. Since SCC measures quality independently of the error quantity which is undoubtedly essential for quality assessment, we propose pairing it with a traditional metric such as ACC. In this way, both the number of erroneous pixels and their spatial locations can be intuitively and independently incorporated into the validation of a segmentation result.



(a) Section image (b) $SCC^{0.5,3}$ (c) $SCC^{1,3}$ (d) $SCC^{2,3}$

(e) Extreme distance (f) $SCC^{0.5,5}$ (g) $SCC^{1,5}$ (h) $SCC^{2,5}$

Figure 4.8: Evaluation of SCC on an image of overlapping spheres with radius 20 pixels and volume density 30% (a). Results for varying parameters a and k (b-d, f-h). Largest distance of an error to the surface for proximate errors and the smallest distance for distant errors (e). See Figure 4.6 (d) for legend.



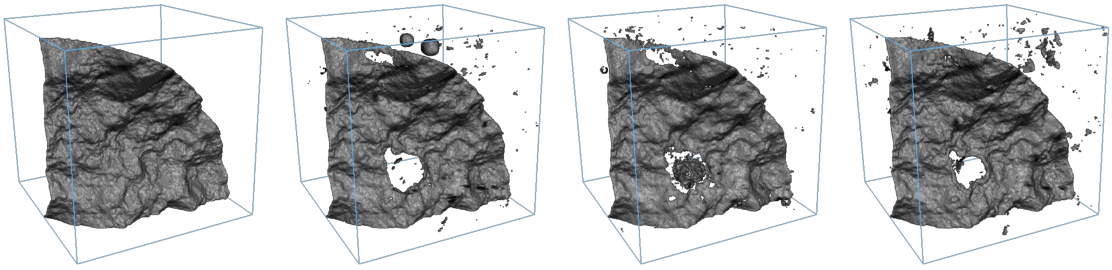
(a) 10% (b) 30% (c) 50% (d) 70%

Figure 4.9: Evaluation of $SCC^{1,5}$ on the entire data set split by volume density. Lines show the mean for each systematic error. See Figure 4.6 (d) for legend.

4.3.3 Application: Cracks in Concrete

In the following, we demonstrate the inclusion and benefits of SCC to the quality assessment of traditional metrics by extending a study of segmentation results performed by Barisin and colleagues [7]. The authors meticulously compared the performance of different methods for the segmentation of cracks in CT images of concrete based on semi-synthetic crack data [65, 66] which provides ground truths. Their quality assessment relied on traditional metrics and visual inspection. In their article, the group found that Hessian-based percolation and the learning-based approaches, random forest and 3D U-Net, produce the best results among the tested methods in terms of visual assessment and using traditional metrics. More precisely, the methods were on par but 3D U-Net was slightly better. For the evaluation with traditional metrics they considered including a tolerance to ignore incorrect labels directly adjacent to the crack. This approach is not

unusual and resembles that capturing the general shape of the crack rather than its precise texture is often more important. The same idea can be found in SCC which penalizes incorrect annotations near the surface less. With the same interpretation small values of SCC would be considered better.



(a) Ground truth (b) Percolation (c) Random forest (d) 3D U-Net

Figure 4.10: Renderings of a synthetic crack image of width 3 pixels (a) and the segmentation results (b-d) obtained from different methods [7].

Method	**Error [%]**	**ACC**$^{\uparrow}$	**DSC**$^{\uparrow}$	**MCC**$^{\uparrow}$	**dAHD**$^{\downarrow}$	**SCC**^{2,3}
Percolation	0.257	0.743	0.868	0.868	0.094	0.317
Random forest	0.231	0.769	0.866	0.866	0.007	0.080
3D U-Net	0.213	0.787	0.892	0.893	0.026	0.177

Table 4.2: Numerical assessment of the segmentation results shown in Figure 4.10.

Table 4.2 displays the numerical assessment of quality without tolerance using the error rate, DSC, MCC, dAHD and SCC on the three methods which were found to perform well. In this example, we use $a = 2$ and $k = 3$ when calculating the SCC due to the small thickness of the crack. Based solely on traditional metrics, we agree with the authors that the three methods perform on par with 3D U-Net marginally better. However, SCC reveals significant differences in the spatial distribution of errors. In particular, the lower value for random forest indicates that it captures the general shape better than the 3D U-Net since errors are more concentrated near the surface. Similarly, the larger value for percolation implies that the surface texture is more precisely reproduced, and that errors are predominantly introduced away from the surface, e.g., by falsely detecting pores or pore boundaries. In such cases, post-processing, such as removing smaller connected components, could be employed to refine the segmentation.
With the help of SCC we were able to improve on the previous quality assessment and determine that the result from 3D U-Net is not universally optimal. Depending on the user's goal to accurately capture the general shape or precise surface texture, random forest or percolation may be the preferred method of choice. In the crack segmentation example, arguments can be made for both sides. Detecting the general areas where cracks are formed is essential for analyzing the structural integrity and mechanical failure of concrete. However, accurate surface texture is required to understand, model, and predict crack formation.

4.3.4 Generalization to Multi-Class Segmentation

We have presented the quality metric SCC in the context of binary segmentation, i.e., considering only two classes, reliant on the distances of incorrectly labeled pixels to the structures surface. This concept can be extended to multi-class segmentation problems as follows. Let $S_{\mathrm{gt}} = \{S^0_{\mathrm{gt}}, \dots, S^m_{\mathrm{gt}}\}$ and $S_{\mathrm{pr}} = \{S^0_{\mathrm{pr}}, \dots, S^m_{\mathrm{pr}}\}$ be the respective partitions of the ground truth and prediction into $m \geq 2$ classes and denote

$$d^i_{\partial \mathrm{gt}}(x) = \begin{cases} \min_{y \in S^i_{\mathrm{gt}}} \|x - y\|_2 & \text{if } x \notin S^i_{\mathrm{gt}} \\ \min_{y \notin S^i_{\mathrm{gt}}} \|x - y\|_2 & \text{if } x \in S^i_{\mathrm{gt}}, \end{cases}$$

as the distances of a pixel to the surface of the structure labeled by class i. Furthermore let $E_i = S^i_{\mathrm{pr}} \backslash S^i_{\mathrm{gt}}$ be the set of incorrectly labeled pixels of class i. Then, the multi-class SSC with m classes is given by

$$\mathrm{SCC}^{a,k}_m(S_{\mathrm{pr}}, S_{\mathrm{gt}}) \coloneqq \frac{1}{\sum_{i=1}^m |E_i|} \sum_{i=1}^m \sum_{x \in E_i} f^{a,k}_{\log}(d^i_{\partial \mathrm{gt}}(x)) \in [0, 1].$$

4.4 Conclusion

In this chapter, we have proposed, analyzed, and validated a novel metric, the surface consistency coefficient (SCC) for assessing image segmentation quality based on geometric relations. Compared to established distance-based metrics, SCC is normalized and yields consistent values that summarize the distribution of errors by their distance from the surface. This makes it easy to interpret and compare on different methods and datasets. Additionally, SCC remains independent of the error quantity, allowing it to be combined with traditional metrics such as accuracy or error rate to quantify the number of errors and their geometric distribution independently. This reduces redundant information found in other distance-based quality metrics enabling a more precise assessment of quality and comparison of segmentations. The validity and interpretability of SCC for segmentation tasks was demonstrated using a synthetic data set consisting of segmentations with highly controlled errors and diverse geometries. By incorporating SCC into a real study of segmentation results we showcased its utility in refining the assessment of segmentation quality.

5 Similarity Measures

5.1 Introduction

The morphological analysis of spatial structures is of interest in various fields of science. From granular assemblies, porous materials, molecular compositions, biological cells to synthetic geometries, the characterization of observed structures using simple and interpretable quantities is vital for understanding and comparing them. They can assist in predicting physical properties of materials [102, 117], classifying the physiological state of cells [53] or improving the training data for AI segmentation tasks [29]. Therefore, it is essential to have well-interpretable measures that apply to a wide variety of geometries and which summarize aspects of interest, e.g., anisotropy.
In the literature, spatial structures are usually assumed to be decomposable into convex and compact grains which are then individually examined, statistically characterized and compared [117]. This typically relies on measurements of the intrinsic volumes, Minkowski tensors or other geometric features such as the mean chord length. In other instances, realizations of Boolean models with specific shapes such as spheres, ellipsoids or polyhedra are considered and estimations for the particles' characteristic parameters, e.g., the sphere's radius, are analytically provided based on Miles' formulae [37, 90]. In [33] general random sets in 2D are investigated and statistically tested for their similarity based on an approximation by disks or their structural skeleton and suitable testing functions. While these approaches provide analytical tools and measurements that summarize essential morphological properties, they remain limited to restricted scenarios. In particular, they cannot be used to characterize and compare granular and porous structures in general.
In this chapter, we conduct an exploratory study of diverse synthetic geometries with the goal of characterizing the morphology of 3D spatial structures in terms of their anisotropy, scale and angularity based on the measurements of common morphological features. Our considered data is derived from stochastic models whose parameters intuitively translate to the studied concepts. The combination of model parameters results in morphologically diverse structures that differ in volume density, size of constituents, directionality and surface texture. Combining the calculation of common morphological features such as the intrinsic volume densities with classical machine learning approaches for dimensionality reduction we obtain a series of measures for anisotropy, scale and angularity based on linear combinations of the considered features. We exemplify and discuss their applicability for the modeling of more general spatial structures on real geometries.

5.2 Methods

5.2.1 Dimensionality Reduction Methods

Dimensionality reduction or manifold learning is a major component of modern machine learning. It captures the idea of reducing complex, high-dimensional data into a more understandable, lower-dimensional format. The concept originates from the assumption that

data points $x_1, \dots, x_n \in \mathbb{R}^d$ are captured via a series of d features which hold redundant information. In particular, we assert that the points can be found on an intrinsic lower dimensional structure.
The application of dimensionality reduction is versatile. It ranges from exploratory tasks such as finding suitable descriptive features [125] or creating interpretable visualizations [8] towards data simplification for subsequent training of supervised methods [49]. The book by Ghojogh and colleagues [51] presents a detailed overview over methods which have emerged over the years under a unified notation and discuss their history, methodology and applications. In this section, we present the most commonly used methods.
Consider the data matrix $X = (x_1, \dots, x_n) = (x_{i,j})_{i=1,\dots,d,j=1,\dots,n} \in \mathbb{R}^{d\times n}$ with $n \gg d$ where the entries $x_{i,j}$ refer to the observation of the i-th feature of the j-th data point. For simplicity, we assume that the data is centered, i.e., that the sample mean $\mu_X = \frac{1}{n}\sum_{j=1}^{n} x_j = 0$. By $\hat{X} = (\hat{x}_1, \dots, \hat{x}_n) = (\hat{x}_{i,j})_{i=1,\dots,p,j=1,\dots,n} \in \mathbb{R}^{p\times n}$ with $p \leq n$ we write the projected data onto a space with dimension $p \leq n$. For supervised methods we also consider the label matrix

$$Y = (y_1, \dots, y_n) = (y_{i,j})_{i=1,\dots,\ell,j=1,\dots,n}$$

where $y_{i,j}$ is the i-th category of the j-th data point. Labels can be numerical $(1, 2, \dots)$ or categorical ('class 1','class 2',...). Let

$$\|X\|_{\text{Fro}} = \sum_{i=1}^{d}\sum_{j=1}^{n} |x_{i,j}|^2$$

denote the Frobenius norm.

5.2.1.1 Principal Component Analysis

The principal component analysis (PCA) [64] is one of the oldest and most well-known techniques for dimensionality reduction. It is a linear, unsupervised method that dates back over 100 years [93] with the goal of preserving most of the data's variability. Several modifications to the original version were proposed which allow learning non-linear manifolds via the kernel trick [115] as kernel PCA or becoming supervised by including class labels [6, 8]. The PCA is a special case of the later-proposed multidimensional scaling [124] using Euclidean distances.
The projected data obtained by the PCA can be expressed as the linear transformation $\hat{X} = U^T X \in \mathbb{R}^{p\times n}$ where $U = (u_1, \dots, u_p) \in \mathbb{R}^{d\times p}$ consists of linearly independent unit vectors which span the projected p-dimensional space, i.e., U is an orthogonal matrix satisfying $U^T U = I$. The PCA can be derived from the optimization problem

$$\min_{\substack{U\in\mathbb{R}^{d\times p}\\ U^TU=I}} \sum_{j=1}^{n} \left\|x_j - UU^T x_j\right\|_2^2 = \min_{\substack{U\in\mathbb{R}^{d\times p}\\ U^TU=I}} \left\|X - U\hat{X}\right\|_{\text{Fro}}^2. \tag{5.1}$$

It aims to find the best reconstruction of the data points from a linear subspace spanned by the columns of U. This can be equivalently interpreted as maximizing the variability

of the reconstructed data., i.e.

$$\max_{\substack{U\in\mathbb{R}^{d\times p}\\ U^TU=I}} \left\|U\hat{X}\right\|_{\text{Fro}}^2. \tag{5.2}$$

This minimization problem is solved using the eigenvectors $u_1,\dots,u_p$ of the p largest eigenvalues $\lambda_1 \geq \dots \geq \lambda_p$ of the covariance matrix $S = XX^T$. We call $u_1,\dots,u_p$ the *principal components* or *directions* of X.

5.2.1.2 Linear Discriminant Analysis

The linear discriminant analysis (LDA) [140] is also among the oldest techniques for dimensionality reduction and was initially proposed by Fisher in 1919 [45]. It is a linear, supervised method based on statistics with the goal of finding a linear subspace in which the projected data has minimal spread within each class and maximum spread between classes. The LDA can be extended to project onto non-linear manifolds via the kernel trick [87] known as kernel LDA.
Analogous to PCA, the projected data obtained by the LDA is a linear transformation $\hat{X} = U^TX \in \mathbb{R}^{p\times n}$ where $U = (u_1,\dots,u_p) \in \mathbb{R}^{d\times p}$ consists of (unit) vectors which span the projected p-dimensional space. We consider the case of a single categorical variable, i.e. $\ell = 1$, with k different classes. Let $\{x_j^{(i)}\}_{j=1}^{n_i}$ be the set of data points which belong to the classes $i = 1,\dots,k$ with corresponding class means

$$\mu_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_j^{(i)}.$$

Then, the total, between and within class covariance matrices are defined by

$$S_T \coloneqq \frac{1}{n}\sum_{j=1}^{n} x_j x_j^T$$

$$S_B \coloneqq \frac{1}{n}\sum_{i=1}^{k} n_i \mu_i \mu_i^T$$

$$S_W \coloneqq \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_j^{(i)} - \mu_i)$$

and fulfill the decomposition equation $S_T = S_B + S_W$. The LDA can be obtained by maximizing the *Fisher criterion*

$$\max_{U\in\mathbb{R}^{d\times p}} \frac{\left\|U^TS_WU\right\|_{\text{Fro}}^2}{\|U^TS_BU\|_{\text{Fro}}^2}. \tag{5.3}$$

The solutions to this problem are given by eigenvectors $u_1,\dots,u_p$ of the p largest eigenvalues $\lambda_1 \geq \dots \geq \lambda_p$ solving the generalized eigenvalue problem $S_Bu = \lambda S_Wu$ [51]. We refer to $u_1,\dots,u_p$ as *Fisher directions* of X.

5.2.1.3 t-distributed Stochastic Neighbor Embedding (t-SNE)

The t-distributed stochastic neighbor embedding (t-SNE) [77] is an unsupervised statistical dimensionality reduction technique based on the stochastic neighbor embedding (SNE) proposed by Hinton [55]. This recently proposed method has become one of the most popular choices for visualizing and exploring high-dimensional data.
The t-SNE performs a non-linear projection to a low-dimensional space with the intent to preserve similarities and dissimilarities between points. To achieve this, we model the conditional probability that a data point x_i considers x_j as similar or its neighbor with a Gaussian distribution

$$p_{j|i} \coloneqq \frac{\exp\left(-\frac{\|x_i - x_j\|_2^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(-\frac{\|x_k - x_i\|_2^2}{2\sigma_i^2}\right)}, \quad i \neq j$$

where $\sigma_i > 0$ controls how quickly this probability falls off with increasing distance. Since we are only interested in modeling pairwise similarities, we set $p_{i|i} = 0$. The parameters σ_i are determined by the equation

$$\text{Perp} = 2^{-\sum_{j \neq i} p_{j|i} \log(p_{j|i})}$$

with *perplexity* parameter Perp that is based on the Shannon entropy and is specified by the user. It can be interpreted as a smoothness measure of the effective number of neighbors ensuring that the probabilities adjust to the density of data points. We assume that the joint probabilities that two points are neighbors to each other can be modeled as $p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}$. In contrast, we model the probability that the projected data points $\hat{x}_i$ and $\hat{x}_j$ are neighbors in the lower-dimensional space with Student t-distributions with one degree of freedom. This prevents the so-called *crowding problem* and makes the embedding more resilient to changes in scale compared to using Gaussian distributions, see [77] for details. The corresponding joint probabilities are given by

$$q_{ij} = \frac{\left(1 + \|\hat{x}_i - \hat{x}_j\|_2^2\right)^{-1}}{\sum_{k \neq l} \left(1 + \|\hat{x}_k - \hat{x}_l\|_2^2\right)^{-1}}.$$

Finally, the t-SNE embedding is defined as a solution to the minimization problem

$$\min_{\substack{\hat{x}_k \in \mathbb{R}^p \\ k=1,\dots,n}} \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) \tag{5.4}$$

that can be interpreted as a symmetric version of the Kullback-Leibler divergence and optimized via an iterative gradient descent algorithm, see [77, Algorithm 1]. Accelerated optimization algorithms are discussed in [67]. A heuristic scaling of the probabilities p_{ij} by a constant factor, e.g. $c = 4$, in the initial stages of the optimization can improve the embedding since it encourages to focus on the modeling of similar points with large p_{ij} by large q_{ij} before the modeling between clusters of similar points. Since t-SNE determines similarity and dissimilarity of data points only through distances between points, it is also possible to calculate an embedding based on different distance functions such as the cosine

similarity instead of the proposed Euclidean distance. While t-SNE is an unsupervised dimensionality method, modifications to the cost function (5.4) can be made to create a supervised version [21].

5.2.1.4 Uniform Manifold Approximation and Projection (UMAP)

The uniform manifold approximation and projection (UMAP) [54] is an unsupervised dimensionality reduction technique that is commonly used for data visualization. The method was first proposed in 2018 by McInnes et al. [83] and follows similar ideas to t-SNE.
UMAP computes an embedding into a lower-dimensional space by preserving the general topological structure of the data. In a first step, a k-nearest neighbors graph with $k \ll n$ is constructed using the Euclidean distance. We denote the set of neighboring points of x_i by $\mathcal{N}_i = \{x_{i,1}, \dots, x_{i,k}\}$ and model the conditional probability that the data point x_j considers x_i as similar with

$$p_{j|i} \coloneqq \begin{cases} \exp\left(-\frac{\|x_i - x_j\|_2 - \rho_i}{\sigma_i}\right) & \text{if } x_j \in \mathcal{N}_i \\ 0 & \text{otherwise} \end{cases}$$

where $\rho_i = \min_{j=1,\dots,k} \|x_i - x_{i,j}\|_2$. $\sigma_i > 0$ is a scale parameter which is calculated such that the total similarity of point x_i to its k nearest neighbors is normalized, i.e., it satisfies the equation

$$\sum_{j=1}^{k} \exp\left(-\frac{\|x_i - x_j\|_2 - \rho_i}{\sigma_i}\right) = \log(k).$$

This normalization adjusts for the density of points so that different neighborhoods are treated similarly in subsequent calculations. To obtain a symmetric measure of similarity we define

$$p_{ij} \coloneqq p_{j|i} + p_{i|j} - p_{j|i}p_{i|j}.$$

In the lower dimensional space, we model the probability that the projected data points $\hat{x}_i$ and $\hat{x}_j$ are similar by

$$q_{ij} \coloneqq \frac{\left(1 + a\,\|\hat{x}_i - \hat{x}_j\|_2^{2b}\right)^{-1}}{\sum_{k \neq l}\left(1 + a\,\|\hat{x}_k - \hat{x}_l\|_2^{2b}\right)^{-1}}$$

where the variables $a, b > 0$ are user-specified hyperparameters. We use the probabilities to establish the minimization problem

$$\min_{\substack{\hat{x}_k \in \mathbb{R}^p \\ k=1,\dots,n}} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(p_{ij} \ln\left(\frac{p_{ij}}{q_{ij}}\right) + (1 - p_{ij}) \ln\left(\frac{1 - p_{ij}}{1 - q_{ij}}\right)\right) \tag{5.5}$$

which can be interpreted as a fuzzy cross-entropy of the nearest neighbors graphs [51]. The first term in (5.5) describes an attractive force towards nearest neighbors while the second term encodes repulsion between non-neighboring points. Problem (5.5) can be solved via a stochastic gradient descent approach [51, Algorithm 17.1]. Analogous to t-SNE, similarity

between data points is determined only by their Euclidean distance which can be replaced by a different distance function such as the cosine similarity. The presented UMAP is an unsupervised method but can be transformed into a supervised method by adding a classification loss to (5.5), e.g., cross-entropy.

Figure 5.1: Illustration of the considered model parameters and their options.

5.2.2 Synthetic Data with Morphological Variability

For the assessment of morphological properties of 3D spatial structures we consider a large dataset containing 6480 different geometries which are generated from Boolean models and rendered in a 512^3 pixel volume. By combining different volume densities, particle shapes, particle sizes and orientations, see Figures 5.1, we obtain diverse geometries as seen in Figures 5.2 and 5.3, whose model parameters intuitively translate to different morphological aspects.
In terms of volume densities, we utilize three values $V_{V,3}(\Xi) \in \{0.3, 0.5, 0.7\}$ to obtain porous, dense and intermediate structures. Each model uses a single type of particle shape which is either spheres, ellipsoids, cylinders, cuboids or cubes. Spheres and cubes can be considered as special cases of ellipsoids and cuboids, respectively. For each shape, various versions with different ratios of their main axes are used. More precisely, cylinders with height-to-diameter ratios of 1:1, 2:1, 5:1, 10:1, 1:2 and 1:5 are considered. Cuboids and ellipsoids are generated for ratios of the main axes 1:1:1, 2:1:1, 5:1:1, 10:1:1, 3:2:1 and 4:2:1. Most notably, we include shapes which are rounded or angular and elongated, isotropic or oblate. The variation in particles is further extended by considering three different sizes obtained by scaling them to match a reference particle volume of $\bar{V} = \frac{4}{3}\pi r^3$ pixels with $r \in \{10, 20, 30\}$, i.e., the reference particles are spheres with radius 10, 20 and 30 pixels. The shapes are combined using one of four different methods to orient them in 3D space. All rotations are applied intrinsically, i.e., each particle is rotated around its own position:

- **Fixed:** All particles are placed without rotations.
- **Unimodal:** For each particle a direction is sampled from the von Mises-Fisher distribution with mean direction $\mu = [0, 0, 1]$ and intensity parameter $\kappa = 10$ to

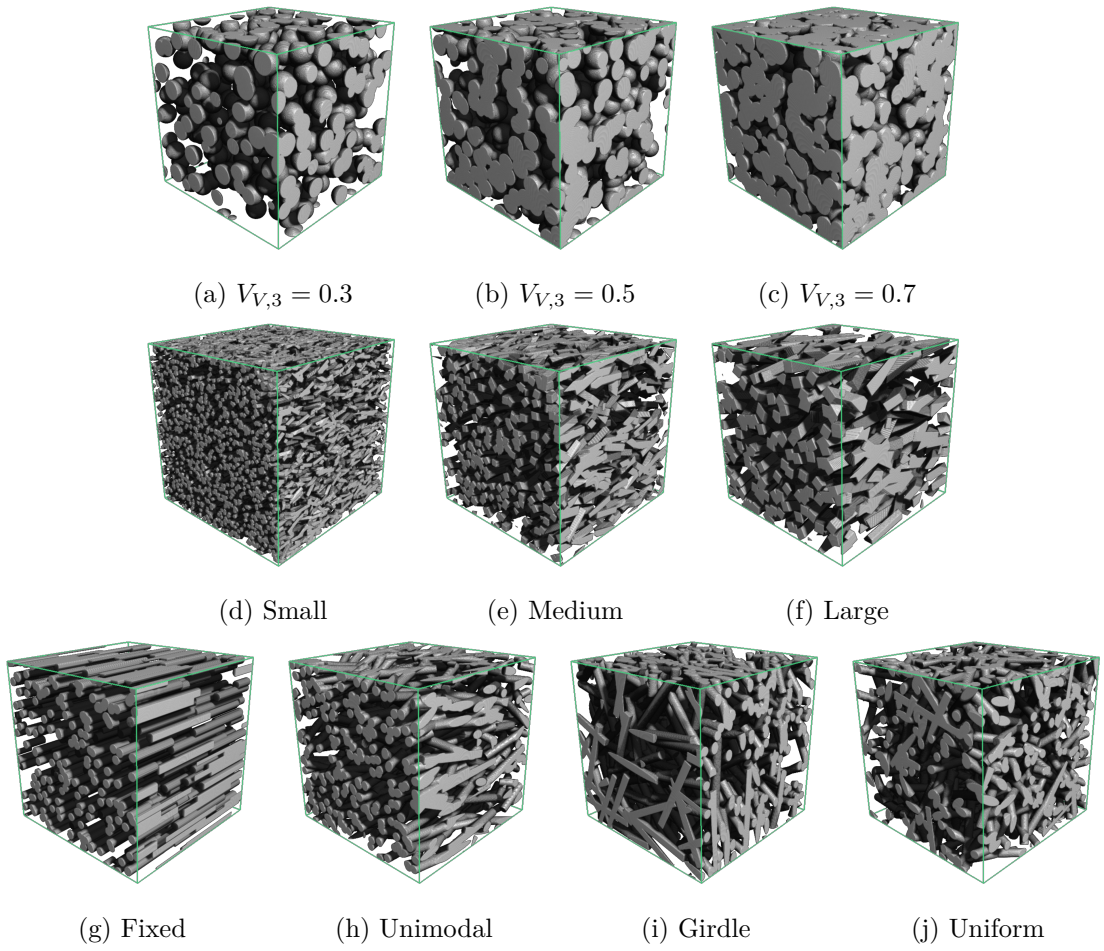
which the main axis of the particle is rotated. Then, each particle is intrinsically rotated around its direction by an angle which is uniformly sampled in $[0, 2\pi)$.

- **Girdle:** For each particle a direction is sampled from the distribution

$$p(\varphi, \theta) = \frac{1}{4\pi} \frac{\beta \sin(\varphi)}{(1 + (\beta^2 - 1)\cos^2(\varphi))^{\frac{3}{2}}}, \quad \varphi \in [0, 2\pi), \theta \in [0, \pi)$$

in polar coordinates, see [112], with parameter $\beta = 20$ to which the main axis of the particle is rotated. The distribution is a special case of an angular central Gaussian distribution [139]. Then, each particle is intrinsically rotated around its direction by an angle which is uniformly sampled in $[0, 2\pi)$.

- **Uniform:** A rotation matrix is sampled uniformly from SO_3 for each particle.



(a) $V_{V,3} = 0.3$ (b) $V_{V,3} = 0.5$ (c) $V_{V,3} = 0.7$

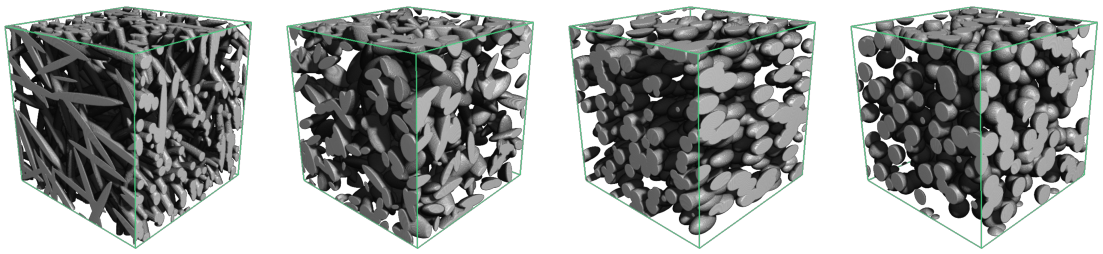
(d) Small (e) Medium (f) Large

(g) Fixed (h) Unimodal (i) Girdle (j) Uniform

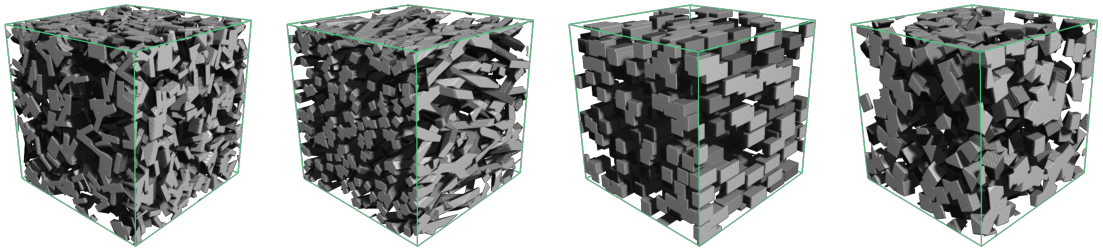
Figure 5.2: Renderings of selected 3D structures in the synthetic data showcasing variations in volume density (top), particle sizes (middle) and orientation distributions (bottom).

For elongated particles, these options can be interpreted as orienting particles in the same direction, around a preferred direction, in a preferential plane or arbitrarily. This directional sampling does not straightforwardly translate to the directionality or anisotropy of

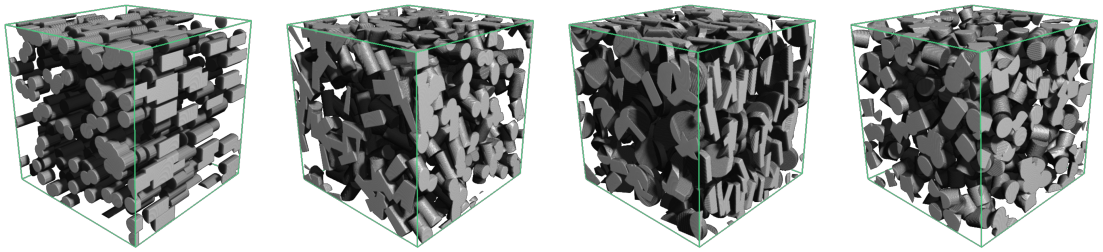
the entire structure since the interpretation varies with the specific shape. For example, oblate particles which have two large main axes behave differently from elongated particles with only one large main axis. Additionally, spheres remain entirely unaffected by rotations. Lastly, we account for the natural variability in stochastic models by generating 10 realizations for each combination of the established model parameters. A summary of all utilized options is shown in Figure 5.1.



(a) Ellipsoids



(b) Cuboids



(c) Cylinders

Figure 5.3: Renderings of selected 3D structures in the synthetic data showcasing variations in shapes and axial ratios for different orientations with $V_{V,3} = 0.3$ and large particles for best visibility.

5.2.3 Extraction of Geometric Features

Large image data, such as 3D geometries, contain vast amounts of redundant and high-dimensional information, most of which is not directly relevant to the task at hand. By extracting features, we can reduce the complex high-dimensional data to a compact and interpretable representation of scalar measurements that summarize essential morphological and geometric characteristics. In the following, we provide a brief overview of the selected characteristics, see Table 5.1, most of which are presented in Section 2.2.
We consider the intrinsic volume densities $\hat{V}_{V,k}$, $k = 0, 1, 2, 3$ since they contain fundamental information about the volume, surface area, curvature and connectivity of a geometry.

We incorporate directional information using the area of total projections $p_{V,1}$, lengths of total projections $p_{V,2}$ and their related directional distributions, the rose of normal directions ρ_1 and rose of intersections ρ_2, see [90]. $p_{V,1}$ and $p_{V,2}$, see Section 2.2.4, are considered as distribution over the corresponding linear subspaces. We extract only the standard deviations of these distributions since their means do not provide additional information beyond the intrinsic volume densities. Additionally, we calculate the box-counting dimensions $D(p_{V,1})$ and $D(p_{V,2})$ [121] of the projection distributions which summarize their local complexity. They are typically found in the context of fractals and in practice, indicate the influence of digitization on the estimation of the respective distributions.
We also incorporate the distribution of mean chord lengths ℓ over all directions via their mean and standard deviation since they intuitively summarize the size and variations of local constituents. These more common features are supplemented by the Minkowski tensors $\Phi_0^{2,0}$, $\Phi_1^{2,0}$ and $\Phi_1^{0,2}$ which can be interpreted as volume moment tensor, surface moment tensor and surface normals moment tensor. In particular, they encode information about the distribution of volume, surface and surface normals. Their interpretation is not straightforward. We condense the complexity of the tensors by utilizing only the ratios of their eigenvalues

$$\alpha_k^{n,m} = \frac{\lambda_3}{\lambda_1}, \quad \beta_k^{n,m} = \frac{\lambda_2}{\lambda_1}, \quad \gamma_k^{n,m} = \frac{\lambda_3}{\lambda_2},$$

where $|\lambda_3| \leq |\lambda_2| \leq |\lambda_1|$ are the eigenvalues of the tensor $\Phi_k^{n,m}$. This captures the essence of the tensors in dimensionless values [118]. The presented features and their estimation for binary images are covered in detail in Section 2.2.7 and 2.2.8, and [90, 91]. In total, we obtain a set of 21 distinct features per geometry. The majority of feature calculations were performed using the ToolIP software from Fraunhofer ITWM. The Minkowski tensors were calculated in MATLAB.

Symbol	**Feature**
$V_{V,3}$	Volume density
$V_{V,2}$	Surface density
$V_{V,1}$	Integral of mean curvature density
$V_{V,0}$	Euler density
$\sigma(p_{V,2})$	Standard deviation of the area of total projections distribution
$D(p_{V,2})$	Box-counting dimensions of the area of total projections
$\sigma(\rho_{V,2})$	Standard deviation of the rose of surface normals
$\sigma(p_{V,1})$	Standard deviation of the lengths of total projections distribution
$D(p_{V,1})$	Box-counting dimensions of the lengths of total projections
$\sigma(\rho_{V,1})$	Standard deviation of the rose of intersections
$\bar{\ell}$	Mean of the mean chord length distribution
$\sigma(\ell)$	Standard deviation of the mean chord length distribution
$\alpha_0^{2,0}$, $\beta_0^{2,0}$, $\gamma_0^{2,0}$	Ratios of eigenvalues of $\Phi_0^{2,0}$
$\alpha_1^{2,0}$, $\beta_1^{2,0}$, $\gamma_1^{2,0}$	Ratios of eigenvalues of $\Phi_1^{2,0}$
$\alpha_1^{0,2}$, $\beta_1^{0,2}$, $\gamma_1^{0,2}$	Ratios of eigenvalues of $\Phi_1^{0,2}$

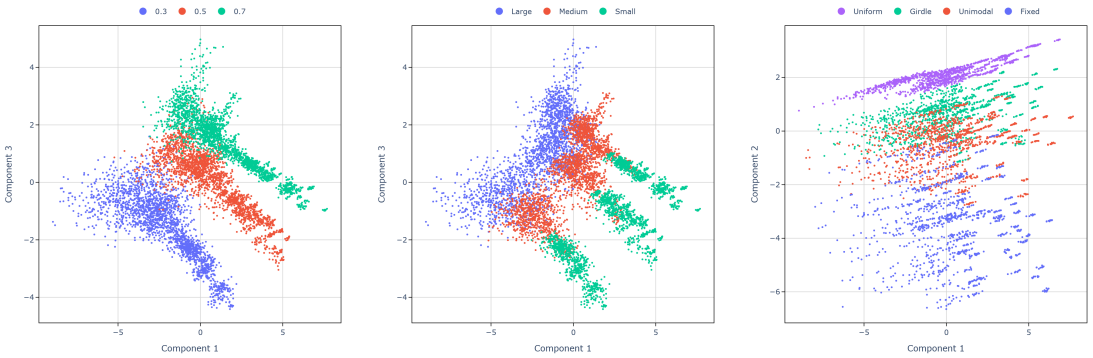
Table 5.1: List of selected morphological features.

5.3 Results and Discussion

5.3.1 Data Exploration

Before we establish quantitative measures that summarize the level of anisotropy, scale and angularity of spatial structures, we first examine the available feature data using dimensionality reduction techniques introduced in Section 5.2.1. The central goal of this analysis is to identify naturally occurring clusterings within the data in a low-dimensional space connected to anisotropy, scale and angularity using the model parameters used to generate the generation of the synthetic data. The linear methods PCA and LDA are of particular interest because they provide embeddings that are easy to interpret, apply to new data and enable the derivation of suitable linear measures from the features. Especially LDA, which is a supervised method that yields an embedding with respect to given labels, allows for the exploration of specific properties. In contrast, the non-linear t-SNE and UMAP are the gold standard for data visualization since they aim to preserve relations between data points found in higher dimensions. However, they do not allow for the inclusion of new data or deduction of relations to the specific features. Nonetheless, their non-linearity can reveal relations between data points that linear methods cannot capture. Thus, they are particularly valuable for addressing the core question of whether a differentiation of the geometric aspects is in principle possible. The calculations of all methods were performed with the implementations provided in the Python packages *scikit-learn* and *umap-learn*. For t-SNE we set the perplexity parameter Perp = 30 and for UMAP we used the number of neighbors $k = 30$ because they yielded the visually best results. The dataset was standardized, i.e., each feature was transformed to zero mean and unit variance. This is crucial since features exist on different scales which otherwise skews them and complicates their comparison. We note that all displayed dimensionality reduction results are shown only as 2D projections in this work. Hence, some nuances found in 3D cannot be fully presented. Additionally, we omit plotting combinations of embedded components and classes that are not significant.

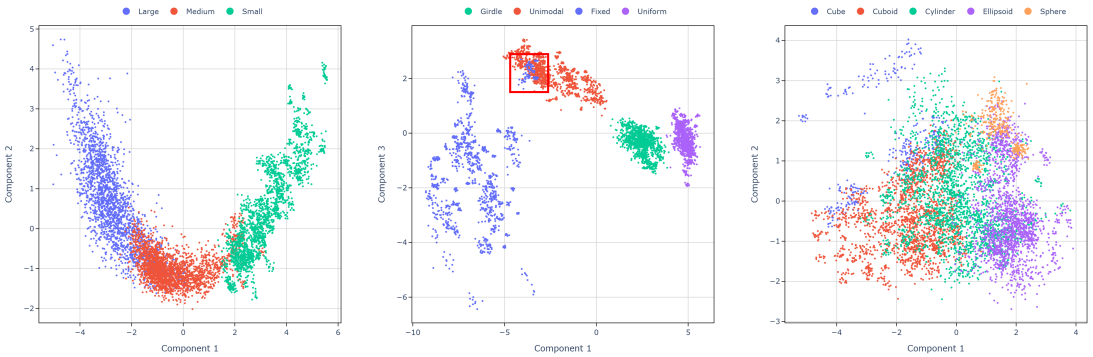
Applying PCA to the feature data, see Figure 5.4, reveals that model parameters such as volume density, orientation and particle size are prominent in the first three components which explain roughly 66% of the variability. This is not particularly surprising since the parameters are directly involved in our generation of geometries. However, it demonstrates two critical points. First, it confirms that the chosen set of features can adequately capture information about directionality and scale of the investigated geometries. In particular, the inclusion of additional morphological features is not necessary. Second, the results indicate that these aspects can be reasonably separated using simple linear combinations of the available features, e.g., by projecting onto combinations of the first three principal components. We remark that considering more principal components or performing PCA on subsets of the available data, e.g., filtered by volume density did not yield further findings. Hence, aspects such as the particle shapes either contribute only weakly to data variability or cannot be distinguished based on the provided features.



(a) Volume density (b) Particle size (c) Orientation

Figure 5.4: Selected dimensionality reduction results of PCA on the dataset colored by different variables. In (c) only geometries with elongated particles are displayed.

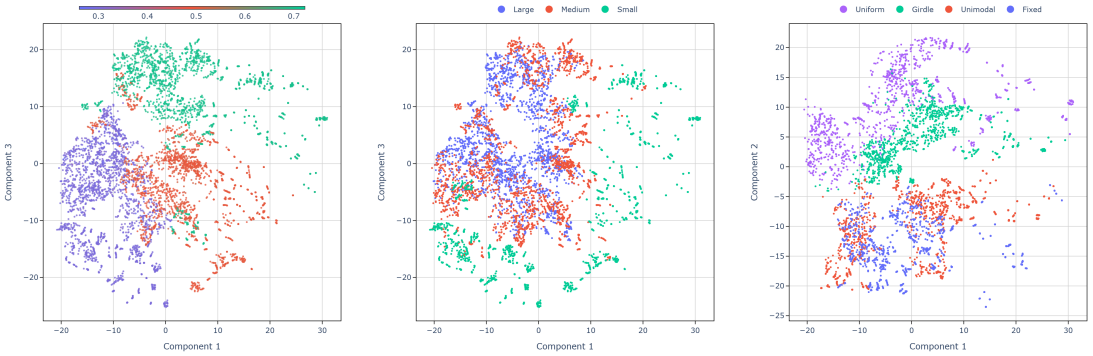
The visualizations of the LDA embeddings calculated with respect to the corresponding model parameters, see Figure 5.5, confirm PCA observations that particle size and orientation can be distinguished based on a linear combination of the features. In particular, the first Fisher direction already provides a sufficient separation. Additionally, they reveal that distinguishing between different particle shapes is indeed possible. More precisely, we observe in (c) that the first component holds information about the particles' angularity separating the angular cubes and cuboids from the rounded particles, ellipsoids and spheres. Cylinders that carry attributes of both extremes can appropriately be found in their center. The second component indicates a separation of particles based on their eccentricities. At least, spheres and cubes are mapped above the remaining ellipsoids and cuboids. The same can be checked for cylinders with a height-to-diameter ratio of 1:1. However, a closer examination of the second component has shown that distinguishing other particle formats was not possible. Nevertheless, we deduce that the available model parameters which are linked to anisotropy, scale and angularity can be distinguished using linear combinations of the calculated features.



(a) Particle size (b) Orientation (c) Particle shape

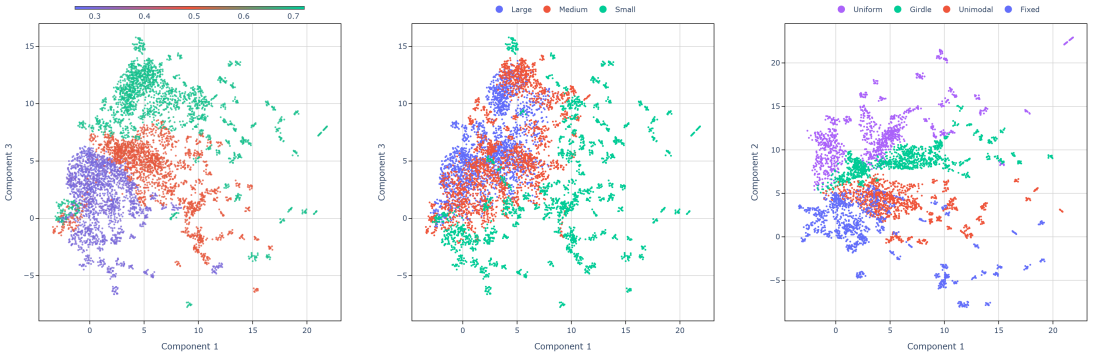
Figure 5.5: Selected dimensionality reduction results of LDA on the dataset calculated and colored with respect to different variables. In (b) LDA is calculated only on geometries with elongated particles. The red box marks a set of particles with the smallest considered axial ratios 2:1 and 2:1:1.

The embeddings calculated by t-SNE and UMAP are displayed in Figures 5.6 and 5.7. While they do not provide any additional clustering of data points beyond our findings from linear dimensionality reduction, they nevertheless substantiate our observations and conclusions about the possibility of measuring anisotropy and scale based on the available features. Surprisingly, these non-linear methods fail to separate the structures based on their particle shapes even when using the corresponding LDA embedding as an initial state in their calculations. This highlights once more that assessing surface texture from spatial structures based on our selected features is difficult as this influence remains low compared to the other factors.



(a) Estimated volume density (b) Particle size (c) Orientation

Figure 5.6: Selected dimensionality reduction results of t-SNE with Perp = 30 on the dataset colored by different variables. In (c) only geometries with elongated particles are displayed.



(a) Estimated volume density (b) Particle size (c) Orientation

Figure 5.7: Selected dimensionality reduction results of UMAP with $k = 30$ on the dataset colored by different variables. In (c) only geometries with elongated particles are displayed.

5.3.2 Development of Morphological Measures

The results presented in Section 5.3.1 reveal that distinguishing between spatial structures in terms of anisotropy, scale and angularity using linear combinations of the calculated features is feasible. To derive appropriate feature combinations that capture the studied

aspect and do not overfit to our data, we employ a systematic greedy feature reduction procedure based on LDA. The remaining features can be interpreted as the most expressive and important for assessing the respective aspects. Our proposed measures are the direct result of this procedure. We begin by considering all features.

1. **Labeling:** A highly simplifying label is assigned to each structure based on its generating parameters and their relation to the selected morphological aspect.

2. **LDA:** The LDA with respect to the assigned labels on the current selection of features is calculated and its Fisher criterion (5.3) evaluated.

3. **Selection:** A candidate feature is selected based on a small contributing weight to the first Fisher direction. Step 2 is performed without the selected feature. Results are numerically and visually compared using estimated densities and bar plots. The feature is removed if the results have not significantly worsened.

4. **Iteration:** Step 3 is repeated until no feature can be removed without visually or numerically impeding the separation of the labeled groups.

5. **Measure:** A linear measure is obtained from the resulting Fisher direction by reverting the feature normalization. The feature weights are the entries of the transformed Fisher direction with minor adjustments such as rounding or scaling to improve interpretability.

The simplicity of the labels chosen in step 1 is essential for the success of our approach for three reasons. First, it enables the LDA to extract a combination of features that encode fundamental information about the investigated property based on a very rough guideline. Second, it prevents the introduction of unnecessary biases and mitigates errors introduced by an inaccurate manual labeling. Third, it allows LDA to find and define a natural ordering within each group of labels and between them. The measures derived by this greedy strategy are not unique. Different combinations of features and weights might be similarly suitable.

5.3.2.1 Anisotropy Measure

The formal definition of isotropy for a random closed set is the rotation invariance of its distribution, see Definition 2.19. For spatial structures such as the realizations of random closed sets, the notion rotation invariance can be characterized by the distribution of surface normals. More precisely, we consider a geometry to be isotropic if all directions appear equally. Conversely, anisotropy refers to cases where at least one direction is more pronounced. Based on this idea, we label the structures according to Figure 5.8 using the interpretation of the orientation distributions given in Section 5.2.2. The assigned labels can be interpreted as an expected dimension of the particles' preferred directions which depends on the selected orientation distribution and particle shape. The labels relate almost one-to-one to the orientation distribution with two exceptions. First, if the particles' axial ratios are the same we always assign dimension three since no preferential direction or plane exists. Second, oblate particles that largely extend in two dimensions themselves are assigned dimension two if the orientation distribution supplies a single preferential direction and one if orientations lie in a preferential plane. Afterwards, the previously discussed feature reduction procedure is applied. For the visualization and

interpretation of results we further mark structures sampled with the fixed orientation distribution opposed to those including randomness.

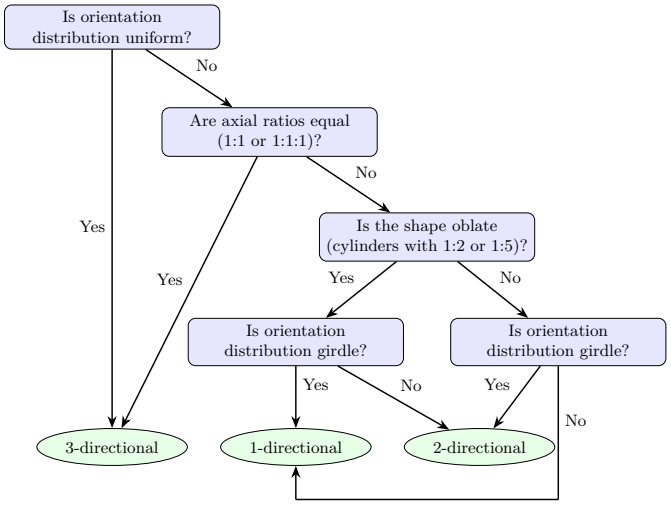
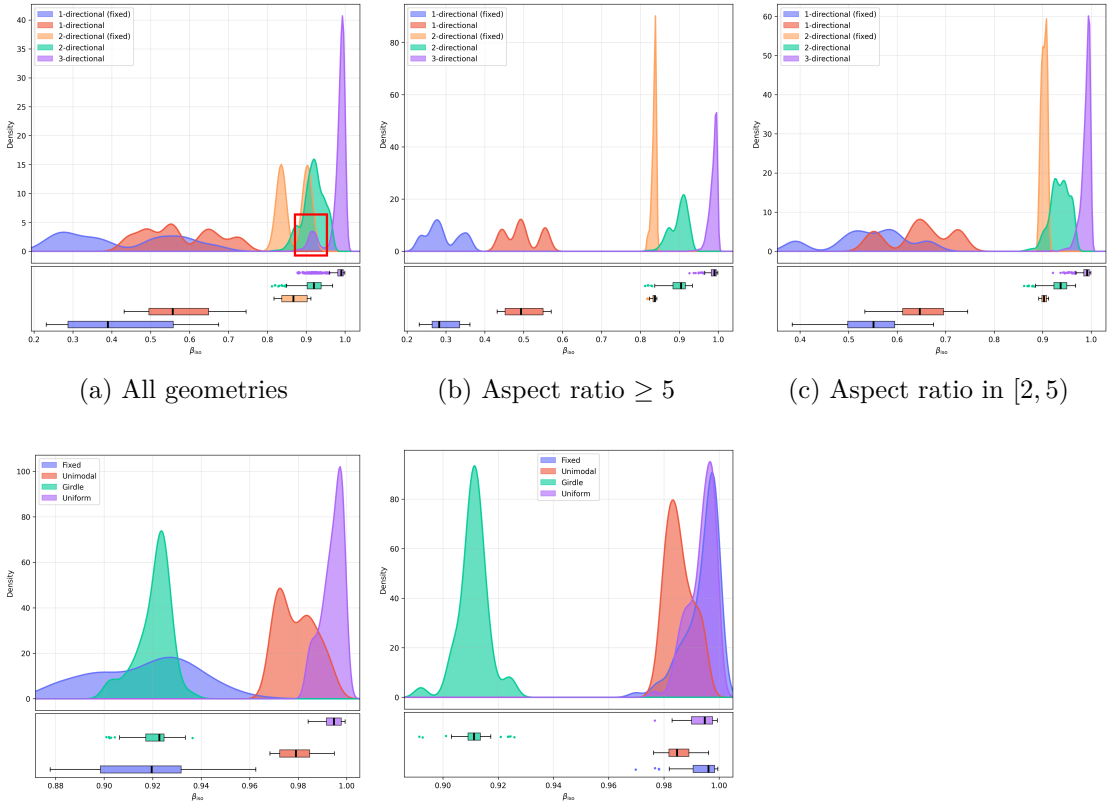


Figure 5.8: Decision tree for assigning simplified labels for isotropy to the geometries based on their generating parameters, see Figure 5.1. An axial ratio with two values implies the particle shape cylinder. Three values are used for cuboids or ellipsoids.

We obtain a measure of anisotropy that is defined by

$$\beta_{\text{iso}} = 0.8\,\beta_1^{0,2} + 0.2\,\gamma_1^{0,2} \in [0, 1].$$

This indicates that the essential information about anisotropy within the structures is captured by the tensor $\Phi_1^{0,2}$. Our measure is conceptually related to the anisotropy measure $\beta_{\text{ST}} = \alpha_1^{0,2}$ which was proposed and validated by Schröder-Turk and colleagues [116, 118] on different spatial structures. In contrast to β_{ST}, our measure β_{iso} combines information from the entire eigenvalue spectrum that encodes the distribution of surface normals rather than considering only its extremes.
The proposed anisotropy measure β_{iso} reliably distinguishes between the assigned labels across the entire dataset, as shown in Figure 5.9. In particular, it accurately captures the varying degrees of anisotropy of the different structures, yielding large values for isotropic and smaller values for increasingly anisotropic structures. Geometries obtained with the fixed orientation distribution consistently yield smaller values than those obtained with randomness. This reflects the reduced orientation variability which corresponds to a stronger anisotropy. When considering geometries which consist of particles with large aspect ratios ≥ 5, that is the ratio of the largest and smallest main axes, we observe an ideal transition from isotropic to highly anisotropic systems according to our assigned labels, see Figure 5.9 (b). A similar trend is visible for particles with smaller aspect ratios in (c). However, values are generally higher indicating that the particles themselves are less anisotropic in comparison. In particular, the measured degree of anisotropy in β_{iso} simultaneously encodes the anisotropy introduced by the orientation distribution and the anisotropy of its particles. The highlighted area in Figure 5.9 (a) is of particular interest. It contains a subset of non-spherical particles with an aspect ratio of 1 which we naively



(a) All geometries (b) Aspect ratio ≥ 5 (c) Aspect ratio in $[2, 5)$

(d) Cylinders of aspect ratio 1 (e) Cubes

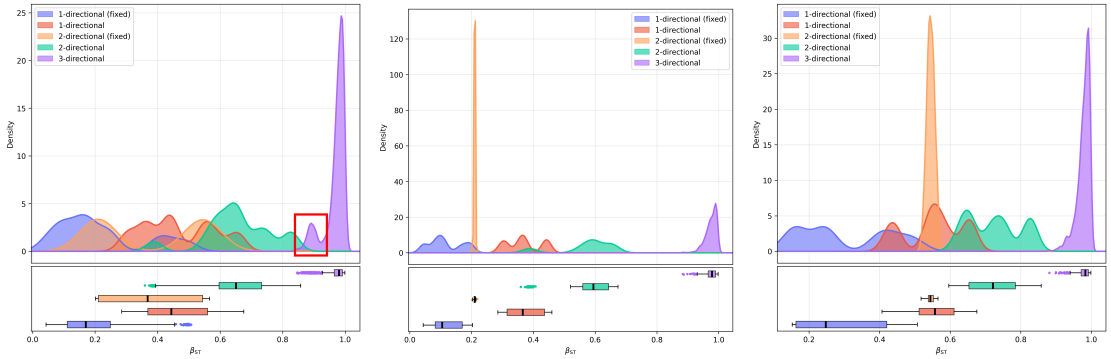
Figure 5.9: Visualization of the anisotropy measure β_{iso} evaluated on different subsets of geometries. The red box in (a) marks the subset of cubes and cylinders with aspect ratio 1 which are displayed in (d) and (e).

labeled as 3-directional. However, Figure 5.9 reveals that not all of these structures are fully isotropic. For cylinders, this is caused by differences in the surface areas of the mantle and bases. For cubes, it is likely caused by their flat faces. Table 5.2 provides a loose interpretation of different values of the anisotropy measure based on our dataset.

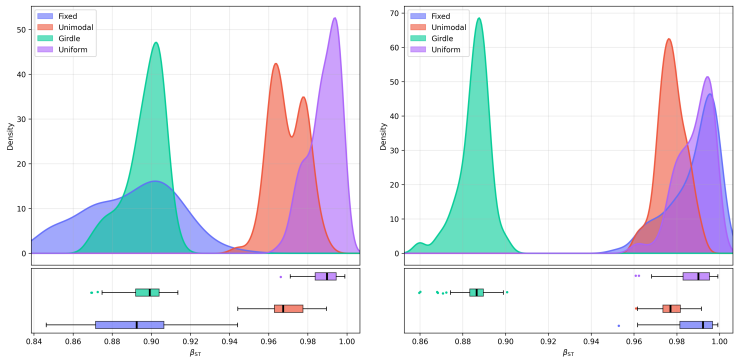
Description	**Extremely anisotropic**	**Strongly anisotropic**	**Moderately anisotropic**	**Weakly anisotropic**	**Isotropic**
Interval	< 0.4	$[0.4, 0.63)$	$[0.63, 0.78)$	$[0.78, 0.96)$	≥ 0.96

Table 5.2: Loose interpretation of the anisotropy measure β_{iso}.

Comparing the derived anisotropy measure β_{iso} with β_{ST}, see Figure 5.10, reveals that while both can differentiate between isotropic and anisotropic structures, our measure provides a clearer and more consistent separation between the assigned labels. Especially for anisotropic structures β_{ST} only weakly relates to our assigned labels and fails to assess the severity of anisotropy. The improved discrimination of β_{iso} compared to β_{ST} arises from including the intermediate eigenvalue λ_2. Since the eigenvalues of the tensor $\Phi_1^{0,2}$ reflect the density of the surface normal distribution along its orthogonal principal axes, β_{iso} can distinguish between the cases of one and two preferred normal directions. In



(a) All geometries (b) Aspect ratio ≥ 5 (c) Aspect ratio in $[2, 5)$



(d) Cylinders of aspect ratio 1 (e) Cubes

Figure 5.10: Visualization of the anisotropy measure β_{ST} evaluated on different subsets of geometries. The red box in (a) marks the subset of cubes and cylinders with aspect ratio 1 which are displayed in (d) and (e).

contrast, β_{ST} cannot as it considers only the strongest and weakest direction.
We argue that the measurement of anisotropy via β_{iso} generalizes to a broad range of porous and spatial structures as it depends entirely on the dimensionless eigenvalue ratios of $\Phi_1^{0,2}$ whose interpretation to quantify the distribution of surface normals remains valid for general spatial structures. Assuming that a structure can be modeled via a Boolean model, β_{iso} can aid in the selection of orientations and the eccentricity of the particles. It becomes most expressive if either of these aspects is known.

5.3.2.2 Scale Measure

In the context of 3D structures, we refer to the scale of a geometry as the characteristic size of geometric features that define its morphology. It describes the spatial extent of structural variations and thus determines the level of geometric detail that is perceptible and analytically relevant. In essence, the scale captures whether morphological variations arise from small or large elements. Related concepts have been studied in the literature, e.g., Ohser et al. [90] and Eremina et al. [37] propose analytical techniques to estimate parameters for the Boolean model in specific scenarios such as spheres or cuboids. However, a general measurement of scale for porous or granular structures has, to our knowledge, not been established in literature.

Since the scale of our geometries strongly depends on the size of their constituent particles, we use this model parameter as the label without further adjustment. Then, feature reduction yields the scale measure

$$\beta_{\mathrm{sc}} = -0.05\, V_{V,3} - 5\, V_{V,2} - 9.25\, V_{V,1} + 0.88\, D(p_{V,2}) + 0.11\, D(p_{V,1}) + 0.008\, \bar{\ell} - 1.28\, \alpha_1^{2,0}.$$

which integrates multiple features that capture distinct aspects of structural scale. We remark that the difference in magnitude of the individual weights is a consequence of the varying scales of the features themselves and does not directly reflect their importance. Furthermore, no theoretical bounds similar to the anisotropy measure can be given. On our data, β_{sc} reliably distinguishes between the broad scale levels represented, see Figure 5.11. Nonetheless, overlap between our assigned labels remains indicating that the separation is not entirely distinct. This can primarily be attributed to the simplicity of our labels which rely solely on particle volumes neglecting any additional shape-related influences on the geometry's scale. In reality, the morphological scale is also influenced by the specific shapes within the structure, that is, thinner structures and finer details correspond to a smaller scale. The results displayed in Figure 5.12 showcase that this aspect is adequately captured by our measure β_{sc}. It yields consistently smaller values across all shapes for larger aspect ratios which relate to thinner structures. At the same time, we observe that the values of cuboids are generally lower than comparable cylinders and ellipsoids, see (a)-(c). This is due to their sharp corners which present finer details and result in a reduced morphological scale. The interpretation of β_{sc} is straightforward. The higher its value, the larger the morphological scale of the structure, i.e., the larger its constituents. The generalization of β_{sc} is not straightforward since it combines multiple morphological features at once. While each individual feature encodes information about the morphological scale, their complex interplay may differ across highly different spatial structures. Hence, it requires further testing beyond the structures included in our study to validate and demonstrate its full potential. However, β_{sc} shows that the assessment of angularity using a linear combination of morphological features is possible. We argue that β_{sc} provides an informed estimation of the local complexity and size of constituents. Especially if spatial structures can be modeled via Boolean models similar to those contained in our study, β_{sc} quantifies particle size. If additional model information about particle shape is provided, this assessment becomes more precise.

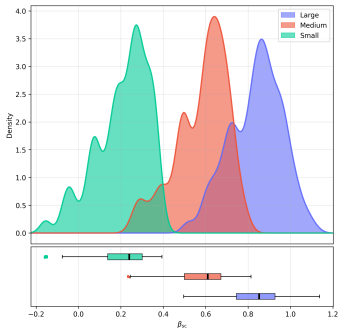
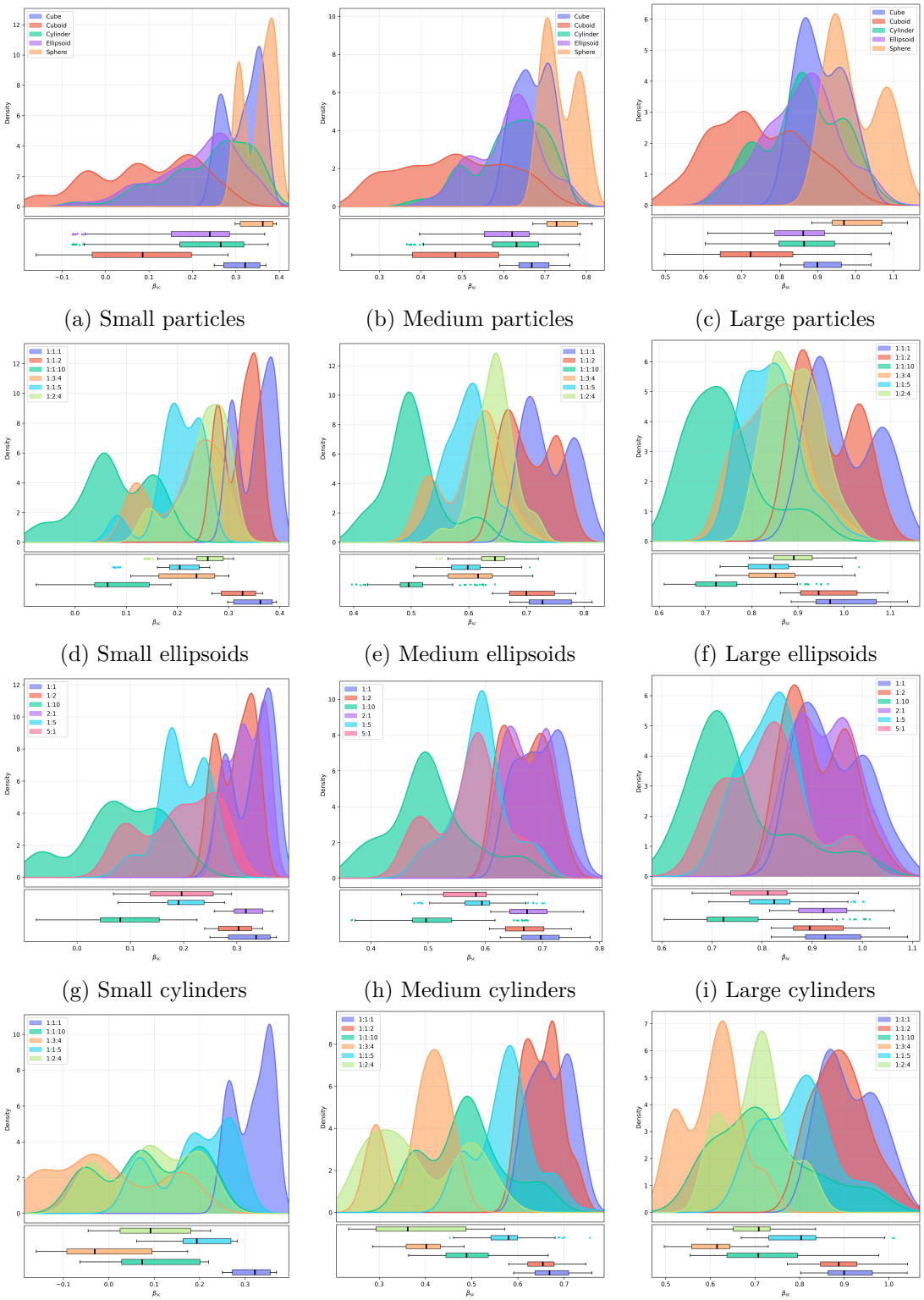


Figure 5.11: Visualization of the scale measure β_{sc} evaluated on all geometries.



(a) Small particles (b) Medium particles (c) Large particles

(d) Small ellipsoids (e) Medium ellipsoids (f) Large ellipsoids

(g) Small cylinders (h) Medium cylinders (i) Large cylinders

(j) Small cuboids (k) Medium cuboids (l) Large cuboids

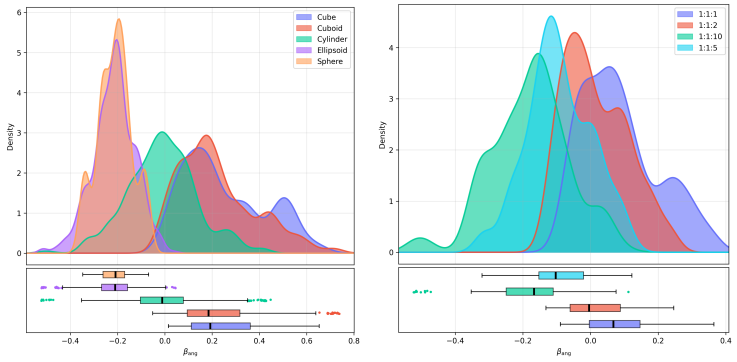
Figure 5.12: Visualization of the scale measure β_{sc} evaluated on the different subsets of geometries.

5.3.2.3 Angularity Measure

The complexity of a structure's surface can be partially characterized by its angularity or smoothness, i.e., the variations in geometrical continuity and curvature across the surface. More precisely, angular surfaces exhibit abrupt changes in their directions of surface normals and curvature compared to smoother surfaces which display more gradual transitions. [53, 136, 141] propose distinct approaches to quantify the shape and surface textures of individual particles. However, they are not applicable to general spatial structures. A general measurement of angularity for porous or granular structures has, to our knowledge, not been established in literature.
The assessment of this property for general spatial structures is difficult as demonstrated in Section 5.3.1 where we found that differentiating between models based on their particle shapes was possible but the descriptive power of this model parameter was low compared to other parameters. We argue that the essence of angularity of a Boolean model is characterized by the shapes of its particles but also depends on more specific factors such as the particles axial ratios. Nonetheless, we utilize the shapes cuboids, cylinders and ellipsoids as simple labels to apply the feature reduction procedure. For the visualization and interpretation we further differentiate spheres and cubes from ellipsoids and cuboids. We obtain and propose the angularity measure

$$\begin{aligned}\beta_{\text{ang}} = &- 0.83\, V_{V,3} - 14.44\, V_{V,2} + 37.27\, V_{V,1} - 364.29\, V_{V,0} \\ &+ 17.82\, D(p_{V,2}) - 3.96 * 10^{-6}\, \sigma(p_{V,1}) + 1.31\, \sigma(\rho_{V,1}) - 35.55\end{aligned}$$

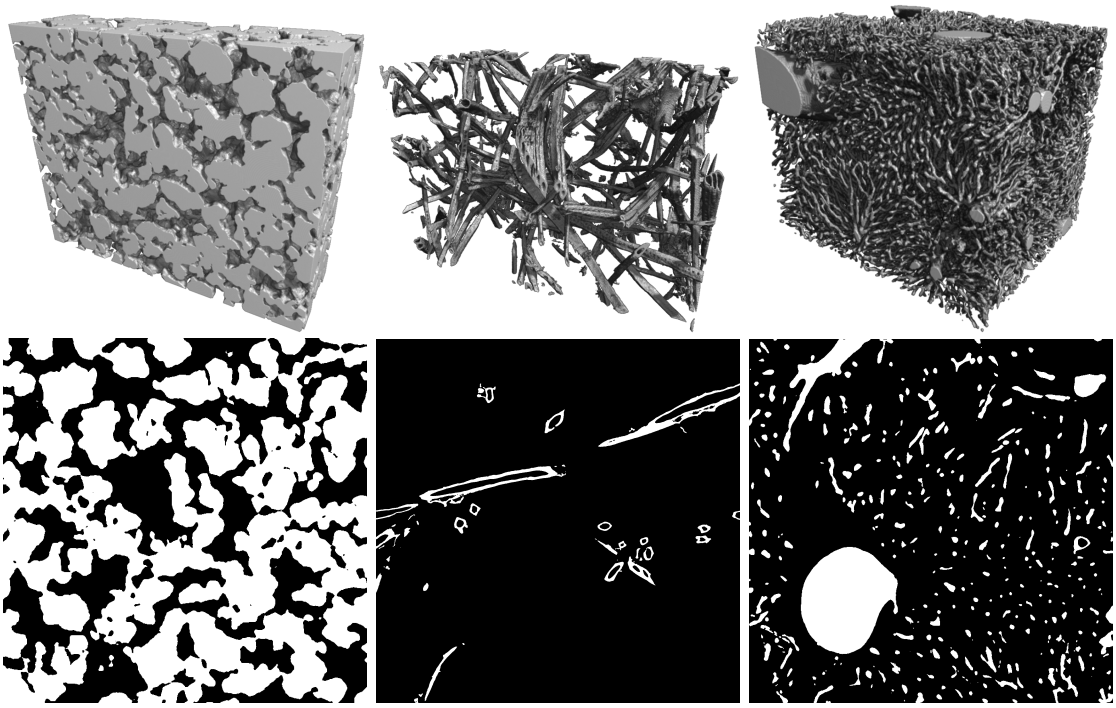
which integrates the intrinsic volume densities and other features that capture information about the surface and curvature. We remark that the intercept is a result of the feature normalization and is included for improved interpretability of β_{ang}. Furthermore, the varying magnitude of weights is primarily due to the different scales of the features themselves and does not reflect their importance. The results in Figure 5.13 show that β_{ang} strongly separates between angular particle shapes and round particle shapes. Cylinders whose circular bases introduce some angularity compared to the otherwise smooth mantle are correctly placed between them. This argument further explains why in (b) elongated cylinders tend to values comparable to ellipsoids and smaller elongations result in values comparable to cuboids. The relatively high uncertainty within each label can be deduced from the complex intersections which are produced by the Boolean models and introduce strong edges despite of otherwise smooth and round surfaces.
The generalization of β_{ang} is not straightforward since it combines multiple morphological features at once. While each individual feature encodes information about the angularity, their complex interplay may differ across highly different spatial structures. Hence, it requires further testing beyond the structures included in our study to validate and demonstrate its full potential. However, β_{ang} shows that the assessment of angularity using a linear combination of morphological features is possible. We argue that β_{ang} provides an informed estimation of the surface texture. If spatial structures can be modeled via Boolean models similar to those contained in our study, β_{ang} quantifies angularity.



(a) All geometries (b) Selected cylinders

Figure 5.13: Evaluation of β_{ang} on selected subsets of geometries and colored by particle shape (a) and axial ratio (b).

5.3.2.4 Application: Modeling of Spatial Structures



(a) ZrO_2 (b) Wood fibers (c) Mouse liver

Figure 5.14: Renderings and sections of real structures with volume sizes 955 x 775 x 270 pixels (a), 1500 x 1500 x 2160 pixels (b) and 905 x 905 x 1080 pixels (c).

In the following, we demonstrate how the inclusion of our derived measures can aid in modeling general spatial structures in real-world applications. We consider three different geometries from materials science and biology, shown in Figure 5.14. The structures are segmentations of a nano-porous ZrO_2 material [106], a highly porous system of wood fibers [1] and a fibrotic mouse liver [130]. Using visual inspection and the assessment

of our metrics, see Table 5.3, we discuss simple stochastic models that approximate the characteristics of these structures.

Structure	β_{iso}	β_{sc}	β_{ang}
ZrO_2	0.92	1.92	-0.55
Wood fibers	0.89	1.4	0.45
Mouse liver	0.95	0.95	0.19

Table 5.3: Evaluation of morphological measures on the data shown in Figure 5.14.

The nano-porous ZrO_2 material exhibits key characteristics that we also find in our synthetic data. It has a volume density of 47% and appears to be a system of slightly elongated constituents with smooth surfaces. However, some differences exist. Constituents are significantly larger than particles considered in our synthetic data and cannot be matched intuitively to one of our primitive shapes. Furthermore, constituents do not overlap but instead form smooth transitions between touching faces. Our measures β_{sc} and β_{ang} align well with our observations, see Table 5.3. β_{sc} indicates a structure with an extremely large morphological scale compared to the synthetic data. A naive estimation of the constituents' widths using manually placed line segments in the section image suggested widths approximately 1.5 to 2 times the diameter of the largest considered sphere model. β_{ang} confirms that the smoothness of the material exceeds even the models with ellipsoidal shapes. The smooth transitions between constituents result in the extreme value since even Boolean models using ellipsoids exhibit sharp edges when particles intersect. Using β_{iso} we find that the structure is weakly anisotropic which agrees with the slightly elongated constituents we observed and also suggests a girdle-like orientation distribution. Combining all observations and measures, we propose modeling the structure using a Boolean model with a volume density of 47%, very large ellipsoids with a small aspect ratio and the girdle directional distribution. We argue that a more refined model could be achieved by using particles that closely match the shapes of the constituents.
In comparison, the highly porous system of wood fibers bears little resemblance to our synthetic data aside from being composed of highly elongated particles. The volume density is very low at only 2.7%. The particles, which appear as highly elongated cuboids, contain a hollow lumen and are larger than any studied particles. Nonetheless, our measurements align well with these observations, see Table 5.3. β_{sc} confirms a large morphological scale but remains lower compared to the ZrO_2 material. Using the same naive estimation method as before yields that the fibers have 3 to 4 times the thickness of any highly elongated particle in our synthetic data. However, the comparably low value can be attributed to the fibers' hollow interiors which significantly reduce the morphological scale. β_{ang} supports our observation that the structures can be modeled via cuboids since the value is consistent with the studied angular particles. The visual assessment of anisotropy is hardly possible but β_{iso} suggests a girdle distribution when considering highly elongated particles. It is important to note that the structures contain few fiber bundles, which consist of multiple parallel and touching particles. These introduce a bias into β_{iso} since they form large surfaces resulting in a lower value. With this in mind, we propose to model the structure using a Boolean model with a volume density of 2.7%, large and highly elongated cuboids and a uniform directional distribution. A more refined model could be achieved by modeling the lumen within the particles and accounting for fiber bundling.
The fibrotic mouse liver also deviates from our synthetic data. Its volume density is only

13%. The structure contains a system of deeply connected cylindrical vessels which are locally aligned but directions change throughout the volume. Additionally, some huge vessels can be found within the structure, see Figure 5.14 (c). However, our measures β_{sc}, β_{ang} and β_{iso} still enable the derivation of an approximating model. β_{sc} indicates a large morphological scale. While the same naive estimation as before yields widths comparable to those of the small cylinders in our synthetic data they extend much farther significantly increasing the assessed morphological scale. In addition, the large vessels also contribute to this. β_{ang} suggests a slight angularity that fits well with the studied cylinders. From β_{iso} we can deduce that the overall structure is isotropic indicating that while vessels are locally aligned over the entire volume they cover all directions equally. Combining our findings and disregarding connectivity, we propose a Boolean model with a volume density of 13%, highly elongated cylinders with an aspect ratio exceeding our studied bounds and a uniform orientation distribution. A more refined model could be achieved by using a unimodal orientation distribution whose preferred direction changes with spatial position though it does not constitute a Boolean model.
The measures we derived from our synthetic data study have been instrumental in identifying suitable models that approximate the complex structures observed in real data. By quantifying anisotropy, scale and angularity, we characterized and improved our understanding of the spatial structures. This enabled the proposition of rough model parameters for an approximating Boolean model. The assessment on real data demonstrated that the measures generalize beyond the studied synthetic geometries and remain interpretable. Nevertheless, visual inspection was crucial for this interpretation as influences on the individual measures remain broad.
Evaluation of the real data has further highlighted the need to extend our study to even more diverse spatial structures that incorporate larger particle sizes and lower volume densities. This will be left for future work along with the corresponding refinement of feature weights based on the extended dataset and their validation on real data.

5.4 Conclusion

In this chapter, we have conducted an exploratory study of a highly diverse set of synthetic geometries using common morphological features and dimensionality reduction techniques. The study has shown that the morphological properties anisotropy, scale and angularity can be assessed via linear combinations of the calculated features. This led us to formulate corresponding measures that intuitively separate the synthetic data based on these aspects. Using real data we demonstrated that these measures generalize towards spatial structures that deviate significantly from the synthetic data. Using the same data we showed the applicability of the measure for deriving informed stochastic models that approximate the observed spatial structures.

6 Conclusion

In this thesis we have demonstrated the importance of synthetic data in image processing as it provides several benefits over real data which make it irreplaceable for developing, validating and analyzing image processing methods and results.
First, we considered the problem of removing stripe artifacts from images obtained using various imaging techniques such as LSFM, FIB-SEM and remote sensing. We have derived and proposed a variational method based on convex optimization that detects and removes the corruptions using an intuitive cost function. Our method was analyzed in great detail by performing a parameter study and evaluating results on real and synthetic images which revealed applicability to a wide variety of different structures and stripes. We compared its performance against state-of-the-art solutions using both real and synthetic data. The latter used physical simulation of light propagation to generate synthetic LSFM images which enabled a quantitative assessment of performance by providing a ground truth for comparison. The quantitative performance evaluation which is impossible with real images aligns with visual inspection, thus proving the superiority of our method.
A different topic in which synthetic data was essential was for the assessment of segmentation quality. We generated and used a large dataset of synthetic segmentation results derived from simple stochastic models along with common morphological operations found in image processing. This combination yielded geometrically diverse segmentations with highly controlled errors that allowed for a precise analysis of the established metrics for assessing segmentation quality. Using the synthetic data, we proposed, analyzed and validated a quality metric that summarizes the spatial distribution of errors but is independent of error quantity. The highly controlled circumstances of the synthetic segmentations allowed us to show the ease of interpretability of our metric and that traditional metrics and the established distance-based metrics are highly redundant. Therefore, they provide very little diversity in information about the actual segmentation performance. Our metric which is independent of error quantity is therefore primed to be combined with an easily interpretable traditional metric to independently quantify error quantity and distribution enabling a precise assessment of quality that can be arbitrarily adjusted for different applications and needs. We also demonstrated its inclusion in real-world segmentation problems on a selected example by refining previously performed quality assessment and revealing nuances to the results that had remained hidden before.
As a final topic we investigated the morphology of spatial structures by conducting an exploratory study of synthetic structures with a large morphological diversity of anisotropy, scale and angularity. The stochastic models used for generating structures provided parameters that intuitively translated to the studied concepts. By calculating common morphological features and using classical machine learning methods for dimensionality reduction, we found that not only can we distinguish between structures based on their morphology, we can also derive a series of measures which are linear combinations of the features that achieve this. We demonstrated that the measures are interpretable and generalize beyond the synthetic data by applying them to derive simple approximating models from real geometries observed in materials science and biology.

In conclusion, synthetic data has been indispensable for our studied topics. It is easy to obtain, cheap to produce in abundance, comes with highly controlled parameters and provides ground truths. These properties are the sole reason we were able to quantify stripe removal performance, develop and analyze a distance-based metric for segmentation quality and study the morphology of spatial structures which provided corresponding measures. This highlights how crucial synthetic data can be, as it unlocks the potential to explore and analyze a wide range of complex issues.

Bibliography

[1] H. Andrä, D. Dobrovolskij, M. Engelhardt, M. Godehardt, M. Makas, C. Mercier, S. Rief, K. Schladitz, S. Staub, K. Trawka, and S. Treml. " Image-based microstructural simulation of thermal conductivity for highly porous wood fiber insulation boards". In: *Wood Science and Technology* 57.1 (2023), pp. 5–31.

[2] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik. " Contour detection and hierarchical image segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.5 (2011), pp. 898–916.

[3] A. N. M. Ashrafuzzaman, S. Y. Lee, and M. K. Hasan. " A self-adaptive approach for the detection and correction of stripes in the sinogram: suppression of ring artifacts in CT imaging". In: *EURASIP Journal on Advances in Signal Processing* 2011.1 (2010), p. 183547.

[4] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. " Structure-texture image decomposition–modeling, algorithms, and parameter selection". In: *International Journal of Computer Vision* 67.1 (2006), pp. 111–136.

[5] O. U. Aydin, A. A. Taha, A. Hilbert, A. A. Khalil, I. Galinovic, J. B. Fiebach, D. Frey, and V. I. Madai. " On the usage of average Hausdorff distance for segmentation performance assessment: hidden error when used for ranking". In: *European Radiology Experimental* 5.1 (2021), p. 4.

[6] E. Bair, T. Hastie, D. Paul, and R. Tibshirani. " Prediction by supervised principal components". In: *Journal of the American Statistical Association* 101.473 (2006), pp. 119–137.

[7] T. Barisin, C. Jung, F. Müsebeck, C. Redenbach, and K. Schladitz. " Methods for segmenting cracks in 3d images of concrete: a comparison based on semi-synthetic images". In: *Pattern Recognition* 129 (2022), p. 108747.

[8] E. Barshan, A. Ghodsi, Z. Azimifar, and M. Zolghadri Jahromi. " Supervised principal component analysis: visualization, classification and regression on subspaces and submanifolds". In: *Pattern Recognition* 44.7 (2011), pp. 1357–1371.

[9] A. Beck. *First-order methods in optimization*. MOS-SIAM series on optimization 25. SIAM, 2017.

[10] M. Boin and A. Haibel. " Compensation of ring artefacts in synchrotron tomographic images". In: *Opt. Express* 14.25 (2006), pp. 12071–12075.

[11] K. Bredies and D. Lorenz. *Mathematische Bildverarbeitung*. Vieweg+Teubner, 2011.

[12] M. Burger, A. Sawatzky, and G. Steidl. " First order algorithms in variational image processing". In: *Splitting Methods in Communication, Imaging, Science, and Engineering*. Ed. by R. Glowinski, S. J. Osher, and W. Yin. Springer International Publishing, 2016, pp. 345–407.

[13] P. L. Butzer and R. J. Nessel. *Fourier analysis and approximation*. Birkhäuser, 1971.

[14] B. Cao, Y. Du, D. Xu, H. Li, and Q. Liu. " An improved histogram matching algorithm for the removal of striping noise in optical remote sensing imagery". In: *Optik* 126.23 (2015), pp. 4723–4730.

[15] A. Chambolle and T. Pock. " A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of Mathematical Imaging and Vision* 40.1 (2011), pp. 120–145.

[16] A. Chambolle and T. Pock. " An introduction to continuous optimization for imaging". In: *Acta Numerica* 25 (2016), pp. 161–319.

[17] Y. Chang, H. Fang, L. Yan, and H. Liu. " Robust destriping method with unidirectional total variation and framelet regularization". In: *Opt. Express* 21.20 (2013), pp. 23307–23323.

[18] Y. Chang, L. Yan, T. Wu, and S. Zhong. " Remote sensing image stripe noise removal: from image decomposition perspective". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.12 (2016), pp. 7018–7031.

[19] S.-W. W. Chen and J.-L. Pellequer. " DeStripe: frequency-based algorithm for removing stripe noises from AFM images". In: *BMC Structural Biology* 11.1 (2011), p. 7.

[20] X. Chen and F. Yang. " Hyperspectral image destriping and denoising using stripe and spectral low-rank matrix recovery and graph Laplacian regularization". In: *2022 34th Chinese Control and Decision Conference (CCDC)*. 2022, pp. 481–486.

[21] Y. Cheng, X. Wang, and Y. Xia. " Supervised t-distributed stochastic neighbor embedding for data visualization and classification". In: *INFORMS journal on computing* 33.2 (2021), pp. 419–835.

[22] D. Chicco and G. Jurman. " The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation". In: *BMC Genomics* 21.1 (2020), p. 6.

[23] D. Chicco and G. Jurman. " The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification". In: *BioData Mining* 16.1 (2023), p. 4.

[24] D. Chicco, N. Tötsch, and G. Jurman. " The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation". In: *BioData Mining* 14.1 (2021), p. 13.

[25] D. Chicco, M. J. Warrens, and G. Jurman. " The Matthews correlation coefficient (MCC) is more informative than Cohen's kappa and Brier score in binary classification assessment". In: *IEEE Access* 9 (2021), pp. 78368–78381.

[26] N. Chinchor. " MUC-4 evaluation metrics". In: *Proceedings of the 4th conference on Message understanding*. MUC4 '92. Association for Computational Linguistics, 1992, pp. 22–29.

[27] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2013.

[28] J. Cohen. “ A coefficient of agreement for nominal scales”. In: *Educational and Psychological Measurement* 20.1 (1960), pp. 37–46.

[29] D. Crampen and J. Blankenbach. “ Data enrichment for semantic segmentation of point clouds for the generation of geometric-semantic road models”. In: *Frontiers in Built Environment* 11 (2025).

[30] Z. Cunha. *Nonsubsampled Contourlet Toolbox*. MATLAB Central File Exchange.

[31] A. Da Cunha, J. Zhou, and M. Do. “ The nonsubsampled contourlet transform: theory, design, and applications”. In: *IEEE Transactions on Image Processing* 15.10 (2006), pp. 3089–3101.

[32] T. Dahmen, M. Kronenberger, N. Rottmayer, K. Schladitz, and C. Redenbach. “ A neural model for high-performance scanning electron microscopy image simulation of porous materials”. In: *Synthetic Data for Computer Vision Workshop @ CVPR 2024*. 2024.

[33] J. Debayle, V. G. Dogaš, K. Helisová, and J. Staněk. “ Methods for assessing similarity of random sets”. In: *2020 10th International Symposium on Signal, Image, Video and Communications (ISIVC)*. 2021, pp. 1–6.

[34] N. Dey, V. Rajinikanth, A. S. Ashour, and J. M. R. S. Tavares. “ Social group optimization supported segmentation and evaluation of skin melanoma images”. In: *Symmetry* 10.2 (2018), p. 51.

[35] D. Di Battista, D. Merino, G. Zacharakis, P. Loza-Alvarez, and O. E. Olarte. “ Enhanced light sheet elastic scattering microscopy by using a supercontinuum laser”. In: *Methods and Protocols* 2.3 (2019).

[36] I. Ekeland and R. Temam. *Convex analysis and variational problems*. 1976.

[37] T. Eremina, J. Debayle, F. Gruy, and J.-C. Pinoli. “ Local measures distribution for the estimation of the elongation ratio of the typical grain in homogeneous Boolean models”. In: *Image Analysis & Stereology* 40.2 (2021), 95 à 103.

[38] P. Escande, P. Weiss, and W. Zhang. “ A variational model for multiplicative structured noise removal”. In: *Journal of Mathematical Imaging and Vision* 57.1 (2017), pp. 43–55.

[39] F. O. Fahrbach, P. Simon, and A. Rohrbach. “ Microscopy with self-reconstructing beams”. In: *Nature Photonics* 4.11 (2010), pp. 780–785.

[40] F. O. Fahrbach, V. Gurchenkov, K. Alessandri, P. Nassoy, and A. Rohrbach. “ Self-reconstructing sectioned Bessel beams offer submicron optical sectioning for large fields of view in light-sheet microscopy”. In: *Opt. Express* 21.9 (2013), pp. 11425–11440.

[41] F. O. Fahrbach and A. Rohrbach. “ Propagation stability of self-reconstructing Bessel beams enables contrast-enhanced imaging in thick media”. In: *Nature Communications* 3.1 (2012), p. 632.

[42] F. O. Fahrbach, F. F. Voigt, B. Schmid, F. Helmchen, and J. Huisken. “ Rapid 3D light-sheet microscopy with a tunable lens”. In: *Opt. Express* 21.18 (2013), pp. 21010–21026.

[43] J. Fehrenbach, P. Weiss, and C. Lorenzo. “ Variational algorithms to remove stationary noise: applications to microscopy imaging”. In: *IEEE Transactions on Image Processing* 21.10 (2012), pp. 4420–4430.

[44] C. Fend, A. Moghiseh, C. Redenbach, and K. Schladitz. “ Reconstruction of highly porous structures from FIB-SEM using a deep neural network trained on synthetic images”. In: *Journal of Microscopy* 281.1 (2021), pp. 16–27.

[45] R. A. Fisher. “ XV.—The correlation between relatives on the supposition of Mendelian inheritance.” In: *Earth and Environmental Science Transactions of The Royal Society of Edinburgh* 52.2 (1919), pp. 399–433.

[46] J. H. Fitschen, J. Ma, and S. Schuff. “ Removal of curtaining effects by a variational model with directional forward differences”. In: *Computer Vision and Image Understanding* 155 (2017), pp. 24–32.

[47] G. B. Folland. *Fourier analysis and its applications*. Pure and Applied Undergraduate Texts v.4. American Mathematical Society, 2009.

[48] E. B. Fowlkes and C. L. Mallows. “ A method for comparing two hierarchical clusterings”. In: *Journal of the American Statistical Association* 78.383 (1983), pp. 553–569.

[49] X. Fu and L. Wang. “ Data dimensionality reduction with application to simplifying RBF network structure and improving classification performance”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 33.3 (2003), pp. 399–409.

[50] F. Garcia-Lamont, J. Cervantes, A. López, and L. Rodriguez. “ Segmentation of images by color features: a survey”. In: *Neurocomputing* 292 (2018), pp. 1–27.

[51] B. Ghojogh, M. Crowley, F. Karray, and A. Ghodsi. *Elements of dimensionality reduction and manifold learning*. Springer International Publishing, 2023.

[52] L. Grafakos et al. *Classical fourier analysis*. Vol. 2. Springer, 2008.

[53] G. Guan, Y. Chen, H. Wang, Q. Ouyang, and C. Tang. “ Characterizing cellular physiological states with three-dimensional shape descriptors for cell membranes”. In: *Membranes* 14.6 (2024), p. 137.

[54] J. Healy and L. McInnes. “ Uniform manifold approximation and projection”. In: *Nature Reviews Methods Primers* 4.1 (2024), p. 82.

[55] G. E. Hinton and S. Roweis. “ Stochastic neighbor embedding”. In: *Advances in Neural Information Processing Systems*. Vol. 15. MIT Press, 2002.

[56] J. Hörrmann, D. Hug, M. A. Klatt, and K. Mecke. “ Minkowski tensor density formulas for Boolean models”. In: *Advances in Applied Mathematics* 55 (2014), pp. 48–85.

[57] J. Hörrmann, D. Hug, M. A. Klatt, and K. Mecke. “ Minkowski tensor density formulas for Boolean models”. In: *Advances in Applied Mathematics* 55 (2014), pp. 48–85.

[58] L. Hubert and P. Arabie. “ Comparing partitions”. In: *Journal of Classification* 2.1 (1985), pp. 193–218.

[59] D. Hug, R. Schneider, and R. Schuster. “ The space of isometry covariant tensor valuations”. In: *St. Petersburg Mathematical Journal* 19.1 (2008), pp. 137–158.

[60] J. Huisken and D. Y. R. Stainier. “ Even fluorescence excitation by multidirectional selective plane illumination microscopy (mSPIM)”. In: *Opt. Lett.* 32.17 (2007), pp. 2608–2610.

[61] P. Jaccard. “ The distribution of the flora in the alpine zone.1”. In: *New Phytologist* 11.2 (1912), pp. 37–50.

[62] E. B. V. Jensen and M. Kiderlen, eds. *Tensor valuations and their applications in stochastic geometry and imaging*. Vol. 2177. Lecture Notes in Mathematics. Springer International Publishing, 2017.

[63] X. Jiang, C. Marti, C. Irniger, and H. Bunke. “ Distance measures for image segmentation evaluation”. In: *EURASIP Journal on Advances in Signal Processing* 2006.1 (2006), p. 035909.

[64] I. T. Jolliffe and J. Cadima. “ Principal component analysis: a review and recent developments”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (2016), p. 20150202.

[65] C. Jung and C. Redenbach. “ Crack modeling via minimum-weight surfaces in 3d Voronoi diagrams”. In: *Journal of Mathematics in Industry* 13.1 (2023), p. 10.

[66] C. Jung, C. Redenbach, and K. Schladitz. “ VoroCrack3d: an annotated semi-synthetic 3d image data set of cracked concrete”. In: *Data in Brief* 54 (2024), p. 110474.

[67] S. Jung, T. Dagobert, J.-M. Morel, and G. Facciolo. “ A review of t-SNE”. In: *Image Processing On Line* 14 (2024), pp. 250–270.

[68] P. Keller and E. Stelzer. “ Digital scanned laser light sheet fluorescence microscopy”. In: *Cold Spring Harbor protocols* 2010 (2010), pdb.top78.

[69] P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. K. Stelzer. “ Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy”. In: *Science* 322.5904 (2008), pp. 1065–1069.

[70] A. Khalilian-Gourtani, M. Tepper, V. Minden, and D. B. Chklovskii. “ Strip the stripes: artifact detection and removal for scanning electron microscopy imaging”. In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 1060–1064.

[71] B. Larsen and C. Aone. “ Fast and effective text mining using linear-time document clustering”. In: *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1999, pp. 16–22.

[72] X. Liang, Y. Zang, D. Dong, L. Zhang, M. Fang, X. Yang, A. Arranz, J. Ripoll, H. Hui, and J. Tian. “ Stripe artifact elimination based on nonsubsampled contourlet transform for light sheet fluorescence microscopy”. In: *Journal of Biomedical Optics* 21.10 (2016), pp. 1–9.

[73] P. Y. Liu, L. K. Chin, W. Ser, H. F. Chen, C.-M. Hsieh, C.-H. Lee, K.-B. Sung, T. C. Ayi, P. H. Yap, B. Liedberg, K. Wang, T. Bourouina, and Y. Leprince-Wang. “ Cell refractive index for cell biology and disease diagnosis: past, present and future”. In: *Lab Chip* 16.4 (2016), pp. 634–644.

[74] S. Liu, L. Sun, and F. Xiao. “ A staged filtering approach to kill curtain noise in FIB-SEM images”. In: 2016.1 (2016), pp. 1–5. DOI: https://doi.org/10.3997/2214-4609.201601063.

[75] X. Liu, X. Lu, H. Shen, Q. Yuan, and L. Zhang. " Oblique stripe removal in remote sensing images via oriented variation". In: *CoRR* abs/1809.02043 (2018).

[76] Y. Liu, K. Weiss, N. Navab, C. Marr, J. Huisken, and T. Peng. " DeStripe: a self2self spatio-spectral graph neural network with unfolded hessian for stripe artifact removal in light-sheet microscopy". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Ed. by L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li. Springer Nature Switzerland, 2022, pp. 99–108.

[77] L. v. d. Maaten and G. Hinton. " Visualizing data using t-SNE". In: *Journal of Machine Learning Research* 9.86 (2008), pp. 2579–2605.

[78] S. U. Mageswari and C. Mala. " Analysis and performance evaluation of various image segmentation methods". In: *2014 International Conference on Contemporary Computing and Informatics (IC3I)*. 2014, pp. 469–474.

[79] D. Martin. " An empirical approach to grouping and segmentation". PhD thesis. University of California, 2002.

[80] D. Martin, C. Fowlkes, D. Tal, and J. Malik. " A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics". In: *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*. Vol. 2. 2001, pp. 416–423.

[81] B. W. Matthews. " Comparison of the predicted and observed secondary structure of T4 phage lysozyme". In: *Biochimica et Biophysica Acta (BBA) - Protein Structure* 405.2 (1975), pp. 442–451.

[82] K. Mcgraw and S. Wong. " Forming inferences about some intraclass correlation coefficients". In: *Psychological Methods* 1 (1996), pp. 30–46.

[83] L. McInnes, J. Healy, and J. Melville. *UMAP: uniform manifold approximation and projection for dimension reduction*. 2020. DOI: 10.48550/arXiv.1802.03426.

[84] P. McMullen. " Isometry covariant valuations on convex bodies". In: *Rend. Circ. Mat. Palermo (2) Suppl* 50 (1997), pp. 259–271.

[85] M. Meilă. " Comparing clusterings by the variation of information". In: *Learning Theory and Kernel Machines*. Ed. by B. Schölkopf and M. K. Warmuth. Lecture Notes in Computer Science. Springer, 2003, pp. 173–187.

[86] M. Meilă. " Comparing clusterings—an information based distance". In: *Journal of Multivariate Analysis* 98.5 (2007), pp. 873–895.

[87] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. Mullers. " Fisher discriminant analysis with kernels". In: *Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468)*. 1999, pp. 41–48.

[88] O. Müller, A. Moghiseh, H. Stephani, N. Rottmayer, and F. Huang. *Application of deep learning for crack segmentation on concrete surface*. 2018. URL: https://publica.fraunhofer.de/handle/publica/403411.

[89] B. Münch, P. Trtik, F. Marone, and M. Stampanoni. " Stripe and ring artifact removal with combined wavelet — Fourier filtering". In: *Opt. Express* 17.10 (2009), pp. 8567–8591.

[90] J. Ohser and F. Mücklich. “ Statistical analysis of microstructures in materials science”. In: *Practical Metallography* 38.9 (2001), pp. 538–539.

[91] J. Ohser and K. Schladitz. *3D images of materials structures: processing and analysis*. John Wiley & Sons, 2009.

[92] A. V. Oppenheim. *Digital signal processing*. Prentice-Hall, 1975.

[93] K. Pearson. “ LIII. On lines and planes of closest fit to systems of points in space”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1901), pp. 559–572.

[94] M. Pietsch, D. Christiaens, J. V. Hajnal, and J.-D. Tournier. “ dStripe: slice artefact correction in diffusion MRI via constrained neural network”. In: *Medical Image Analysis* 74 (2021), p. 102255.

[95] J. Pont-Tuset and F. Marques. “ Measures and meta-measures for the supervised evaluation of image segmentation”. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition*. 2013, pp. 2131–2138.

[96] S. Preibisch, S. Saalfeld, J. Schindelin, and P. Tomancak. “ Software for bead-based registration of selective plane illumination microscopy data”. In: *Nature Methods* 7.6 (2010), pp. 418–419.

[97] P. Rakwatin, W. Takeuchi, and Y. Yasuoka. “ Stripe noise reduction in MODIS data by combining histogram matching with facet filter”. In: *IEEE Transactions on Geoscience and Remote Sensing* 45.6 (2007), pp. 1844–1856.

[98] J. Ramanathan. *Methods of applied Fourier analysis*. Birkhäuser, 1998.

[99] W. M. Rand. “ Objective criteria for the evaluation of clustering methods”. In: *Journal of the American Statistical Association* 66.336 (1971), pp. 846–850.

[100] A. Reinke, M. D. Tizabi, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, A. E. Kavur, T. Rädsch, C. H. Sudre, L. Acion, M. Antonelli, T. Arbel, S. Bakas, A. Benis, F. Buettner, M. J. Cardoso, V. Cheplygina, J. Chen, E. Christodoulou, B. A. Cimini, K. Farahani, L. Ferrer, A. Galdran, B. van Ginneken, B. Glocker, P. Godau, D. A. Hashimoto, M. M. Hoffman, M. Huisman, F. Isensee, P. Jannin, C. E. Kahn, D. Kainmueller, B. Kainz, A. Karargyris, J. Kleesiek, F. Kofler, T. Kooi, A. Kopp-Schneider, M. Kozubek, A. Kreshuk, T. Kurc, B. A. Landman, G. Litjens, A. Madani, K. Maier-Hein, A. L. Martel, E. Meijering, B. Menze, K. G. M. Moons, H. Müller, B. Nichyporuk, F. Nickel, J. Petersen, S. M. Rafelski, N. Rajpoot, M. Reyes, M. A. Riegler, N. Rieke, J. Saez-Rodriguez, C. I. Sánchez, S. Shetty, R. M. Summers, A. A. Taha, A. Tiulpin, S. A. Tsaftaris, B. Van Calster, G. Varoquaux, Z. R. Yaniv, P. F. Jäger, and L. Maier-Hein. “ Understanding metric-related pitfalls in image analysis validation”. In: *Nature Methods* 21.2 (2024), pp. 182–194.

[101] P. Ricci, V. Gavryusev, C. Müllenbroich, L. Turrini, G. d. Vito, L. Silvestri, G. Sancataldo, and F. S. Pavone. “ Removing striping artifacts in light-sheet fluorescence microscopy: a review”. In: *Progress in Biophysics and Molecular Biology* 168 (2022), pp. 52–65.

[102] S. P. Rigby. “ The anatomy of amorphous, heterogeneous catalyst pellets”. In: *Materials* 16.8 (2023), p. 3205.

[103] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*. Vol. 317. Springer Science & Business Media, 2009.

[104] R. T. Rockafellar. *Convex analysis*. 10. print. and 1. paperback print. Princeton landmarks in mathematics and physics. Princeton Univ. Press, 1997.

[105] A. Rohrbach. " Artifacts resulting from imaging in scattering media: a theoretical prediction". In: *Opt. Lett.* 34.19 (2009), pp. 3041–3043.

[106] D. Roldán, C. Redenbach, K. Schladitz, M. Klingele, and M. Godehardt. " Reconstructing porous structures from FIB-SEM image data: optimizing sampling scheme and image processing". In: *Ultramicroscopy* 226 (2021), p. 113291.

[107] D. Roldán, C. Redenbach, K. Schladitz, C. Kübel, and S. Schlabach. " Image quality evaluation for FIB-SEM images". In: *Journal of Microscopy* 293.2 (2024), pp. 98–117.

[108] N. Rottmayer. *General stripe removal*. GitHub repository. 2024.

[109] N. Rottmayer and C. Redenbach. " A novel distance-based metric for quality assessment in image segmentation". In: *IET Image Processing* 20.1 (2026), e70296.

[110] N. Rottmayer, C. Redenbach, and F. O. Fahrbach. " A universal and effective variational method for destriping: application to light-sheet microscopy, FIB-SEM, and remote sensing images". In: *Optics Express* 33.3 (2025), pp. 5800–5809.

[111] S. M. Salili, M. Harrington, and D. J. Durian. " Note: Eliminating stripe artifacts in light-sheet fluorescence imaging". In: *Review of Scientific Instruments* 89.3 (2018), p. 036107.

[112] K. Schladitz, S. Peters, D. Reinel-Bitzer, A. Wiegmann, and J. Ohser. " Design of acoustic trim based on geometric modeling and flow simulation for non-woven". In: *Computational Materials Science* 38.1 (2006), pp. 56–66.

[113] R. Schneider and W. Weil. *Stochastic and integral geometry*. Springer Berlin / Heidelberg, 2008.

[114] R. Schneider and W. Weil. *Stochastische Geometrie*. Ed. by U. Gather, J. Lehn, N. Schmitz, and W. Weil. Teubner Skripten zur Mathematischen Stochastik. Vieweg+ Teubner Verlag, 2000.

[115] B. Schölkopf, A. Smola, and K.-R. Müller. " Kernel principal component analysis". In: *Artificial Neural Networks — ICANN'97*. Ed. by W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud. Springer, 1997, pp. 583–588.

[116] G. . Schröder-Turk, S. Kapfer, B. Breidenbach, C. Beisbart, and K. Mecke. " Tensorial Minkowski functionals and anisotropy measures for planar patterns". In: *Journal of Microscopy* 238.1 (2010), pp. 57–74.

[117] G. E. Schröder-Turk, W. Mickel, S. C. Kapfer, M. A. Klatt, F. M. Schaller, M. J. F. Hoffmann, N. Kleppmann, P. Armstrong, A. Inayat, D. Hug, M. Reichelsdorfer, W. Peukert, W. Schwieger, and K. Mecke. " Minkowski tensor shape analysis of cellular, granular and porous structures". In: *Advanced Materials* 23.22-23 (2011), pp. 2535–2553.

[118] G. E. Schröder-Turk, R. Schielein, S. C. Kapfer, F. M. Schaller, G. W. Delaney, T. Senden, M. Saadatfar, T. Aste, and K. Mecke. " Minkowski tensors and local structure metrics: amorphous and crystalline sphere packings". In: *AIP Conference Proceedings* 1542.1 (2013), pp. 349–352.

[119] J. Schwartz, Y. Jiang, Y. Wang, A. Aiello, P. Bhattacharya, H. Yuan, Z. Mi, N. Bassim, and R. Hovden. " Removing stripes, scratches, and curtaining with non-recoverable compressed sensing". In: *Microscopy and Microanalysis* 25.S2 (2019), pp. 174–175.

[120] P. E. Shrout and J. L. Fleiss. " Intraclass correlations: uses in assessing rater reliability". In: *Psychological Bulletin* 86.2 (1979), pp. 420–428.

[121] D. Stoyan and H. Stoyan. *Fractals, random shapes, and point fields: methods of geometrical statistics*. Wiley series in probability and mathematical statistics Applied probability and statistics. Wiley, 1994.

[122] J. Swoger, P. Verveer, K. Greger, J. Huisken, and E. H. K. Stelzer. " Multi-view image fusion improves resolution in three-dimensional microscopy". In: *Opt. Express* 15.13 (2007), pp. 8029–8042.

[123] A. A. Taha and A. Hanbury. " Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool". In: *BMC Medical Imaging* 15.1 (2015), p. 29.

[124] W. S. Torgerson. " Multidimensional scaling: I. Theory and method". In: *Psychometrika* 17.4 (1952), pp. 401–419.

[125] F. S. Tsai and K. L. Chan. " Dimensionality reduction techniques for data exploration". In: *2007 6th International Conference on Information, Communications & Signal Processing*. 2007, pp. 1–5.

[126] F. Tsai, S.-Q. Lin, J.-Y. Rau, L.-C. Chen, and G.-R. Liu. " Destriping hyperion imagery using spline interpolation". In: *Proc. 26th Asian Conf. Remote Sensing, November, Hanoi, Vietnam*. 2005.

[127] L. A. Vese and C. Le Guyader. *Variational methods in image processing*. CRC Press Boca Raton, FL, 2016.

[128] P. Viola and W. M. Wells III. " Alignment by maximization of mutual information". In: *International Journal of Computer Vision* 24.2 (1997), pp. 137–154.

[129] A. Vretblad. *Fourier analysis and its applications*. Ed. by S. Axler, F. W. Gehring, and K. A. Ribet. Vol. 223. Graduate Texts in Mathematics. Springer, 2003.

[130] W. L. Wagner, S. Föhst, J. Hock, Y. O. Kim, Y. Popov, D. Schuppan, K. Schladitz, C. Redenbach, and M. Ackermann. " 3D analysis of microvasculature in murine liver fibrosis models using synchrotron radiation-based microtomography". In: *Angiogenesis* 24.1 (2021), pp. 57–65.

[131] S. Wang, X. Liu, Y. Li, X. Sun, Q. Li, Y. She, Y. Xu, X. Huang, R. Lin, D. Kang, X. Wang, H. Tu, W. Liu, F. Huang, and J. Chen. " A deep learning-based stripe self-correction method for stitched microscopic images". In: *Nature Communications* 14.1 (2023), p. 5393.

[132] Z. Wang, E. Simoncelli, and A. Bovik. " Multiscale structural similarity for image quality assessment". In: *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*. Vol. 2. 2003, pp. 1398–1402.

[133] Z. Wang, E. Wang, and Y. Zhu. " Image segmentation evaluation: a survey of methods". In: *Artificial Intelligence Review* 53.8 (2020), pp. 5637–5674.

[134] Z. Wei, X. Wu, W. Tong, S. Zhang, X. Yang, J. Tian, and H. Hui. “ Elimination of stripe artifacts in light sheet fluorescence microscopy using an attention-based residual neural network”. In: *Biomed. Opt. Express* 13.3 (2022), pp. 1292–1311.

[135] M. Weigert, K. Subramanian, S. T. Bundschuh, E. W. Myers, and M. Kreysing. “ Biobeam—multiplexed wave-optical simulations of light-sheet microscopy”. In: *PLoS Computational Biology* 14.4 (2018), pp. 1–11.

[136] S. Weis, G. E. Schröder-Turk, and M. Schröter. “ Structural similarity between dry and wet sphere packings”. In: *New Journal of Physics* 21.4 (2019), p. 043020.

[137] D. Wilken, P. Feldens, T. Wunderlich, and C. Heinrich. “ Application of 2D Fourier filtering for elimination of stripe noise in side-scan sonar mosaics”. In: *Geo-Marine Letters* 32.4 (2012), pp. 337–347.

[138] F. Yan, S. Wu, Q. Zhang, Y. Liu, and H. Sun. “ Destriping of remote sensing images by an optimized variational model”. In: *Sensors* 23.17 (2023).

[139] N. Zhang. “ Fibre Processes and their Applications”. PhD thesis. Technische Universität Kaiserslautern, 2013.

[140] S. Zhao, B. Zhang, J. Yang, J. Zhou, and Y. Xu. “ Linear discriminant analysis”. In: *Nature Reviews Methods Primers* 4.1 (2024), p. 70.

[141] B. Zhou, J. Wang, and H. Wang. “ Three-dimensional sphericity, roundness and fractal dimension of sand particles”. In: *Géotechnique* 68.1 (2017), pp. 18–30.

[142] J. Zhou, A. Cunha, and M. Do. “ Nonsubsampled contourlet transform: construction and application in enhancement”. In: *IEEE International Conference on Image Processing 2005*. Vol. 1. 2005, pp. I–469.

[143] X. Zhou, Y. Zhang, J. Liu, and Y. Hu. “ A novel two-stage destriping algorithm based on MWIR energy separation and image guidance (MES-IG)”. In: *IEEE Transactions on Geoscience and Remote Sensing* 61 (2023), pp. 1–16.

Scientific Career

[109] N. Rottmayer and C. Redenbach. “ A novel distance-based metric for quality assessment in image segmentation”. In: *IET Image Processing* 20.1 (2026). _eprint: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/ipr2.70296, e70296

[110] N. Rottmayer, C. Redenbach, and F. O. Fahrbach. “ A universal and effective variational method for destriping: application to light-sheet microscopy, FIB-SEM, and remote sensing images”. In: *Optics Express* 33.3 (2025). Publisher: Optica Publishing Group, pp. 5800–5809

[32] T. Dahmen, M. Kronenberger, N. Rottmayer, K. Schladitz, and C. Redenbach. “ A neural model for high-performance scanning electron microscopy image simulation of porous materials”. In: *Synthetic Data for Computer Vision Workshop @ CVPR 2024*. 2024

[88] O. Müller, A. Moghiseh, H. Stephani, N. Rottmayer, and F. Huang. *Application of deep learning for crack segmentation on concrete surface*. 2018. URL: `https://publica.fraunhofer.de/handle/publica/403411`

Curriculum Vitae

since 11.2021	PhD in Mathematics, Scientific Assistant at Mathematics Departments, RPTU Kaiserslautern-Landau
04.2019-10.2021	M. Sc. in Industrial Mathematics, TU Kaiserslautern
04.2016-03.2019	B. Sc. in Mathematics, TU Kaiserslautern

Akademischer Werdegang

seit 11.2021	Promotionsstudium in Mathematik, Wissenschaftlicher Mitarbeiter im Fachbereich Mathematik, RPTU Kaiserslautern-Landau
04.2019-10.2021	M. Sc. in Technomathematik, TU Kaiserslautern
04.2016-03.2019	B. Sc. in Mathematik, TU Kaiserslautern