



UNIVERSITÄT ZU LÜBECK  
INSTITUTE OF MATHEMATICS AND  
IMAGE COMPUTING

# Cepstralanalyse und energiebasierte Modellierungsansätze zur Reduzierung von Gitterstrukturen auf Röntgenbildern

*Cepstral analysis and energy-based modeling approaches to reduce lattice structures on X-ray images*

## Masterarbeit

im Rahmen des Studiengangs  
Mathematik in Medizin und Lebenswissenschaften  
der Universität zu Lübeck

## Vorgelegt von

Nathalie Weber

## Ausgegeben und betreut von

Prof. Dr. rer. nat. Jan Modersitzki  
Institute of Mathematics and Image Computing

## Mit Unterstützung von

Olaf Schmidt, Dipl.-Inf.

Die Arbeit ist im Rahmen einer Tätigkeit bei der Firma  
VisiConsult X-ray Systems & Solutions GmbH entstanden.

Lübeck, den 25.07.2023



## Eidesstattliche Erklärung

Ich versichere an Eides statt, die vorliegende Arbeit selbstständig und nur unter Benutzung der angegebenen Quellen und Hilfsmittel angefertigt zu haben.

Lübeck,  
25.07.2023

---

Autor



## Kurzfassung

Die Segmentierung von Fehlern auf Röntgenbildern in der Industrie wird häufig durch das Überdecken von Lufteinschlüssen in den Gussteilen durch Gitterstrukturen erschwert, was die Qualität der Gussteile beeinträchtigen kann. Diese Arbeit schlägt die Anwendung der Cepstralanalyse und einem energiebasierten Modellierungsansatz zur Reduktion der Gitterstrukturen vor und untersucht ihre Eignung für diese Aufgabe. Beide Verfahren wurden ursprünglich für andere Bildverarbeitungsprobleme entwickelt. Dabei betritt diese Arbeit unbekanntes Terrain, da es keine vorherigen Forschungsarbeiten zu dieser speziellen firmenspezifischen Problemstellung gibt.

Zusätzlich zu den vorgeschlagenen Ansätzen zur Reduktion der Gitterstrukturen beschäftigt sich diese Arbeit mit der Untersuchung einer geeigneten Metrik zur Messung der Abschwächung, die eine Grundlage zur Bewertung der Methoden auf ihre Eignung bilden. Die Ergebnisse zeigen, dass weiterer Bedarf an Forschung besteht, um eine optimale Metrik zu finden. Die detaillierte Untersuchung und Bewertung der vorgeschlagenen Ansätze sowie die Diskussion der Herausforderungen bei der Bewertung der Reduktionsmethoden liefern wertvolle Erkenntnisse für die weitere Entwicklung von effektiven Fehlersegmentierungstechniken in dieser speziellen industriellen Anwendung.

## Abstract

The segmentation of defects in X-ray images in the industry often encounters the situation where air inclusions in cast parts are obscured by lattice structures, potentially compromising the quality of the cast component. In this work, the application of Cepstral analysis and two energy-based modeling approaches for reducing lattice structures is proposed and examined for their suitability. Both methods were originally introduced for other image processing problems, and this work ventures into uncharted territory as there are no previous research papers addressing this specific company-specific issue.

In addition to the proposed lattice reduction approaches, this work investigates the quest for an appropriate metric to measure attenuation, upon which the methods can be evaluated. The findings reveal that further research is needed to identify an optimal metric. The comprehensive examination and evaluation of the proposed approaches, along with the discussion of challenges in evaluating the reduction methods, provide valuable insights for the further development of effective defect segmentation techniques in this specific industrial application.



## Danksagung

Ein besonderer Dank gilt meinem Betreuer, Prof. Dr. rer. nat. Jan Modersitzki, der sich viel Zeit für anregende Gespräche genommen hat und mich mit seiner Expertise und Ermutigung während des gesamten Prozesses begleitet hat.

Ich möchte auch der Firma VisiConsult X-ray Systems & Solutions GmbH herzlich danken, die mir dieses interessante Forschungsthema ermöglicht hat. Ihre offenen Ohren und Unterstützung waren von unschätzbarem Wert und haben mir geholfen, meine Arbeit voranzubringen.

Ein weiterer Dank gebührt all meinen Freunden, die während der Schreibphase immer ein offenes Ohr für mich hatten und mich ermutigt haben. Ihre Unterstützung und aufmunternden Worte haben mir Kraft gegeben, um Hindernisse zu überwinden und das Beste aus meiner Arbeit herauszuholen.

Vielen herzlichen Dank an alle, die mich auf meinem Weg begleitet und unterstützt haben. Ihre Hilfe hat einen entscheidenden Beitrag zu meinem Erfolg geleistet.





# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>2</b>
1.1	Motivation zur Reduktion von Gitterstrukturen . . . . .	2
1.2	Zielsetzung dieser Arbeit . . . . .	4
1.3	Aufbau der Arbeit . . . . .	5
<b>2</b>	<b>Bilder in dieser Arbeit</b>	<b>7</b>
2.1	Bilder in diskreter Form . . . . .	7
2.2	Fehler und Gitterstrukturen in Röntgenbildern . . . . .	9
2.3	Synthetische Bilder für die Analyse der Methoden . . . . .	11
<b>3</b>	<b>Cepstralanalyse</b>	<b>15</b>
3.1	Die diskrete zweidimensionale Fourier-Transformation . . . . .	16
3.2	Definition und Eigenschaften des Power-Cepstrums . . . . .	18
3.3	Transformation vom Quefreny-Raum in den Bildraum . . . . .	24
3.4	Anwendungen der Cepstralanalyse zur Gitterreduktion von Röntgenbildern	26
3.5	Implementierung mit der Fast-Fourier-Transformation . . . . .	29
<b>4</b>	<b>Energiebasierter Modellierungsansatz</b>	<b>30</b>
4.1	Beschreibung der energiebasierten Modellierung . . . . .	31
4.2	Definition der Totalen Variation und der Funktionsräumen $L^2$ und BV . .	32
4.3	Formulierung der Modelle . . . . .	37
<b>5</b>	<b>Metriken zur Qualitätsanalyse</b>	<b>38</b>
5.1	Distanzmaße . . . . .	38
5.2	Ähnlichkeitsanalyse . . . . .	40
5.3	Kantenfilter . . . . .	41
5.4	Beispiele . . . . .	42
<b>6</b>	<b>Ergebnisse</b>	<b>44</b>
6.1	Erzielte Ergebnisse mit der Cepstralanalyse . . . . .	44
6.2	Erzielte Ergebnisse mit dem Modellierungsansatz . . . . .	45
6.3	Erzielte Ergebnisse mit den Modellierungsansätzen auf Röntgenbildern . .	47
<b>7</b>	<b>Diskussion und Schlussfolgerung</b>	<b>49</b>
<b>A</b>	<b>Anhang</b>	<b>55</b>



# Abbildungsverzeichnis

1.1	Exemplarische Darstellung eines Röntgenbildes mit der zugehörigen Segmentierung nach der automatischen Fehlererkennung: (a) zeigt ein exemplarisches Röntgenbild eines Gussteils für das VAIP, während (b) die resultierende Segmentierung in Form einer Maske anzeigt. Die Fehler sind dabei in Rot markiert [1]. . . . .	3
1.2	Vereinfachte Darstellung eines synthetischen Objektes: (a) zeigt eine Kreisstruktur im Zentrum des Bildes, während bei (b) diese durch eine Gitterstruktur überlappt wird, (c) zeigt die Kreisstruktur in vier Regionen unterteilt, während bei (d) die Kreisstruktur als eine Region gesehen wird. Interpretation: Das VAIP trifft die Entscheidung bei (c) das Gussteil fälschlicherweise beizubehalten, da die gefundenen Fehlergrößen und Häufigkeiten noch im Toleranzbereich liegen und bei (d) wird das Gussteil aufgrund der Größe des Fehlers verworfen [2]. . . . .	4
1.3	Sechs Röntgenbilder mit unterschiedlichen Situationen von der Lage der Fehler auf Gussteilen [3] . . . . .	5
1.4	Exemplarische Gitterstrukturen auf einem Gussteil [3]. . . . .	6
2.1	Modell eines zellzentrierten Gitters mit $\Omega = (a_1, b_1) \times (a_2, b_2)$ und $m_1 = 4$ , $m_2 = 5$ (vgl. Bredits et al. und Modersitzki [4, 5]) . . . . .	8
2.2	3D-Oberflächendarstellung mit (a) einem Fehler in Form eines Maximums, (b) Gitterstrukturen in Form von Vertiefungen, die regelmäßig verlaufen, (c) einem Fehler, der durch ein Gitterelement in zwei Regionen getrennt wird und (d) einem Fehler, der am Rande eines Gitterelements liegt [6]. . . . .	10
2.3	Ein Beispiel eines synthetischen Eingabebildes für die Röntgenprüfung mit zwei verschiedenen Darstellungsmöglichkeiten. Links: der Fehler ist hellgrau, und durch ein dunkelgraues Gitterelement in zwei Hälften unterteilt. Rechts: andere Fensterung des linken Bildes. Der Fehler ist in Weiß dargestellt und wird ebenfalls durch ein Gitterelement geteilt. Dabei nimmt das Gitterelement von einer schwarzen Färbung eine graue Färbung an, wo sich der Fehler und das Gitterelement schneiden [6]. . . . .	10
2.4	Darstellung der Schichten des synthetischen Bildes: (a) zeigt das Grundlagenbild $u_{\text{Base}}$ mit Rauschen, in (b) sind zusätzlich zum Grundlagenbild Gitterstrukturen mit Gitterweite $h = (30, 30)$ eingerechnet und ergibt das Gittermusterbild $u_{\text{Pat}}$ ; (c) zeigt das Fehlerbild $u_{\text{Err}}$ mit sechs kreisförmigen Lufteinschlüsse; (d) zeigt das Fehlerbild mit den Gitterstrukturen in einem mit Beschriftung der Fehlernummerierung. Dieses wird als Eingabebild $u_{\text{In}}$ bezeichnet [6]. . . . .	12

2.5	Oben: Darstellung des Grundlagenbildes $u_{\text{Base}}$ als 3D-Oberflächenplot; unten links: der Ausschnitt zeigt detailliert die gitterförmigen Strukturen in Form von Minima auf einem 3D-Oberflächenplot; unten Mitte: in diesem Ausschnitt ist ein Fehler in Form eines Maximums zu sehen, der sich an einem Gitterelement befindet auf einem 3D-Oberflächenplot und unten rechts: das Gitterelement teilt den Fehler in zwei Teile [6]. . . . .	13
3.1	Oben: Ausschnittes vom Zentrum ausgehende dritte Quadranten von $u_{\text{Pat}}$ mit markierten Nebenmaxima auf den Koordinatenachsen. Die Abstände der markierten Maxima entspricht der Gitterweite $h_1$ in horizontaler bzw. $h_2$ in vertikaler Richtung. Unten: Ausschnitt des vom Zentrum ausgehende dritte Quadranten von $c_{\text{Diff}}$ . Es treten gitterförmige Nebenmaxima um das Zentrum herum auf [6]. . . . .	28
4.1	Motivation der energiebasierten Modellierung: Links zeigt das Eingabebild mit Fehlern und Gitterstrukturen. Bei den beiden darauffolgenden Bildern wurden 1. die Frequenzen der Fehler geglättet und 2. die Fehler der Gitterstrukturen geglättet. Rechts zeigt das Ergebnis [6]. . . . .	30
4.2	Funktion $f$ auf dem Intervall $[-\pi, \pi]$ . . . . .	34
4.3	Funktion $f$ auf dem Intervall $[-4, 4]$ . . . . .	36
5.1	Oben links: Bild Lena512.bmp aus der MATLAB Image Processing Toolbox. Oben Mitte: Originalbild mit eingerechneter Gitterstruktur (Pixelwerte sind 10% abgeschwächt) mit Gitterweite $h = (30, 30)$ . Oben rechts: Originalbild mit eingerechneter Gitterstruktur (Pixelwerte sind 90% abgeschwächt) mit Gitterweite $h = (30, 30)$ . Die untere Reihe zeigt die zugehörigen Kantenbilder mit dem Canny-Edge-Detection-Algorithmus zu dem oberen Bild [6]. . . . .	43
6.1	Links zeigt das Ergebnis $u_{\text{TCep}}$ mit $\tau_{\text{Red}} = 0.55$ , $r_{\text{Center}} = 10$ , $\tau_{\text{Diff}} = 0.001$ mit dem zugehörigen Kantenbild mittels des Canny-Edge-Detection Algorithmus [6]. . . . .	44
6.2	Links zeigt das Ergebnis $u_{\text{MCep}}$ mit $r_{\text{Center}} = 1$ , $\tau_{\text{Red}} = 0.45$ und $R_{x_1} = R_{x_2} = 5$ und dem zugehörigen Kantenbild mittels des Canny-Edge-Detection Algorithmus [6]. . . . .	45
6.3	Links zeigt das Ergebnis $u_{\text{ROF}}$ (oben) mit $\lambda_{\text{ROF1}} = 0.1$ , $\lambda_{\text{ROF2}} = 2.5$ und Anzahl der Iterationen = 900. Die Fehler sind für das menschliche Auge schwer zu erkennen, weshalb sie mit einem roten Kreis markiert sind. Rechts zeigt das zugehörige Kantenbild mittels des Canny-Edge-Detection Algorithmus [6]. . . . .	46
6.4	Links zeigt das Ergebnis $u_{L1}$ mit $\lambda_{L1} = 0.315$ , $\lambda_{L2} = 0.05$ und Anzahl der Iterationen 500 bzw. 800. Die Fehler sind für das menschliche Auge schwer zu erkennen, weshalb sie mit einem roten Kreis markiert sind. Rechts zeigt das zugehörige Kantenbild mittels des Canny-Edge-Detection Algorithmus [6]. . . . .	46
6.5	Erzielte Ergebnisse unter der Verwendung des Modellierungsansatzes. Links zeigt die Originalbilder (aus [3]), Mitte zeigt $u_{\text{ROF}}$ und rechts zeigt $u_{L1}$ [6].	47

## Tabellenverzeichnis

2.1	Position und Größe der Luftpfeinschlüsse in den synthetischen Bildern $u_{\text{Err}}$ und $u_{\text{In}}$ . . . . .	14
5.1	Ergebnisse der Metriken $\ \cdot\ _F$ , SSIM, MSE und PSNR. Bei der Frobeniusnorm handelt es sich um das Differenzbild beider Bilder. . . . .	42
7.1	Übersicht der Mittelwerte der jeweiligen Fehler und der Gitterstruktur von $u_{\text{Err}}$ (bei den Fehlern 1-6) bzw. $u_{\text{Pat}}$ (bei den Gitterstrukturen) und prozentualen Anteil der Intensitätswerte nach dem Schwellenansatz. . . .	50
7.2	Ergebnisse der Metriken zur Qualitätsanalyse von $u_{\text{In}}$ , $u_{\text{TCep}}$ , $u_{\text{MCep}}$ , $u_{\text{ROF}}$ und $u_{L^1}$ mit $u_{\text{Err}}$ . . . . .	50
7.3	Übersicht der Messungen der erkannten Fehlergrößen des ADRs im Vergleich für die verschiedenen Ergebnisbilder $u_{\text{TCep}}$ , $u_{\text{MCep}}$ , $u_{\text{ROF}}$ und $u_{L^1}$ . Bei „-“ gibt es keine Messungen. Die Größenangaben sind in $\text{pix}^2$ (Pixel zum Quadrat) gegeben. . . . .	51
A.1	Übersicht aller in Kapitel 6 und 7 verwendeten Bilder und Power-Cepstra. . . . .	55
A.2	Übersicht der in Kapitel 6 und 7 verwendeten Parametern. . . . .	56

---

# Kapitel 1: Einleitung

## 1.1 Motivation zur Reduktion von Gitterstrukturen

In der Industrie wird zur Herstellung von Bauteilen unter anderem der sogenannte *Aluminiumguss* verwendet. Bei diesem Verfahren entsteht durch Gießen von spezifischen Metallen ein neues Gussteil. Beim Gießen in die Form kann es zu Materialfehlern kommen, die sich z.B. durch Fremdkörper- oder Lufteinschlüsse auszeichnen. Diese können zur Folge haben, dass sich die Qualität des Gussteils – je nach Art, Beschaffenheit und Beanspruchung des Bauteils – ändert. In der Industrie werden solche Einschlüsse bis zu einem gewissen Grad toleriert, jedoch führen diese bei einer zu großen Häufigkeit und Größe zu Stabilitätsproblemen, die sicherheitsrelevant sind [7]. Materialfehler aller Art werden im Folgenden auch als „Fehler“ bezeichnet. Allgemeine Details zu Fehlern in Röntgenbildern liefert Kapitel 2.2.

Solche sicherheitsrelevanten Bauteile werden in der Regel verschiedener Materialprüfungsverfahren unterzogen. Eines dieser Verfahren ist die Röntgenprüfung. Hierbei werden Röntgenaufnahmen des Gussteils von verschiedenen Positionen aus unterschiedlichen Winkelbetrachtungen ausgewertet. Die Firma VisiConsult X-ray Systems & Solutions GmbH hat das Tool VAIP (VisiConsult Advanced Image Processing) entwickelt, welches eine vollautomatische Bildverarbeitung ermöglicht. Mit diesem Tool kann die automatische Fehlererkennung (englisch Automated Defect Recognition, kurz ADR) durchgeführt werden. Nachdem ein diskretes Röntgenbild als Eingabe in das VAIP eingegeben wurde, führt es eine interne Prüfung durch. Das Ergebnis ist eine Segmentierung, die Materialfehler farblich markiert in Form einer Maske und eine Entscheidung über das Beibehalten oder Verwerfen des Gussteils trifft [7]. Die Abb. 1.1 zeigt ein exemplarisches Röntgenbild und seine Segmentierung nach Prüfung. Diese Bilder stammen von der Internetseite von VisiConsult X-ray Systems & Solutions GmbH [8, 9]. Solche automatische Bildverarbeitungsverfahren sind in der industriellen Massenfertigung bereits Stand der Technik und werden häufig eingesetzt. Dennoch gibt es hier noch weiteren Bedarf an Verbesserungen [7].

In dieser Arbeit sind die relevanten Röntgenbilder von Gussteilen zu sehen, auf deren Oberfläche Gitterstrukturen oder Gittermuster aufgebracht sind. Diese Strukturen haben je nach Gussteil verschiedene Ursachen. Bei allen Gussteilen bewirken sie jedoch eine Verlangsamung der Auskühlung, was insbesondere in dünnwandigen Bereichen der Gussteile für zusätzliche Stabilität sorgt. Diese Gitterstrukturen stellen für die automatische Fehlererkennung eine besondere Herausforderung dar, da zusätzliche Strukturen in das Bild eingefügt werden, die die Detektion der Materialfehler erschweren; siehe Abb. 1.2 für eine schematische Darstellung und Abb. 1.3 für reale Daten. Letztere Abbildung zeigt sechs verschiedene Situationen, die häufig in der Praxis vorkommen: Abb. 1.3a zeigt einen Fehler, der sich an der Grenze zu der Gitterstruktur befindet und bei einer Prüfung übersehen werden kann. Abb. 1.3b zeigt vereinzelte Materialfehler. Abb. 1.3c

---

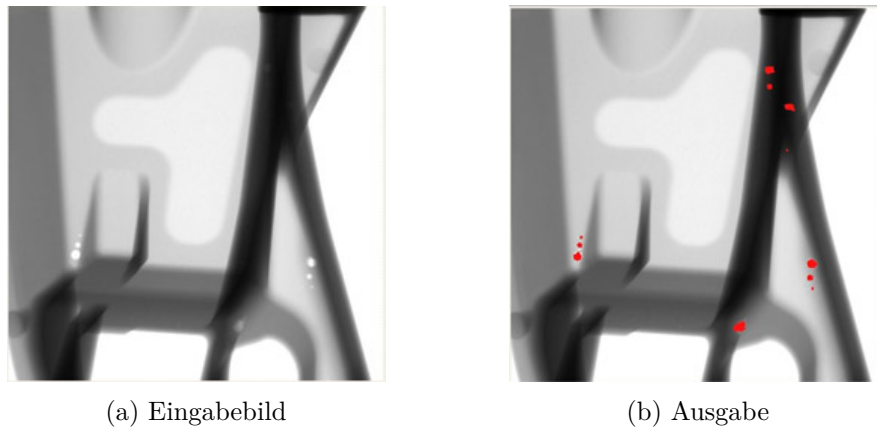


Abbildung 1.1: Exemplarische Darstellung eines Röntgenbildes mit der zugehörigen Segmentierung nach der automatischen Fehlererkennung: (a) zeigt ein exemplarisches Röntgenbild eines Gussteils für das VAIP, während (b) die resultierende Segmentierung in Form einer Maske anzeigt. Die Fehler sind dabei in Rot markiert [1].

zeigt einen Materialfehler, der von der Gitterstruktur überdeckt wird, und bei Abb. 1.3d liegt ein Materialfehler komplett auf der Gitterstruktur. Die letzten beiden Bilder zeigen auf, wie sich ein Materialfehler über zwei Gitterstreben ausstreckt (vgl. Abb. 1.3e) oder wie dieser in (hier zwei) separate Regionen aufgeteilt wird (vgl. Abb. 1.3f).

In Abb. 1.2 wird das Problem anhand eines synthetischen Bildes dargestellt und liefert ein erstes Verständnis für das eigentliche Problem bei der Röntgenprüfung. Dabei zeigt Abb. 1.2a eine Kreisstruktur im Zentrum des Bildes, welches zum Verständnis einen Materialfehler darstellt. Abb. 1.2b hingegen überdeckt diesen „Fehler“ mit einer Gitterstruktur. Bei der Prüfung basiert die Entscheidung vom Beibehalten oder Verwerfen des Gussteils darauf, in welchem Ausmaß ein Materialfehler gefunden wird. Die Entscheidung bei Abb. 1.2c ist das Beibehalten des Gussteils, denn die Größe der Materialfehler liegt noch im Toleranzbereich. Die Entscheidung bei Abb. 1.2d hingegen fällt auf das Verwerfen des Bauteils, da die Fehlergröße den Toleranzbereich übersteigt. Beiden Entscheidungen liegt ein gemeinsames Eingabebild zugrunde, obwohl die tatsächliche Größe des Fehlers den Toleranzbereich überschreitet. In Zukunft sollen solche Gitterstrukturen die Entscheidung nicht mehr beeinflussen. Diese Abbildung zeigt eine idealisierte Situation, die Variabilität bei Realdaten ist erheblich komplexer.

Die Herausforderung in dieser Arbeit ist die Reduktion der Gitterstrukturen, sodass eine Detektion des Materialfehlers in seiner ganzen Größe möglich ist. Bei der Materialprüfung werden Größe und Häufigkeiten der Materialfehler bis zu einem gewissen Grad toleriert, damit die Stabilität und Qualität des Bauteils garantiert werden kann. Die Gitterstrukturen erschweren die Detektion der Fehler und können zu Fehlentscheidungen führen und somit die Stabilität und Qualität gefährden. In der Realität ist das

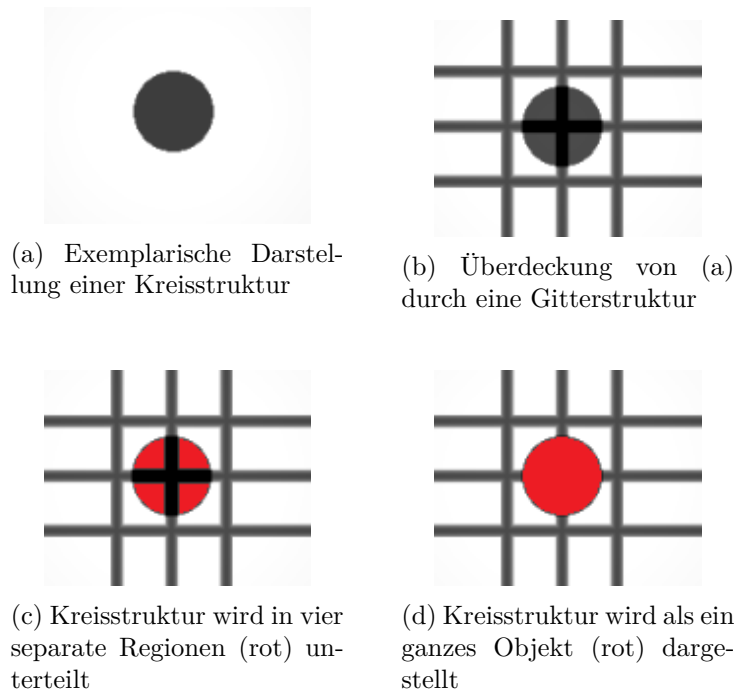


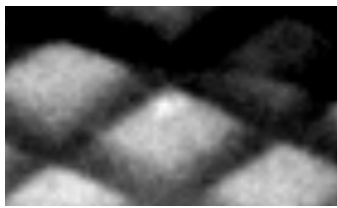
Abbildung 1.2: Vereinfachte Darstellung eines synthetischen Objektes: (a) zeigt eine Kreisstruktur im Zentrum des Bildes, während bei (b) diese durch eine Gitterstruktur überlappt wird, (c) zeigt die Kreisstruktur in vier Regionen unterteilt, während bei (d) die Kreisstruktur als eine Region gesehen wird. Interpretation: Das VAIP trifft die Entscheidung bei (c) das Gussteil fälschlicherweise beizubehalten, da die gefundenen Fehlergrößen und Häufigkeiten noch im Toleranzbereich liegen und bei (d) wird das Gussteil aufgrund der Größe des Fehlers verworfen [2].

Verhalten solcher Gitterstrukturen komplexer. Dabei entstehen bei gekrümmten Oberflächen des Bauteils verzerrte Gitterstrukturen (vgl. Abb. 1.4a) oder weiter können Gitterstrukturen an mehreren Stellen der Oberfläche des Bauteils sein, was sich durch die Projektionscharakteristik der Röntgentechnologie als eine Überlagerung der Gittermuster im anschließenden Röntgenbild widerspiegelt (vgl. Abb. 1.4b). Diese Gründe machen eine Vorverarbeitung für das VAIP zur automatischen Fehlererkennung für Bilder mit Gitterstrukturen notwendig. Zwei Ansätze werden in dieser Masterarbeit erforscht.

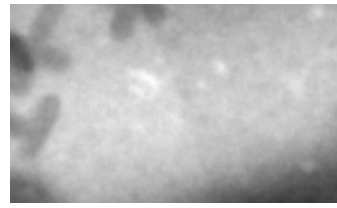
## 1.2 Zielsetzung dieser Arbeit

Bei einem Röntgenprüfprozess in der Industrie werden Gussteile mittels Röntgenaufnahmen aus verschiedenen Positionen und Winkelbetrachtungen auf ihre Stabilität überprüft. Diese Arbeit fokussiert sich auf einen Vorverarbeitungsprozess für das spezifische Tool zur vollautomatischen Bildverarbeitung der Firma VisiConsult X-ray Systems &

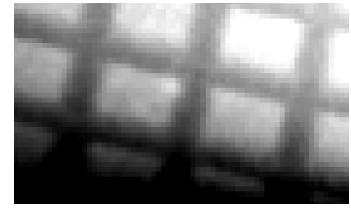




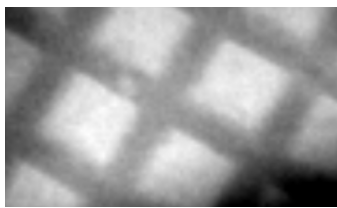
(a) Fehler (weiß) unter einem Gitter (dunkelgrau bis schwarz) versteckt



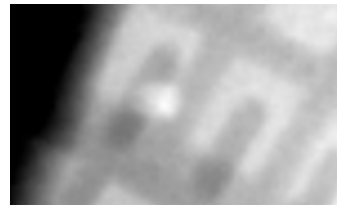
(b) Mehrere Luftporen (weiß); die dunkelgrauen Strukturen (links) sind vom Gitter des Gussteils



(c) Ein weiterer Fehler (weiß, Mitte), welcher von einer Gitterstruktur (dunkelgrau) überdeckt wird



(d) Fehler (hellgrau) befindet sich komplett auf dem Gitter



(e) Fehler (weiß) liegt auf zwei Gitterstreben



(f) Fehler wird durch eine Gitterstrebe geteilt

Abbildung 1.3: Sechs Röntgenbilder mit unterschiedlichen Situationen von der Lage der Fehler auf Gussteilen [3]

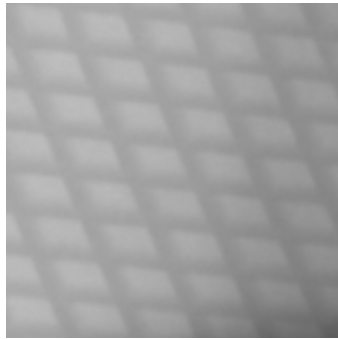
Solutions GmbH, sodass die störenden Gitterstrukturen im Röntgenbild zukünftig die Entscheidung zum Beibehalten oder Verwerfen der Gussteile nicht beeinflussen. Eine derartige Vorverarbeitung ist sehr firmenspezifisch und es wurden keine Forschungsarbeiten zu diesem speziellen Thema gefunden. An einem verwandten Themengebiet forschten Çakir und Çetin in ihren Arbeiten [10, 11], die sich mit der Extraktion von Bildmerkmalen befassen. Diese Arbeit stellt in Kapitel 3 einen verwandten Ansatz vor.

Bei der Analyse der Methoden zur Reduzierung von Gitterstrukturen werden folgende Leitfragen untersucht:

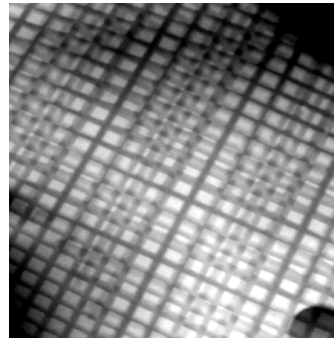
- Eignen sich die hier vorgestellten Methoden zur Reduktion der Gitterstrukturen im Bildraum, sodass die Informationen der darunterliegenden Strukturen beibehalten werden?
- Wie kann eine Reduktion der Gitterstrukturen objektiv gemessen werden?

### 1.3 Aufbau der Arbeit

Kapitel 2 fasst die benötigten mathematischen Kenntnisse für Bilder für diese Masterarbeit zusammen. Der Ausgangspunkt ist ein Röntgenbild, wobei in diesem Kapitel Bilder allgemein definiert und die Unterschiede zwischen diskreten und kontinuierlichen Bildern



(a) Röntgenbild mit einer verzerrten Gitterstruktur.



(b) Röntgenbild mit einer Überlagerung des Gitters.

Abbildung 1.4: Exemplarische Gitterstrukturen auf einem Gussteil [3].

erläutert werden. Weiter werden auch die Auswirkungen von Fehlern und Gitterstrukturen (oder auch Gittermuster) in den Röntgenbildern erläutert. Das Ende des Kapitels beschreibt die Erstellung synthetischer Bilder. Kapitel 3 und 4 stellen zwei Ansätze zur Reduktion der Gitterstrukturen vor. Dabei handelt es sich um Methoden, die ursprünglich für andere Herausforderungen in der Bildverarbeitung eingeführt wurden. In Kapitel 5 werden verschiedene Möglichkeiten vorgestellt, eine objektive Messung der Reduktion der Gitterstrukturen zu erhalten. Neben Messungen der Reduktion von Gitterstrukturen wird ein Algorithmus vorgestellt, der eine Beurteilung der Gitterreduktion ermöglicht. Dabei ist das Vorhandensein eines Kontrolleurs unvermeidbar. Für die Praxis kommt eine solche Bewertungsmethode nicht infrage, für die Bewertung der Ergebnisse ist dieser jedoch von Vorteil. Kapitel 6 stellt die Ergebnisse der Masterarbeit dar und in Kapitel 7 werden sie diskutiert.

---

## Kapitel 2: Bilder in dieser Arbeit

Dieses Kapitel befasst sich mit den grundlegenden Konzepten und Zusammenhängen, die für den weiteren Verlauf der Arbeit von Bedeutung sind. Hauptaufgabe dieser Arbeit ist ein Vorverarbeitungsprozess zu entwickeln, der zur Bearbeitung eines Röntgenbildes für die Röntgenprüfung dient. Die Bildverarbeitung ist ein wichtiges mathematisches Konzept, das in vielen Bereichen Anwendung findet. Sie hilft beim Verständnis und zur Visualisierung von komplexen Zusammenhängen und Beziehungen. Es gibt verschiedene Betrachtungsweisen und Arten von Bildern, die sich von ihrer Entstehung unterscheiden. Im Allgemeinen schließt diese Arbeit dabei an Bredits et al., Pöschel, Burger et al. und Elstrodt [4, 12, 13, 14] an.

Nach Bredits et al. können Bilder auf verschiedene Arten interpretiert werden: Die einen beschreiben Bilder als ein Abbild einer Szene in der realen Welt, andere bezeichnen ein Bild als ein Foto oder Scan [4]. Formal lässt sich ein Bild formulieren als:

**Definition 2.1.** *Ein Bild  $u : \Omega \rightarrow X$  sei eine Abbildung, wobei  $\Omega \subseteq \mathbb{R}^d$ ,  $\Omega \neq \emptyset$  das Bildgebiet,  $d \in \mathbb{N}$  die Bilddimension und  $X \subseteq \mathbb{R}^k$ ,  $k \in \mathbb{N}$  der Wertebereich ist.*

Für  $d = 1$  heißt  $u$  ein 1D-Signal, für  $d = 2$  bezeichnet man  $u$  als ein 2D-Bild. Die Wahl für den Parameter  $k$  entscheidet über die Erscheinung des Bildes bzgl. seiner möglichen Werte, d.h. für  $k = 1$  handelt es sich um ein Grauwertbild, bzw. Intensitätsbild, für  $k = 3$  wird ein RGB-Farbbild betrachtet.

Methoden zur mathematischen Bildverarbeitung werden häufig im Kontinuierlichen definiert (vgl. Bredits et al. und Pöschel [4, 12]). Dies hat den großen Vorteil, dass diese Methoden unabhängig von der Diskretisierung und dessen Parametern angewendet werden können. In dieser Arbeit liegen die Eingabebilder jedoch in diskreter Form vor, weshalb eine Definition eines diskreten Bildes von Interesse ist.

### 2.1 Bilder in diskreter Form

In der digitalen Bildverarbeitung wird das Bildgebiet  $\Omega$  in kleine Zellen unterteilt, denen ein einheitlicher Wert zugeordnet wird (Pixel). Üblich werden diese Werte in einem 2D oder 3D Array kompakt zusammengefasst (vgl. Abb. 2.1). Die Transformation zur Digitalisierung eines kontinuierlichen Bildes erfolgt in der Regel über das Abtasten. Dabei gibt es mehrere Wege, die nicht Hauptbestandteil dieser Arbeit ist. Dazu wird auf Bredits et al. [4] und Erhardt [15] verwiesen.

In der Regel werden Bilder auf einem Gitter (Englisch *Grid*), bestehend aus *Zellen*, diskretisiert. Abb. 2.1 stellt das in der Arbeit verwendete Modell zur Diskretisierung dar. Wenn  $m = (m_1, m_2)$  die Anzahl der Pixel in den ersten beiden Dimensionen bezeichnen, dann kann der räumliche Abstand (und damit die Pixellänge zweier benachbarter

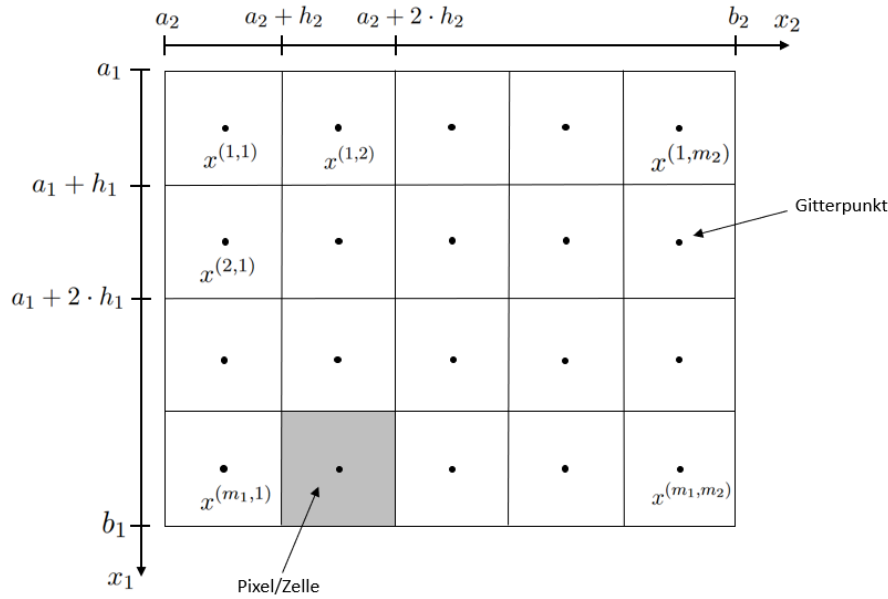


Abbildung 2.1: Modell eines zellzentrierten Gitters mit  $\Omega = (a_1, b_1) \times (a_2, b_2)$  und  $m_1 = 4$ ,  $m_2 = 5$  (vgl. Bredits et al. und Modersitzki [4, 5])

Gitterpunkte) berechnet werden durch

$$h_1 = \frac{b_1 - a_1}{m_1} \text{ und } h_2 = \frac{b_2 - a_2}{m_2} \text{ f\u00fcr } \Omega = (a_1, b_1) \times (a_2, b_2).$$

Den Vektor  $h = (h_1, h_2)$  nennt man auch *Gitterweite* oder *Zellgr\u00f6\u00dfe*. In dieser Arbeit wird von einem zellzentrierten Gitter ausgegangen. Jeder Pixelwert wird von dem Intensit\u00e4tswert des Zentrums repr\u00e4sentiert. Die Pixel k\u00f6nnen mit einem *Multiindex*  $i = (i_1, i_2) \in I_h := \{1, \dots, m_1\} \times \{1, \dots, m_2\}$  nummeriert werden, f\u00fcr die Koordinaten der Zellenmittelpunkte  $x^{\bar{i}} = x^{(i_1, i_2)}$  gilt somit

$$(x^{\bar{i}})_j = a_j + (i_j - \frac{1}{2})h_j \text{ f\u00fcr } j \in \{1, 2\}.$$

Die Menge  $\Omega_h := \{x^{\bar{i}} | i \in I_h\}$  hei\u00dft dann ein *zellzentriertes Gitter*. Es ist zu beachten, dass in dieser Arbeit der Begriff *Gitter* zwei Bedeutungen hat:  $\Omega_h$  beschreibt ein Gitter, auf dem das diskrete Bild definiert ist. Wenn im Folgenden von einem Gittermuster oder einer Gitterstruktur die Rede ist, dann ist das in Kapitel 1 beschriebene Gitter auf der Oberfl\u00e4che der Gussteile gemeint.

Formal l\u00e4sst sich ein diskretes Bild formulieren als:

**Definition 2.2.** Ein diskretes Bild sei eine Abbildung  $u_h : \Omega_h \rightarrow \mathbb{R}^k$ , wobei  $\Omega_h$  das diskrete Bildgebiet von  $\Omega$  und  $\mathbb{R}^k$  den Wertebereich mit  $k \in \mathbb{N}$  bezeichnet. Die Menge  $\Omega$

wird dabei in endliche Abschnitte unterteilt, die auf einem Gitter angeordnet sind. Diese Abschnitte nennt man auch Pixel oder Zelle.

Jedes diskrete Bild  $u_h$  kann in Matrixschreibweise als

$$u_h = \begin{pmatrix} u_h^{(1,1)} & \dots & u_h^{(1,m_2)} \\ \vdots & \ddots & \vdots \\ u_h^{(m_1,1)} & \dots & u_h^{(m_1,m_2)} \end{pmatrix} \quad (1)$$

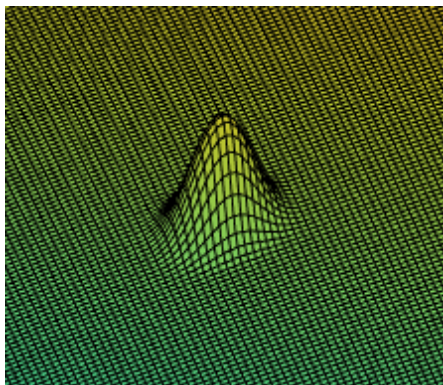
formuliert werden, wobei  $u_h^i \in \mathbb{R}^k$ .

Im Folgenden bezeichne  $u_{i,j} := u(x_i, x_j)$  die diskrete Information an dem entsprechenden Gitterpunkt  $x^{(i,j)}$ .

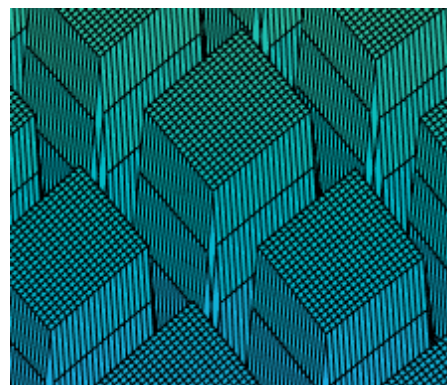
In dieser Arbeit werden Röntgenbilder betrachtet, welche eine Klasse der Grauwertbilder bilden und für die  $d = 2$  und  $k = 1$  gilt.

## 2.2 Fehler und Gitterstrukturen in Röntgenbildern

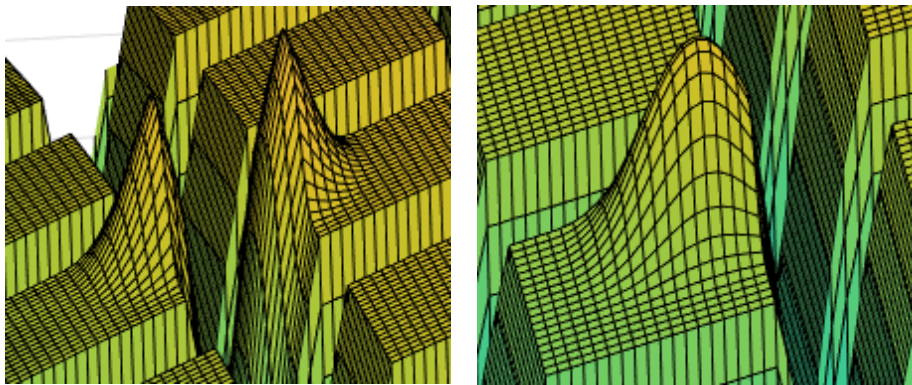
Kommt es beim Gießen des Gussteils zu einem Materialfehler, handelt es sich um einen Materialdefizit. Das führt bei der Röntgenprüfung zu weniger Absorption an der entsprechenden Stelle, sodass mehr Strahlung den Detektor erreicht und der Pixel im resultierenden Bild somit an dieser Stelle höher ist als in seiner lokalen Umgebung. In der Praxis äußert sich das Materialdefizit häufig durch einen Anstieg des Intensitätswertes um 2 % verglichen mit seiner lokalen Umgebung (vgl. Abb. 1.3b) [7]. Neben Materialfehlern und Gitterstrukturen sind weitere Strukturen im Röntgenbild, weshalb ein Schwellenwertansatz in der Praxis nicht zielführend ist und zu einer Fehldetektion führt. Die erhöhten Intensitätswerte bei den Fehlern zeichnen sich in der 3D-Darstellung des Bildes als Maxima aus.



(a) Der Fehler manifestiert sich in Form eines Maximums



(b) Regelmäßige Gitterstrukturen in Form von Vertiefungen



(c) Gitterstrukturen und Fehler mit großer Schnittfläche

(d) Gitterstrukturen und Fehler mit kleiner Schnittfläche

Abbildung 2.2: 3D-Oberflächendarstellung mit (a) einem Fehler in Form eines Maximums, (b) Gitterstrukturen in Form von Vertiefungen, die regelmäßig verlaufen, (c) einem Fehler, der durch ein Gitterelement in zwei Regionen getrennt wird und (d) einem Fehler, der am Rande eines Gitterelements liegt [6].

Anders als bei den Materialfehlern erscheinen die Gitterstrukturen im Röntgenbild dunkler, da sie eine höhere Absorption der Röntgenstrahlung an den entsprechenden Stellen verursachen. Dadurch gelangt weniger Strahlung zum Detektor (vgl. Abb. 2.2b).

Eine große Schnittfläche zwischen einem Fehler und einem Gitterelement führt neben der Teilung des Fehlers auch zu einem höheren Intensitätswert seitens des Gitterelements; siehe Abb. 2.3. Für die Analyse der Methoden in Kapitel 3 und 4 sind Fehlerpixel mit einer erhöhten Intensität von 5% ausreichend.

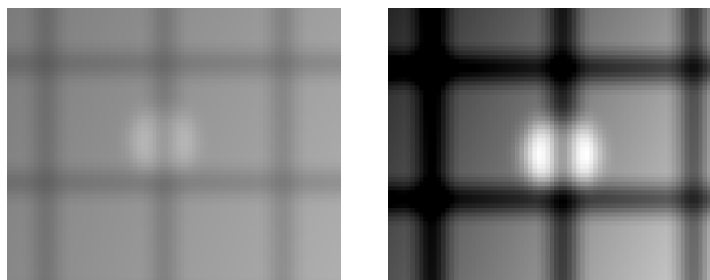


Abbildung 2.3: Ein Beispiel eines synthetischen Eingabebildes für die Röntgenprüfung mit zwei verschiedenen Darstellungsmöglichkeiten. Links: der Fehler ist hellgrau, und durch ein dunkelgraues Gitterelement in zwei Hälften unterteilt. Rechts: andere Fensterung des linken Bildes. Der Fehler ist in Weiß dargestellt und wird ebenfalls durch ein Gitterelement geteilt. Dabei nimmt das Gitterelement von einer schwarzen Färbung eine graue Färbung an, wo sich der Fehler und das Gitterelement schneiden [6].

Diese Arbeit fokussiert sich auf eine spezifische Fragestellung aus der Industrie, weshalb Röntgenbilder kundenspezifischer Gussteile vorliegen. Denen ist man zur Geheimhaltung verpflichtet, weshalb für die Analyse andere Bilder hinzugezogen werden muss. Weiter gestaltet sich die Analyse als schwierig, wenn nur Ausschnitte von den Röntgenbildern betrachtet werden, da diese Ausschnitte nicht ausreichend viele Gitterstrukturen enthalten. Ein Röntgenbild von einem Gussteil enthält neben möglichen Fehlern und Gitterstrukturen noch weitere Strukturen. Diese haben somit einen zu starken Einfluss auf die Interpretation und können eine präzise Analyse der Methoden beeinträchtigen. Aus diesem Grund findet die Analyse der Methoden auf einem synthetischen Bild statt.

### 2.3 Synthetische Bilder für die Analyse der Methoden

Die Erstellung des synthetischen Bildes in dieser Arbeit – auf der die Analyse der Methoden in Kapitel 7 erfolgt – erfordert die Erstellung von drei Schichten:

1. Grundlage, mit oder ohne Objekt schaffen,
2. Fehler einrechnen,
3. Gitterstrukturen einrechnen.

Als Grundlage für ein Röntgenbild wurde ein Bild mit einem Programm namens *Analytical RT Inspection Simulation Tool* (aRTist) generiert. Nach eigenen Angaben handelt es sich um ein Computersimulationsprogramm zur quantitativen Beschreibung der Röntgenprüfung [16]. Das Bild wurde in der Version 2.12.6 64 Bit erstellt. In dieser Arbeit wurde sich für eine Oberfläche entschieden, die in der 3D-Ansicht ein elliptisches Paraboloid ähnelt (vgl. Abb. 2.5). Das bedeutet für das Röntgenbild, dass im Zentrum die höchste Intensität ist und kreisförmig nach außen hin die Intensität sinkt. Dieses Bild hat durch aRTist eine Größe von  $1000 \times 1000$  Pixeln mit Werten zwischen 3000 und 10000. Für eine bessere Darstellungsmöglichkeit ist ein  $512 \times 512$  Pixel großer Ausschnitt ausreichend. Weiter enthält das Bild schon Rauschen, was intern von aRTist stammt. Beim Rauschen handelt es sich um zufällige Intensitätssprünge im Bild, die nach Rudin et al. z.B. während der Aufnahme, der Entstehung oder der Übertragung auftreten kann und in der Realität unvermeidlich ist [17]. Dieses Grundlagenbild ist in Abb. 2.4a zu sehen und wird als  $u_{\text{Base}}$  bezeichnet.

Basierend auf dem Grundlagenbild  $u_{\text{Base}}$  wurden insgesamt sechs kreisförmige Strukturen mit einer erhöhten Intensität erstellt, die Lufteinschlüsse simulieren. Dazu wurde in MATLAB ein Programmcode geschrieben, wo an sechs Stellen im Bild mit jeweils Radius  $r$  die Pixelwerte mit dem Faktor 1.05 multipliziert wurden. Dieser Faktor steht für eine Erhöhung der Intensität um 5%, wie gefordert wird. Diese Lufteinschlüsse manifestieren sich als lokale Maxima im Bild. Details zu den Fehlern sind in Tabelle 2.1.

Der letzte Schritt ist das Einrechnen der Gitterstrukturen. Dazu wurde im Fehlerbild  $u_{\text{Err}}$  ein Gittermuster eingerechnet, mit einer Gitterweite von  $h = (30, 30)$  und einer

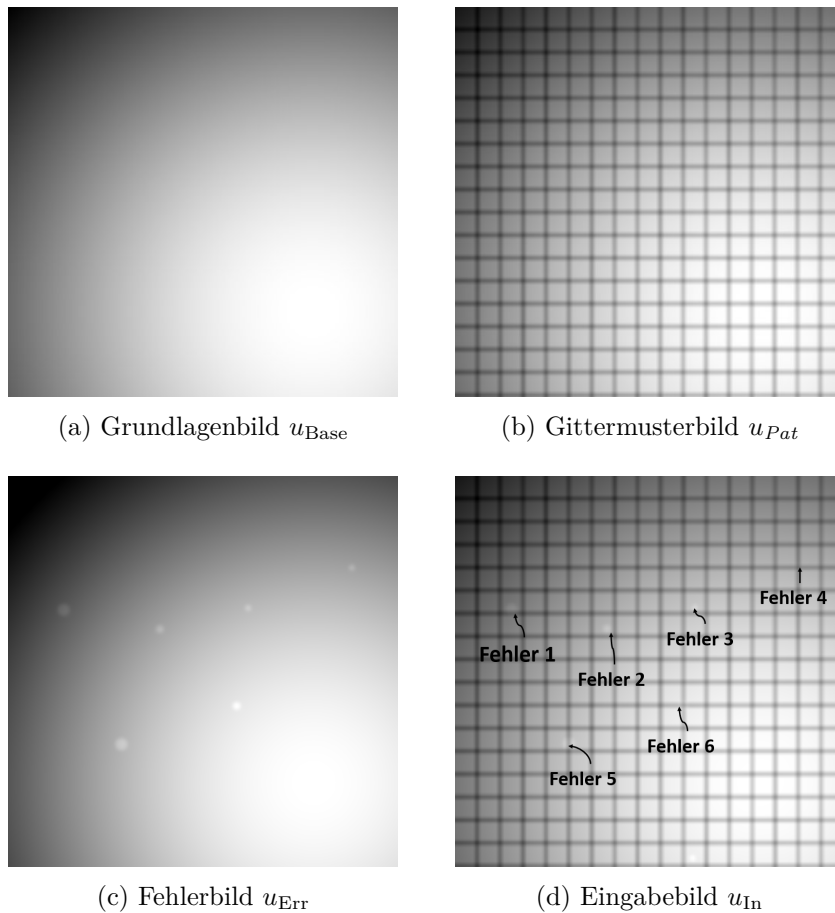


Abbildung 2.4: Darstellung der Schichten des synthetischen Bildes: (a) zeigt das Grundlagenbild  $u_{\text{Base}}$  mit Rauschen, in (b) sind zusätzlich zum Grundlagenbild Gitterstrukturen mit Gitterweite  $h = (30, 30)$  eingerechnet und ergibt das Gittermusterbild  $u_{\text{Pat}}$ ; (c) zeigt das Fehlerbild  $u_{\text{Err}}$  mit sechs kreisförmigen Lufteneinschlüsse; (d) zeigt das Fehlerbild mit den Gitterstrukturen in einem mit Beschriftung der Fehlernummerierung. Dieses wird als Eingabebild  $u_{\text{In}}$  bezeichnet [6].



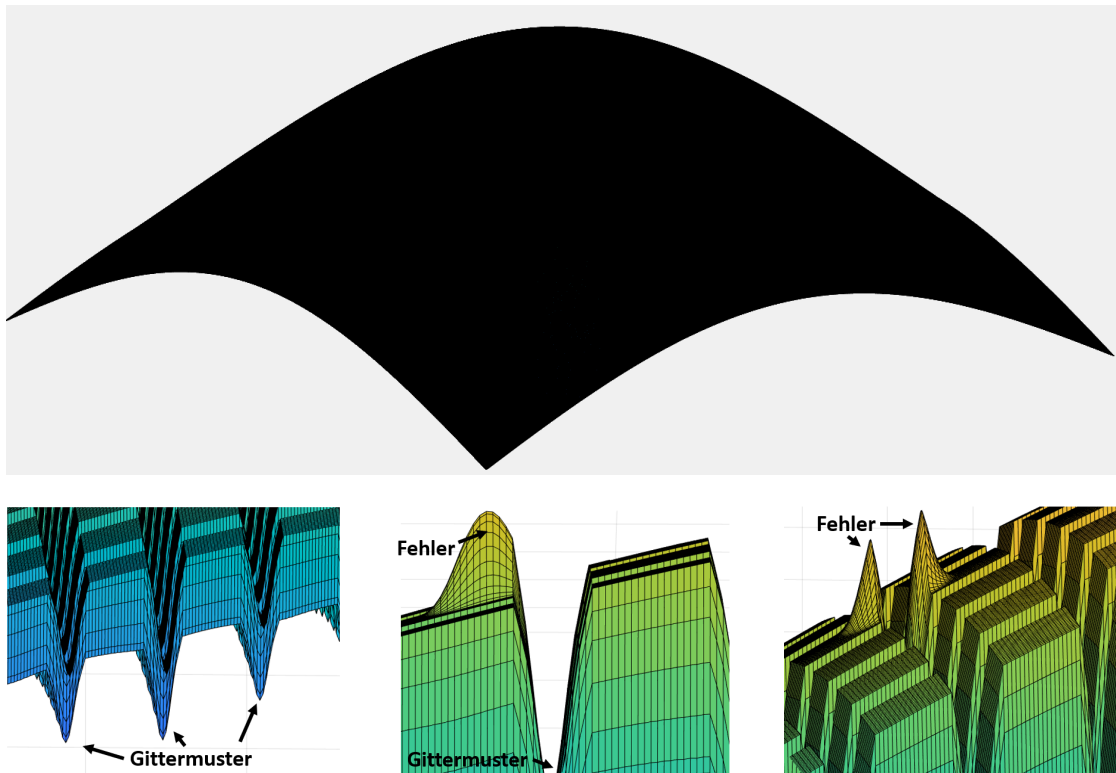


Abbildung 2.5: Oben: Darstellung des Grundlagenbildes  $u_{\text{Base}}$  als 3D-Oberflächenplot; unten links: der Ausschnitt zeigt detailliert die gitterförmigen Strukturen in Form von Minima auf einem 3D-Oberflächenplot; unten Mitte: in diesem Ausschnitt ist ein Fehler in Form eines Maximums zu sehen, der sich an einem Gitterelement befindet auf einem 3D-Oberflächenplot und unten rechts: das Gitterelement teilt den Fehler in zwei Teile [6].

	Mittelpunkt	Radius $r$ [pix]
Fehler 1	(175, 75)	8
Fehler 2	(200, 200)	5
Fehler 3	(172, 315)	4
Fehler 4	(120, 450)	4
Fehler 5	(350, 150)	8
Fehler 6	(300, 300)	5

Tabelle 2.1: Position und Größe der Luftschlüsse in den synthetischen Bildern  $u_{\text{Err}}$  und  $u_{\text{In}}$

Gitterbreite von 1. Dafür wurde jeder dreißigste Wert von  $u_{\text{Base}}$  mit dem Faktor 0.55 multipliziert und dieses Ergebnis wurde am entsprechenden Pixel von  $u_{\text{Err}}$  eingesetzt. Dieser Schritt führt dazu, dass im Fall einer Überdeckung der Fehler durch das Gittermuster, das Gittermuster nicht von den schon erhöhten Intensitätswerten der Fehler beeinflusst wurde. In Abb. 2.5 wird anschaulich das Verhalten der Gittermuster und die überdeckten Fehler in 3D-Oberflächenplots dargestellt. Das Ergebnis wird als Eingabebild  $u_{\text{In}}$  bezeichnet und ist in Abb. 2.4d zu sehen.

Für eine detaillierte Analyse für das Verhalten der Gitterstrukturen wird das Gittermusterbild  $u_{\text{Pat}}$  eingeführt, das vor allem bei der Analyse der Cepstralanalyse in Kapitel 3 Anwendung findet. Dabei handelt es sich um das Grundlagenbild  $u_{\text{Base}}$  mit eingerechneten Gittermustern wie eben beschrieben und ist in Abb. 2.4b dargestellt. Damit starke Intensitätssprünge bei der Erstellung der Bilder vermieden werden, wurde anschließend eine Glättung mit einem Gaußfilter und einer Standardabweichung von  $\sigma = 2.5$  über das gesamte Bild angewendet (die Filterung sorgt für eine veränderte Form und Größe der Fehler). Das Bild  $u_{\text{In}}$  ist das Eingabebild für die Analyse der in Kapitel 6.1 und 6.2 vorgestellten Methoden.

---

## Kapitel 3: Cepstralanalyse

Bei der Segmentierung von Fehlern in Röntgenbildern ist ein häufig auftretendes Problem, dass vor allem kleinere Fehler von anderen Strukturen überdeckt werden. Ein Beispiel dieser Struktur ist ein Gittermuster, wie es in Abb. 1.2, 1.3, 1.4 oder 2.3 dargestellt ist. Durch die Bearbeitung der Gittermuster im Bildraum gehen die Informationen über kleine überdeckte Fehler verloren. Ein effektiverer Ansatz besteht darin, die Bearbeitung der Gittermuster im Frequenzraum durchzuführen. Hier lassen sich die größeren Frequenzen der Gittermuster von den kleineren Frequenzen der Fehler getrennt bearbeiten. Eine auf der Fouriertransformation basierende Transformation ist das Cepstrum, das die inverse Fouriertransformation auf einem logarithmierten Spektrum des Bildes berechnet.

Das Cepstrum wurde von Bogert, Healy und Tukey in ihrer Arbeit „The Quefreny Alalysis of Time Series for Echos: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking“ eingeführt und 1963 in *Journal of the Acoustical Society of America* veröffentlicht [18]. Bogert, Healy und Tukey waren in der Forschungsabteilung Bell Laboratories in den Vereinigten Staaten von Amerika tätig und entwickelten das Cepstrum als eine Methode zur Analyse von eindimensionalen Echosignalen. Der Bedarf entstand durch die Analyse seismischer Daten.

Seit seiner Einführung hat sich das Cepstrum zu einem wichtigen Werkzeug in verschiedenen Bereichen der Signalverarbeitung entwickelt, einschließlich der Spracherkennung, Bildverarbeitung und Akustik [18, 19, 20, 21]. Bei dem Cepstrum handelt es sich um eine Transformation des Spektrums eines Signals, welches ursprünglich mit der diskreten Fourier-Transformation und der inversen diskreten Fourier-Transformation berechnet wurde. Das Cepstrum ermöglicht es, verschiedene Eigenschaften eines Signals zu analysieren, wie zum Beispiel die Periodizität und die Frequenzverteilung.

Durch die Transformation eines Spektrums wird ein neuer Raum betrachtet, der in diesem Fall als *Quefreny* bezeichnet wird. Die Namensgebung des Cepstrums und Quefreny entstanden nicht willkürlich, sondern wurden jeweils als Anagramme der englischen Wörter, mit denen sie in Analogie stehen, gebildet: Cepstrum leitet sich aus dem englischen Wort *Spectrum* ab, sowie Quefreny aus dem englischen Wort *Frequency*. Noch im 20. Jahrhundert gab es viele Debatten über die Aussprache der neu eingeführten Begriffe. Einige Autoren entschieden sich deshalb für die Schreibweise "Kepstrum", da sie damit die Aussprache verdeutlichen wollten [22]. Bogert et al. führten neben dem Cepstrum noch weitere Anagramme ein, welche sich abgesehen vom *Liftering* (Analogie zum *filtering*) nicht etabliert haben.

Die anfänglichen Ergebnisse basierten auf eindimensionalen Signalen, jedoch nahm auch das Interesse an der Erforschung von zweidimensionalen Signalen im Laufe der Zeit zu. Die Analyse in dieser Arbeit bezieht sich auf die zweidimensionale Cepstralanalyse.

Das Cepstrum ist in vielen Bereichen der Forschung von Bedeutung, weshalb verschiedene Definitionen in der Literatur auftauchen. In dieser Arbeit wird die Definition des zweidimensionalen Cepstrums von Çakir und Çetin [23] verwendet. Weiter sagt [24] aus, dass das Cepstrum im Allgemeinen für jede lineare separable Transformation anwendbar ist.

Die zweidimensionale diskrete Fouriertransformation spielt eine zentrale Rolle in dieser Arbeit. Aus diesem Grund kommt zuerst eine kurze Einführung in die Fourieranalyse.

### 3.1 Die diskrete zweidimensionale Fourier-Transformation

Bei der Fouriertransformation handelt es sich um eine Abbildung, die ein Signal vom Ortsraum in den Frequenzraum transformiert. Sie basiert auf dem Konzept, dass jedes periodische Signal als eine Summe von Sinus- und Kosinusfunktionen unterschiedlicher Frequenzen dargestellt werden kann. Diese Frequenzkomponenten können verwendet werden, um verschiedene Merkmale des Bildes zu analysieren und zu manipulieren, wie zum Beispiel den Kontrast, die Schärfe oder die Textur. Die zweidimensionale Fouriertransformation ist ein wichtiger Bestandteil vieler Bildverarbeitungsverfahren und ermöglicht es, komplexe Datenanalysen durchzuführen, um Informationen in Bildern zu extrahieren und zu manipulieren.

Dazu wird die Definition nach Bredis et al. [4]. betrachtet:

**Definition 3.1.** *Es sei  $u \in \mathbb{C}^{m_1 \times m_2}$  in diskreter Darstellung gegeben. Dann ist die zweidimensionale diskrete Fouriertransformation von  $u$  ist  $\hat{u} \in \mathbb{C}^{m_1 \times m_2}$  definiert durch:*

$$\begin{aligned} \hat{u}_{k,l} &:= \frac{1}{m_1 \cdot m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot e^{-2\pi i \cdot \frac{l \cdot n}{m_2}} \cdot u_{m,n} \\ &= \frac{1}{m_1 \cdot m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} e^{-2\pi i \cdot \left(\frac{k \cdot m}{m_1} + \frac{l \cdot n}{m_2}\right)} \cdot u_{m,n} \end{aligned} \quad (2)$$

für  $k = 0, \dots, m_1 - 1$  und  $l = 0, \dots, m_2 - 1$ .

$\hat{u}$  kann invertiert werden durch:

$$u_{m,n} := \frac{1}{m_1 \cdot m_2} \sum_{k=0}^{m_1-1} \sum_{l=0}^{m_2-1} e^{2\pi i \cdot \frac{m \cdot k}{m_1}} \cdot e^{2\pi i \cdot \frac{n \cdot l}{m_2}} \cdot \hat{u}_{k,l}$$

$$= \frac{1}{m_1 \cdot m_2} \sum_{k=0}^{m_1-1} \sum_{l=0}^{m_2-1} e^{2\pi i \cdot (\frac{m \cdot k}{m_1} + \frac{n \cdot l}{m_2})} \cdot \hat{u}_{k,l} \quad (3)$$

für  $m = 0, \dots, m_1 - 1$  und  $n = 0, \dots, m_2 - 1$ .

Die Fouriertransformation ist eine lineare Transformation, das bedeutet speziell für die diskrete Fouriertransformation  $\mathcal{F}_n : \mathbb{C}^n \rightarrow \mathbb{C}^n$  dass die sich als Matrix-Vektor-Produkt formulieren lässt:

$$\mathcal{F}_n(u) = \frac{1}{\sqrt{n}} \underbrace{\begin{pmatrix} \omega_n^0 & \omega_n^0 & \omega_n^0 & \dots & \omega_n^0 \\ \omega_n^0 & \omega_n^{1 \cdot 1} & \omega_n^{1 \cdot 2} & \dots & \omega_n^{1 \cdot (n-1)} \\ \omega_n^0 & \omega_n^{2 \cdot 1} & \omega_n^{2 \cdot 2} & \dots & \omega_n^{2 \cdot (n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_n^0 & \omega_n^{(n-1) \cdot 1} & \omega_n^{(n-1) \cdot 2} & \dots & \omega_n^{(n-1) \cdot (n-1)} \end{pmatrix}}_{=\mathcal{F}_n} \cdot u \quad (4)$$

für einen Vektor  $u \in \mathbb{C}^n$  und  $\omega_n^k = e^{-2 \cdot \pi \cdot i \frac{k}{n}}$ .

Die zweidimensionale diskrete Fouriertransformation von  $u \in \mathbb{R}^{m_1 \times m_2}$  lässt sich als Matrix-Vektor-Produkt formulieren:

$$\hat{u}^* = \mathcal{F}_{m_1} \otimes \mathcal{F}_{m_2} \cdot \tilde{u} \quad (5)$$

dabei sei  $\tilde{u}$  eine vektorisierende Darstellung von  $u$  mittels der *Column-Major-Anordnung* (vgl. Bem. 3.4) und  $\otimes$  bezeichne das *Kronecker-Produkt* (vgl. Bem. 3.2). Bei  $\hat{u}_h^*$  handelt es sich um einen komplexen Vektor der Länge  $m_1 \cdot m_2$ . Um wieder das Bild in der Darstellung  $\hat{u} \in \mathbb{R}^{m_1 \times m_2}$  zu erhalten, muss die Column-Major-Anwendung invertiert werden.

**Bemerkung 3.2.** Das Kroneckerprodukt  $\otimes : \mathbb{C}^{m \times n} \times \mathbb{C}^{p \times r} \rightarrow \mathbb{C}^{m \cdot p \times n \cdot r}$  mit  $(A, B) \mapsto A \otimes B$  ist eine Abbildung. Dann gilt für zwei Matrizen  $A \in \mathbb{R}^{m \times n}$  und  $B \in \mathbb{R}^{p \times r}$  mit  $a_{i,j} \in A$ ,  $i = 1, \dots, m$  und  $j = 1, \dots, n$ :

$$A \otimes B := \begin{pmatrix} a_{1,1} \cdot B & \dots & a_{1,n} \cdot B \\ \vdots & \ddots & \vdots \\ a_{m,1} \cdot B & \dots & a_{m,n} \cdot B \end{pmatrix} \in \mathbb{R}^{m \cdot p \times n \cdot r} \quad (6)$$

Für ein besseres Verständnis ein Beispiel:

**Beispiel 3.3.** Es seien  $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{pmatrix} \in \mathbb{R}^{3 \times 2}$  und  $B = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ .

Dann ist

$$A \otimes B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{pmatrix} \otimes \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} & 1 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \\ 2 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} & 2 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \\ 3 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} & 3 \cdot \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 4 & 2 & 4 \\ 6 & 8 & 6 & 8 \\ 4 & 8 & 4 & 8 \\ 12 & 16 & 12 & 16 \\ 6 & 12 & 6 & 12 \\ 18 & 24 & 18 & 24 \end{pmatrix}$$

**Bemerkung 3.4.** Für eine Matrix  $A \in \mathbb{C}^{m \times n}$ ,  $A = \begin{pmatrix} a_{1,1} \cdot B & \dots & a_{1,n} \cdot B \\ \vdots & \ddots & \vdots \\ a_{m,1} \cdot B & \dots & a_{m,n} \cdot B \end{pmatrix}$  heißt

die Darstellung in  $A = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \\ a_{1,2} \\ \vdots \\ a_{m,2} \\ \vdots \\ a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}$  die Column-Major Darstellung der Matrix  $A$ .

Zum Verständnis ein Beispiel:

**Beispiel 3.5.** Für  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$  ist die Column-Major Darstellung von  $A = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 5 \\ 3 \\ 6 \end{pmatrix}$

### 3.2 Definition und Eigenschaften des Power-Cepstrums

Wie bereits erwähnt, gibt es in der Literatur verschiedene Definitionen des Cepstrums. In dieser Arbeit lehnt sich an die Definition, wie sie Çakir et al. [10] einführen:

**Definition 3.6.** Es sei  $u \in \mathbb{R}^{m_1 \times m_2}$  und  $\mathcal{F} : \mathbb{C}^{m_1 \times m_2} \rightarrow \mathbb{C}^{m_1 \times m_2}$  sei die zweidimensionale diskrete Fourier-Transformation. Dann heißt  $\mathcal{C} : \mathbb{C}^{m_1, m_2} \rightarrow \mathbb{C}^{m_1, m_2}$  mit

$$\mathcal{C}(u) = \mathcal{F}^{-1}(\ln(|\mathcal{F}(u)|^2)) \quad (7)$$

das zweidimensionale Power-Cepstrum von  $u$ .

Dabei ist zu verstehen, dass im gesamten Kapitel der Betrag, das Quadrat und der natürliche Logarithmus elementweise angewendet werden, d.h. für eine Matrix

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n} \quad (8)$$

gilt

$$|A| = \begin{pmatrix} |a_{1,1}| & \dots & |a_{1,n}| \\ \vdots & \ddots & \vdots \\ |a_{m,1}| & \dots & |a_{m,n}| \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (9)$$

und

$$A^2 = \begin{pmatrix} a_{1,1}^2 & \dots & a_{1,n}^2 \\ \vdots & \ddots & \vdots \\ a_{m,1}^2 & \dots & a_{m,n}^2 \end{pmatrix} \in \mathbb{C}^{m \times n} \quad (10)$$

und

$$\ln(A) = \begin{pmatrix} \ln(a_{1,1}) & \dots & \ln(a_{1,n}) \\ \vdots & \ddots & \vdots \\ \ln(a_{m,1}) & \dots & \ln(a_{m,n}) \end{pmatrix} \in \mathbb{C}^{m \times n} \quad (11)$$

Der natürliche Logarithmus einer komplexen Zahl  $r \cdot e^{i\varphi} \in \mathbb{C}$ ,  $r \in \mathbb{R}$ ,  $\varphi \in (-\pi, \pi]$  ist:

$$\begin{aligned} \ln(r \cdot e^{i\varphi}) &= \ln(r) + \ln(e^{i\varphi}) \\ &= \left\{ \ln(r) + i(\varphi + 2k\pi), k \in \mathbb{Z} \right\} \end{aligned} \quad (12)$$

Anwendung an einem Beispiel:

Für  $A = \begin{pmatrix} 3 + 4i & 2 - i & 4 \\ 2 + 3i & 4i & 1 - 5i \end{pmatrix} \in \mathbb{C}^{2 \times 3}$  gilt in dieser Arbeit

$$|A| = \begin{pmatrix} 5 & \sqrt{3} & 4 \\ \sqrt{13} & 4 & \sqrt{26} \end{pmatrix} \in \mathbb{R}^{2 \times 3}.$$

und

$$A^2 = \begin{pmatrix} -7 + 24i & 3 - 4i & 16 \\ -5 + 12i & -16 & -24 - 10i \end{pmatrix} \in \mathbb{C}^{2 \times 3}.$$

und

$$\begin{aligned} \ln(A) &= \begin{pmatrix} \ln(5) + \arg(3 + 4i)i & \ln(\sqrt{3}) + \arg(2 - i)i & \ln(4) \\ \ln(\sqrt{13}) + \arg(2 + 3i)i & \ln(4) + \pi/2i & \ln(\sqrt{26}) + \arg(1 - 5i)i \end{pmatrix} \\ &= \begin{pmatrix} 1.6094 + 0.9273i & 0.8047 - 0.4636i & 1.3863 \\ 1.2825 + 0.9828i & 1.3863 + 1.5708i & 1.6290 - 1.3734i \end{pmatrix} \in \mathbb{C}^{2 \times 3} \end{aligned}$$

Basierend auf den Eigenschaften der Fouriertransformation, hat auch das Cepstrum verschiedene Eigenschaften. Bei der Literaturrecherche konnte nur eine Arbeit gefunden werden, die die Untersuchung der Eigenschaften des zweidimensionalen Cepstrums vorstellt. Rom hat in seiner Arbeit [24] acht Eigenschaften des Cepstrums vorgestellt, ohne Beweise und auf die Zwischenschritte einzugehen. Die Definition des Cepstrums in dieser Arbeit beruht auf der von Çakir et al. [10] weshalb es zu Abweichungen zu Roms Arbeit kommt.

Bevor es zu den Eigenschaften geht, werden vorher noch zwei Abbildungen benötigt, die zum einen die Addition einer Matrix mit einem Skalar ermöglichen und eine Multiplikation zweier Vektoren aus unterschiedlichen Dimensionen:

### Elementweise Addition von Matrix und Skalar

Es sei  $\oplus : \mathbb{R} \times \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{m \times n}$  mit  $(\alpha, A) \mapsto \alpha \oplus A$  eine Abbildung. Dann gilt für eine Matrix  $A \in \mathbb{C}^{m \times n}$  und ein Skalar  $\alpha \in \mathbb{R}$ :

$$\alpha \oplus A = \begin{pmatrix} \alpha + a_{1,1} & \alpha + a_{1,2} & \cdots & \alpha + a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha + a_{m,1} & \cdots & \cdots & \alpha + a_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n} \quad (13)$$

Mit der Kommutativität der Skalar Addition folgt  $\alpha \oplus A = A \oplus \alpha$ .



Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathbb{C}^{3 \times 3}, \alpha = 5$$

dann ist

$$\alpha \oplus A = \begin{pmatrix} 6 & 7 & 8 \\ 9 & 10 & 11 \\ 12 & 13 & 14 \end{pmatrix}$$

### Multiplikation von zwei Vektoren aus unterschiedlichen Dimensionen

Es sei  $\cdot : \mathbb{C}^m \times \mathbb{C}^n \rightarrow \mathbb{C}^{m \times n}$  mit  $(u, v) \mapsto u \cdot v^T$  eine Abbildung. Dann gilt für zwei Vektoren  $u \in \mathbb{R}^m, v \in \mathbb{R}^n$ :

$$u \cdot v^T = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \cdot \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} u_1 \cdot v_1 & u_1 \cdot v_2 & \cdots & u_1 \cdot v_n \\ u_2 \cdot v_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ u_m \cdot v_1 & \cdots & \cdots & u_m \cdot v_n \end{pmatrix} \in \mathbb{R}^{m_1 \times m_2} \quad (14)$$

Beispiel:

Für  $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$  und  $v = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \in \mathbb{R}^2$ . Dann gilt

$$u \cdot v^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{pmatrix}$$

**Satz 3.7.** Für  $f, g \in \mathbb{C}^{m_1 \times m_2}$ ,  $h \in \mathbb{C}^{m_1}$ ,  $k \in \mathbb{C}^{m_2}$  und  $\alpha \in \mathbb{R}$  gelten:

1. wenn  $f$  separable ist mit  $f = h \cdot k^T \Rightarrow \mathcal{C}(f) = \mathcal{C}(h) \oplus \mathcal{C}(k)^T$
2.  $\mathcal{C}(a \cdot f) = \mathcal{C}(f) \oplus \ln|a|^2$
3.  $\mathcal{C}(f * g) = \mathcal{C}(f) + \mathcal{C}(g)$

**Beweis:** zu 1.

Es sei  $f \in \mathbb{C}^{m_1 \times m_2}$ ,  $h \in \mathbb{C}^{m_1}$ ,  $k \in \mathbb{C}^{m_2}$  mit  $f = h \otimes k^T$  und  $\mathcal{F}$  bezeichne die zweidimensionale Fourier-Transformation, dann gilt:

$$\begin{aligned}
 [\mathcal{F}(f)]_{k,l} &= [\mathcal{F}(h \otimes k^T)]_{k,l} \\
 &= \frac{1}{m_1 \cdot m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot e^{-2\pi i \cdot \frac{l \cdot n}{m_2}} \cdot (h_m \otimes k_n^T) \\
 &= \frac{1}{m_1} \cdot \frac{1}{m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} (e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot h_m) \otimes (e^{-2\pi i \cdot \frac{l \cdot n}{m_2}} \cdot k_n^T) \\
 &= \underbrace{\left[ \frac{1}{m_1} \cdot \sum_{m=0}^{m_1-1} (e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot h_m) \right]}_{[\mathcal{F}(h)]_k} \otimes \underbrace{\left[ \frac{1}{m_2} \sum_{n=0}^{m_2-1} (e^{-2\pi i \cdot \frac{l \cdot n}{m_2}} \cdot k_n^T) \right]}_{[\mathcal{F}(k^T)]_l} \\
 &= \underbrace{[\mathcal{F}(h)]_k \otimes [\mathcal{F}(k^T)]_l}_{(*)}
 \end{aligned}$$

für  $k = 1, \dots, m_1$  und  $l = 1, \dots, m_2$ . Diese Umformungsschritte sind möglich, denn  $\otimes$  ist nach Horn [25] assoziativ. Weiter gilt für ein Signal  $h \in \mathbb{C}^{m_1}$  (genauer  $h \in \mathbb{C}^{m_1 \times 1}$ ):

$$\begin{aligned}
 \hat{h}_{k,l} &= \frac{1}{m_1 \cdot 1} \sum_{m=0}^{m_1-1} \sum_{n=0}^0 e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot e^{-2\pi i \cdot \frac{l \cdot n}{m_2}} \cdot (h_{m,n}) \\
 &= \frac{1}{m_1 \cdot 1} \sum_{m=0}^{m_1-1} e^{-2\pi i \cdot \frac{k \cdot m}{m_1}} \cdot 1 \cdot (h_m)
 \end{aligned}$$

$\Rightarrow$  Die zweidimensionale DFT von  $h$  liefert dasselbe Ergebnis wie die eindimensionale DFT von  $h$  für ein  $h \in \mathbb{C}^{m_1}$ . Aus diesem Grund muss in (\*) keine weitere Annahme für  $\mathcal{F}$  getroffen werden.

Dann gilt weiter:

$$\begin{aligned}
 \mathcal{C}(f) &= \mathcal{F}^{-1}(\ln(|\mathcal{F}(f)|^2)) \\
 &\stackrel{(*)}{=} \mathcal{F}^{-1}(\ln(|\mathcal{F}(h) \cdot \mathcal{F}(k)|^2)) \\
 &= \mathcal{F}^{-1}(\ln(|\mathcal{F}(h)|^2 \cdot |\mathcal{F}(k)|^2)) \\
 &= \mathcal{F}^{-1}(\ln(|\mathcal{F}(h)|^2) + \ln(|\mathcal{F}(k)|^2)) \\
 &= \mathcal{F}^{-1}(\ln|\mathcal{F}(h)|^2 + \ln|\mathcal{F}(k)|^2) \\
 &= \mathcal{F}^{-1}(\ln|\mathcal{F}(h)|^2) + \mathcal{F}^{-1}(\ln|\mathcal{F}(k)|^2) \\
 &= \mathcal{F}^{-1}(\ln|\mathcal{F}(h)|^2) + \mathcal{F}^{-1}(\ln|\mathcal{F}(k)|^2)
 \end{aligned}$$

$$\begin{aligned}
 &= \underbrace{\mathcal{F}^{-1}(\ln|\mathcal{F}(h)|^2)}_{\mathcal{C}(h)} + \underbrace{\mathcal{F}^{-1}(\ln|\mathcal{F}(k)|^2)}_{\mathcal{C}(k)} \\
 &= \mathcal{C}(h) + \mathcal{C}(k)
 \end{aligned}$$

□

**Beweis:** zu 2.

Es seien  $f \in \mathbb{C}^{m_1 \times m_2}$ ,  $\alpha \in \mathbb{R}$  und  $\mathcal{F}$  bezeichne die zweidimensionale Fourier-Transformation. Dann gilt:

$$\begin{aligned}
 [\mathcal{F}(\alpha \cdot f)]_{k,l} &= \frac{1}{m_1 \cdot m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} e^{-2\pi i \cdot (\frac{k \cdot m}{m_1} + \frac{l \cdot n}{m_2})} \cdot \alpha \cdot f_{m,n} \\
 &= \alpha \cdot \underbrace{\frac{1}{m_1 \cdot m_2} \sum_{m=0}^{m_1-1} \sum_{n=0}^{m_2-1} e^{-2\pi i \cdot (\frac{k \cdot m}{m_1} + \frac{l \cdot n}{m_2})} \cdot f_{m,n}}_{\mathcal{F}(f)} \\
 &= \underbrace{\alpha \cdot \mathcal{F}(f)}_{(*)}
 \end{aligned}$$

Weiter sei nun:

$$\begin{aligned}
 \mathcal{C}(\alpha \cdot f) &= \mathcal{F}^{-1}(\ln(|\mathcal{F}(\alpha \cdot f)|^2)) \\
 &\stackrel{(*)}{=} \mathcal{F}^{-1}(\ln(|\alpha \cdot \mathcal{F}(f)|^2)) \\
 &= \mathcal{F}^{-1}(\ln(|\alpha|^2 \cdot |\mathcal{F}(f)|^2)) \\
 &= \mathcal{F}^{-1}(\ln(|\alpha|^2) \oplus \ln(|\mathcal{F}(f)|^2)) \\
 &= \mathcal{F}^{-1}(\ln(|\alpha|^2)) \oplus \mathcal{F}^{-1}(\ln(|\mathcal{F}(f)|^2)) \\
 &= \ln(|\alpha|^2) + \mathcal{C}(f)
 \end{aligned}$$

□

**Beweis:** zu 3.

Seien  $f, g \in \mathbb{C}^{m_1 \times m_2}$ ,  $\mathcal{F}$  bezeichne die zweidimensionale Fourier-Transformation und  $*$  bezeichne die Faltung. Dann gilt:

$$\begin{aligned}
 \mathcal{C}(f * g) &= \mathcal{F}^{-1}(\ln|\mathcal{F}(f) \cdot \mathcal{F}(g)|^2) \\
 &= \mathcal{F}^{-1}(\ln|\mathcal{F}(f)|^2) + \ln(|\mathcal{F}(g)|^2) \\
 &= \mathcal{F}^{-1}(\ln|\mathcal{F}(f)|^2) + \mathcal{F}^{-1}(\ln|\mathcal{F}(g)|^2) \\
 &= \mathcal{C}(f) + \mathcal{C}(g)
 \end{aligned}$$

□

### 3.3 Transformation vom Quefrequency-Raum in den Bildraum

Der Quefrequency-Raum bietet eine andere Darstellung des Signals und enthält andere Informationen als der Bildraum. Hier werden spezifische Merkmale oder Muster des Signals analysiert, betont oder verarbeitet, die im Bildraum möglicherweise nicht so leicht erkennbar sind [26].

Eine Transformation vom Quefrequency-Raum in den Bildraum, und damit eine Inverse des Power-Cepstrums, so wie sie in 3.6 definiert ist, ist aufgrund des Informationsverlustes der Phase nicht möglich. Auch eine alternative Transformation vom Quefrequency-Raum in den Bildraum ist während der Literaturrecherche nicht gefunden worden. Diese Transformation stellt einen wichtigen Aspekt dar, weshalb diese Arbeit eine Möglichkeit vorstellt, in der eine Transformation von dem Quefrequency-Raum eines modifizierten Power-Cepstrums in den Bildraum möglich ist. Dazu wird folgende Modifizierung des Power-Cepstrums aus 3.6 betrachtet:

**Definition 3.8.** Sei  $u \in \mathbb{C}^{m_1 \times m_2}$ ,  $\mathcal{F}$  bezeichne die zweidimensionale Fourier-Transformation und  $\Phi_u$  das Phasenspektrum von  $u$ .

Dann heißt  $\mathcal{E}_{mod} : \mathbb{R}^{m_1 \times m_2} \rightarrow \mathbb{C}^{m_1 \times m_2} \times \Phi_u$  mit

$$\mathcal{E}_{mod}(u) = (\mathcal{F}^{-1}(\ln(|\mathcal{F}(u)|^2)), \Phi_u) \quad (15)$$

das modifizierte Power-Cepstrum von  $u$ .

Für die Konstruktion eines approximativen Inversen des modifizierte Power-Cepstrums sind noch folgenden Anmerkungen zu beachten:

#### Elementweise Anwendung der Exponentialfunktion

Für eine Matrix

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n}$$

gilt

$$e^A = \begin{pmatrix} e^{a_{1,1}} & \dots & e^{a_{1,n}} \\ \vdots & \ddots & \vdots \\ e^{a_{m,1}} & \dots & e^{a_{m,n}} \end{pmatrix} \in \mathbb{C}^{m \times n} \quad (16)$$

### Punktweise Multiplikation zweier Matrizen

Für zwei Matrizen  $A, B \in \mathbb{C}^{m \times n}$  mit

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n}$$

und

$$B = \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n}$$

bezeichne  $\odot : \mathbb{C}^{m \times n} \times \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{m \times n}$ ,  $(A, B) \mapsto A \odot B$  eine Abbildung, für die dann gilt:

$$A \odot B = \begin{pmatrix} a_{1,1} \cdot b_{1,1} & \dots & a_{1,n} \cdot b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} \cdot b_{m,1} & \dots & a_{m,n} \cdot b_{m,n} \end{pmatrix} \quad (17)$$

Beispiel:

Für  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$  und  $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$  gilt

$$A \odot B = \begin{pmatrix} 5 & 12 \\ 21 & 32 \end{pmatrix}$$

Die punktweise Multiplikation wird auch *Hadamard-Produkt* genannt.

**Definition 3.9.** Für ein gegebenes Power-Cepstrum und Phasenspektrums eines Bildes  $u_h$  mit den Dimensionen  $m_1 \times m_2$  definieren wir  $\hat{\mathcal{C}} : \mathbb{C}^{m_1 \times m_2} \times \Phi_u \rightarrow \mathbb{C}^{m_1 \times m_2}$  mit

$$\hat{\mathcal{C}}(\mathcal{C}(u), \Phi_u) := \mathcal{F}^{-1}(\sqrt{e^{\mathcal{F}(\mathcal{C}(u))}} \odot e^{i \cdot \Phi_u}) \quad (18)$$

Der Beweis, dass es sich bei  $\hat{\mathcal{C}}$  um das Inverse des modifizierten Power-Cepstrum liefert diese Arbeit nicht. Sie zeigt, dass es sich bei  $\hat{\mathcal{C}}$  um einen möglichen Kandidaten für das Inverse des modifizierten Power-Cepstrum handelt.

**Satz 3.10.**  $\hat{\mathcal{C}}$  ist ein möglicher Kandidat für eine approximative Inverse des modifizierten Power-Cepstrums  $\mathcal{C}_{mod}^{-1}$ .

**Beweis:** Es gilt zu zeigen  $(\hat{\mathcal{C}} \circ \mathcal{C}_{\text{mod}})(u) = u$  für ein diskretes Grauwertbild  $u$ , wobei  $\circ$  für die allgemeine Operation einer Verknüpfung steht.

Es sei  $u$  ein beliebiges diskretes Grauwertbild gegeben und  $\mathcal{C}, \mathcal{C}_{\text{mod}}$  und  $\hat{\mathcal{C}}$  sind wie in 3.6, 3.8 und 3.9 definiert. Dann gilt:

$$\begin{aligned} (\hat{\mathcal{C}} \circ \mathcal{C}_{\text{mod}})(u) &= \mathcal{F}^{-1}(\sqrt{e^{\mathcal{F}(\mathcal{F}^{-1}(\ln(|\mathcal{F}(u)|^2))}}) \odot e^{i \cdot \Phi_u}) \\ &= \mathcal{F}^{-1}(\sqrt{e^{\ln(|\mathcal{F}(u)|^2)}} \odot e^{i \cdot \Phi_u}) \\ &= \mathcal{F}^{-1}(\sqrt{|\mathcal{F}(u)|^2} \odot e^{i \cdot \Phi_u}) \end{aligned}$$

$$Z := \mathcal{F}(u), z \in Z$$

$$= \mathcal{F}^{-1}(|Z| \odot e^{i \cdot \Phi_u})$$

$$|z| = \sqrt{\Re(z)^2 + \Im(z)^2} = r_z, \quad z = r_z e^{i \cdot \varphi_z}, \quad \varphi \in \Phi_u$$

$$= \mathcal{F}^{-1}(r_z \odot e^{i \cdot \Phi_u})$$

$$= \mathcal{F}^{-1}(Z)$$

$$\mathcal{F}(u) = Z$$

$$= \mathcal{F}^{-1}(\mathcal{F}(u))$$

$$= u$$

$\Rightarrow \hat{\mathcal{C}}$  ist eine Inverse des modifizierten Power-Cepstrums  $\mathcal{C}_{\text{mod}}$ , also gilt  $\hat{\mathcal{C}} = \mathcal{C}_{\text{mod}}^{-1}$ . □

### 3.4 Anwendungen der Cepstralanalyse zur Gitterreduktion von Röntgenbildern

Bei der 2D-DFT wird das Bild vom Ortsbereich in den Frequenzbereich transformiert und erzeugt somit ein Spektrum, was die Beträge der verschiedenen Frequenzen im Bild darstellt. Nach Burger et al. [13] kann die spektrale Energie als Amplitude der Frequenzen interpretiert werden. In den meisten natürlichen Bildern ist die spektrale Energie vorwiegend in den niedrigen Frequenzen konzentriert, wobei ein Maximum am Koordinatenursprung auftritt. Da sich im ursprünglichen Ergebnis der 2D-DFT der Koordinatenursprung in der linken oberen Ecke befindet und das Spektrum eine periodische Funktion ist, d.h.  $\hat{u}(u, v) = \hat{u}(u + p \cdot m_1, v + q \cdot m_2)$  für  $p, q \in \mathbb{Z}$  gilt, ist der Koordinatenursprung zudem an den anderen Eckpunkten. Die höchsten Frequenzen hingegen befinden sich im Zentrum des Spektrums [13]. Aus diesem Grund ist eine *zentrierte Darstellung* des Spektrums üblich, indem die vier Quadranten geflippt werden. Somit sind

die niedrigen Frequenzen im Zentrum konzentriert, während sich die hohen Frequenzen am Rand befinden. Dieser Quadrantenshift findet auch bei der Darstellung der Queffrenzen im Queffrenzy-Raum statt.

Das Ziel dieser Arbeit ist ein Bild, welches das Fehlerbild  $u_{\text{Err}}$  ähnelt und als Vorverarbeitung für das VAIP zur automatischen Fehlererkennung dient, indem die Luftfeinschlüsse auch als diese erkannt werden. Der Unterschied zwischen dem Eingabebild  $u_{\text{In}}$  und dem optimalen Ergebnis  $u_{\text{Err}}$  ist das Vorkommen der Gitterstrukturen. Diese machen sich im Queffrenzy-Raum als Nebenmaxima entlang der Koordinatenachsen bemerkbar. Bei Betrachtung des Power-Cepstrums stellt man fest, dass sich der Abstand der Maxima der einzelnen Peaks mit der Gitterweite  $h$  in horizontaler und vertikaler Richtung übereinstimmt, beginnend im Zentrum (vgl. die Abb.3.1). Das globale Maximum ist der Ursprung. Bei Abschwächung des Zentrums und Rücktransformation in den Bildraum mittels  $\mathcal{C}_{\text{mod}}^{-1}$  bleibt das Verhältnis zu den Gitterstrukturen bei, das gesamte Bild verliert jedoch an Intensitätswerten.

Betrachtet wird nun das Grundlagenbild  $u_{\text{Base}}$  und das Gitterbild  $u_{\text{Pat}}$ . Die Power-Cepstra in zentrierter Darstellung werden als  $c_{\text{Base}}$  bzw.  $c_{\text{Pat}}$  bezeichnet. Der Unterschied beider Bilder  $u_{\text{Base}}$  und  $u_{\text{Pat}}$  liegt im Vorhandensein der Gitterstrukturen. Zur Analyse der Gitterstrukturen wird das absolute Differenzbild von  $c_{\text{Base}}$  und  $c_{\text{Pat}}$  betrachtet und im Folgenden als  $c_{\text{Diff}}$  bezeichnet.

### 1. Ansatz: Schwellenwertansatz

Die Idee hinter dem Schwellenwertansatz besteht darin, dass ausgewählte Frequenzen eines Bildes reduziert werden können. Dieser Faktor wird in Form eines Parameters mit  $\tau_{\text{Red}} \in [0, 1]$  definiert.

Beim Schwellenwertansatz wird das Power-Cepstrum des Differenzbildes vom Gittermusterbild  $u_{\text{Pat}}$  und dem Grundlagenbild  $u_{\text{Base}}$  betrachtet und ein Schwellenwert  $\tau_{\text{Diff}}$  bestimmt, um die Regionen mit höherer Differenz zu reduzieren. Dieses Power-Cepstrum des Differenzbildes wird im Folgenden  $u_{\text{Diff}}$  genannt. Der Faktor, um das diese Regionen reduziert werden, wird mit einem weiteren Schwellenwert  $\tau_{\text{Red}}$  angegeben. Die höchste Differenz ist im Ursprung gegeben. Bei seiner Reduzierung verliert das Bild an Intensitätswerten, weshalb dieser Bereich mit Radius  $r_{\text{Center}}$ ,  $r_{\text{Center}} \in \mathbb{N}_0$  bei den folgenden Analysen beibehalten wird.

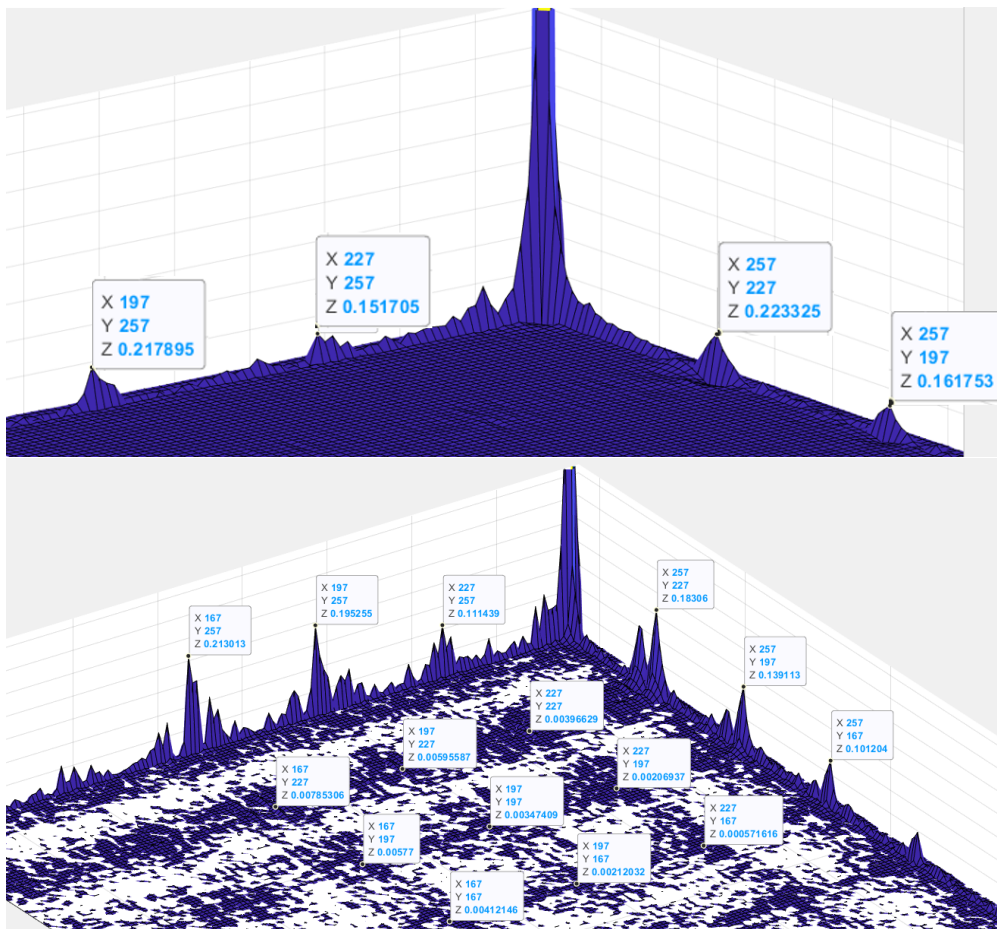


Abbildung 3.1: Oben: Ausschnittes vom Zentrum ausgehende dritte Quadranten von  $u_{\text{Pat}}$  mit markierten Nebenmaxima auf den Koordinatenachsen. Die Abstände der markierten Maxima entspricht der Gitterweite  $h_1$  in horizontaler bzw.  $h_2$  in vertikaler Richtung. Unten: Ausschnitt des vom Zentrum ausgehende dritte Quadranten von  $c_{\text{Diff}}$ . Es treten gitterförmige Nebenmaxima um das Zentrum herum auf [6].

## 2. Ansatz: Maskenansatz

Das Power-Cepstrum des Differenzbildes  $c_{\text{Diff}}$  hat beim genaueren Betrachten neben den Maxima entlang der Koordinatenachsen weitere Nebenmaxima, die sich regelmäßig um den Ursprung befinden (vgl. Abb. 3.1) in Anordnung der Gittermuster. Der zweite Ansatz zielt darauf ab, nicht nur die Maxima entlang der Koordinatenachsen zu redu-



zieren, sondern auch diejenigen, die sich regelmäßig um den Ursprung herum befinden. Hierbei werden Regionen um die jeweiligen Nebenmaxima ausgewählt und anschließend reduziert. Die Größe dieser ausgewählten Regionen wird durch die Parameter  $R_{x_1}$  (für die  $x_1$  Richtung),  $R_{x_1} \in \mathbb{N}_0$  und  $R_{x_2}$  (für die  $x_2$  Richtung),  $R_{x_2} \in \mathbb{N}_0$  gesteuert, wodurch die Effektivität der Reduktion beeinflusst werden kann. Die Stärke der Reduktion wird durch den Parameter  $\tau_{\text{Red}}$ ,  $\tau_{\text{Red}} \in [0, 1]$  bestimmt.

### 3.5 Implementierung mit der Fast-Fourier-Transformation

Die Laufzeit des Cepstrums ist primär von der Laufzeit der Fouriertransformation abhängig. Die Gleichung (5) lässt erahnen, dass die Berechnung der 2D diskreten Fouriertransformation für Bilder mit großen Dimensionen rechentechnisch aufwendig ist, vor allem für 2D-Signale mit hohen Dimensionen. Die Laufzeit für ein 2D-Bild der Dimension  $M \times N$  beträgt  $\mathcal{O}(N^2 \cdot M^2)$ . Ein Trick ist, die Zwischenergebnisse der Berechnung der DFT mehrfach wiederverwenden, um die Anzahl der notwendigen Operationen zu reduzieren. Diesen Trick verwendet der FFT-Algorithmus (Fast-Fourier-Algorithmus), indem er spezielle Eigenschaften der periodischen Funktionen ausnutzt. Dieser wurde 1965 von Cooley und Tukey vorgestellt. Der gesamte Algorithmus ist in ihrer Arbeit ausgeführt [27]. Insgesamt liefert der FFT-Algorithmus eine Laufzeit von  $\mathcal{O}(M \log_2 M \cdot N \log_2 N)$ . Das bedeutet für zum Beispiel  $M = N = 512$  erhält man eine Beschleunigung von  $\approx 3250$  bei Anwendung der FFT im Vergleich zur Berechnung der diskreten Fouriertransformation. Für  $M, N \gg 512$  ist die Beschleunigung deutlich größer.

Für die Implementierung des Power-Cepstrums wurde aus diesem Grund die zweidimensionale diskrete Fouriertransformation durch die FFT ersetzt. Diese ist in MATLAB (Version R2022b) mittels der Funktion „fft“ aufrufbar.

---

## Kapitel 4: Energiebasierter Modellierungsansatz

In diesem Kapitel wird ein Ansatz vorgestellt, bei dem gezielte Frequenzen in einem Bild reduziert oder eliminiert werden. Dieser Weg folgt der Idee, dass insgesamt zwei Bilder rekonstruiert werden, wobei das:

- erste Bild die Frequenzen der Fehler glättet
- zweite Bild die Frequenzen der Gitterstrukturen eliminiert

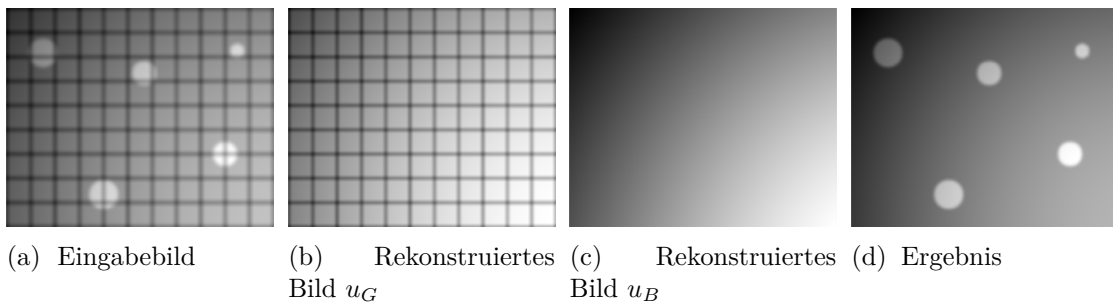


Abbildung 4.1: Motivation der energiebasierten Modellierung: Links zeigt das Eingabebild mit Fehlern und Gitterstrukturen. Bei den beiden darauffolgenden Bildern wurden 1. die Frequenzen der Fehler geglättet und 2. die Fehler der Gitterstrukturen geglättet. Rechts zeigt das Ergebnis [6].

Anschließend wird aus einer Verknüpfung des Eingabebildes mit den beiden rekonstruierten Bildern ein gesamtes Bild berechnet, beidem die Gitterstrukturen reduziert sind und die Materialfehler beibehalten. Unter der Voraussetzung, dass sich die Gitterstrukturen über das gesamte Bild erstrecken und seine Frequenzen größer sind als die der Fehler, geht man in dieser Arbeit davon aus, dass bei der Glättung der Frequenzen für die Gitterstrukturen die Frequenzen der Fehler ebenfalls mit geglättet werden. Eine Visualisierung der Bilder zeigt Abb. 4.1. Dabei ist ganz links ein synthetisches Bild mit fünf Fehlern und Gitterstrukturen zu erkennen. Darauf folgt ein Bild  $u_G$ , bei dem die Frequenzen der Fehler geglättet wurden (vgl. Bild 4.1b). Das Bild  $u_B$  zeigt ein Bild nach der Eliminierung der Gitterstrukturen (vgl. Bild 4.1c). Das Ergebnis ist eine Verknüpfung der drei Bilder und zeigt die Fehler in seiner Gesamtgröße und ohne die Gitterstrukturen. Das ist ein Idealfall und gilt nur zur Visualisierung der Motivation hinter diesem Ansatz.

Beim Ansatz handelt es sich um eine energiebasierten Modellierung, diese wird im Folgenden kurz vorgestellt. Üblicherweise erfolgt die Formulierung der energiebasierten Modellierung und der Modelle (mehr dazu in Kapitel 4.3) auf kontinuierlichen Bildern.

---

## 4.1 Beschreibung der energiebasierten Modellierung

Zur Beschreibung der energiebasierten Modellierung wird die folgende Energie betrachtet:

$$\underbrace{F(u)}_{\text{Zielfunktion}} \tag{19}$$

bzw.

$$\underbrace{F(u)}_{\text{Zielfunktion}} + \underbrace{\lambda \cdot J(u)}_{\text{Regularisierer}} \tag{20}$$

für ein Bild  $u \in \mathcal{U}$ ,  $F, J : \mathcal{U} \rightarrow \mathbb{R}_{\geq 0}$ ,  $\lambda \in \mathbb{R}$ . Für die Berechnung von  $u_G$  und  $u_B$  wird das zugehörige Optimierungsproblem betrachtet:

$$\inf_{u \in \mathcal{U}} F(u) \tag{21}$$

bzw.

$$\inf_{u \in \mathcal{U}} F(u) + \lambda \cdot J(u) \tag{22}$$

für  $F, J : \mathcal{U} \rightarrow \mathbb{R}$ ,  $\lambda \in \mathbb{R}_{\geq 0}$ . Der Minimierer der Gleichung (21) bzw. (22) wird mit  $u^*$  bezeichnet.

Die Idee hinter der Modellierung ist das Berechnen eines Modells, das dem Objekt sehr nahe kommt, die Berechnungen jedoch vereinfacht. Informell gesprochen: *Es wird dann ein Modell gebildet, wenn das reale Objekt zu kompliziert ist* [28]. In dieser Arbeit liegt der Fall vor, dass das reale Objekt das Eingabebild  $u_{\text{In}}$  darstellt und ein Modell soll dabei sowohl für  $u_G$  als auch für  $u_B$  gebildet werden.

Eine Möglichkeit für die Wahl der Zielfunktion ist die Frobeniusnorm  $\|\cdot\|_F$  des Differenzbildes eines Ausgangsbildes  $g \in \mathcal{U}$  und dem gesuchten Bild  $u^*$ . Dann gilt:

$$\inf_{u \in \mathcal{U}} F(u) \Leftrightarrow \inf_{u \in \mathcal{U}} \int_{\Omega} \|u - g\|_F^2 \, dx \tag{23}$$

Ein Minimierer  $u^*$  von Gleichung (23) ist das Ausgangsbild  $g$  selbst, denn die Frobeniusnorm ist nur dann Null, wenn es sich um das Nullbild selbst handelt laut Definition. Aus diesem Grund müssen noch mehr Informationen – in Form eines Regularisierers – hinzugenommen werden.

Dann ist:

$$\inf_{u \in \mathcal{U}} \int_{\Omega} \|u - g\|_F \, dx + \lambda \cdot J(u). \quad (24)$$

Chambolle et al. [29] analysieren verschiedene Typen von Regularisierer und gehen tiefer in die Herleitung ein. In dieser Arbeit wird als Regularisierer die *Totale Variation* (TV) vorgestellt. Die Motivation hinter der Totalen Variation ist die Möglichkeit, dass Sprünge in Bildern erlaubt werden. Anschaulich für eine eindimensionale Funktion  $u$  summiert die  $TV(u)$  die Differenz der lokalen Minima und Maxima auf, wenn  $u$  monoton ist. Dabei werden Sprünge ignoriert. Sprünge tauchen in Bildern vor allem entlang von Kanten auf und wenn viel Struktur im Bild vorhanden ist. Aus diesem Grund ist die TV scheint ein guter Kandidat für die Wahl als Regularisierer in dieser Arbeit zu sein, denn Gitterstrukturen in Röntgenbildern weisen viele Sprünge auf.

Eine der Hauptanwendungsgebiete der Totalen Variation in der Bildverarbeitung ist die Bildentrauschung. Als Pioniere gelten Leonid Rudin, Stanley Osher und Emad Fatemi [17], die 1992 das nach ihnen benannte ROF-Modell einführten. Bevor es zur Definition der Modelle kommt, gibt es einen Überblick über die verschiedenen Formulierungen der Totalen Variation und die zugehörigen Funktionsräume.

## 4.2 Definition der Totalen Variation und der Funktionsräumen $L^2$ und BV

Bilder in der mathematischen Bildverarbeitung werden häufig als kontinuierliche Objekte definiert, genauer gesagt als Objekte aus dem  $L^2(\Omega)$ . Dazu eine kurze Einführung zu den  $L^p$  Räumen mit ihrer zugehörigen Norm nach Pöschel in [12].

**Definition 4.1.** Sei  $n \in \mathbb{N}$ ,  $\Omega \subseteq \mathbb{R}^n$ ,  $1 \leq p < \infty$  und  $\mu$  beschreibe das Lebesgue-Maß. Für eine  $\mu$ -messbare Funktion  $f : \Omega \rightarrow \mathbb{R}$  ist die  $p$ -Norm definiert als

$$\|f\|_p := \left( \int_{\Omega} |f|^p d\mu \right)^{1/p}. \quad (25)$$

Dann heißt der Unterraum

$$\mathcal{L}^p(\mu) = \mathcal{L}^p(\Omega, \mu) := \left\{ f \text{ ist } \mu \text{ messbar} \mid \|f\|_p < \infty \right\}$$

Eine Funktion  $f : X \rightarrow Y$  heißt dann  $\mu$  messbar (oder *Lebesgue* messbar), wenn für jede messbare Menge  $I \subseteq Y$  das Urbild  $f^{-1}(I)$  in  $X$  messbar ist [14].

Der Vollständigkeit halber ist noch angemerkt, dass es sich bei  $\|\cdot\|_p$  um keine vollständige Norm auf  $\mathcal{L}^p$  handelt, denn das erste Axiom, die Definitheit ( $\forall f \in \Omega : \|f\| \geq$

$0 \wedge (\|f\| = 0 \Leftrightarrow f = 0)$  ist nicht erfüllt. Das bedeutet es existiert eine Funktion  $f$ , die das Maß Null hat, aber nicht die Nullfunktion ( $f(x) \equiv 0$ ) ist. Um vollständige Normen zu erhalten, werden alle Funktionen, die fast überall dieselben Funktionswerte aufweisen, zu einer Klasse zusammengefasst. Daher werden die  $p$ -Normen auf äquivalenten Klassen von Funktionen definiert und zwei Funktionen werden als äquivalent betrachtet, wenn sie fast überall gleich sind. Dies geschieht durch die Division mit dem Nullraum

$$N(\mu) := \left\{ f \text{ ist } \mu\text{-messbar} \mid f =_{\mu} 0 \right\}.$$

Formal bedeutet das nach Pöschel und Storch et al. [12, 30]:

**Definition 4.2.** Für ein  $1 \leq p < \infty$  heißt

$$L^p(\mu) := L^p(\Omega, \mu) := \frac{\mathcal{L}^p(\mu)}{N(\mu)} \quad (26)$$

der Lebesgueraum  $L^p$  auf  $\Omega$  bzgl.  $\mu$ .

Der Beweis, dass es sich bei dem Vektorraum  $L^p(\mu)$  mit der Norm  $\|\cdot\|_p$  für  $p \geq 1$  um einen normierten Vektorraum handelt, ist in [12] auf Seite 134 f. zu finden.

Die Totale Variation einer Funktion misst die Abweichung (Variation) lokaler Funktionswerte auf einem beschränkten Intervall. Wenn eine Funktion stetig differenzierbar ist, dann lässt sich die Variation ihrer Funktionswerte durch den absoluten Betrag des Gradienten beschreiben. Das heißt nach [29, 31] gilt:

**Satz 4.3.** Für ein  $u \in \mathcal{C}^1(\Omega, \mathbb{R}) \cap L^1(\Omega, \mathbb{R})$ ,  $\Omega \subseteq \mathbb{R}^d$ ,  $d = 1, 2$ ,  $TV : \mathcal{C}^1(\Omega, \mathbb{R}) \cap L^1(\Omega, \mathbb{R}) \rightarrow \mathbb{R}$  gilt

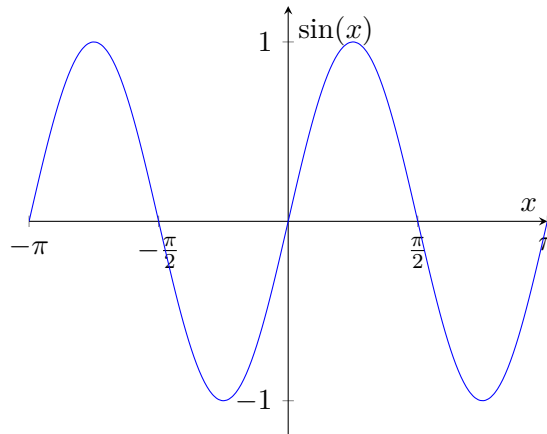
$$TV(u) = \int_{\Omega} |\nabla u(x)| \, dx \quad (27)$$

Der Fokus dieser Arbeit liegt auf der Vorstellung der TV als Regularisierer im energiebasierten Modell und aus diesem Grund wird für den Beweis auf [31] verwiesen. Zur Visualisierung ein Beispiel:

**Beispiel 4.4.** Es sei  $f : \Omega \rightarrow [-2, 2]$  mit  $\Omega = [-\pi, \pi]$  und  $x \in \Omega$ . Dann ist für

$$f(x) = \sin(2x)$$

die Totale Variation nach der Formel (27)

Abbildung 4.2: Funktion  $f$  auf dem Intervall  $[-\pi, \pi]$ 

$$\begin{aligned}
TV(f) &= \int_{-\pi}^{\pi} |\sin'(2x)| dx \\
&= 2 \cdot \int_{-\pi}^{\pi} |\cos(2x)| dx \\
&= 2 \left( \int_{-\pi}^{-\frac{3}{4}\pi} |\cos(2x)| dx + \int_{-\frac{3}{4}\pi}^{-\frac{1}{4}\pi} |\cos(2x)| dx + \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} |\cos(2x)| dx \right. \\
&\quad \left. + \int_{\frac{1}{4}\pi}^{\frac{3}{4}\pi} |\cos(2x)| dx + \int_{\frac{3}{4}\pi}^{\pi} |\cos(2x)| dx \right) \\
&= 2 \left( \left| \frac{\sin(-2\pi) - \sin(-\frac{6}{4}\pi)}{2} \right| + \left| \frac{\sin(-\frac{6}{4}\pi) - \sin(-\frac{1}{2}\pi)}{2} \right| \right. \\
&\quad \left. + \left| \frac{\sin(\frac{1}{2}\pi) - \sin(-\frac{1}{2}\pi)}{2} \right| + \left| \frac{\sin(\frac{6}{4}\pi) - \sin(\frac{1}{2}\pi)}{2} \right| + \left| \frac{\sin(2\pi) - \sin(\frac{6}{4}\pi)}{2} \right| \right) \\
&= 2 \left( \frac{1}{2} + 1 + 1 + 1 + \frac{1}{2} \right) \\
&= 8
\end{aligned}$$

Das bedeutet, dass die Funktionswerte von  $f$  auf dem Intervall  $[-\pi, \pi]$  um 8 variieren. Dies kann mit Abb. 4.2 bestätigt werden.

Diese Formulierung eignet sich für Bilder, dessen Ableitungen existieren, das bedeutet für unstetige Funktionen muss eine weitere Definition her, denn die hier betrachteten Röntgenbilder weisen in der Regel Unstetigkeiten auf. Aus diesem Grund ist eine Definition der Totalen Variation von unstetigen Funktionen von Interesse in dieser Arbeit. Dazu wird auf die duale Formulierung der Totalen Variation nach Chambolle et al. [29] betrachtet:

**Definition 4.5.** Für  $\Omega \subseteq \mathbb{R}^d$  offen und  $u \in L^1(\Omega, \mathbb{R}^k)$  heißt die Abbildung  $TV : L^1(\Omega) \rightarrow \mathbb{R}^+$  mit

$$TV(u) := \sup \left\{ - \int_{\Omega} \langle u(x), \operatorname{Div} \varphi(x) \rangle dx \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^{d \times k}), \|\varphi(x)\|_F \leq 1 \forall x \in \Omega \right\} \quad (28)$$

die Totale Variation von  $u$ .

Dabei ist

$$\varphi = (\varphi_1, \varphi_2, \dots, \varphi_k), \operatorname{Div} \varphi(x) = \begin{pmatrix} \operatorname{div} \varphi_1(x) \\ \vdots \\ \operatorname{div} \varphi_k(x) \end{pmatrix}$$

mit der Divergenz von  $\varphi$  für  $x \in \Omega$  (vgl. [32]):

$$\operatorname{div} \varphi(x) := \sum_{i=1}^d \frac{\partial}{\partial x_i} \varphi_i(x)$$

für die Frobeniusnorm  $\|\cdot\|_F$  von  $\varphi(x)$  gilt:

$$\|\varphi(x)\|_F = \left( \sum_{i=1}^d \sum_{j=1}^k |\varphi_{i,j}(x)|^2 \right)^{\frac{1}{2}}$$

für  $k=1$  gilt:  $\|\varphi(x)\|_F = \|\varphi(x)\|_2$

Mit  $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^{d \times k})$ , was bedeutet, dass  $\varphi$  ein mal stetig differenzierbar und der Träger eine kompakte Teilmenge von  $\Omega$  ist.

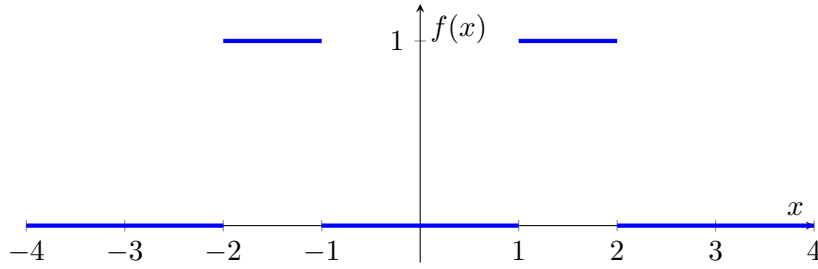
Die Formulierung von Definition 4.5 hat den Vorteil, dass an  $u$  wenige Voraussetzung erfüllt sein müssen. Die Berechnung ist dagegen etwas aufwendiger. Für ein besseres Verständnis der Formulierung in Definition 4.5 ein Beispiel:

**Beispiel 4.6.** Es sei  $f : (-4, 4) \rightarrow \{0, 1\}$  eine stückweise konstante Funktion mit

$$f(x) = \begin{cases} 1 & \text{für } x \in [-2, -1] \cup [1, 2] \\ 0 & \text{sonst} \end{cases}$$

Da  $d = 1 \Rightarrow \operatorname{div} \varphi(x) = \varphi'(x)$  und die Totale Variation von  $f$  ist:

$$TV(f) = \sup_{\|\varphi(x)\|_F \leq 1} \left\{ - \int_{\Omega} \langle f(x), \operatorname{div} \varphi(x) \rangle dx \right\}$$


 Abbildung 4.3: Funktion  $f$  auf dem Intervall  $[-4,4]$ 

$$\begin{aligned}
 &\Leftrightarrow \sup_{\|\varphi(x)\|_F \leq 1} \left\{ - \int_{-4}^4 f(x) \cdot \operatorname{div} \varphi(x) \, dx \right\} \\
 &\Leftrightarrow \sup_{\|\varphi(x)\|_F \leq 1} \left\{ - \int_{-4}^4 f(x) \cdot \varphi'(x) \, dx \right\} \\
 &\Leftrightarrow \sup_{\|\varphi(x)\|_F \leq 1} \left\{ - \int_{-4}^{-2} f(x) \cdot \varphi'(x) \, dx - \int_{-2}^{-1} f(x) \cdot \varphi'(x) \, dx - \int_{-1}^1 f(x) \cdot \varphi'(x) \, dx \right. \\
 &\quad \left. - \int_1^2 f(x) \cdot \varphi'(x) \, dx - \int_2^4 f(x) \cdot \varphi'(x) \, dx \right\} \\
 &\Leftrightarrow \sup_{\|\varphi(x)\|_F \leq 1} \left\{ - \int_{-2}^{-1} \varphi'(x) \, dx - \int_1^2 \varphi'(x) \, dx \right\} \\
 &\Leftrightarrow \sup_{\|\varphi(x)\|_F \leq 1} \left\{ \underbrace{-\varphi(-2) + \varphi(-1) - \varphi(2) + \varphi(1)}_{(*)} \right\}
 \end{aligned}$$

(\*) wird unter der Bedingung  $\|\varphi(x)\|_F \leq 1$  maximal für  $\varphi(-2) = \varphi(2) = -1$  und  $\varphi(-1) = \varphi(1) = 1$ .

$$\Rightarrow TV(f) = -(-1) + 1 - (-1) + 1 = 4$$

Dies kann mit Abb. 4.3 bestätigt werden.

Der Funktionenraum, auf dem sich Funktionen mit endlicher Variation befinden, ist nach Pöschel [12] definiert als:

**Definition 4.7.** Die Menge

$$BV(\Omega, \mathbb{R}^k) := \{u \in L^1(\Omega, \mathbb{R}^k) \mid TV(u) < +\infty\} \quad (29)$$

ist der Raum der Funktionen beschränkter Variation (engl: bounded variation) auf  $\Omega$ .

Das energiebasierte Modell lautet nun:

$$\inf_{u \in \mathcal{U}} F(u) + \lambda \cdot J(u) \quad (30)$$



mit  $\mathcal{U} = BV(\Omega, \mathbb{R})$ ,  $J(u) = TV(u)$  und  $\lambda \in \mathbb{R}$ .

### 4.3 Formulierung der Modelle

Für die Berechnung von  $u_G$  und  $u_B$  werden im Folgenden zwei verschiedene Modelle vorgestellt, die in [17] und [29] definiert sind.

**Definition 4.8.** Für  $g \in L^2(\Omega, \mathbb{R}^k)$  und  $\lambda > 0$  ist

$$\inf_{u \in BV(\Omega, \mathbb{R}^k)} \frac{1}{2} \int_{\Omega} \|u(x) - g(x)\|_2^2 dx + \lambda TV(u) \quad (31)$$

das Rudin-Osher-Fatemi-Modell (oder kurz ROF-Modell).

**Definition 4.9.** Für  $g \in L^1(\Omega, \mathbb{R}^k)$  und  $\lambda > 0$  ist

$$\inf_{u \in BV(\Omega, \mathbb{R}^k)} \underbrace{\int_{\Omega} \|u(x) - g(x)\|_2 dx}_{\|u-g\|_{L^1}} + \lambda TV(u) \quad (32)$$

das  $L^1$ -TV-Modell.

Das Ergebnis in Kapitel 6 ist dann das Modell selbst, d.h. der Minimierer  $u^*$  ist dann je nach Fall das ROF-Modell oder das  $L^1$ -TV-Modell.

Für die Algorithmen zur Berechnung und Diskretisierungen der Modelle werden auf die Arbeiten von Lorakis et al. [33], Magiera et al. [34], Goldstein et al. [35], Rudin et al. [17] und Chambolle et al. [29, 36, 37, 38, 39] verwiesen.

---

## Kapitel 5: Metriken zur Qualitätsanalyse

Die visuelle Bewertung von Bildern ist für Menschen eine natürliche und intuitive Aufgabe. Sie können schnell Unterschiede in Bildern erkennen. Beim Versuch Unterschiede komplexer Bilder mit und ohne spezifischen Strukturen oder Mustern objektiv zu bewerten, kann es bei einer eindeutigen Bewertung zu Herausforderungen kommen.

Ein solches Szenario trifft in dieser Arbeit ein, indem die Ergebnisse der Cepstralanalyse aus Kapitel 3 und die Ergebnisse der energiebasierten Modellierung aus Kapitel 4 objektiv (bzgl. der Reduktion der Gitterstrukturen) bewertet werden. Da Bilder mit Gitterstrukturen für die Eingabe in die Röntgenprüfung ungeeignet sind, werden sie vorher reduziert oder bestenfalls eliminiert. Herausforderung ist dabei eine Möglichkeit, diese Reduktion objektiv zu messen. Hier kommt die Notwendigkeit eines geeigneten Messverfahrens ins Spiel. Ein solches Messverfahren wird in der Diskussion in Kapitel 7 auf das jeweilige Ergebnisbild mit dem gewünschten Fehlerbild (vgl. Bild 2.4c, mehr dazu in Kapitel 2.3) angewendet, sodass eine objektive Messung diskutiert werden kann. Dazu werden in diesem Kapitel Distanz- und Ähnlichkeitsmaße vorgestellt.

Zuerst eine kurze Übersicht:

- **Distanzmaße:** Bei den Distanzmaßen handelt es sich um Maße, die den Unterschied zwischen den betrachteten Objekten quantifizieren. Dabei liegt der Fokus auf der Berechnung der *Distanz*. Dies bedeutet für zwei Bilder, dass der Unterschied der Pixelwerte der beiden Bilder betrachtet wird. In der Regel gilt, je größer die Distanz, umso unterschiedlicher sind die Bilder.
- **Ähnlichkeitsanalyse:** Die Ähnlichkeitsanalyse konzentriert sich darauf, die *Ähnlichkeit* oder *Gleichartigkeit* zwischen den untersuchten Objekten zu bewerten. Dabei gilt, je größer die Ähnlichkeit ist, desto ähnlicher sind die Objekte zueinander.

Seien nun  $u, v : \Omega_h \rightarrow \mathbb{R}$  zwei diskrete Bilder, die miteinander verglichen werden sollen. Dabei seien  $u, v \in \mathbb{R}^{m_1 \times m_2}$  und  $u_{i,j}, v_{i,j} \in [0, 65535]$ .

### 5.1 Distanzmaße

Das erste hier vorgestellte Distanzmaß ist die Berechnung der Frobeniusnorm des Differenzbildes von  $u$  und  $v$ . Die Definition stammt aus dem Buch von [40].

**Definition 5.1.** Die Frobeniusnorm ist eine Abbildung  $\|\cdot\|_F : \mathbb{C}^{m_1 \times m_2} \rightarrow \mathbb{R}_0^+$ , mit

$$\|A\|_F = \sqrt{\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} |a_{i,j}|^2} \quad (33)$$

---

für eine Matrix  $A \in \mathbb{C}^{m_1 \times m_2}$  und  $a_{i,j}$  seien seine Einträge für  $i = 1, \dots, m_1$  und  $j = 1, \dots, m_2$ .

Die Frobeniusnorm des Differenzbildes von  $u$  und  $v$  lautet dann:

$$\|u - v\|_F = \sqrt{\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} |u_{i,j} - v_{i,j}|^2}.$$

Ein weiterer Ansatz ist die Berechnung des durchschnittlichen quadratischen Fehlers zweier Bilder. Das bedeutet, es werden durchschnittlich, pixelweise die Differenzen der Bilder  $u$  und  $v$  berechnet. Dieses Maß heißt der *Mittlere quadratische Fehler* (im Englischen: Mean Squared Error, oder kurz MSE) und ist ein weit verbreitetes Maß zur Bewertung der Unterschiede zwischen zwei Signalen. Der MSE findet häufig Anwendung bei der Bewertung des Ergebnisses nach dem Entrauschen (vgl. [4]). Dabei wird das entrauschte Bild mit seinem Originalbild verglichen (weitere Informationen dazu in [4]).

**Definition 5.2.** (*Mittlerer quadratische Fehler, MSE*)

Für zwei diskrete Bilder  $u, v$ , wie oben definiert, ist der MSE definiert durch:

$$MSE(u, v) = \frac{1}{m_1 \cdot m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (u_{i,j} - v_{i,j})^2 \quad (34)$$

mit  $MSE: \mathbb{R}^{m_1 \times m_2} \times \mathbb{R}^{m_1 \times m_2} \rightarrow \mathbb{R}$ .

Der MSE misst den durchschnittlichen quadratischen Fehler zwischen den Pixelwerten. Je niedriger der MSE-Wert ist, desto geringer sind die Unterschiede zwischen den beiden Bildern und desto höher wird die Ähnlichkeit oder Genauigkeit bewertet. Ein MSE-Wert von 0 bedeutet, dass die beiden Bilder identisch sind.

In der Bildkompression kommt häufig das sogenannte „Peak-Signal-to-Noise Ratio“ (oder kurz: PSNR) zum Einsatz. Hier wird das unkomprimierte Bild oft mit dem komprimierten Bild verglichen. Bei dem PSNR handelt es sich um eine skalierte Variante des MSE, die dazu dient, das Verhältnis der maximal möglichen Energie des Signals zur Energie des vorhandenen Rauschens zu messen (weitere Informationen dazu in [4]).

**Definition 5.3.** (*Peak-Signal-to-Noise-Ration, PSNR*)

Für zwei diskrete Bilder  $u, v$ , wie oben definiert, ist der PSNR definiert durch:

$$PSNR(u, v) = 10 \log_{10} \left( \frac{65535^2}{MSE(u, v)} \right) \text{ db} \quad (35)$$

mit PSNR:  $\mathbb{R}^{m_1 \times m_2} \times \mathbb{R}^{m_1 \times m_2} \rightarrow \mathbb{R}$ . Die Einheit des PSNR wird in der Regel in Dezibel (db) angegeben (vgl. [4]).

Ein höherer PSNR-Wert deutet auf eine bessere Bildqualität hin, denn dies weist auf ein stärker dominierendes Signal im Verhältnis zum Rauschen. Ein niedrigerer PSNR-Wert hingegen impliziert eine geringere Bildqualität, da das Rauschen einen größeren Einfluss auf das Signal hat.

Sowohl das MSE als auch das PSNR sind Metriken, die sich ohne hohen Aufwand implementieren und berechnen lassen. Sie haben jedoch einige Einschränkungen: Das MSE berücksichtigt nicht die Wahrnehmung des menschlichen Auges und kann daher Unterschiede in der visuellen Qualität nicht immer genau erfassen. Insbesondere bei Bildern können kleine Unterschiede im MSE zu subjektiv wahrnehmbaren Unterschieden führen. Die Autoren gehen in [41] genauer darauf ein. Ihren Ansatz zur Bewertung der Unterschiede und Ähnlichkeit zweier Bilder wird im Folgenden genauer betrachtet.

## 5.2 Ähnlichkeitsanalyse

Natürliche Bildsignale zeichnen sich durch eine starke Struktur aus. Die Pixel in diesen Bildern weisen insbesondere dann starke Abhängigkeiten auf, wenn sie räumlich nah beieinander liegen. Diese Abhängigkeiten enthalten wichtige Informationen über die Struktur der Objekte in der visuellen Szene. Die Autoren in [41] beschreiben ein neues Konzept, um Strukturen zweier Bilder zu vergleichen. Das vorgeschlagene neue Konzept für Bildqualitätsmaße basiert auf der Annahme, dass das menschliche visuelle System strukturelle Informationen aus dem Sichtfeld extrahieren kann. Auf dieser Grundlage begründen sie, dass eine Bewertung der Veränderung der Strukturinformationen eine gute Näherung für wahrgenommene Bildverzerrungen sein kann. Im Gegensatz zur Fehlerempfindlichkeit, die auf wahrgenommenen Fehlern basiert, betrachtet diese Philosophie Bildverschlechterungen als Veränderungen der Strukturinformationen. Dadurch können Qualitätseinschätzungen besser mit wahrgenommenen Unterschieden übereinstimmen und komplexe Bildstrukturen direkt bewertet werden.

Der Übersicht halber werden im Folgenden diskrete Bilder  $u, v$ , in Column-Major-Anordnung (vgl. Bem. 3.4) betrachtet. Das bedeutet  $u, v \subseteq \mathbb{R}^N$  mit  $u_i, v_i \in \{0, \dots, 65535\}$  für  $i = 1, \dots, N$ .

**Definition 5.4.** (*Structural Similarity, SSIM*)

Es seien  $u, v$  wie in 5.2 gegeben. Dann ist der SSIM

$$SSIM(u, v) = \left( l(u, v)^\alpha \cdot c(u, v)^\beta \cdot s(u, v)^\gamma \right) \quad (36)$$

eine Funktion aus einer Kombination von  $l$ : Leuchtdichtvergleichsfunktion,  $c$ : Kontrastvergleichsfunktion und  $s$ : Strukturvergleichsfunktion mit  $\alpha, \beta, \gamma > 0$ . Dabei ist:

$$l(u, v) = \frac{2\mu_u\mu_v + C_1}{\mu_u^2 + \mu_v^2 + C_1} \quad (37)$$

$$c(u, v) = \frac{2\sigma_u\sigma_v + C_2}{\sigma_u^2 + \sigma_v^2 + C_2} \quad (38)$$

$$s(u, v) = \frac{\sigma_{uv} + C_3}{\sigma_u\sigma_v + C_3} \quad (39)$$

mit  $C_i = (K_i \cdot L)^2$ ,  $K_i \ll 1$ ,  $i \in \{1, 2, 3\}$ ,  $L = 65535$ ,  $\mu_i = \frac{1}{N} \sum_{k=1}^N i_k$  und  $\sigma_i = (\frac{1}{N-1} \sum_{k=1}^N (i_k - \mu_i)^2)^{1/2}$ ,  $i \in \{u, v\}$ .

Die Maße Forbeniusnorm, MSE, PSNR und SSIM bestimmen die Unterschiede bzw. Ähnlichkeit der beiden Bilder  $u$  und  $v$ . Das bedeutet, dass diese Maße eine Bewertung in Form eines skalaren Wertes zurückgibt. In Kapitel 7 wird gezeigt, wo eine solche Bewertung ihre Grenzen hat. Ein weiterer Ansatz ist die Beurteilung der Gitterreduktion anhand eines Kantenbildes durch das Detektieren der Fehler im Bild. Dieser Ansatz liefert weder einen skalaren Wert, noch eine objektive Beurteilung, die für einen direkten Vergleich der Methoden Cepstralanalyse und energiebasierte Modellierung ermöglicht. Jedoch kann mit einem Kantenbild die Eignung der Methoden für den Vorverarbeitungsprozess des ADRs überprüft werden (mehr dazu in Kapitel 7).

### 5.3 Kantenfilter

Für die Vorverarbeitung für die automatische Fehlererkennung im VAIP ist es wichtig, dass Materialfehler in ihrer ursprünglichen Größe als solche erkannt werden. Aus diesem Grund ist für die Beurteilung der Methoden und Analysen der beiden Verfahren in Kapitel 3 und 4 auch der Kantenfilter von Interesse. Diese Arbeit konzentriert sich dabei auf den sogenannten *Canny-Edge-Detection-Algorithmus*. Dieser Algorithmus besteht aus mehreren Schritten, um eine genaue Darstellung der Kanten im Bild zu liefern. Die Schritte sind nach den Autoren von [42, 43] zusammengefasst in:

1. Glättung zur Rauschunterdrückung: erfolgt durch die Anwendung eines Gauß-Filters
2. Berechnung des Gradienten: Der Gradient des Bildes wird berechnet, um Bereiche mit starker Intensitätsänderung zu identifizieren.
3. Unterdrückung nicht-maximaler Werte: In diesem Schritt werden die Gradientenstärken bereinigt, um nur die lokalen Maxima beizubehalten.

4. Hysterese-Schwellwertung: Bestimmung der endgültigen Kanten, wofür zwei Schwellenwerte benötigt werden.

Eine Methode zur Bestimmung der Schwellenwerte ist die sogenannte *Otsu-Methode*. Weitere Details sind in [42] und [44] zu finden.

## 5.4 Beispiele

Für ein besseres Verständnis der Frobeniusnorm, MSE, PSNR, SSIM und dem Kantenfiter Canny-Edge-Detection folgt ein Beispiel des Testbildes „Lena512.bmp“ aus der MATLAB-Toolbox Image Processing.

Die Abb. 5.1 zeigt in der ersten Zeile ein Ausgangsbild 5.1a und zwei Variationen des Bildes. Dabei wurden bei den beiden überarbeiteten Bildern Gitterstrukturen mit Gitterweite  $h = (30, 30)$  eingerechnet. Der Unterschied liegt in ihrer Intensität. Die Abb. 5.1c simuliert das Ursprungsbild mit einer Gitterstruktur, welche in Abb. 5.1b) abgeschwächt dargestellt ist. Die untere Reihe zeigt die zugehörigen Kantenbilder mittels des Canny-Edge-Detection-Algorithmus. Die Wahl des oberen Schwellenwertes  $\tau$  wurde mithilfe der Otsu-Methode bestimmt. Der untere Schwellenwert wurde nach Definition des Algorithmus von MATLAB durch  $0.4 \cdot \tau$  festgelegt. Die Tabelle 5.1 gibt die Maße, wie sie in Kapitel 5.1 und Kapitel 5.2 definiert wurden, wieder. Dabei steht die erste Zeile für die Maße des Originalbildes (Bild 5.1a) und des Bildes mit den vollen Gitterstrukturen (Bild 5.1c). Die zweite Zeile stellt die Ergebnisse der Maße für das Originalbild und das Bild mit den abgeschwächten Gitterstrukturen (Bild 5.1b) dar. Sowohl die Frobeniusnorm des Differenzbildes, als auch der MSE und der PSNR weisen bei der Bewertung des Vergleichs mit dem abgeschwächten Gitterstrukturen niedrigere Werte auf, als beim Vergleich mit dem ursprünglichen Gitter. Daraus lässt sich schließen, dass eine Reduktion der Gitterstrukturen objektiv von den hier vorgestellten Metriken berechnet werden kann. Auch bei der Ähnlichkeitsanalyse hat die Stärke der Gitterstrukturen Einfluss.

	$\ \cdot\ _F$	SSIM	MSE	PSNR
(Bild 5.1a, Bild 5.1b)	6.9794	0.9396	1.8582e-04	85.4399
(Bild 5.1a, Bild 5.1c)	61.6130	0.6281	0.9396	48.4014

Tabelle 5.1: Ergebnisse der Metriken  $\|\cdot\|_F$ , SSIM, MSE und PSNR. Bei der Frobeniusnorm handelt es sich um das Differenzbild beider Bilder.

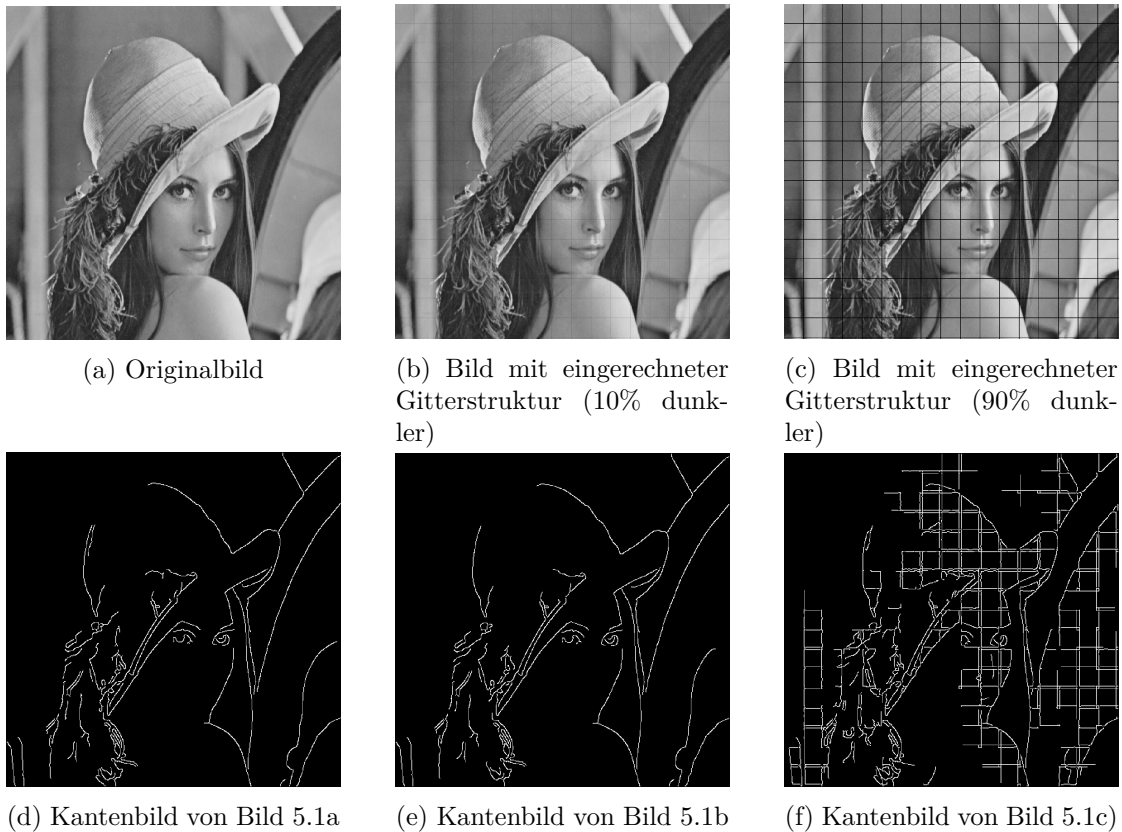


Abbildung 5.1: Oben links: Bild Lena512.bmp aus der MATLAB Image Processing Toolbox. Oben Mitte: Originalbild mit eingerechneter Gitterstruktur (Pixelwerte sind 10% abgeschwächt) mit Gitterweite  $h = (30, 30)$ . Oben rechts: Originalbild mit eingerechneter Gitterstruktur (Pixelwerte sind 90% abgeschwächt) mit Gitterweite  $h = (30, 30)$ . Die untere Reihe zeigt die zugehörigen Kantenbilder mit dem Canny-Edge-Detection-Algorithmus zu dem oberen Bild [6].

---

## Kapitel 6: Ergebnisse

Dieses Kapitel befasst sich mit den Ergebnissen und anschließender Diskussionen dieser Arbeit. Damit keine Rückschlüsse auf die Kunden gezogen werden können, findet die Analyse der in Kapitel 3 und Kapitel 4 dargestellten Methoden auf synthetischen Bildern statt. Das Unterkapitel 6.3 zeigt die erzielten Ergebnisse mit dem energiebasierten Modellierungsansatz für die Realdatenausschnitte. Eine Übersicht über die in diesem und nächsten Kapitel verwendeten Bilder, Power-Cepstra und Parametern liefern die Tabellen A.1 und A.2 im Anhang.

### 6.1 Erzielte Ergebnisse mit der Cepstralanalyse

Die Implementierungen des Power-Cepstrums und der inversen modifizierten Cepstrum erfolgen in Matlab (Version R2022b).

#### Ergebnisse mit dem Schwellenwertansatz

Betrachtet wird der Schwellenwertansatz, wie er in Kapitel 3.4 vorgestellt wurde. Für die Anwendung des Schwellenwertes auf das synthetische Bild  $u_{\text{In}}$  zeigt Abb. 6.1 das erzielte Ergebnis mit den Parametern  $\tau_{\text{Red}} = 0.55$ ,  $r_{\text{Center}} = 10$ ,  $\tau_{\text{Diff}} = 0.001$ . Das Ergebnis wird mit  $u_{\text{TCep}}$  bezeichnet.

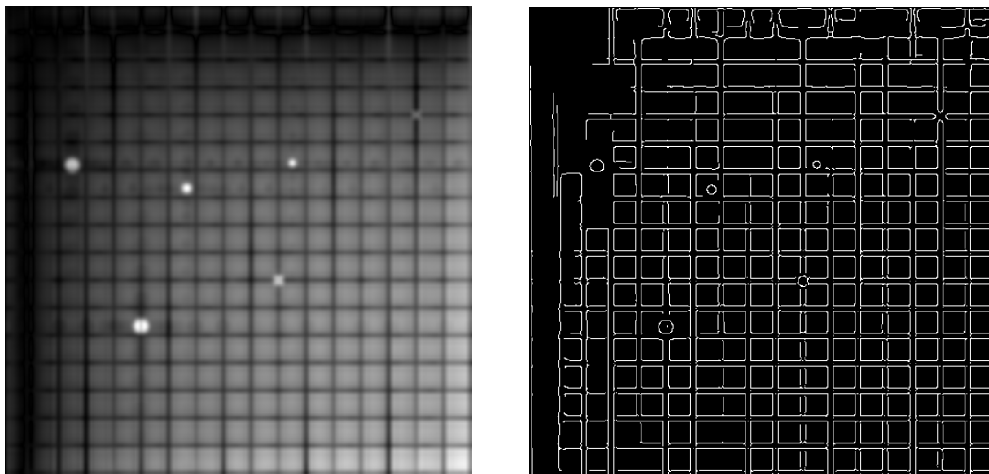


Abbildung 6.1: Links zeigt das Ergebnis  $u_{\text{TCep}}$  mit  $\tau_{\text{Red}} = 0.55$ ,  $r_{\text{Center}} = 10$ ,  $\tau_{\text{Diff}} = 0.001$  mit dem zugehörigen Kantenbild mittels des Canny-Edge-Detection Algorithmus [6].

#### Ergebnisse mit dem Masken-Ansatz

Die erzielten Ergebnisse mit dem Maskenansatz, wie er in Kapitel 3.4 vorgestellt wurde, liefert Abb. 6.2 für das Eingabebild  $u_{\text{In}}$  und den Parametern  $r_{\text{Center}} = 1$ ,  $\tau_{\text{Red}} = 0.45$

---



und  $R_{x_1} = R_{x_2} = 5$ . Das Ergebnis wird mit  $u_{\text{MCep}}$  bezeichnet.

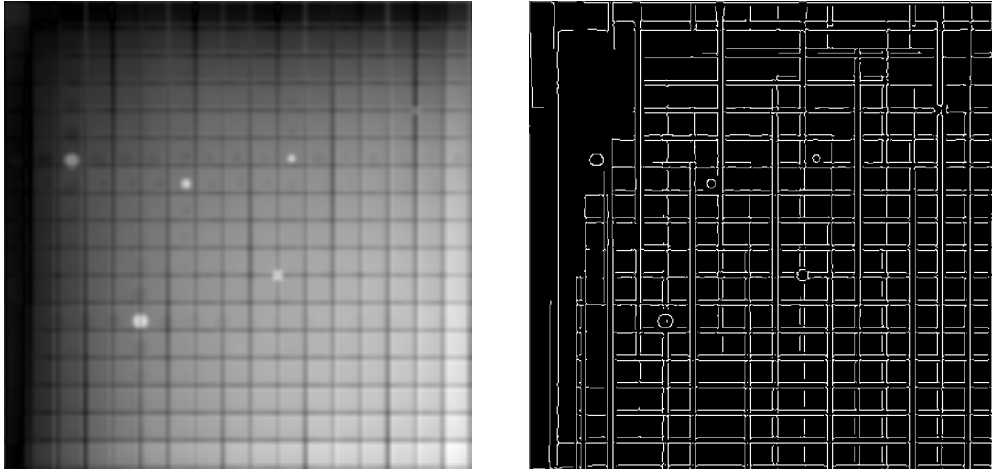


Abbildung 6.2: Links zeigt das Ergebnis  $u_{\text{MCep}}$  mit  $r_{\text{Center}} = 1$ ,  $\tau_{\text{Red}} = 0.45$  und  $R_{x_1} = R_{x_2} = 5$  und dem zugehörigen Kantenbild mittels des Canny-Edge-Detection Algorithmus [6].

## 6.2 Erzielte Ergebnisse mit dem Modellierungsansatz

Die Implementierung des ROF-Modells und des  $L^1$ -TV-Modells erfolgen in MATLAB (Version R2022b) mit den Funktionen „ROFdenoise.m“ von Philippe Magiera und Carl Löndahl, Version 1.0.0.0 aus dem Jahr 2008 [34] und der Funktion „TVL1denoise.m“ von Manolis Lourakis, Version 1.0.0.0 aus dem Jahr 2017 [33]. Die Algorithmen basieren auf den Arbeiten von Rudin et al. [17] und Chambolle et al. [29].

### ROF-Modell

Die Abb. 6.3 zeigt das Gesamtergebnis aus der Kombination vom Eingabebild  $u_{\text{In}}$ , dem ersten ROF-Modell  $u_{\text{ROF1}}^*$  mit dem Parameter  $\lambda_{\text{ROF1}} = 0.1$  und insgesamt 900 Iterationsschritten und dem zweiten ROF-Modell  $u_{\text{ROF2}}^*$  mit den Parameter  $\lambda_{\text{ROF2}} = 2.5$  und ebenfalls 900 Iterationsschritten. Die Anzahl der Iterationsschritte wurden der Funktion nicht mit übergeben, sondern sind im Algorithmus auf 5 gesetzt. Dies wurde in der Funktion geändert. Das Gesamtergebnis wird mit  $u_{\text{ROF}}$  bezeichnet.

### $L^1$ -TV Modell

Das Gesamtergebnis  $u_{L^1}$ , basierend auf dem Eingabebild  $u_{\text{In}}$ , dem ersten  $L^1$ -TV-Modell  $u_{L^1}^*$  mit dem Parameter  $\lambda_{L^1} = 0.315$  und insgesamt 500 Iterationen und dem zweiten  $L^1$ -TV-Modell  $u_{L^2}^*$  mit  $\lambda_{L^2} = 0.05$  und 800 Iterationen zeigt die Abb. 6.4.

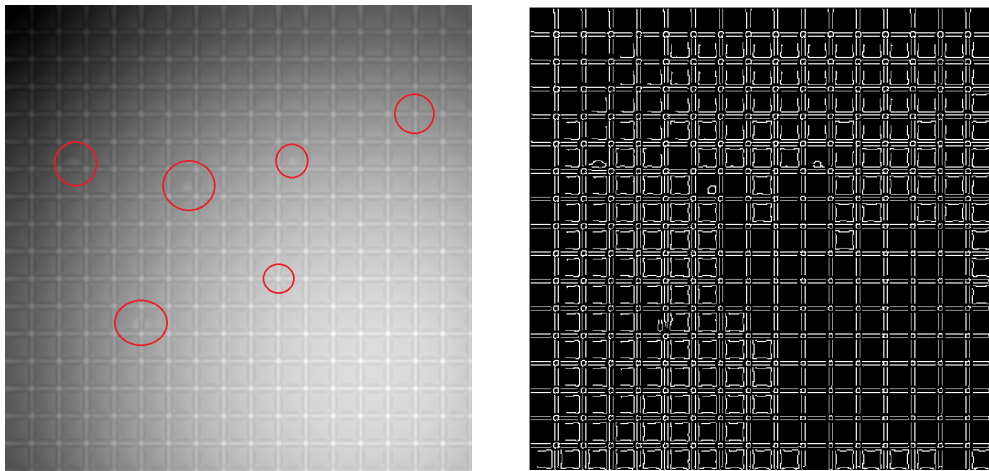


Abbildung 6.3: Links zeigt das Ergebnis  $u_{\text{ROF}}$  (oben) mit  $\lambda_{\text{ROF}1} = 0.1$ ,  $\lambda_{\text{ROF}2} = 2.5$  und Anzahl der Iterationen = 900. Die Fehler sind für das menschliche Auge schwer zu erkennen, weshalb sie mit einem roten Kreis markiert sind. Rechts zeigt das zugehörige Kantenbild mittels des Canny-Edge-Detection Algorithmus [6].

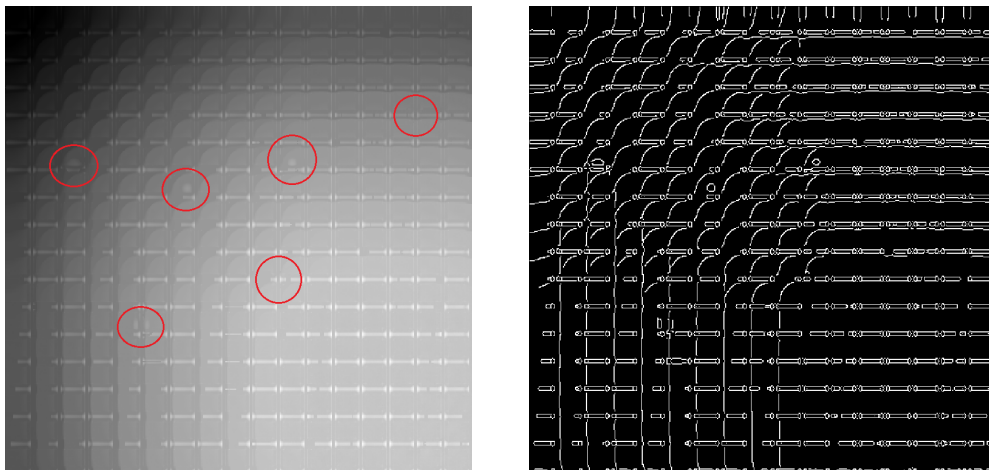


Abbildung 6.4: Links zeigt das Ergebnis  $u_{L1}$  mit  $\lambda_{L1} = 0.315$ ,  $\lambda_{L2} = 0.05$  und Anzahl der Iterationen 500 bzw. 800. Die Fehler sind für das menschliche Auge schwer zu erkennen, weshalb sie mit einem roten Kreis markiert sind. Rechts zeigt das zugehörige Kantenbild mittels des Canny-Edge-Detection Algorithmus [6].

### 6.3 Erzielte Ergebnisse mit den Modellierungsansätzen auf Röntgenbildern

Die Ergebnisse mit dem Modellierungsansatz mit dem ROF-Modell und dem  $L^1$ -TV-Modell sind in der Abb. 6.5 dargestellt. In der linken Spalte sind die Realbildausschnitte aus der firmeninternen Bildersammlung von VisiConsult X-ray Systems & Solutions GmbH [3] und wurden für die Analysen der Cepstralanalyse und dem Modellierungsansatz zugeschnitten. In der mittleren Spalte sind die Ergebnisse  $u_{\text{ROF}}$  dargestellt. Die rechte Spalte repräsentiert die Ergebnisse  $u_{L^1}$ . Mit der Cepstralanalyse konnten keine repräsentativen Bilder erstellt werden.

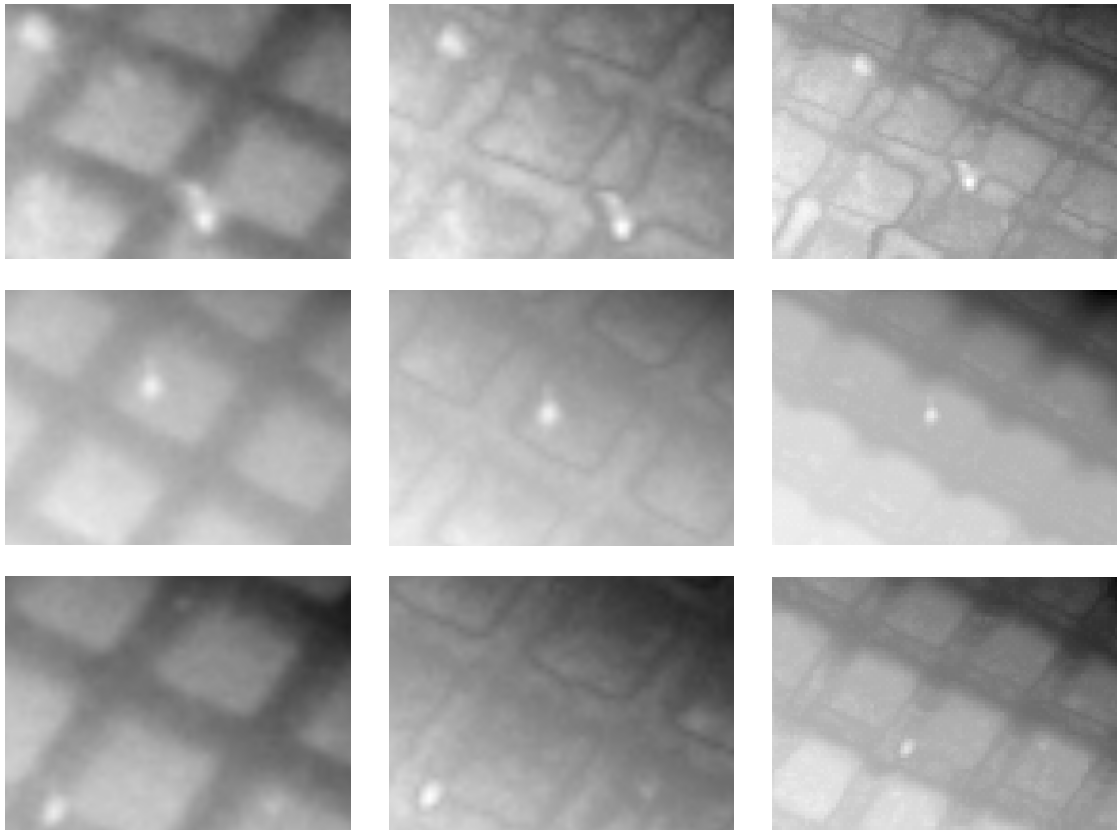


Abbildung 6.5: Erzielte Ergebnisse unter der Verwendung des Modellierungsansatzes. Links zeigt die Originalbilder (aus [3]), Mitte zeigt  $u_{\text{ROF}}$  und rechts zeigt  $u_{L^1}$  [6].



---

## Kapitel 7: Diskussion und Schlussfolgerung

Diese Arbeit entsteht in Zusammenarbeit mit der Firma VisiConsult X-ray Systems & Solutions GmbH aus Stockelsdorf. Dabei handelt die Aufgabenstellung von der Erstellung eines Vorverarbeitungsprozesses, für eine automatische Fehlererkennung im firmeninternen Tool VAIP. Dabei soll die Entscheidung des VAIPs über das Beibehalten oder Verwerfen von Gussteilen mit Gitterstrukturen auf der Oberfläche nicht von den Gitterstrukturen beeinflusst werden. Dabei können die Gitterstrukturen wie in Kapitel 1 in allen möglichen Krümmungen, Verzerrungen und generell nicht linearen Darstellungen vorkommen. Eine umfangreiche Literaturrecherche blieb ergebnislos, was vermutlich auf die spezifische Formulierung der Aufgabenstellung zurückzuführen ist. Vor diesem Hintergrund werden in dieser Arbeit zwei potenzielle Ansätze zur Reduktion von Gitterstrukturen vorgestellt und auf ihre Anwendbarkeit in der Praxis überprüft. Es ist wichtig anzumerken, dass diese Arbeit keine Lösung für das Problem liefert, sondern lediglich auf vereinzelte Bilder angewendet wurde und somit als eine explorative Grundlage für weiterführende Forschungen dient. Die vorgestellten Ansätze bieten wertvolle Denkanstöße und zeigen Möglichkeiten auf, wie das Problem möglicherweise angegangen werden kann. Weitere Untersuchungen sind erforderlich, um eine vollständig auf das spezifische Problem zugeschnittene Lösung zu entwickeln.

Die Cepstralanalyse ist ein Verfahren, das von Bogert et al. [18] zur Erforschung von seismischen Daten entwickelt wurde. In bisherigen Anwendungen auf zweidimensionalen Bildern wird sie beispielsweise zur Extraktion von Merkmalen [10], zur Schattendetektion [21] oder zur Bewertung der visuellen Qualität [23] eingesetzt. Die Cepstralanalyse eignet sich als Instrument zur Erkennung von Periodizität in einem Spektrum, und daher erscheinen bei synthetischen Bildern periodische Gitterstrukturen in gleichmäßigen Abständen als Nebenmaxima auf den Hauptachsen. Dieser Abstand entspricht der Gitterweite  $h$ .

Die Idee hinter der Anwendung der Cepstralanalyse in dieser Arbeit besteht darin, dass durch die Abschwächung dieser Nebenmaxima auch eine Abschwächung des Gittermusters auf dem Bild erreicht werden kann. Allerdings ist das Cepstrum aufgrund seiner Definition mit den Beträgen nicht invertierbar. Daher wird ein eigens erstelltes modifiziertes Cepstrum eingeführt, um eine Transformation vom Quefrenzy-Raum in den Bildraum zu ermöglichen (vgl. Definition 3.8 und 3.9). Auf diese Weise kann eine gezielte Modifikation des Gittermusters erreicht werden, indem die entsprechenden Nebenmaxima im Cepstrum geschwächt werden. Dies eröffnet neue Möglichkeiten zur Reduktion von Gitterstrukturen und zur Verbesserung der visuellen Qualität der Bilder in der vorliegenden Arbeit.

Die Ergebnisse in Abb. 6.1 und Abb. 6.2 deuten darauf hin, dass sowohl der Schwellenwertansatz als auch der Maskenansatz vielversprechend zu sein scheinen, da sie Ausgabebilder liefern, in denen sich die Fehler visuell deutlich sichtbar vom Hintergrund und

den Gitterstrukturen unterscheiden. Durch das Logarithmieren und die anschließende Reduzierung unterscheiden sich die Intensitätswerte stark von denen des Ausgangsbildes (Mittelwerte des gesamten Bildes: 195.0 beim Schwellenwertansatz, 477.9 beim Maskenansatz und 7630.3 bei  $u_{\text{In}}$ , vgl. Tabelle 7.1). Die Ergebnisse zeigen, dass die Anwendung der Cepstralanalyse in dieser Arbeit (vgl. Kapitel 6.1) nicht nur zu einer Reduktion der Gitterstrukturen führt, sondern auch zu einer erheblichen Veränderung der Intensitätswerte. Das Ergebnis unterscheidet sich stark vom gewünschten Ergebnis  $u_{\text{Err}}$  (SSIM beider Ansätze sinken im Vergleich zum Eingabebild  $u_{\text{In}}$ , der MSE und die Frobeniusnorm werden jeweils um mehr als den Faktor 100 größer, vgl. Tabelle 7.2).

	Mittelwert $u_{\text{Err}}$ bzw. $u_{\text{Pat}}$	Mittelwert $u_{\text{TCep}}$ mit $r_{\text{Center}} = 10$	Anteil [%]
Fehler 1	6727.1	115.0564	1.7
Fehler 2	8322.2	599.9141	7.2
Fehler 3	8357.0	251.2324	3.0
Fehler 4	7963.2	127.2598	1.6
Fehler 5	8249.7	169.2912	2.1
Fehler 6	9523.7	464.4916	4.9
Gitterstrukturen	7018.7	147.3875	2.1

Tabelle 7.1: Übersicht der Mittelwerte der jeweiligen Fehler und der Gitterstruktur von  $u_{\text{Err}}$  (bei den Fehlern 1-6) bzw.  $u_{\text{Pat}}$  (bei den Gitterstrukturen) und prozentualen Anteil der Intensitätswerte nach dem Schwellenansatz.

Jedoch zeigen die Abbildungen 6.1 und 6.2, dass die Cepstralanalyse als Vorverarbeitung für das firmeninterne Tool VAIP zur automatischen Fehlererkennung für synthetische Daten geeignet ist, da sie Fehler erkennt, die zuvor aufgrund von Gitterstrukturen nicht erkannt wurden (vgl. das Kantenbild, das fünf der sechs Fehler auch als einzelne Fehler erkennt, obwohl die Gitterstrukturen zuvor zu zwei separaten Fehlern geführt hat und die Fehlergrößenmessung in Tabelle 7.3).

	$\ \cdot\ _F$	SSIM	MSE	PSNR
$u_{\text{In}}$	2.7359e+05	0.1981	2.8554e+05	41.7728
$u_{\text{TCep}}$	4.0197e+06	0.0029	6.1639e+07	18.4309
$u_{\text{MCep}}$	3.8675e+06	0.0190	5.7058e+07	18.7663
$u_{\text{ROF}}$	11.9799	0.9121	5.4748e-04	32.6163
$u_{L^1}$	13.3611	0.9268	6.8100e-04	31.66850

Tabelle 7.2: Ergebnisse der Metriken zur Qualitätsanalyse von  $u_{\text{In}}$ ,  $u_{\text{TCep}}$ ,  $u_{\text{MCep}}$ ,  $u_{\text{ROF}}$  und  $u_{L^1}$  mit  $u_{\text{Err}}$ .

In Anbetracht dieser Erkenntnisse lässt sich schlussfolgern, dass die hier vorgestellten Metriken allein keine angemessene Beurteilung der Eignung der Cepstralanalyse ermöglichen, denn die Bilder  $u_{\text{TCep}}$  und  $u_{\text{MCep}}$  wurden so stark verändert, dass die Interpretati-

on der Metriken nicht sinnvoll erscheint. Die Interpretation der Metriken würde nämlich bedeuten, dass die Cepstralanalyse zur Reduktion der Gitterstrukturen ungeeignet ist. Jedoch liefern sowohl Tabelle 7.3, als auch die Abbildungen 6.1 und 6.2 deutliche Hinweise darauf, dass nach der Verarbeitung des Eingabebildes  $u_{\text{In}}$  mit der Cepstralanalyse fast alle Fehler in ihrer ursprünglichen Größe gefunden werden. Bei der durchgeführten Messung der Fehlergrößen lieferte der Maskenansatz bessere Ergebnisse als der Schwellenwertansatz.

	Größe [pix <sup>2</sup> ]	$u_{\text{TCep}}$	$u_{\text{MCep}}$	$u_{\text{ROF}}$	$u_{L1}$
Fehler 1	241	225	228	-	-
Fehler 2	115	149	114	61	71
Fehler 3	61	86	54	54	46
Fehler 4	55	-	-	-	-
Fehler 5	336	279	271	-	-
Fehler 6	172	55	100	-	-

Tabelle 7.3: Übersicht der Messungen der erkannten Fehlergrößen des ADRs im Vergleich für die verschiedenen Ergebnisbilder  $u_{\text{TCep}}$ ,  $u_{\text{MCep}}$ ,  $u_{\text{ROF}}$  und  $u_{L1}$ . Bei „-“ gibt es keine Messungen. Die Größenangaben sind in pix<sup>2</sup> (Pixel zum Quadrat) gegeben.

In der Praxis soll das Detektieren der Fehler in „so schnell wie möglich“ laufen. Dank der effizienten Berechnung von der diskreten Fouriertransformation mittels der FFT (vgl. Kapitel 3.5) scheint die Cepstralanalyse diese Anforderungen zu genügen [7]. Die tatsächliche Laufzeit ist von vielen Faktoren abhängig und wird aus diesem Grund hier nicht näher erläutert. Es reicht zu sagen, dass das Programm auf einem firmenüblichen Laptop in wenigen Millisekunden durchläuft.

Damit der Vorteil der FFT, und damit die effiziente Laufzeit, genutzt werden kann, sollte das Eingabebild  $u_{\text{In}}$  nach Çakir et al. [10, 11] die Dimension einer Größenordnung einer Zweierpotenz  $2^N$  haben,  $N \in \mathbb{N}$ .

Für eine Einbindung in die firmeninterne Pipeline ist die Cepstralanalyse zur Vorverarbeitung zum jetzigen Stand (25.07.2023) noch nicht geeignet, denn es sind mehrere bilderabhängige Parameter zu bestimmen. Weiter konnten keine präsentierbaren Ergebnisse bei Anwendung der Cepstralanalyse auf den Realbilddatenausschnitten möglich. Das wird wahrscheinlich daran liegen, dass das Gittermuster in diesen Ausschnitten in zu weniger Frequenz auftreten, sodass sich keine Nebenmaxima bildeten, die bei Dämpfung relevante Ergebnisse lieferten. Jedoch ist das Cepstrum ein gewaltiges Werkzeug und seine Analyse ist bis hier nicht beendet.

Ein weiterer vorgestellter Ansatz basiert auf der Voraussetzung, dass die Gitterstrukturen einen Großteil des Bildes einnehmen. Im Kontext von zweidimensionalen Funktionen beschreibt die Totale Variation einer charakteristischen Funktion auf einer gegebenen

Menge die Länge der Kurve, die diese Menge umschließt. Dies wird auch als Perimeter (Umfang einer Menge) bezeichnet [28]. Die Totale Variation eignet sich in der Bildverarbeitung als Regularisierer, denn diese erlauben Sprünge im Bild. Diese treten häufig in Röntgenbildern auf, wenn es zu großen Intensitätsveränderungen von einem Pixel zum nächsten kommt. Unter der Voraussetzung, dass die Gitterstrukturen sich über einen Großteil des Bildes verlaufen, können diese mittels der Totalen Variation geschwächt werden, denn die Fehler treten in der Regel mit kleineren Perimetern auf im Vergleich zu den Gitterstrukturen. So können die Informationen von Gitterstrukturen und Fehler getrennt behandelt werden. Die Ergebnisse zeigen, dass die Anwendung des modellbasierten Ansatzes für die Anwendung der synthetischen Bilder weniger gut geeignet ist als die Cepstralanalyse (vgl. die Abbildungen 6.3 und 6.4). Sowohl beim ROF-Modell als auch beim  $L^1$ -TV-Modell wurden nur zwei der sechs Fehler gefunden, siehe Tabelle 7.3. Die Intensitätswerte von  $u_{\text{ROF}}$ ,  $u_{L^1}$  hingegen variieren nicht so stark zu  $u_{\text{In}}$ , weshalb auch ein deutlich höherer SSIM (91.2% bzw. 92.7% im Vergleich zu 0.3% bzw. 1.9% bei der Cepstralanalyse (vgl. Tabelle 7.2)) erscheint. Andererseits lieferte die Anwendung des Modellierungsansatzes beeindruckende Ergebnisse bei der Reduktion der Gitterstrukturen bei den Realdatenausschnitten, siehe Abb. 6.5. Auch bei diesem Ansatz ist eine weitere Untersuchung erstrebenswert, z.B. auch mit einer weighted-Total-Variation als Regularisierer. Ein Vorteil des Modellierungsansatzes mit der Totalen Variation als Regularisierer gegenüber der Cepstralanalyse liegt darin, dass weitere Kenntnisse über die Gitterstrukturen nicht notwendig sind. Ein weiterer Schritt ist die Untersuchung eines Algorithmus, der die benötigten Parameter bestimmt. Auch eine Kombination der Zwischenergebnisbilder  $u_{\text{ROF1}}^*$ ,  $u_{\text{ROF2}}^*$ ,  $u_{L^1}^*$  und  $u_{L^2}^*$  ist denkbar.

Betrachten wir zum Ende noch einmal die Leitfragen:

- Eignen sich die hier vorgestellten Methoden zur Reduktion der Gitterstrukturen im Bildraum, sodass die Informationen der darunterliegenden Strukturen beibehalten werden?
- Wie kann eine Reduktion der Gitterstrukturen objektiv gemessen werden?

Die hier vorgestellten Methoden haben beide jeweils ihre Vor- und Nachteile. Die Ergebnisse bei den synthetischen Bildern scheinen bei der Cepstralanalyse geeigneter als Vorverarbeitungsprozess zur automatischen Fehlererkennung zu sein. Die Anwendung auf den Realdaten hingegen spricht für den modellbasierten Ansatz (vgl. Kapitel 6.3). Ein Problem mit der Cepstralanalyse ist, dass die Nebenmaxima bei der Transformation der Realdatenausschnitte in den Quefrequency-Raum auf den Hauptachsen nicht so stark ausgeprägt sind, wie es bei den synthetischen Bildern der Fall war. Das kann daran liegen, dass die Gitterstrukturen auf einem Bildabschnitt nicht mit ausreichender Frequenz vorhanden sind.

Auch die Frage nach einer objektiven Messung bleibt noch aus. Die hier genannten Metriken sind nicht für die Beurteilung auf Eignung der hier vorgestellten Methoden



---

geeignet, denn entweder wurde das Bild so stark verändert, dass eine Interpretation von SSIM, MSE, PSNR oder der Frobeniusnorm nicht sinnvoll ist. Oder die Messungen scheinen vielversprechend zu sein, aber das Detektieren der Fehler war nicht erfolgreich.

Der nächste Schritt ist eine Analyse des energiebasierten Modells, die klären soll, weshalb es zu nicht befriedigenden Ergebnissen auf den synthetischen Bildern gekommen ist. Weiter sollen Analysen beider Methoden auf weiteren synthetischen Bildern erfolgen, denn in dieser Arbeit basieren alle Ergebnisse (außer den auf den Realbildern) auf nur einem Bild. Auch fand das komplexe Cepstrum bisher keine Anwendung, was nachgeholt wird. Weiter wurde von Çakir und Çetin [10, 11] eine Methode zur Extraktion von Merkmalen mittels des 2D-Mel-Cepstrums vorgestellt. Dieser Ansatz basiert auf der 2D Cepstralanalyse und ist eine weitere Analyse wert. Zusammenfassend kann man sagen, dass es noch weitere Analysen der Methoden bedarf, aber bereits gute, aber noch nicht praxistaugliche Ergebnisse erzielt wurden.



---

## Kapitel A: Anhang

	Beschreibung
$u_{\text{Base}}$	Unkalibriertes Detektorbild mit Rauschen, $u_{\text{Base}} \in [0, 65535]^{512 \times 512}$
$u_{\text{Err}}$	Synthetisches Bild mit eingerechneten Lufteinschlüssen (5% höhere Intensität) und Rauschen, $u_{\text{Err}} \in [0, 65535]^{512 \times 512}$
$u_{\text{Pat}}$	Synthetisches Bild mit Gitterstrukturen (55% geringere Intensität) und Rauschen, $u_{\text{Pat}} \in [0, 65535]^{512 \times 512}$
$u_{\text{In}}$	Synthetisches Bild mit Lufteinschlüssen, Gitterstrukturen und Rauschen, $u_{\text{In}} \in [0, 65535]^{512 \times 512}$
$c_{\text{Base}}$	Power Cepstrum von $u_{\text{Base}}$
$c_{\text{Pat}}$	Power Cepstrum von $u_{\text{Pat}}$
$c_{\text{In}}$	Power Cepstrum von $u_{\text{In}}$
$c_{\text{Diff}}$	Betrag des Differenzbildes von $c_{\text{Pat}}$ und $c_{\text{Base}}$
$u_{\text{ROF1}}^*$	ROF-Modell 1 (Ohne Materialfehler), $u_{\text{ROF1}}^* \in [0, 1]^{512 \times 512}$
$u_{\text{ROF2}}^*$	ROF-Modell 2 (Ohne Gitterstrukturen), $u_{\text{ROF2}}^* \in [0, 1]^{512 \times 512}$
$u_{L1}^*$	$L^1$ -TV-Modell 1 (Ohne Materialfehler), $u_{L1}^* \in [0, 1]^{512 \times 512}$
$u_{L2}^*$	$L^1$ -TV-Modell 2 (Ohne Gitterstrukturen), $u_{L2}^* \in [0, 1]^{512 \times 512}$
$u_{\text{TCep}}$	Ergebnis der Cepstralanalyse mit dem Schwellenwertansatz, $u_{\text{TCep}} \in [0, 65535]^{512 \times 512}$
$u_{\text{MCep}}$	Ergebnis der Cepstralanalyse mit dem Maskenansatz, $u_{\text{MCep}} \in [0, 65535]^{512 \times 512}$
$u_{\text{ROF}}$	Ergebnis des ROF-Modells, $u_{\text{ROF}} \in [0, 1]^{512 \times 512}$
$u_{L1}$	Ergebnis des $L^1$ -TV-Modells, $u_{L1} \in [0, 1]^{512 \times 512}$

Tabelle A.1: Übersicht aller in Kapitel 6 und 7 verwendeten Bilder und Power-Cepstra.

	Beschreibung
$\tau_{\text{Red}}$	Schwellenwert, dient zur Steuerung der Stärke der Reduktion mit $\tau_{\text{Red}} \in [0, 1]$
$\tau_{\text{Diff}}$	Schwellenwert zur Bestimmung, ab wann die Differenz in $c_{\text{Diff}}$ als groß genug gilt für eine Reduktion mit $\tau_{\text{Diff}} \in \mathbb{R}_{\geq 0}$
$r_{\text{Center}}$	Radius für die Umgebung um das Zentrum, der erhalten bleibt mit $r_{\text{Center}} \in \mathbb{N}$
$R_{x_1}$	Parameter, der die Umgebung in $x_1$ Richtung um die Regionen der Nebenmaxima herum bestimmt, die gitterförmig um den Ursprung liegen, mit $R_{x_1} \in \mathbb{N}_0$
$R_{x_2}$	Parameter, der die Umgebung in $x_2$ Richtung um die Regionen der Nebenmaxima herum bestimmt, die gitterförmig um den Ursprung liegen, mit $R_{x_2} \in \mathbb{N}_0$
$h$	Gitterweite von $u_{\text{Pat}}$ und $u_{\text{In}}$ mit $h = (30, 30)$
$\lambda_{\text{ROF1}}$	Gewichtungsparameter für $u_{\text{ROF1}}^*$ mit $\lambda_{\text{ROF1}} \in \mathbb{R}_{\geq 0}$
$\lambda_{\text{ROF2}}$	Gewichtungsparameter für $u_{\text{ROF2}}^*$ mit $\lambda_{\text{ROF2}} \in \mathbb{R}_{\geq 0}$
$\lambda_{\text{L1}}$	Gewichtungsparameter für $u_{\text{L1}}^*$ mit $\lambda_{\text{L1}} \in \mathbb{R}_{\geq 0}$
$\lambda_{\text{L2}}$	Gewichtungsparameter für $u_{\text{L2}}^*$ mit $\lambda_{\text{L2}} \in \mathbb{R}_{\geq 0}$

Tabelle A.2: Übersicht der in Kapitel 6 und 7 verwendeten Parametern.

## Literatur

- [1] VisiConsult X ray Systems Solutions GmbH. zuletzt aufgerufen 07.2023. URL <https://visiconsult.de/technologien/automated-defect-recognition-adr/>.
- [2] Nathalie Weber. Eigene Bildererstellung mit aRTist. 2023.
- [3] VisiConsult X-ray Systems Solutions GmbH. Firmeninterne Bildersammlung. 2023.
- [4] Kristian Bredis and Dirk Lorenz. *Mathematische Bildverarbeitung-Einführung in Grundlagen und moderne Theorie*. Vieweg+Teubner, 2015.
- [5] Jan Modersitzki. *FAIR Flexible Algorithms for Image Registration*. Society for Industrial Applied Mathematics,U.S., 2009.
- [6] Nathalie Weber. Eigene Bildererstellung mit Matlab. 2023.
- [7] VisiConsult X-ray Systems Solutions GmbH. Private Diskussion. 2023.
- [8] VisiConsult X-ray Systems Solutions GmbH. Automatische Auswertung. Internetseite wurde zuletzt aufgerufen im Juli 2023. URL <https://visiconsult.de/technologien/automated-defect-recognition-adr/>.
- [9] VisiConsult X ray Systems Solutions GmbH. Automated defect recognition for 24/7 automated in-line inspection. zuletzt aufgerufen 07.2023. URL <https://visiconsult.de/documents/Technology/ADR/VC-ADR.pdf>.
- [10] Serdar Çakir and A. Enis Çetin. Image Feature Extraction Using 2D Mel-Cepstrum. *IEEE*, 2010.
- [11] Serdar Çakir and A. Enis Çetin. Two-Dimensional Mellin and Mel-Cepstrum for Image Feature Extraction. *Springer Science+Business Media B.V.*, page 271–276, 2011.
- [12] Jürgen Pöschel. *Noch mehr Analysis*. Springer-Verlag GmbH, 2015.
- [13] Wilhelm Burger und Mark James Burge. *Digitale Bildverarbeitung*. Springer-Verlag, 2009.
- [14] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer Spektrum, 2008.
- [15] Angelika Erhardt. *Einführung in die Digitale Bildverarbeitung*. Vieweg+Teubner, 2008.
- [16] Bundesanstalt für Materialforschung und -prüfung (BAM). Internetseite wurde zuletzt aufgerufen im Juli 2023. URL <https://artist.bam.de/>.

- [17] Leonid I. Rudin, Stanley Osher und Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 1992.
- [18] Bruce P. Bogert, M. J. R. Healy und John W. Tuckey. The Quefreny Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking. 1963.
- [19] R. C. Kemerait und D. G. Childers. Signal Detection and extraction by Cepstrum Techniques. *IEEE Transactions on Information Theory*, 1972.
- [20] Thomas M.Lehmann, Carsten Goerke, A. Kaupp, Walter Schmitt und Rudolf Repges. A Rotation-extended Cepstrum Technique and Its Application to Medical Images. *Digital Signal Processing*, 1996.
- [21] B. Ugur Toreyin und A. Enis Çetin. Shadow Detection Using 2D Cepstrum. *SPIE Digital Library*, 2009.
- [22] A. V. Oppenheim. From Frequency to Quefreny: A History of the Cepstrum. *IEEE Signal Processing Magazine*, 2004.
- [23] Serdar Çakir and A. Enis Çetin. Image quality assessment using two-dimensional complex mel-cepstrum. *Journal of Electronic Imaging*, 2016.
- [24] Raphael Rom. On the Cepstrum of Two-Dimensional Functions. *IEEE Transactions on Information Theory*, 1975.
- [25] Roger A. Horn. The Hadamard Product. 1990.
- [26] R.B. Randall, B. Tech., B.A. Bruel Kjaer. Cepstrum Analysis and Gearbox Fault Diagnosis. 2007.
- [27] James Cooley und John Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, 1965.
- [28] Prof. Dr. rer. nat. Jan Moderitzki. Private Diskussion. 2023.
- [29] Antonin Chambolle, Viscent Caselles, Matteo Novaga, Daniel Cremers und Thomas Pock. *An introduction to Total Variation for Image Analysis*. HAL open science, 2009.
- [30] Uwe Storch, Hartmut Wiebe und Claas Becker. *Maß- und Integrationstheorie*. Springer Verlag GmbH, 2020.
- [31] Luigi Ambrosio, Nicola Fusco und Diego Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.
- [32] Friedrich Sauvigny. *Analysis*. Springer Spektrum, 2014.
- [33] Manolis Lourakis. zuletzt aufgerufen 07.2023. URL <https://de.mathworks.com/matlabcentral/fileexchange/57604-tv-l1-image-denoising-algorithm>.

- [34] Philippe Magiera Carl Löndahl. aus dem Jahr 2008. URL <https://de.mathworks.com/matlabcentral/fileexchange/22410-rof-denoising-algorithm>.
- [35] Tom Goldstein und Stanley Osher. The split Bregman method for L1 regularized Problems. *SIAM Journal on Imaging Sciences*, pages 323–343, 2009.
- [36] Antonin Chambolle und Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, pages 120–145, 2010.
- [37] Antonin Chambolle. An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision* 20, pages 89–97, 2004.
- [38] Laurent Condat. Discrete Total Variation: New Definition and Minimization. *HAL open science*, 2016.
- [39] Antonin Chambolle und Thomas Pock. *Learning Consistent Discretizations of the Total Variation*. Society for Industrial and Applied Mathematics, 2021.
- [40] Gene Golub und Charles van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.
- [41] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh und Eero P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on image processing*, 2004.
- [42] John Canny. A Computational Approach to Edge Detection. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 1986.
- [43] Richard Szeliski. *Computer Vision: Algorithms and Applications 2nd Edition*. Springer, 2022.
- [44] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on systems*, 1979.

