

Super-Resolution-Techniken zur Steigerung des lateralen Auflösungsvermögens von 3D-PMD-Kameras

Dissertation
zur Erlangung des akademischen Grades

Doktoringenieur
(Dr.-Ing.)

vorgelegt der
Fakultät für Maschinenbau der
Technischen Universität Ilmenau

von Herrn

M.Sc. Henrik Lietz
geboren am 13.01.1988 in Reutlingen

1. Gutachter: Univ.-Prof. Dr. rer. nat. habil. Stefan Sinzinger
2. Gutachter: Univ.-Prof. Dr. rer. nat. Gunter Notni
3. Gutachter: Prof. Dr. Jörg Eberhardt

Tag der Einreichung: 20.04.2021

Tag der wissenschaftlichen Aussprache: 29.03.2022

DOI: 10.22032/dbt.51968

URN: urn:nbn:de:gbv:ilm1-2022000142

Für Fenia und Jonte

Danksagung

An dieser Stelle bedanke ich mich bei den zahlreichen Personen, die mich während der Promotionsphase direkt oder indirekt unterstützt haben. Ohne deren Beiträge wäre diese Arbeit nicht zustande gekommen.

Dazu zählt in erster Linie meine liebe Frau Nadja. Ihr danke ich von Herzen für die außerordentliche Unterstützung, die sie mir abseits der Wissenschaftswelt entgegengebracht hat, ihre Geduld - insbesondere während der schriftlichen Ausarbeitung - sowie für die unzähligen Stunden, in denen sie mit der Kinderbetreuung auf sich alleine gestellt war.

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. rer. nat. habil. Stefan Sinzinger für die wissenschaftliche Betreuung dieser Arbeit, seinen konstruktiven Diskussionen an der Technischen Universität Ilmenau vor Ort, bei Konferenzen und per Telefon. Egal über welchen Kanal, er hatte sich stets Zeit für meine Anliegen genommen und unterstützte mich dabei meinen Weg zu finden und zu gehen.

Mein nicht minderer Dank gilt Herrn Prof. Dr. Jörg Eberhardt für die fachliche Betreuung im wissenschaftlichen Umfeld der Hochschule Ravensburg-Weingarten. Durch seine Expertise und Erfahrungen im Bereich der digitalen Bildverarbeitung war er mein Hauptansprechpartner bei fachlichen Belangen. Nicht zuletzt seine kritische gleichwohl motivierende Art, trugen einen wesentlichen Beitrag zum erfolgreichen Abschluss dieser Arbeit bei.

Ganz herzlich danke ich Herrn Prof. Dr. sc. techn. Michael Pfeffer, der ein erstes Kennenlernen mit meinem Doktorvater ermöglicht hatte und so den Grundpfeiler für diese wissenschaftliche Ausarbeitung legte.

Abschließend danke ich meiner Familie sowie den zahlreichen Freunden, Kollegen, Studierenden und sonstigen Personen, die mich, in welcher Form auch immer, beim Gehen dieses Weges begleiteten und mich bis zuletzt unterstützten, motivierten und gelegentlich auch ablenkten.

Herzlichen Dank!

Kurzfassung

PMD-Kameras ermöglichen die 3D-Bilderfassung in Echtzeit und liefern robuste Aufnahmen auch bei schwierigen Lichtverhältnissen. Allerdings limitiert deren geringe Bildauflösung die Detektion kleiner Objekte erheblich. Unter dem Einsatz rechnergestützter Bildverarbeitung verschieben Super-Resolution-Techniken (SR) diese Grenze hin zu einem höheren lateralen Auflösungsvermögen, ohne dabei die spezifischen Eigenschaften des Sensors zu verändern. In der Literatur wird die Leistungsfähigkeit von SR-Methoden mit Rauschanalysen bewertet, indem die SR-Ergebnisse mit ihrem fehlerfreien Ground-Truth-Referenzbild (GT) verglichen werden. Dabei beschränken sich die SR-Methoden zumeist auf definierte Anwendungsfälle und deren Auswertungen auf synthetisch erzeugte Datensätze, ohne die Auflösungssteigerung am realen Aufbau zu verifizieren.

Mit dieser Arbeit wird eine generische SR-Methode auf eine echte, niedrig auflösende PMD-Kamera appliziert, die für den Einsatz im Außenbereich konzipiert ist und eine Bildauflösung von lediglich 64 x 16 Bildpunkten aufweist. Zunächst erfolgen simulative Untersuchungen von drei aktuellen SR-Techniken hinsichtlich deren Eignung zur generischen Steigerung der lateralen Ortsauflösung einer PMD-Kamera:

- 1) KI-lernbasierte Single Image SR (SISR)
- 2) Inkohärente Fourier Ptychographie (FP)
- 3) Subpixel-Shift SR (SpS-SR)

Ein Vergleich der Simulationsergebnisse zeigt Vorteile für die SpS-SR, sodass diese mit einem neuartigen Lens-Shift-Aufbau in die praktische Anwendung überführt wird.

Die geringe Bildauflösung der eingesetzten PMD-Kamera schließt die Bewertung der SR-Ergebnisse nach einer klassischen Rauschanalyse aus. Stattdessen wird die erreichte Bildqualität im SR-Amplitudenbild nach dem optischen Auflösungsvermögen in Linienpaaren pro Pixel [LP/mm] bewertet. Dieses Absolutmaß gewährleistet eine Vergleichbarkeit der SR-Ergebnisse aus Simulation und Messung auch für Kameras, von denen keine hochauflösenden GT-Bilder existieren.

Letztlich zeigt diese Arbeit die Möglichkeiten und Grenzen der Auflösungssteigerung mit SR-Methoden in der praktischen Anwendung für den spezifischen Anwendungsfall bei PMD-Kameras auf. Darüber hinaus bildet der hier vorgestellte Lens-Shift-Aufbau die Basis zur flexiblen und kostengünstigen Auflösungssteigerung digitaler Kameras im Allgemeinen, wie sie die bildverarbeitende Optik-Industrie immer häufiger fordert.

Abstract

PMD cameras enable real-time 3D image acquisition and deliver robust images even under poor lighting conditions. However, their low spatial resolution significantly limits the detection of small objects. Super resolution (SR) techniques shift this limit towards a higher image resolution without changing the sensor's specific characteristics by using computational image processing. In the literature, noise analyses evaluate the performance in increasing image resolution by comparing the SR result with its underlying ideal ground-truth image (GT). SR methods are mostly limited to defined applications while their evaluation is limited to synthetically generated data sets without verifying resolution enhancement on a real setup.

In this work, a generic SR method is applied to a real low-resolution PMD camera, which is designed for outdoor applications and has a spatial resolution of only 64 x 16 pixels. First, simulative investigations of three state-of-the-art SR techniques for the generic increase of a PMD sensor's lateral resolution are performed:

- 1) AI-based single image SR (SISR)
- 2) Incoherent Fourier Ptychography (FP)
- 3) Subpixel shift SR (SpS-SR)

A comparison of the simulation results shows advantages for the SpS method, so that it is transferred to practical application with a novel lens-shift setup.

The low spatial resolution of the PMD camera used excludes the evaluation of the SR results according to a classical noise analysis. Instead, the evaluation of the super resolved amplitude image achieved is performed according to the optical resolving power in line pairs per pixel [LP/mm]. This absolute measure ensures comparability of SR results from simulation and measurement data even for cameras from which high-resolution GT images does not exist.

In the end, this work demonstrates the possibilities and limitations of resolution enhancement with SR methods in practical applications for the specific use case with PMD cameras. In addition, the lens-shift setup presented here provides the basis for flexible and cost-effective resolution enhancement of digital cameras in general, as increasingly demanded by the image-processing optics industry.

Inhaltsverzeichnis

Danksagung	v
Kurzfassung	vii
Abstract	ix
1 Einleitung	1
2 3D-PMD-Kameras	7
2.1 Funktionsweise von PMD-Kameras	7
2.2 Charakterisierung der verwendeten PMD-Kamera	11
3 Rechnerische Erhöhung der Bildauflösung	13
3.1 PCHIP-Interpolation	13
3.2 Geometrische Super-Resolution	15
3.3 Super-Resolution Beobachtungsmodell	16
3.4 Super-Resolution Problemstellungen	18
4 Lernbasierte Single Image Super-Resolution	21
4.1 Super-Resolution Convolutional Neural Networks	22
4.2 EnhanceNet Super-Resolution	25
5 Multi-Frame Super-Resolution	29
5.1 Inkohärente Fourier Ptychographie	29
5.1.1 Funktionsweise der inkohärenten Fourier Ptychographie	30
5.1.2 FP-Simulationsmodell für PMD-Kameras	32
5.2 Subpixel-Shift Super-Resolution	34
5.2.1 Funktionsweise der Subpixel-Shift Super-Resolution	35
5.2.2 Iterativley Re-weighted Minimization Algorithmus	36
6 Methoden zur Bewertung von Auflösungssteigerungen	39
6.1 Ortsfrequenzganganalyse nach der Schräge-Kante-Methode	40
6.2 USAF-1951 Auflösungsstesttafel	41
6.3 PSNR Rauschanalyse	42
7 Simulationen	43
7.1 Erzeugung synthetischer Testbilder	43
7.2 Umsetzung der Simulationen	46
7.3 Simulationsergebnisse auf Basis der synthetischen Testbilder	48
7.3.1 Interpretation der Simulationsergebnisse	53

7.3.2	Zwischenfazit zu den Simulationsergebnissen	56
8	Demonstratoren	59
8.1	Sensor-Shift Verfahren	59
8.2	Lens-Shift Verfahren	60
8.3	Messbedingungen	62
9	Experimentelle Ergebnisse	63
9.1	Messergebnisse	63
9.2	Anpassung des Super-Resolution Beobachtungsmodells	67
9.3	Simulationsergebnisse der SpS-SR mit angepasstem Beobachtungsmodell	70
9.3.1	Einfluss der Parameterwahl auf das Simulationsergebnis	74
9.3.2	Interpretation der Simulationsergebnisse mit angepasstem Beobachtungsmodell	74
9.3.3	Optimierungspotenzial am angepassten Simulationsmodell	78
9.4	Rauschanalyse	78
9.5	Anwendung der Subpixel-Shift SR auf Entfernungsbilder	79
9.6	Abschließende Diskussion	82
10	Zusammenfassung und Ausblick	85
A	Codeverzeichnis	89
A.1	Simulationsmodell zur inkohärenten Fourier Ptychographie	89
A.2	Erzeugen einer LR-Bilderfolge nach dem allgemeinen Beobachtungsmodell	92
A.3	Erzeugen einer LR-Bilderfolge nach dem angepassten Beobachtungsmodell	93
	Literaturverzeichnis	95
	Abbildungsverzeichnis	105
	Tabellenverzeichnis	109
	Abkürzungs- und Symbolverzeichnis	113
	Publikationsliste	119
	Fachvorträge	121
	Betreuung studentischer Abschlussarbeiten	123
	Erklärung	125

Kapitel 1

Einleitung

Mit dem ersten Kapitel wird die Umsetzung moderner Super-Resolution-Techniken zur generischen Steigerung des lateralen Auflösungsvermögens niedrig auflösender 3D-PMD-Kameras in ihren wissenschaftlichen Kontext eingeordnet. Dazu erfolgt zunächst eine Beschreibung zum Stand der Technik mit seinen Einschränkungen, bevor daraus die Arbeitsziele abgeleitet werden. Anschließend wird der Aufbau dieser Arbeit erläutert und die Vorgehensweise zum Erreichen der Ziele beschrieben. Abschließend sind die Forschungsfragen dieser Arbeit in Form von Hypothesen zusammengefasst.

Stand der Technik

Mit fortschreitender Digitalisierung steigt der Bedarf an einer schnellen und robusten Erfassung dreidimensionaler Objekte. Dabei fordern Einsatzgebiete unter anderem in der Robotik, Automobilindustrie, industriellen Fertigung und der Unterhaltungselektronik eine immer höhere Detailtreue bei der Umfelderkennung. Während PMD-Kameras (Photonic mixing device, dt. Photomischdetektor) die dreidimensionale Objekterfassung in Echtzeit bereitstellen und robuste Aufnahmen auch unter schwierigen Lichtverhältnissen ermöglichen, grenzen ihre im Vergleich zu 2D-Bildsensoren großen Pixelabmessungen die Detektion kleiner Objekte stark ein. Da sie zur Bildaufnahme die Objektszene mit moduliertem Licht aktiv beleuchten, verstärkt sich diese Limitierung zusätzlich für Einsatzbereiche, wo ein hoher Anteil an Umgebungslicht vorherrscht. Dies gilt insbesondere für den Einsatz im Außenbereich bei Tageslicht.

Die in vorliegender Arbeit eingesetzte PMD-Kamera ist für den Einsatz im Außenbereich konzipiert und stellt eine Bildauflösung von 64 x 16 Pixeln bereit. Mit ihren verhältnismäßig großen Pixelabmessungen im Submillimeterbereich ist sie ausreichend empfindlich, um Objekte in Entfernungen bis 30 m auch noch bei Sonnenschein mit Beleuchtungsstärken bis $\sim 120 \text{ kLux}$ detektieren zu können. Eine physische Verkleinerung der Pixelfläche wirkt sich unmittelbar auf ihre sensorspezifischen Eigenschaften aus und würde beispielsweise die Messreichweite der PMD-Kamera und ihre Robustheit gegenüber Fremdlicht reduzieren. Folglich ist sicherzustellen, dass sich auflösungssteigernde Maßnahmen nicht auf die spezifischen Sensoreigenschaften auswirken, um die Einsatzfähigkeit der Kamera nicht zu gefährden. Überall dort, wo eine physische Reduktion der Pixelflächen nicht möglich ist, bieten sich auflösungssteigernde Techniken aus der computergestützten Bildgebung an. Solche *Super-Resolution (SR)* genannten Methoden erhöhen die Bildauflösung niedrig auflösender Bilder rechnerisch, ohne Einfluss auf die physikalischen Eigenschaften des zugrunde liegenden optischen Systems zu nehmen.

Zur Bewertung der Leistungsfähigkeit solcher SR-Techniken haben sich in der Literatur Rauschanalysen durchgesetzt, die Abweichungen vom SR-Ergebnis zu seinem fehlerfreien Ground-Truth-Bild (GT) messen. Dabei beschränkt sich die Auswertung für gewöhnlich auf synthetisch erzeugte Datensätze, ohne deren Ergebnis an einem realen Aufbau mit echten Messungen zu verifizieren. Existieren dazuhin keine hochauflösenden Referenzbilder, sind solche relativ messenden Bewertungsverfahren nicht anwendbar - so wie im Fall der hier verwendeten PMD-Kamera. Ihre Bildauflösung ist zu gering, um ein eigenes GT-Bild bereitstellen zu können, das als Referenzbild herangezogen werden könnte. Auf der anderen Seite gibt es derzeit aber auch keine PMD-Kameras mit höherer Bildauflösung, die für Anwendungen im Außenbereich konzipiert sind. Erschwerend kommt hinzu, dass SR-Techniken in der Regel Eingangsbilder mit Auflösungen $> 640 \times 480$ Bildpunkten (VGA) optimieren und damit nicht sichergestellt ist, dass ein SR-Verfahren imstande ist auch für Eingangsbilder in geringerer Auflösung (64×16 Bildpunkte) plausible SR-Ergebnisse zu rekonstruieren.

Arbeitsziele

In der vorliegenden Arbeit werden SR-Techniken zur Steigerung der lateralen Ortsauflösung einer realen PMD-Kamera untersucht, primär unter Berücksichtigung ihrer Amplitudenbilder. Dabei soll die Umsetzung einer generischen SR-Methode gewährleisten, dass die Auflösungssteigerung unabhängig vom späteren Einsatzgebiet der Kamera anwendbar ist. Dazu werden zunächst drei aktuelle SR-Methoden implementiert und deren Simulationsergebnisse auf Basis synthetisch erzeugter LR-Bilder (low resolution, dt. niedrig auflösend) verglichen:

- 1) Die KI-lernbasierte *Single Image SR (SISR)* rekonstruiert mit Hilfe eines trainierten künstlichen neuronalen Netzes das SR-Ergebnis aus einem einzelnen LR-Eingangsbild.
- 2) Der Ansatz der *inkohärenten Fourier Ptychographie (FP)* nutzt die Wechselwirkung eines intensitätsmodulierten Beleuchtungsmusters mit der Objektszene, um durch die Verwertung von Mehrfachaufnahmen die numerische Apertur der Kamera synthetisch zu erweitern.
- 3) Die *Subpixel-Shift (SpS) SR* verwertet eine Folge niedrig auflösender LR-Bilder, die im Bruchteil ihrer Pixelgröße zueinander verschoben sind. Deren komplementäre Bildinformationen sorgen für ein gesteigertes Auflösungsvermögen in der SR-Rekonstruktion.

Zur Bewertung der Bildqualität im SR-Ergebnisbild dient mit dem optischen Auflösungsvermögen in LP/px eine optische Kenngröße als absolut messendes Gütekriterium, welches nach zwei unterschiedlichen Methoden ermittelt wird: Die 1) Schräge-Kante-Methode misst den Bildkontrast an einer schrägen Kante in Abhängigkeit der Ortsfrequenz und trägt den Frequenzgang in einem Diagramm auf. Dessen MTF_{50} -Wert dient dem schnellen Vergleich des erreichten Auflösungsvermögens im SR-Ergebnis, ohne den gesamten Frequenzgang berücksichtigen zu müssen. Er gibt die Ortsfrequenz in LP/px an, bei welchem der Bildkontrast an der schrägen Kante auf 50% abgefallen ist. Nach Anwendung dieser Methode auf die Simulationsergebnisse, wird sie später ebenfalls auf die Messergebnisse angewandt und ermöglicht damit den direkten Vergleich der SR-Ergebnisse aus Simulations- und Messdaten. Das 2) USAF-1951-Testbild dagegen dient der Visualisierung potenzieller SR-Artefakte und zeigt auf einfache Weise die erreichte Grenzauflösung des optischen Systems auf. Seine komplexeren Strukturen und Geometrien erfordern

eine höhere Bildauflösung, weshalb diese Auswertung lediglich auf die Simulationsdaten angewendet werden kann. Für die finalen Simulationsergebnisse wird mit der PSNR-Auswertung zusätzlich eine etablierte Rauschanalyse zur Bewertung der erreichten Bildqualität in den SR-Rekonstruktionen herangezogen, um die Erkenntnisse aus den vorgeschlagenen Bewertungsmethoden auf Plausibilität hin zu überprüfen.

Nach Auswertung der Simulationsergebnisse zeigt sich, dass zur Umsetzung einer generischen Auflösungssteigerung von PMD-Kameras die SpS-Methode besser geeignet ist als die beiden anderen SR-Verfahren. Darum sind die experimentellen Versuche unter Verwendung realer Bildaufnahmen mit Bezug auf die SpS-SR ausgelegt, wobei zwei unterschiedliche Aufbauten umgesetzt werden. Ein konventioneller Sensor-Shift-Demonstrator verschiebt den Bildsensor im Bruchteil seiner Pixelgröße und dient als Referenzaufbau für einen neuartigen Lens-Shift-Ansatz. Bei diesem Aufbau führen transversale Linsenbewegungen im optischen System zu einem Subpixel-Versatz in der Bildebene. Schließlich werden mit dem Lens-Shift-Demonstrator Amplituden- und Entfernungsbildpaare einer dreidimensionalen Objektszene aufgenommen und damit für jeden Bildtyp separat die SR-Ergebnisse rekonstruiert, um deren erzielte Auflösungssteigerungen in Abhängigkeit des Bildtyps zu vergleichen.

Es sei darauf hingewiesen, dass in dieser Arbeit SR-Algorithmen als Handwerkszeug betrachtet werden und es kein Ziel darstellt deren Leistungsfähigkeit oder Effizienz zu steigern. Stattdessen erfolgt die Überführung von bestehenden, dem Stand der Technik entsprechenden SR-Techniken in die praktische Anwendung unter Verwendung einer realen, niedrig auflösenden PMD-Kamera.

Zusammenfassend lässt sich sagen, dass das laterale Auflösungsvermögen von PMD-Kameras, die für den Einsatz im Außenbereich konzipiert sind, durch große Pixelabmessungen stark limitiert ist. Der in der Literatur zumeist simulativ untersuchte SpS-SR-Algorithmus wird mit dieser Arbeit in die praktische Anwendung überführt, wofür ein neuartiger Lens-Shift-Aufbau entwickelt wird. Die damit erlangten SR-Ergebnisse werden nach einem absolut messenden Verfahren bewertet, das den direkten Vergleich von Theorie und Praxis erlaubt.

Die Arbeitsziele lassen sich wie folgt stichpunktartig zusammenfassen:

- Applikation einer generischen SR auf niedrig auflösende PMD-Kameras.
- Umsetzung eines absolut messenden Bewertungsverfahrens, das einen direkten Vergleich von SR-Ergebnissen aus Simulations- und Messdaten erlaubt.
- Untersuchung und Vergleich von aktuellen SR-Methoden für den spezifischen Anwendungsfall bei PMD Kameras.
- Beschreibung eines Simulationsmodells zur realitätsnahen Vorhersage von SR-Rekonstruktionen nach der SpS-Methode.
- Demonstration der Leistungsfähigkeit der SpS-Methode, primär am Lens-Shift-Aufbau.
- Verständnis für die Möglichkeiten und Grenzen der Auflösungssteigerung mit SR-Methoden in der praktischen Anwendung.
- Anwendung der SpS-SR auf Amplituden- und Entfernungsbilder einer dreidimensionalen Objektszene.

Aufbau der Arbeit

Die vorliegende Arbeit umfasst 10 Kapitel und lässt sich, abgesehen von der Einleitung in diesem Kapitel und dem Schluss in Kapitel 10, in zwei Bereiche unterteilen. In den Kapiteln 2 bis 6 werden die Grundlagen zum besseren Verständnis des Beitrags dieser Arbeit in den Kapiteln 7 bis 9 geschaffen.

Einführend in den Grundlagen-Teil erfolgt in Kapitel 2 eine Beschreibung zum Funktionsprinzip von PMD-Kameras im Allgemeinen sowie die Charakterisierung der für diese Arbeit verwendete PMD-Kamera im Speziellen. Anschließend wird in Kapitel 3 der Begriff der *Super-Resolution (SR)* in den Kontext der geometrischen Auflösung eingeordnet und das allgemeine SR-Beobachtungsmodell beschrieben. Darüber hinaus erfolgt eine Auflistung der mathematischen Problemstellungen, die es mit den SR-Techniken zu lösen gilt. In den darauffolgenden Kapiteln 4 und 5 werden die drei untersuchten SR-Methoden der SISR, inkohärenten FP und SpS erläutert und in Abhängigkeit der benötigten Anzahl an Eingangsbildern der Single Image SR oder der Multi-Frame SR zugeordnet. Das letzte Grundlagenkapitel 6 schließt mit einer Vorstellung der in dieser Arbeit angewandten Bewertungsverfahren ab. Dazu zählen in erster Linie die Ortsfrequenzganganalyse und die USAF-1951-Auflösungstesttafel, aber auch die aus der Literatur bekannte Rauschanalyse nach dem Spitzen-Signal-Rausch-Verhältnis.

Mit Kapitel 7 wird in den Beitragsteil dieser Arbeit übergeleitet, der zunächst die Datengrundlage in Form der Schräge-Kante- und USAF-1951-Testbilder erläutert. Anschließend erfolgen Simulationen für die drei SR-Techniken auf Basis von synthetisch erzeugten LR-Amplitudenbildern. Deren SR-Rekonstruktionen werden mit Hilfe der Ortsfrequenzganganalyse und der USAF-1951-Auflösungstesttafel bewertet und auf bestmögliche Eignung zur Auflösungssteigerung von PMD-Kameras hin überprüft. Abschließend erfolgt in Kapitel 7 eine Zusammenfassung der wesentlichen Erkenntnisse zu den drei untersuchten SR-Methoden, woraus die SpS-SR als die Methode der Wahl abgeleitet wird. Dessen Umsetzung folgt der Beschreibung in Kapitel 8 und mündet in zwei Demonstratoren; dem Sensor-Shift- und dem Lens-Shift-Aufbau. Darauf folgt in Kapitel 9 ein Vergleich der SR-Rekonstruktionen aus den Messaufnahmen beider Aufbauten mit denen aus den synthetisch erzeugten Simulationsdaten unter Verwendung der Ortsfrequenzganganalyse. Die mitunter großen Diskrepanzen können auf ein unzureichendes Simulationsmodell zurückgeführt werden, dessen Anpassung Abhilfe schafft, wie abschließende Simulationen aufzeigen. Schließlich erfolgt für die finalen SR-Ergebnisse nach dem angepassten Simulationsmodell eine Rauschanalyse, um die Aussagekraft der beiden Bewertungsverfahren nach der Ortsfrequenzganganalyse und der USAF-1951-Auflösungstesttafel zu bestätigen. Die letzten Untersuchungen bezüglich der vorgeschlagenen SpS-SR erfolgen an realen Amplituden- und Entfernungsbildaufnahmen der in dieser Arbeit eingesetzten, niedrig auflösenden PMD-Kamera. Dazu wird mit einem Gradientenvergleich die laterale Auflösungssteigerung in den SR-Ergebnissen beider Bildtypen bewertet. Den Schluss des Beitragsteils bildet eine Diskussion zu den wesentlichen Forschungsfragen, wie sie im nächsten Abschnitt in Form von Hypothesen definiert sind.

Mit einer Zusammenfassung und einem Ausblick in weiterführende Forschungsthemen schließt Kapitel 10 diese Arbeit ab.

Hypothesen

Nachfolgend sind die Forschungsfragen der vorliegenden Arbeit als Hypothesen ausformuliert und in Kurzform erläutert:

Hypothese 1: SR-Techniken sind in der Lage SR-Ergebnisse gesteigerter Bildqualität zu rekonstruieren, unabhängig von der bereitgestellten Bildauflösung der LR-Eingangsbilder.

In der Regel steigern SR-Methoden die Bildqualität von LR-Eingangsbildern, deren Bildauflösung größer ist als der VGA-Standard mit seinen 640 x 480 Bildpunkten. Es wird davon ausgegangen, dass sich SR-Techniken auch auf LR-Bilder einer deutlich kleineren Ausgangsgröße von 16 x 16 Bildpunkten adaptieren lassen.

Hypothese 2: Zur Bewertung der erreichten Bildqualität im SR-Ergebnis liefert die Ortsfrequenzganganalyse qualitativ dieselben Erkenntnisse wie die etablierte Rauschanalyse.

In der Literatur erfolgt die Bewertung der erreichten Bildqualität im SR-Ergebnis zumeist nach einer Rauschanalyse mit Bezug zu deren GT-Referenzbild. Die Ortsfrequenzganganalyse dagegen misst das erreichte Auflösungsvermögen im SR-Ergebnis absolut mit Hilfe eines Schräge-Kante-Testbildes. Es wird erwartet, dass beide Bewertungsverfahren qualitativ dieselben Erkenntnisse über die Leistungsfähigkeit einer SR-Technik liefern.

Hypothese 3: Für den SR-Rekonstruktionsprozess nach der Subpixel-Shift-Methode ist der Entstehungsort der Subpixel-Bewegung nicht relevant.

Für gewöhnlich geht die Literatur davon aus, dass Subpixel-Bewegungen über den Versatz des Bildsensors direkt umgesetzt werden. Sorgt dagegen die Verschiebung einer optischen Linse für den Subpixel-Versatz in der Bildebene, ist davon auszugehen, dass sich eine dezentrierte Linsenposition negativ auf die Abbildungsleistung des optischen Systems auswirkt und unter Umständen die maximal erreichbare Auflösungssteigerung limitiert. Wird dies jedoch bereits bei dessen Auslegung berücksichtigt, sind dieselben SR-Resultate wie bei der Umsetzung nach dem direkten Sensor-Shift-Verfahren zu erwarten.

Hypothese 4: Bei der Subpixel-Shift SR begrenzt das Auflösungsvermögen der abbildenden Optik die maximal erreichbare Auflösungssteigerung im SR-Ergebnis.

Das Auflösungsvermögen im rekonstruierten SR-Ergebnis kann beim Subpixel-Shift-Verfahren das der abbildenden Optik nicht übersteigen. Im Umkehrschluss wird eine theoretisch erreichbare Auflösungssteigerung erwartet, die betragsmäßig dem Verhältnis der Auflösungsgrenzen von abbildender Optik zu Bildsensor entspricht.

Hypothese 5: Für die Subpixel-Shift SR ist die Steigerung des lateralen Auflösungsvermögens in beiden Bildtypen, dem Amplituden- wie auch dem Entfernungsbild, betragsmäßig gleich.

Beide Bildtypen der PMD-Kamera beschreiben ihre Messwerte als Grauwertbilder, allerdings auf Basis unterschiedlicher Datengrundlagen. Beim Amplitudenbild entsprechen die Grauwerte der Helligkeitsverteilung in der Objektszene. Das Entfernungsbild dagegen stellt die Entfernung zur Objektoberfläche in Graustufen dar. Es wird davon ausgegangen, dass sich die Steigerungsfaktoren des Auflösungsvermögens - unabhängig vom Bildtyp - in Amplituden- und zugehörigem Entfernungsbild identisch verhalten.

Kapitel 2

3D-PMD-Kameras

Im zweiten Kapitel wird in Abschnitt 2.1 zunächst das Funktionsprinzip von PMD-Kameras erläutert, bevor in Abschnitt 2.2 die in dieser Arbeit verwendete PMD-Kamera charakterisiert wird. Hinsichtlich der Funktionsweise wird auf die 4-Phasen-Messung eingegangen, anhand derer die Amplituden- und Entfernungsbilder erzeugt werden. Außerdem sind mehrere Auflösungstypen aufgeführt, mit denen PMD-Kameras üblicherweise charakterisiert werden. Dabei liegt das Hauptaugenmerk auf der lateralen Ortsauflösung und es werden diejenigen Einflussfaktoren beschrieben, die zu dessen Limitierung beitragen.

2.1 Funktionsweise von PMD-Kameras

PMD-Kameras sind monokular-optische Phasenmesssysteme, die gleichzeitig Amplituden- und ihre zugehörigen Entfernungsbilder in Echtzeit ausgeben.

Als ToF-System (Laufzeit, engl. Time-of-Flight; ToF) gehören PMD-Kameras (Photomischdetektor, engl. photonic mixing device; PMD) zur Familie der Laufzeit messenden Verfahren. Sie rekonstruieren ihr Umfeld im dreidimensionalen Raum und geben zu jedem Ortspunkt einen Helligkeitswert aus. Dabei liefern sie auch unter schwierigen Lichtverhältnissen robuste Aufnahmen in Form von Amplituden- und Entfernungsbildern. Im Gegensatz zu LIDAR-Systemen rastern sie die Objektszene nicht punktuell ab, sondern erfassen diese als Ganzes. Dabei misst jedes PMD-Pixel zusätzlich zum Helligkeitswert seinen Objektabstand. Das Grundprinzip geht auf Schwartes Patent zum PMD-Sensor aus dem Jahre 1996 [1] zurück.

Eine PMD-Kamera besteht aus Sende- und Empfangseinheit mit festem Positionsbezug zueinander, wobei im einfachsten Fall beide Einheiten in derselben Ebene liegen. Der aktive Sender emittiert Strahlung im nahen Infrarot (engl. Near-infrared; NIR) mit periodisch modulierter Amplitude. Als Strahlungsquelle dienen in der Regel LEDs oder Laserdioden, deren Ein- und Ausschaltverhalten näherungsweise einen sinusförmigen Signalverlauf erzeugen. In Abhängigkeit vom Reflexionsgrad der Oberflächen, wirft die Objektszene einen Teil der Strahlung zurück in Richtung Kamera, wo es eine abbildende Optik auf den Sensor fokussiert. Aus dem eintreffenden Licht generiert jedes Pixel für sich die Ladungsmenge $Q(n, m)$. Die Signallaufzeit von der Kamera zum Messobjekt - und wieder zurück - ist proportional zur Objektentfernung. Nach folgender Gleichung ergeben sich die Entfernungswerte $d(n, m)$ für jedes Pixel indirekt aus der Phasendifferenz φ_{mod} vom ausgesandten zum reflektierten Signal [2–5]:

$$d(n, m) = \frac{\varphi_{mod} \cdot c}{2 \cdot 2\pi \cdot f_{mod}} \quad (2.1)$$

Die Periodizität des Nutzsignals geht mit einem Eindeutigkeitsbereich von 2π einher. Damit bestimmt die Modulationsfrequenz f_{mod} die maximale Mess-Reichweite der Kamera maßgeblich, weil Phasenverschiebungen nur innerhalb dieser einen Schwingungsperiode eindeutig messbar sind. Die Lichtgeschwindigkeit c ist eine Naturkonstante und beträgt $2,998 \cdot 10^8$ m/s, der Faktor 2 im Nenner ergibt sich aus dem doppelten Laufweg zwischen Kamera und Objekt. Das periodisch modulierte Signal ermöglicht die Unterdrückung konstanter Umgebungsstrahlung [6, 7], wie zum Beispiel Tageslicht, und macht die PMD-Kamera robust für Aufnahmen unter schwierigen Lichtverhältnissen. Somit sind sie über die 3D-Umfelderfassung hinaus auch im Kontext der 2D-Bilderfassung interessant, wo aufgrund der Beleuchtungssituation konventionelle Sensoren an ihre Grenzen stoßen, sei es durch Sättigungseffekte bei zu viel Licht oder geringe Bildkontraste aufgrund zu wenig Licht.

Bei sinusförmiger Modulation eignet sich die 4-Phasen-Messung [2] zur robusten Rekonstruktion des reflektierten Signals, wie in Abbildung 2.1 skizziert. Demnach gibt jedes PMD-Pixel (n, m) gleichzeitig zu je vier Zeitpunkten in äquidistantem Abstand Δt seine bis dahin erzeugte Ladungsmenge $Q_1(n, m)$ bis $Q_4(n, m)$ aus. Alle Einzelpixel zusammengenommen ergeben für jeden Zeitpunkt ein Phasenbild. Die Kenntnisse über den sinusförmigen Signalverlauf sowie die Ladungsverteilungen innerhalb jedes Pixels erlauben die Rekonstruktion des Modulationssignals aus dem reflektierten Signal. Optische Verluste mindern die Amplitude des Nutzsignals (blaue Linie) gegenüber dem emittierten Signal (rote Linie).

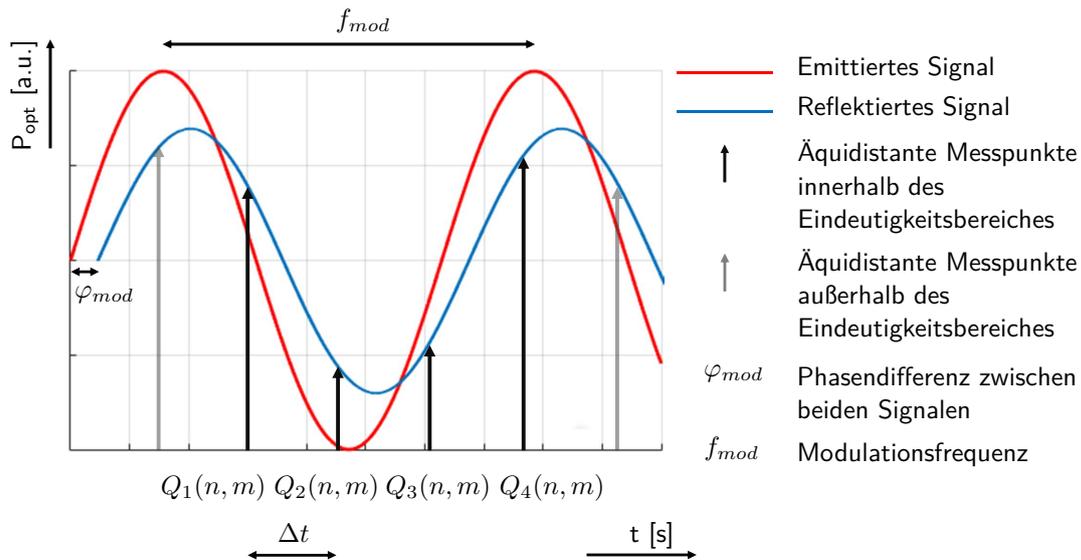


Abbildung 2.1: Die 4-Phasen-Messung rekonstruiert das reflektierte, sinusförmige Nutzsignal (blaue Kurve) aus vier äquidistanten Einzelmessungen $Q_1(n, m)$ bis $Q_4(n, m)$ innerhalb einer Schwingungsperiode f_{mod} .

Jedes Pixel ermittelt die Objektentfernung aus der Phasendifferenz φ_{mod} beider Signale zueinander. Nach [2] ergibt sich dann die Phasendifferenz aus den vier Phasenmessungen zu:

$$\varphi_{mod}(n, m) = \arctan \left(\frac{Q_3(n, m) - Q_1(n, m)}{Q_2(n, m) - Q_4(n, m)} \right) \quad (2.2)$$

Genauer betrachtet besitzt ein einzelnes PMD-Pixel zwei voneinander getrennte lichtempfindliche Photoflächen. Beide besitzen einen eigenen Auslesekanal, die im Gegenteil zueinander, jedoch mit derselben Frequenz wie das Modulationssignal, schwingen. Dieses Prinzip der Ladungsschaukel verteilt die detektierten Ladungsträger auf die beiden Auslesekanäle. Jedes Pixel rekonstruiert das reflektierte Nutzsignal über diese korrelierte Ladungsträgerseparation, wie Heinol in seiner Dissertation [2] ausführlich erläutert.

Neben den Entfernungsbildern geben PMD-Kameras die zugehörigen Amplitudenbilder in Grauwertstufen aus. Sie beschreiben damit die Helligkeitsverteilung in der Objektszene - also die Lichtmenge, die von den Objektoberflächen reflektiert wird. Für das Amplitudenbild mittelt jedes Pixel die Ladungsmenge $Q_0(n, m)$ aus allen vier Phasenmessungen wie folgt [2]:

$$Q_0(n, m) = \frac{Q_1(n, m) + Q_2(n, m) + Q_3(n, m) + Q_4(n, m)}{4} \quad (2.3)$$

Bei der Spezifikation von PMD-Kameras kann die allgemeine Bezeichnung der *Auflösung* mehrere Bedeutungen einnehmen. Zum besseren Verständnis spezifiziert nachfolgende Übersicht die für die PMD-Kamera relevanten Auflösungstypen.

- Die *axiale Ortsauflösung* oder Tiefenauflösung beschreibt den kleinstmöglich darstellbaren Entfernungswert. Neben der Modulationsfrequenz des Beleuchtungssignals beeinflusst die Oberflächenreflektanz der Objekte die axiale Ortsauflösung. Typischerweise erreichen PMD-Kameras eine Tiefenauflösung im Bereich weniger Zentimeter [2–9] bei Messentfernungen bis rund 10 m.
- Die *Bildauflösung* oder - meist im englischsprachigen Raum - Pixelauflösung kennzeichnet die Größe eines digitalen Bildes über dessen Anzahl an Bildpunkten. In der Regel erfolgt die Angabe über die Spaltenanzahl multipliziert mit der Zeilenanzahl in der Form 640 x 480.
- Die *laterale Ortsauflösung*, auch als räumliche oder geometrische Auflösung bezeichnet, ergibt sich aus Pixelgröße und -abstand (engl. Pixel-Pitch). Sie begrenzt die kleinstmögliche Strukturgröße, die der Bildsensor darstellen kann. Neben dem Abstand in Längeneinheiten ist ebenso die Angabe dessen Kehrwert als Ortsfrequenz gebräuchlich.
- Das *optische Auflösungsvermögen* beschreibt die Fähigkeit eines optischen Systems zwei dicht beieinander liegende Objektpunkte - sei es als räumlicher Abstand oder unter einer Winkelbetrachtung - getrennt voneinander abzubilden [10]. Die Bezeichnung der Auflösungsgrenze erfolgt wahlweise als Längenmaß, Winkelangabe oder Ortsfrequenz.
- Die *radiometrische Auflösung* ist die Quantisierung von Messwerten in 2^n Grauwertstufen. Amplitudenbilder quantisieren kontinuierliche Helligkeitsverteilungen, Entfernungsbilder dagegen Distanzen.
- Die *spektrale Auflösung* liegt bei PMD-Kameras für gewöhnlich bei einem Spektralkanal im NIR, weshalb lediglich Grauwertbilder ausgegeben werden können. Die spektrale Bandbreite hängt von der Kombination aus IR-Lichtquelle und IR-Bandpassfilter ab. Aufgrund der vorherrschenden stimulierten Emission weisen Halbleiterlaser geringere Spektralbreiten auf als LEDs, bei denen die spontane Emission zu einem flachen und breiten Spektralverlauf um die mittlere Wellenlänge herum beiträgt [11]. Bei einer mittleren Wellenlänge von 850 nm liegen die Halbwertsbreiten (engl. full width at half maximum; FWHM)

für Halbleiterlaser bei 1 nm bis 2 nm und für LEDs bei 50 nm bis 60 nm [11]. Ein IR-Bandpassfilter vor dem Sensor filtert das Nutzsignal aus dem Umgebungslicht. Je besser das Transmissionsverhalten des Bandpassfilters zur spektralen Bandbreite der Lichtquelle passt, desto effizienter ist die Detektion des Nutzsignals.

- Die *temporale Auflösung* beschreibt die Bildwiederholrate oder Bildwiederholfrequenz (engl. frames per second; fps) in Bildern pro Sekunde (Hz). Abhängig von Bewegungsstärke und Bildgröße nimmt das menschliche Auge bereits ab 12 fps bis 15 fps aufeinanderfolgende Einzelbilder als flüssige Bewegungen wahr [12]. Klassische Filmkameras nehmen Bewegungsbilder mit rund 24 fps auf [12]. Bezogen auf die 4-Phasen-Messung einer PMD-Kamera ist die Bildwiederholrate der Phasenbilder um den Faktor 4 höher als die der ausgegebenen Amplituden- und Entfernungsbilder.

Das Primärziel dieser Arbeit liegt in der Erhöhung der *lateralen Ortsauflösung* im Amplitudenbild. Dabei soll die Detailtreue im Ausgabebild erhöht werden, ohne die sensorspezifischen Eigenschaften wie Pixel- oder Sensorgröße zu verändern.

Ganz allgemein begrenzt die Bildauflösung von PMD-Kameras nicht nur deren laterales Auflösungsvermögen, sondern ebenso ihre maximale Messreichweite. Im Groben lassen sich die Kameras nach ihrem späteren Einsatzzweck in Anwendungen für den Innenraum und Außenbereich eingruppiert. Für Innenraum-Anwendungen bildet Microsofts Kinect v4 *Azure Kinect DK* [13] den Stand der Technik. Bei einer Bildwiederholrate von 30 Hz und einem Entfernungsbereich von 0,5 m bis 3,86 m übersteigt deren Bildauflösung mit 640 x 576 Bildpunkten den VGA-Standard [14] von 1987 (640 x 480 Pixel) nur knapp.

Der Einsatz im Außenbereich reduziert das maximal erreichbare laterale Auflösungsvermögen nochmals drastisch. Zum einen sorgt das Sonnenlicht für höhere Bestrahlungsstärken auf Objektoberflächen und einen höheren IR-Anteil im Umgebungslicht als in geschlossenen Räumen, wodurch die PMD-Pixel schneller in Sättigung geraten und infolgedessen das Nutzsignal von der Umgebungsstrahlung nicht mehr eindeutig unterscheiden können. Messfehler bis hin zu Pixelausfällen sind die Folge [15]. Zum anderen sind Objektentfernungen im Innenraum für gewöhnlich deutlich geringer als im Außenbereich. Nach dem photometrischen Entfernungsgesetz $E_e = I_e/r^2 \cdot \cos(\epsilon)$ [16] ist die Strahlungsstärke I_e umgekehrt proportional zum Quadrat der Entfernung r . Das bedeutet, mit doppelter Objektentfernung viertelt sich die Bestrahlungsstärke und weniger Licht trifft auf den Sensor. Der Einstrahlwinkel ϵ berücksichtigt die Strahlungsrichtung relativ zur Flächennormale. Beiden Störeinflüssen wirkt eine physische Vergrößerung der Pixelfläche entgegen, indem sie deren Empfindlichkeit erhöht. Dies gewährleistet auch für Anwendungen im Außenbereich eine robuste Rekonstruktion des Nutzsignals.

Messreichweite und Tiefenaufklärung korrelieren mit der Modulationsfrequenz in der Form, dass die Tiefenaufklärung mit steigender Modulationsfrequenz zunimmt und die Messreichweite dagegen abnimmt. Eine Vergrößerung der Messreichweite bei Beibehaltung der absoluten Genauigkeit verspricht der Heterodyn-Betrieb [3, 4]. Dieser erweitert den Eindeutigkeitsbereich von 2π , indem mehrere unterschiedliche Modulationsfrequenzen genutzt werden [5].

Abgesehen von den kameraspezifischen Eigenschaften, wirken sich zudem Oberflächenbeschaffenheiten auf die Messung mit einer PMD-Kamera aus [17]. Je geringer der Reflexionskoeffizient einer Oberfläche, umso mehr Licht absorbiert diese. Folglich reduziert sich die Lichtmenge im Nutzsignal und vermindert die Messgenauigkeit der Kamera.

2.2 Charakterisierung der verwendeten PMD-Kamera

Die für diese Arbeit zugrunde liegende PMD-Kamera ist der 3D-Sensor *O3M151* [18] der Firma *ifm electronic gmbh* nach Abbildung 2.2. Dessen Einsatzgebiet umfasst die 3D-Bilderzeugung im Außenbereich, wie beispielsweise die Objekterkennung, Distanzüberwachung und Kollisionswarnung [19]. Die wesentlichen Kameraeigenschaften fasst Tabelle 2.1 zusammen. Aus den, verglichen mit den typischen Pixelgrößen aus der 2D-Bilderfassung, großen Bildpunkten ihres PMD-Sensors resultiert eine geringe Bildauflösung von lediglich 64 x 16 Bildpunkten. Andererseits ermöglichen sie große Messreichweiten und robuste Messungen unter schwierigen Lichtverhältnissen, wie sie bei hoher Sonneneinstrahlung im Sommer vorliegen können.



Abbildung 2.2: Der 3D-Sensor *O3M151* [18] der Firma *ifm electronic gmbh*, bestehend aus dem Empfänger (links) und Emitter (rechts), arbeitet nach dem PMD-Prinzip. Die Fotoaufnahme entstammt der Werbeschrift in [20].

Tabelle 2.1: Eigenschaften der PMD-Kamera *O3M151*.

Axiale Ortsauflösung	5 cm bis 10 cm im Nahbereich (< 5 m) [21]
Bildauflösung	64 x 16 Bildpunkte [18]
Radiometrische Auflösung	max. 16 bit (unsigned integer) [21]
Bildwiederholraten	25 Hz, 33 Hz und 50 Hz [18]
Öffnungswinkel der Optik	70° x 23° (hor. x vert.) [18]
Wellenlänge	850 nm [18]
Messreichweite	0,25 m bis 30 m [19]
Einsatzbedingung	Außenbereich, sonnig (bis ~120 kLux) [19]

Kapitel 3

Rechnerische Erhöhung der Bildauflösung

In folgendem Kapitel werden die Grundlagen zum Verständnis auflösungssteigernder Verfahren gebildet. Es startet mit der Beschreibung einer klassischen Interpolationsmethode, hier vertreten durch die PCHIP-Interpolation (stückweise kubisch interpolierendes Hermite-Polynom, engl. piecewise cubic Hermite interpolating polynomial; PCHIP) in Abschnitt 3.1, als eine der direktesten Arten zur Erhöhung der Bildauflösung. Prinzipbedingt sind Interpolationsmethoden nicht in der Lage hochfrequente Ortsinformationen zu rekonstruieren und zählen deshalb streng genommen nicht zur zweiten Rubrik [22, 23]: der Super-Resolution (SR). Als deren Vertreter wird die geometrische SR mit ihren verschiedenartigen SR-Techniken in Abschnitt 3.2 vorgestellt. Mit Bezug auf das allgemeingültige lineare Beobachtungsmodell beschreibt Abschnitt 3.3 den Prozess der SR-Rekonstruktion und Abschnitt 3.4 die mathematischen Problemstellungen, die von SR-Verfahren zu lösen sind.

Wie vorhergehend in Abschnitt 2.1 beschrieben, tragen eine Zunahme der geforderten Messreichweite sowie der Anteil an vorherrschendem Störlicht in der Messszene zur Vergrößerung der Pixelabmessung des PMD-Sensors bei. Dies begrenzt deren Bildauflösung insbesondere bei Einsatzgebieten im Außenbereich und damit deren Fähigkeit feine Objektstrukturen aufzulösen. Dort wo eine physische Reduktion der Pixelfläche nicht umsetzbar ist, ermöglichen rechnergestützte Methoden die Erhöhung der lateralen Ortsauflösung. Dabei variieren der Programmieraufwand und die resultierende Bildqualität in Abhängigkeit der gewählten Methode erheblich.

3.1 PCHIP-Interpolation

Interpolationsverfahren fügen einem Einzelbild zusätzliche Stützstellen zwischen bestehende Bildpunkte ein und berechnen deren Helligkeitswerte unter Berücksichtigung seiner Nachbarpixel.

Im Allgemeinen liefern Interpolationsmethoden gute Ergebnisse für gleichförmige Bildinhalte. An Diskontinuitäten führen sie jedoch zu unerwünschten Artefakten, wie blockige Pixel oder Unschärfe [24–27]. Dennoch erfreuen sie sich wegen ihrer Recheneffizienz und einer einfachen Umsetzung großer Beliebtheit, wenn es darum geht die Anzahl an Bildpunkten in einem Einzelbild zu erhöhen.

In der Bildverarbeitung zählen die Nearest Neighbor (NN) Interpolation sowie die bilineare und bikubische Interpolation zu den gebräuchlichsten Verfahren. Dennoch wird in vorliegender Arbeit die PCHIP-Interpolation verwendet. In Abbildung 3.1 sind die Interpolationsergebnisse der drei Methoden der PCHIP-Interpolation gegenübergestellt. Zum Vergleich dient eine

eindimensionale Messreihe, dessen Wertebereich von benachbarten Datenpunkten sich sprunghaft ändert. Das heißt, ein angenommener Helligkeitswert von 210 fällt auf 34 ab. Die NN-Interpolation, dargestellt als lilafarben gestrichelte Linie, kopiert für einen neuen Datenpunkt seinen nächstgelegenen Helligkeitswert. Dies führt zwar zu steilen Kontrastübergängen, übertragen auf ein zweidimensionales (2D-) Bild allerdings auch zu ungewollten blockigen und treppentufenartigen Artefakten. Bei der linearen Interpolation (blau gestrichelte Linie) erfolgt die Berechnung neuer Helligkeitswerte nach einem linearen Zusammenhang zwischen benachbarten Datenpunkten. Der Kontrastübergang verläuft flacher als bei der NN-Interpolation, was dem blockartigen Erscheinungsbild im zweidimensionalen Interpolationsergebnis entgegenwirkt. Allerdings verursacht dies Unschärfe. Je flacher der Gradient beim Kontrastübergang ist, desto unschärfer wirkt das Ergebnisbild. Diesem Unschärfeverhalten wirkt die kubische Interpolation mit seinem steileren Kontrastübergang entgegen, dargestellt als grün gestrichelte Linie. Diese Interpolationsmethode berücksichtigt für seine Helligkeitsberechnungen Werte über die direkt angrenzenden Datenpunkte hinaus. Die hier gezeigte kubische Interpolation setzt einen separablen Catmull-Rom-Filter [28] ein, der zu einem Überschwingen an Diskontinuitäten führt. Im Falle eines 2D-Bildes entspricht dies einer Kontrastüberhöhung, was den Eindruck einer höheren Bildschärfe erweckt [29]. Allerdings mit dem Nachteil, dass sich dadurch an kontrastreichen Kanten visuell wahrnehmbare Doppelkanten ergeben können. Dies verhindert die Interpolation mit einem stückweise kubisch interpolierenden Hermite-Polynom, kurz PCHIP-Interpolation. Sie interpoliert lediglich in den Teilbereichen zwischen benachbarten Datenpunkten nach einem kubischen Zusammenhang und wirkt damit formbeständig auf den Signalverlauf [30]. In diesem Vergleich erzeugt somit die PCHIP-Interpolation den steilsten Gradientenverlauf an Diskontinuitäten, ohne dabei den Wertebereich durch eine Kontrastüberhöhung zu manipulieren.

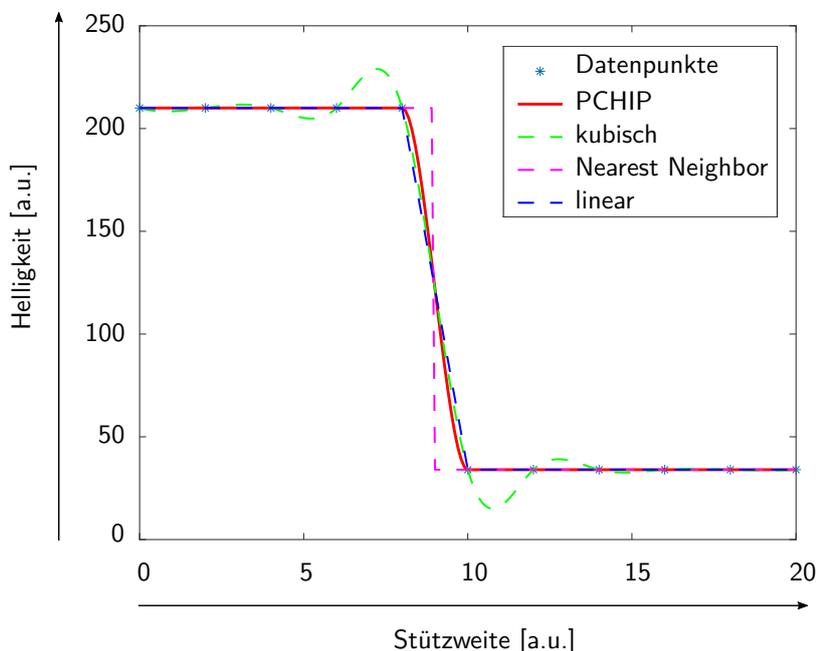


Abbildung 3.1: Unter den Interpolationsergebnissen mit kontinuierlichem Signalverlauf liefert das PCHIP-Verfahren den steilsten Gradienten für sprunghafte Änderungen im Wertebereich, hier in Form ganzzahliger Helligkeitswerte.

3.2 Geometrische Super-Resolution

Die geometrische Super-Resolution bezeichnet eine softwarebasierte Erhöhung der Bildauflösung digitaler Bildsensoren durch die Rekonstruktion hochfrequenter Bildinhalte mit Hilfe zusätzlicher Objektinformationen.

Wie im vorherigen Abschnitt dargelegt, sind Interpolationsverfahren hinsichtlich ihrer erreichbaren Bildqualität begrenzt. Insbesondere an Diskontinuitäten, wo sich die Gradientensteigung beim Kontrastübergang auf die Bildschärfe auswirkt, sind fehlende Objektinformationen visuell wahrnehmbar. Entweder durch blockartige Artefakte, wie bei der NN-Interpolation, oder durch eine verminderte Bildschärfe, wie bei den übrigen vorgestellten Interpolationsmethoden. Diesem Informationsmangel nehmen sich die Super-Resolution (SR) Methoden an. Die Bezeichnung der Super-Resolution (SR) ist nicht eindeutig definiert und erstreckt sich über weite Bereiche wissenschaftlicher Disziplinen, wie u.a. der Optik, Informatik, Medizin und den Geowissenschaften [31]. In vorliegender Arbeit steht *Super-Resolution (SR)* für die geometrische SR zur Steigerung der *Bildauflösung*. Dabei vereinen Algorithmen den physisch begrenzten Informationsgehalt niedrig auflösender, digitaler Bildsensoren mit zusätzlich gewonnenen Informationen weiterer Quellen [32]. Letztlich erhöhen sie die Pixeldichte im SR-Ergebnisbild rechnerisch, um hochfrequente Informationen zu rekonstruieren, die in den zugrundeliegenden Einzelbildern per se nicht enthalten sind [31]. Durch die reine Anwendung auf Softwareebene bleiben die sensorspezifischen Eigenschaften wie z.B. Sensor- und Pixelgröße, Empfindlichkeit, Dynamikbereich etc. vom SR-Prozess unberührt.

Entsprechend der Übersicht in Abbildung 3.2, lässt sich die geometrische SR in fünf Ebenen klassifizieren, wobei die erste Ebene alle geometrischen SR-Methoden umfasst. Die zweite Ebene beinhaltet die mathematische Betrachtung, die in den Ortsfrequenzraum (auch Fourierraum) und den Ortsraum unterteilt. Historisch betrachtet entwickelten sich die ersten SR-Algorithmen im Fourierraum und erreichten durch die Publikation von Tsai und Huang [33] bereits im Jahr 1984 erstmals ein größeres Publikum. Letztlich haben sich die SR-Berechnungen im Ortsraum durchgesetzt, weil diese eine präzisere Formulierung des mathematischen Modells erlauben und somit reale Anwendungen genauer beschreiben können [25, 34]. Dennoch gelten die Klassifikationen in den nachfolgenden Ebenen für beide Domänen gleichermaßen. So unterscheidet die dritte Ebene in Mehrfach- und Einzelbildaufnahmen, beziehungsweise auf die Anzahl der niedrig auflösenden Eingangsbilder. Verallgemeinert liefern Mehrfachaufnahmen (engl. multi-frame SR; MFSR) mit ihren zusätzlichen Bildern einen höheren Beitrag komplementärer Ortsinformationen. Dagegen sind Einzelbild-Verfahren (single image SR; SISR) typischerweise darauf ausgelegt, Kantenartefakte zu reduzieren und feine Objektstrukturen über zusätzliche externe Informationen zu schätzen [23]. Die vierte Ebene fasst einen Auszug der am häufigsten beschriebenen SR-Techniken in Gruppen zusammen. MFSR-Algorithmen ziehen komplementäre Bildinformationen stets aus mehreren, zueinander geringfügig abweichenden Einzelaufnahmen, um im SR-Ergebnis ein gesteigertes Auflösungsvermögen zu erzielen. Subpixel-Shift-Verfahren [22, 23, 31, 34–44] vereinen eine Bilderfolge aus Einzelbildern eines einzelnen Bildsensors, die im Bruchteil dessen Pixelgröße zueinander verschoben sind. Fusionsansätze [45–48] dagegen kombinieren Bildinformationen aus zwei oder mehr voneinander getrennten Sensoren. Bei der strukturierten Beleuchtung [49–55] variiert die Beleuchtungssituation jedes einzelnen Bildes und verschiebt damit dessen Objektinformationen im Ortsfrequenzraum. Bei der SISR erlangten rekonstruktionsbasierte [56–59] und lernbasierte [24, 60–67] Algorithmen zuletzt die größte

Aufmerksamkeit. Wobei sich mit Einführung der künstlichen Intelligenz die lernbasierten SISR-Methoden endgültig durchgesetzt haben [23]. Diese eignen sich in einem Lernprozess typische Verhaltensmuster bei der Auflösungssteigerung an und greifen bei der SR-Rekonstruktion auf diese Informationen zurück. Die fünfte Ebene ist exemplarisch für die lernbasierte SISR aufgeführt, gilt aber prinzipiell für jede SR-Technik. Sie unterteilt die übergeordnete SR-Methode nach deren Anwendungsfall. Die generischen Methoden haben zum Ziel möglichst unabhängig vom späteren Einsatzzweck die Auflösung zu steigern, während die spezifischen Methoden bei einem bestimmten Gebiet, wie beispielsweise der Gesichtserkennung, Texterkennung, etc. Anwendung finden.

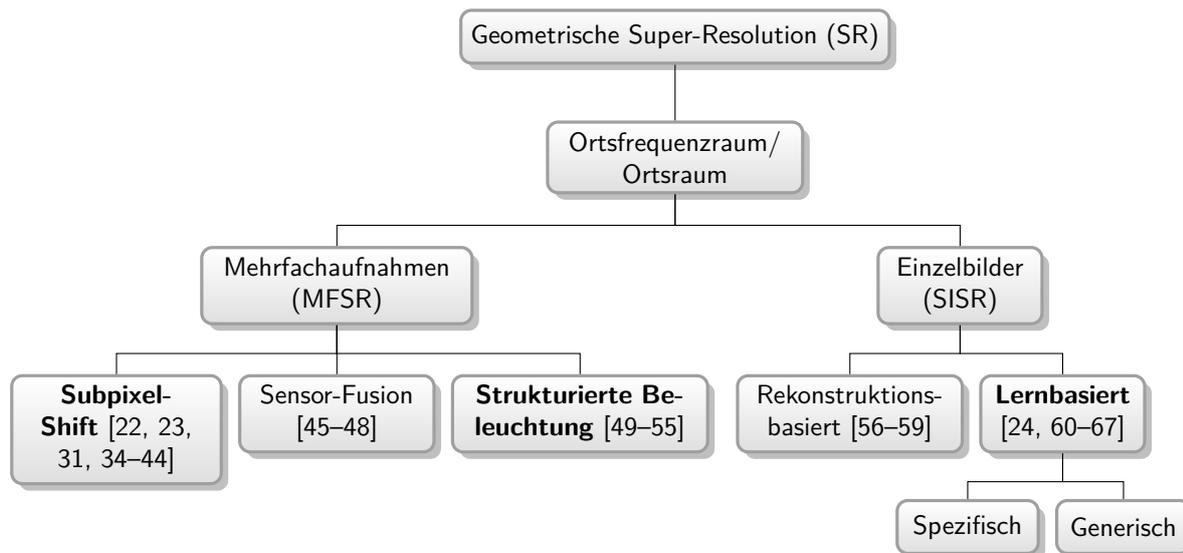


Abbildung 3.2: Die Übersicht untergliedert sich in eine Auswahl der gängigsten Techniken (in Ebene 4) zur geometrischen SR (Ebene 1) nach ihrem Berechnungsraum (Ebene 2) sowie der benötigten Anzahl an niedrigauflösenden Eingangsbildern (Ebene 3). Ebene 5 deutet den Einsatzzweck nach einer generischen oder spezifischen SR am Beispiel der lernbasierten SISR an. In Fettdruck sind die in dieser Arbeit untersuchten Verfahren hervorgehoben.

3.3 Super-Resolution Beobachtungsmodell

Die Beschreibung der SR-Rekonstruktion erfolgt am Vorwärtsmodell [31, 41, 68] nach Gleichung 3.1 und Abbildung 3.3. Dieses allgemeingültige lineare Beobachtungsmodell [22, 23, 25, 31, 36, 44, 68] beschreibt, wie aus einem hochauflösenden Ursprungsbild eine Bilderfolge mit mehreren degradierten, niedrig auflösenden Einzelbildern entsteht. Dabei steht der Index k für die Anzahl der niedrig auflösenden Bilder.

$$LR_k = D \cdot B_k \cdot M_k \cdot GT + n_k \quad \text{für } 1 \leq k \leq K \quad (3.1)$$

Ein als ideal angenommenes Ausgangsbild GT (engl. ground truth; GT) beschreibt die so genannte Grundwahrheit einer kontinuierlichen Objektszene diskretisiert, jedoch in beliebig hoher Bildauflösung. Vier Operationen wirken sich auf dessen Bildqualität aus und führen sukzessive

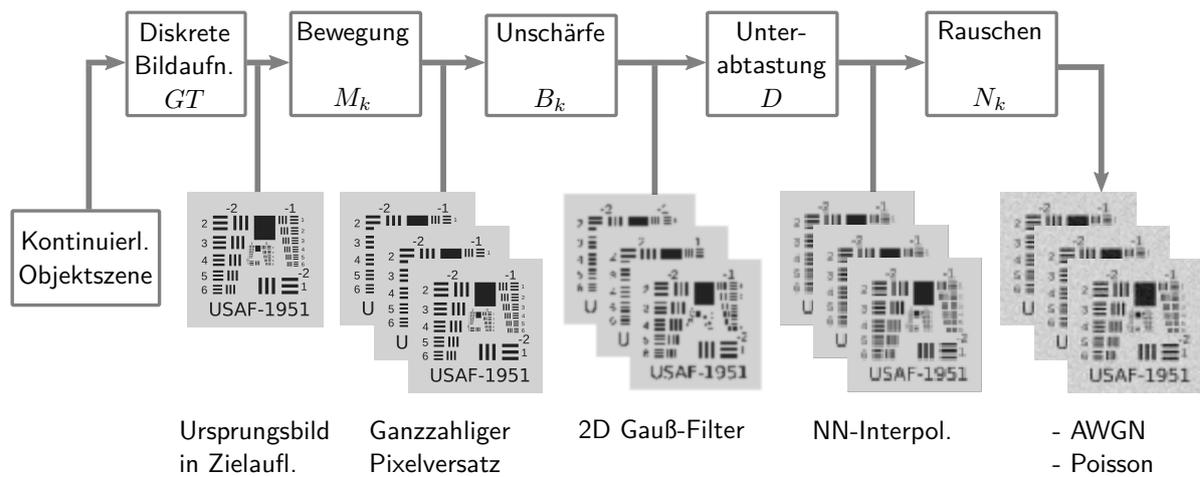


Abbildung 3.3: Das allgemeine Beobachtungsmodell beschreibt den Prozess von einem hochauflösenden Ursprungsbild zu mehreren degradierten Bildern geringerer Bildauflösung. Zu den berücksichtigten Degradationsmechanismen zählen in sequentieller Reihenfolge Bewegung, Unschärfe, Unterabtastung und Rauschen. Letzteres in Form von additivem weißen Gauß'schen Rauschen (AWGN) und Poisson-verteilterm Schrotrauschen.

zu der degradierten, niedrig auflösenden LR-Bilderfolge (engl. low resolution; LR). In sequenzieller Reihenfolge sind dies Relativbewegungen M_k zu einem Referenzbild GT , Unschärfe B_k , Unterabtastung (engl. Downsampling) D und Rauschen n_k . Die Bewegungen [22] aufeinanderfolgender LR-Bilder können globaler und/oder lokaler Natur sein und dürfen sowohl in der Objektszene als auch in der Bildebene vorliegen. In der Bildebene betrachtet, müssen sie jedoch im Bruchteil der Pixelgröße vorliegen, also kleiner sein als die LR-Pixelkanten lang sind [22, 42]. Während frequenzbasierte SR-Methoden auf globale translatorische Bewegungen für aufeinanderfolgende LR-Bilder limitiert sind [22, 34, 36, 41, 69], erlauben SR-Verfahren im Ortsraum eine nahezu beliebige Bewegungs-Modellierung [34]. Der Bewegungs-Operator M_k modelliert Translations- und Rotationsbewegungen [23] sowie Verzerrungen [22, 31, 36]. Der Unschärfe-Operator B_k beinhaltet jeglichen Einfluss von Unschärfe auf das LR-Bild. Dazu zählen beispielsweise die Sensor- und Linsenunschärfe [22, 36, 41, 43, 44] des optischen Systems, aber auch - sofern vorhanden - äußere Einflüsse wie Bewegungsunschärfe [22, 31, 41, 44, 46], Defokus [44] und atmosphärische Effekte [31, 41]. Für gewöhnlich modelliert eine räumlich mittelnende Punktspizfunktion (PSF) die Unschärfe [22], dessen PSF-Filter mit dem (gegebenenfalls verschobenen) GT-Bild gefaltet wird. Auch hinsichtlich der Unschärfe-Modellierung zeigt sich die frequenzbasierte SR mit der Limitierung auf ausschließlich rauminvariante Unschärfe [34, 36] weniger flexibel als die SR-Verfahren im Ortsraum. Dort sind sowohl linear rauminvariante als auch linear raumvariante Unschärfe-Modelle modellierbar [22]. Die Bewegungs- und Unschärfe-Operatoren sind kommutativ, wenn für alle LR-Bilder dieselbe rauminvariante Unschärfe gilt und lediglich translatorische Bewegung zwischen den LR-Bildern vorliegt [31, 68, 70]. Der Downsampling-Operator D definiert die Abtastrate, mit der das GT-Bild heruntergerechnet wird [31] und legt somit den Skalierungsfaktor fest. Wegen seiner Unabhängigkeit vom GT-Bild ist der Rauschparameter n_k als additiver Operator berücksichtigt [25]. Er modelliert das Rauschverhalten des Sensors, das bei der Bilderzeugung entsteht.

Einen Sonderfall beschreibt die SISR [23, 46, 68], wenn $k = 1$ gilt. Dann fällt der Bewegungsoperator M_k weg und das allgemeine Beobachtungsmodell vereinfacht sich zu:

$$LR = D \cdot B \cdot GT + n \quad (3.2)$$

Im Wesentlichen durchläuft der SR-Prozess drei Schritte [22, 25]:

1. Bildregistrierung (nur bei der MFSR): Subpixelgenaue Positionserfassung aller LR-Bilder zu einem Referenzbild (z.B. das erste Bild der LR-Bilderfolge). Sind die Bildverschiebungen nicht bekannt, werden sie durch Algorithmen geschätzt.
2. Interpolation: Eine ungleichförmige Interpolation überträgt die unter Umständen willkürlich verteilten LR-Bilder auf ein gemeinsames äquidistantes Pixelraster, dessen Größe der Zielauflösung entspricht.
3. Bildrestauration: Verbesserung der Bildqualität durch die Beseitigung von Unschärfe und Unterdrückung von Rauschen.

Je genauer die bereitgestellten oder geschätzten Informationen vorliegen, desto besser ist die Bildqualität im SR-Ergebnis zu erwarten.

3.4 Super-Resolution Problemstellungen

Bei der geometrischen SR erfolgt die Schätzung eines hochauflösenden HR-Bildes (engl. high resolution; HR) aus einem oder mehreren niedrig auflösenden LR-Bildern durch die Umkehrung des eben beschriebenen Beobachtungsmodells. Dabei wird von einer *Wirkung* - dem aufgenommenen LR-Bild - auf deren *Ursache* - dem zugrundeliegenden GT-Bild geschlossen [70]. Nach diesem Zusammenhang stellt die SR per Definition ein *inverses Problem* dar [71]. Außerdem handelt es sich nach den Bedingungen von Hadamard [62, 72] und Tikhonov [72] zudem um ein *schlecht gestelltes Problem*, da für das SR-Ergebnis keine eindeutige Lösung existiert [26, 73–77]. Neben der Korrektheit eines Problems formuliert Tikhonov die Stabilität eines Problems über die Stetigkeit der Lösung: Unendlich kleine Änderungen in den Eingangsdaten sollen zu unendlich kleinen Änderungen in der Lösung führen [72]. Im Kontext der SR wird dies in der Literatur oftmals als *schlecht konditioniertes Problem* [22, 31, 32, 68, 78, 79] beschrieben. Die schlecht konditionierten Probleme beziehen sich dabei insbesondere auf nicht exakt bekannte Degradationsoperatoren. Das heißt, kleinste Abweichungen im Degradationsmodell können zu großen Abweichungen im SR-Ergebnis führen. Konkret beschreiben die Gleichungen 3.1 und 3.2 ein unterbestimmtes lineares Gleichungssystem, das eine unendlich große Anzahl möglicher inverser Lösungen für das GT-Bild ausgibt [62]. Oder anders ausgedrückt: Die Anzahl unbekannter Variablen im HR-Rekonstruktionsbild übersteigt die Anzahl bekannter Variablen, die durch die LR-Eingangsdaten bereitgestellt werden [22, 78, 80]. In der Literatur wird dies auch als *One-to-many-Abbildungsproblem* [76, 81, 82] oder *Inferenzproblem* [76, 81] bezeichnet, welches sich mit größerem Skalierungsfaktor verschlechtert [73, 82]. Zusammengefasst wird die SR mit der Lösung der folgenden mathematischen Probleme konfrontiert:

- Inverses Problem: Die Schlussfolgerung von einem oder mehreren aufgenommenen LR-Bild(ern) auf das zugrundeliegende GT-Bild.

- Schlecht gestelltes Problem: Es existiert keine eindeutige SR-Lösung.
- Schlecht konditioniertes Problem: Geringfügige Änderungen in den Eingangsdaten können starke Änderungen im SR-Ergebnis herbeiführen.
- Inferenzproblem: Die Anzahl unbekannter Variablen im SR-Bild übersteigt die Anzahl an bekannten Variablen aus den LR-Bildern.

Zur eindeutigen und zuverlässigen Lösung der aufgeführten Probleme nutzen SR-Algorithmen oftmals Regularisierungsverfahren unter Verwendung von so genanntem a-priori-Wissen (Vorwissen). Zu den am häufigsten eingesetzten Regularisierungsmethoden zählen die verallgemeinerte Inverse, die Tikhonov-Regularisierung sowie probabilistische Ansätze wie Maximum-Likelihood (ML) oder Maximum-a-Posteriori (MAP) [31, 44, 62]. Dabei nutzen Regularisierungsterme a-priori-Wissen über das optische System, um den Lösungsraum einzuschränken [34] und damit das schlecht gestellte Problem in ein gut gestelltes Problem umzuwandeln [32, 62]. Je genauer die Degradationsoperatoren das optische System beschreiben, desto weniger potenzielle SR-Ergebnisse existieren. Ziel der SR ist die Rekonstruktion eines SR-Bildes, das möglichst wenig vom idealen GT-Bild abweicht. Dafür bestimmt eine Kostenfunktion die Abweichung von der SR-Rekonstruktion zum GT-Bild, beispielsweise über deren Differenzbild. Das finale SR-Bild resultiert aus einem Optimierungsprozess, der die minimale Abweichung findet.

Kapitel 4

Lernbasierte Single Image Super-Resolution

In diesem Kapitel wird mit der lernbasierten SISR eine erste Hauptgruppe der geometrischen SR-Techniken aus der Übersicht in Abbildung 3.2 vorgestellt. Einführend werden drei generische, lernbasierte SISR-Typen nach ihren Bildprioren, also den kanten- bzw. bildstatistischen, den beispelbasierten und den CNN-basierten Verfahren erläutert. Als aktueller Stand der Technik bei der Auflösungssteigerung mit nur einem LR-Eingangsbild, folgt in Abschnitt 4.1 eine nähere Beschreibung der CNN-basierten SISR. Unter einer Vielzahl möglicher Algorithmen, wird in dieser Arbeit der *EnhanceNet* SISR-Algorithmus zur Umsetzung der lernbasierten SISR gewählt und dessen Implementierung in Abschnitt 4.2 erläutert.

Eine Schlüsselannahme vieler SISR-Techniken ist, dass ein Großteil der hochfrequenten Informationen redundant sind und aus niederfrequenten Komponenten eines einzelnen LR-Bildes rekonstruiert werden können [81]. Im Fall der lernbasierten SISR lernen Algorithmen implizite Redundanz im LR-Bild anhand lokaler räumlicher Korrelation in LR-/HR-Bildpaaren [81]. Dazu lernen sie Priore (a-priori-Wissen) expliziter Bildeigenschaften aus Trainings-Datensätzen, welche Bildpaare aus einem hochauflösenden GT-Bild(-Ausschnitt) und seinem korrespondierenden niedrig auflösenden LR-Bild(-Ausschnitt) enthalten. Dieses erlernte Wissen wird bei der SR-Rekonstruktion genutzt, um das GT-Bild aus einem einzelnen LR-Bild abzuschätzen.

Nach ihren Bildprioren unterscheidet die generische SISR drei lernbasierte Typen [67]. Die 1.1) kantenbasierte (engl. edge based) und 1.2) bildstatistische (engl. image statistical) SR lernt Priore von Bildeigenschaften wie Gradientenprofile [60, 61] oder Gradientenverteilungen [62–65]. Dagegen lernen 2) beispelbasierte (engl. example-based oder patch based) SR-Methoden Abbildungsfunktionen mit Hilfe von Teilbildern. Diese stammen größtenteils aus externen Datensätzen [24, 63, 66, 73, 78, 83–93], teilweise aber auch aus dem Eingangsbild selbst [94–98]. Bei der beispelbasierten SR werden Eingangsbilder in eine Vielzahl kleiner Teilbilder (engl. Patches) zerlegt, deren Inhalt einfachste Bildinformationen, wie beispielsweise vertikale, horizontale und diagonale Kanten oder Kreise darstellen. Ein Trainingsprozess speichert die LR-Patches und deren korrespondierende HR-Patches in einer Datenbank ab, auf die der Algorithmus während dem SR-Prozess zugreift. Dieser vergleicht schließlich die LR-Patches aus dem Eingangsbild mit jenen, die in der Datenbank hinterlegt sind und wählt für die weitere Verarbeitung jenes HR-Patch, dessen LR-Patch am besten zu dem LR-Patch im Eingangsbild passt. Dies erfolgt für jedes einzelne Patch, bis das gesamte SR-Bild rekonstruiert ist. Bis zum Jahre 2014 stellte die beispelbasierte SR mit sparse coding [66, 78, 83, 99] den Stand der Technik dar, danach läuteten Dong et al. mit ihrer Publikation [84] das Zeitalter der 3) CNN-basierten SR [73, 84–93] ein. Diese erweitern gewissermaßen die beispelbasierte SR um neuronale Netze, mit denen sie die Abbil-

dungsfunktionen von LR- zu HR-Bild effizienter lernen. Eine tiefere Zusammenfassung der SISR-Methoden bietet die Publikation von Yang et al. in [67].

4.1 Super-Resolution Convolutional Neural Networks

Faltende neuronale Netze (engl. convolutional neural networks; CNNs) gehören den künstlichen neuronalen Netzen (engl. artificial neural networks; ANNs) [100] an und werden im Kontext der geometrischen SR derzeit in überwiegender Weise zum Lösen der SR-Problemstellungen eingesetzt. In der Regel lernen SR CNNs eine nichtlineare Abbildung von LR- zu GT-Bild [81, 88, 93] und wenden diese bei der SR-Rekonstruktion eines hochauflösenden Bildes aus einem einzelnen LR-Eingangsbild an [73, 84–93]. Dabei passt ein iterativer Trainingsprozess die variabel definierten Parameter, wie Filter und Funktionsparameter, im Netzwerk so lange an, bis das SR-Ergebnisbild möglichst wenig vom bekannten GT-Bild abweicht.

In Abbildung 4.1 ist der Aufbau eines einfachen SR CNNs skizziert. Das Netzwerk baut sich aus Gewichten (engl. weights) W (Rechtecke), Neuronen (Kreise) mit ihren inhärenten Aktivierungsfunktionen f_{akt} und Verbindern (Pfeile) auf, wobei es beliebig viele Schichten, mit wiederum beliebig vielen Neuronen, enthalten kann. Die Anzahl der Schichten beschreibt dessen Tiefe und die Anzahl der darin enthaltenen Neuronen seine Weite. Bei dem dargestellten Netz handelt es sich um ein vollständig verbundenes Netz, ersichtlich an den vollständig miteinander verbundenen Neuronen aufeinanderfolgender Schichten, die jedoch innerhalb ihrer Schichten voneinander getrennt sind [26].

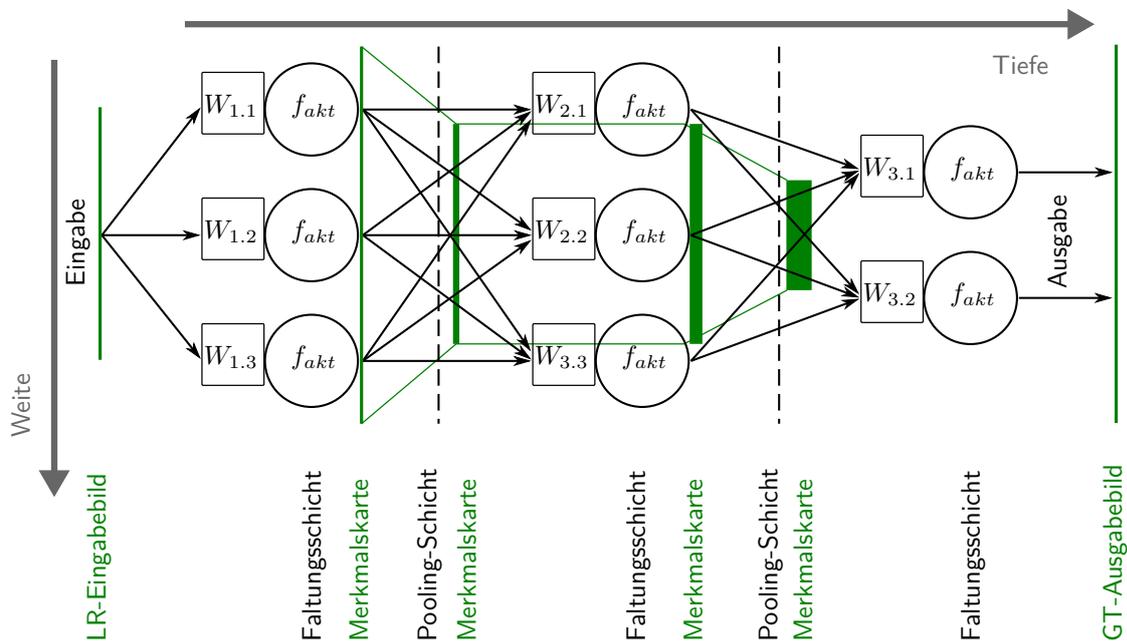


Abbildung 4.1: In einem SR CNN gewichten Filter W die eingelesenen Bildinformationen, bevor sie diese an die Neuronen übergeben und dort eine Aktivierungsfunktion f_{akt} über dessen Aktivierung entscheidet. In einem Trainingsprozess werden sämtliche Filter und Funktionsparameter so lange angepasst, bis die gewünschte Bildausgabe erreicht ist.

Typischerweise handelt es sich bei SISR-Anwendungen um zwei- oder dreidimensionale Eingangsdaten, deren Matrizen die Helligkeitswerte niedrig auflösender Grau- oder Farbwertbilder beinhalten. Gewichte W verstärken oder dämpfen das Eingangssignal jedes Neurons [26, 101], bevor eine Aktivierungsfunktion f_{akt} im Neuron über dessen Aktivierung entscheidet [101]. Optional gewichtet ein Bias, teilweise auch als Offset bezeichnet, die Eingangsdaten [26, 101] und stellt damit einen zusätzlichen Schwellwert zur Aktivierung eines Neurons dar. Lediglich bei seiner Aktivierung gibt ein Neuron seinen Wert zur weiteren Verarbeitung aus. Je höher der ausgegebene Aktivierungswert ist, desto genauer stimmt das durch das Filter abgefragte Merkmal mit den Eingangsdaten überein. Zu den gebräuchlichsten Aktivierungsfunktionen bei den SISR-CNNs zählen die tanh-, Sigmoid- und ReLU-Funktionen (engl. Rectified Linear Unit, deutsch gleichgerichtete Lineareinheit) [26], welche die nichtlineare Abbildung von Eingangs- zu Ausgangsdaten umsetzen [88].

In den namensgebenden Faltungsschichten der CNNs falten Filtermatrizen die Eingangsdaten der einzelnen Neuronen mit dessen vorgeschalteten Gewichten [77]. Deren Filtergröße, auch rezeptives Feld genannt, ist kleiner als das LR-Eingangsbild [27, 75, 84, 89, 90, 93] und muss deshalb sukzessive über das Bild gerastert werden, um dieses vollständig zu erfassen. Dabei überlappen sich die aufeinanderfolgenden Teilausschnitte, aus deren Gesamtheit der Algorithmus schlussendlich das SR-Ergebnisbild rekonstruiert [84].

Die Ausgaben der Faltungsschichten münden in Merkmalskarten (engl. feature maps oder activation maps), die in Matrizendarstellung beschreiben, wie gut die Filterabfrage mit der Eingabematrix (Teilausschnitt) übereinstimmt. Für jeden Teilausschnitt entspricht die Anzahl der Merkmalskarten der Filteranzahl (und damit der Neuronenanzahl) in der betrachteten Schicht [84]. Optional reduzieren bündelnde Pooling-Schichten die Informationsmenge durch Unterabtastung (engl. subsampling), woraus sich für tiefer gelegene Faltungsschichten eine Vergrößerung des rezeptiven Feldes ergibt und damit das Extrahieren komplexerer Merkmale ermöglicht wird [82, 102].

Bei mehreren Hundert oder Tausend Neuronen ist ein händisches Anpassen von Gewichtungen nicht mehr sinnvoll möglich, weshalb dies von Algorithmen übernommen wird [101]. Dabei steuert die Wahl der Gewichtungen die Ausgabedaten des gesamten Netzwerks auf Basis seiner Eingabedaten [101]. Im Falle der SISR sucht der Algorithmus eine Konstellation von Gewichtungen, mit denen aus dem niedrig auflösenden Eingangsbild ein hochauflösendes SR-Ergebnisbild resultiert, das möglichst wenig vom idealen GT-Bild abweicht. Das Anpassen der Gewichte - und auch von optionalen Parametern, wie beispielsweise dem Bias - nennt sich Lernen oder Trainieren [101]. Ein Trainingsprozess passt sämtliche optimierbare Parameter iterativ durch das Netzwerk selbst an. Dazu erfolgt die Optimierung auf eine Zielfunktion (engl. objective function) hin, die bei SR-Anwendungen im Kern einer Verlustfunktion entspricht [75]. Dabei minimiert der Optimierungsprozess die Fehlerabweichung vom SR-Ergebnis zu seinem GT-Bild [84]. In Gleichung 4.1 [27] ist eine solche Optimierungsfunktion $\hat{\theta}$ zusammengefasst. Basierend auf dem allgemeinen Beobachtungsmodell der SISR nach Gleichung 3.2, sucht die Verlustfunktion das Fehlerminimum zwischen I^{SR} und I^{GT} . Der zweite Term $\Phi(\theta)$ beschreibt den Regularisierungsterm [27] mit seinem Regularisierungsparameter λ [31] und seinem Argument θ . Der Regularisierungsparameter wägt zwischen einer Bildglättung und dem Erhalt von Diskontinuitäten (z.B. Kanten) ab [31] und wird daher auch als Kompromissparameter bezeichnet [27, 92].

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \mathcal{L}(I^{SR}, I^{GT}) + \lambda \Phi(\theta) \quad (4.1)$$

In der Literatur findet die Verlustfunktion nach Gleichung 4.2 [27], die den mittleren quadratischen Fehler (engl. mean squared error; MSE) vom Schätzwert (I^{SR}) zum Zielwert (I^{GT}) pixelweise im Bildraum minimiert [27, 75, 76, 81, 84, 89–92, 103], häufig Anwendung. Sie entspricht dem arithmetischen Mittel der quadrierten l2-Norm (auch euklidischer Abstand genannt), wie sie in Gleichung 4.3 formuliert ist.

$$MSE(I^{SR}, I^{GT}) = \frac{1}{M} \|I^{SR} - I^{GT}\|_2^2 = \frac{1}{M} \sum_i^M (I_i^{SR} - I_i^{GT})^2 \quad (4.2)$$

$$l2 - Norm = \|I^{SR} - I^{GT}\|_2 = \sqrt{\sum_i^M (I_i^{SR} - I_i^{GT})^2} \quad (4.3)$$

Mit ihrem SRCNN-Algorithmus begründeten Dong et al. im Jahr 2014 [84] die Ära der auf künstlicher Intelligenz basierten SISR. Seitdem gingen zahlreiche weitere Publikationen hervor, die sich in der erreichten Bildqualität bei den SR-Ergebnissen gegenseitig aufs Neue übertrafen. In Abbildung 4.2 ist ein Auszug einiger SR CNNs gegenübergestellt, deren erreichte Bildqualität als Spitzen-Signal-Rausch-Verhältnis (engl. peak signal-to-noise ratio; PSNR) über die Netzwerktiefe aufgetragen ist. Die gestrichelten Linien unterteilen die Netzwerktiefe in drei Bereiche. Schmale Netze enthalten weniger als 5 Schichten, tiefe Netze zwischen 5 und 30 Schichten und sehr tiefe Netze enthalten mehr als 30 Schichten. Neben den Kurzbezeichnungen und dem Veröffentlichungsjahr in Klammern, gibt die zweite Zeile für jedes Netzwerk das PSNR in Dezibel [dB] an. Diese Angaben gelten für den Trainings-Datensatz *Set5* nach Bevilacqua et al. [104] bei vierfacher Auflösungssteigerung und pixelweiser Fehlerminimierung beim Optimierungsprozess. Die Gegenüberstellung erhebt keinen Anspruch auf Vollständigkeit und stellt viel mehr eine Momentaufnahme der bekanntesten bis ins Jahr 2017 publizierten CNN-basierten SISR-Verfahren dar.

Als Ausgangspunkt der Gegenüberstellung ist das SRCNN von Dong et al. [84] zu sehen, das mit seinen drei Schichten ein verhältnismäßig schmales Netz darstellt und ein PSNR von 30,09 dB erreicht. Bei gleicher Anzahl an Schichten liefert das ESPCN-Netzwerk von Shi et al. [81] bereits ein PSNR von 30,90 dB. Das CSCN-Netzwerk von Wang et al. [85] schafft mit fünf Schichten den Sprung in die Kategorie der tiefen Netze und kommt auf ein PSNR von 30,86 dB. Darauf folgt das EEDS-Netzwerk von Wang et al. [88], das ein 13 Schichten tiefes Netzwerk mit einem dreischichtigen schmalen Netzwerk kombiniert und damit ein PSNR von 31,14 dB erreicht. Die beiden nächsten Netze kommen auf 20 Schichten, wobei das VDSR-Netzwerk von Kim et al. [89] ein PSNR von 31,35 dB liefert und das DRCN-Netzwerk von Wang et al. [90] 31,53 dB. Mit bis zu 52 Faltungsschichten zählt das DRRN-Netzwerk von Wang et al. [93] zu den sehr tiefen Netzen und erreicht ein PSNR von 31,68 dB. Das beste SR-Ergebnis in dieser Gegenüberstellung liefert das ENet-Netzwerk von Sajjadi et al. [82], in Fettdruck hervorgehoben, das mit insgesamt 35 Schichten (20 Schichten im Generator-Netz sowie 15 Schichten im Diskriminator-Netz) den sehr tiefen Netzen zuzuordnen ist. Sein SR-Ergebnis liefert ein PSNR von 31,74 dB und wird deshalb in vorliegender Arbeit als Vertreter der lernbasierten SISR gewählt. Im nachfolgenden Kapitel folgt eine nähere Beschreibung zu dessen Funktionsweise.

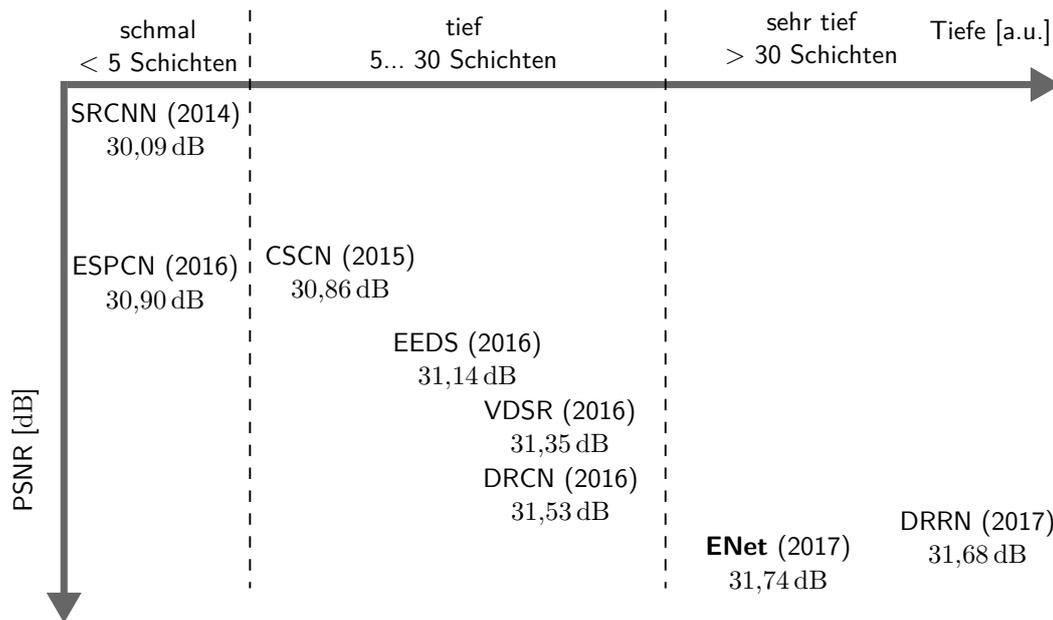


Abbildung 4.2: Die Gegenüberstellung trägt die erreichte Bildqualität im SR-Ergebnis als PSNR über der Netzwerktiefe als Anzahl der enthaltenen Schichten für einige bekannte Vertreter der SR CNNs auf. Alle Angaben beziehen sich auf den Set5-Datensatz nach Bevilacqua et al. [104] bei 4-facher Vergrößerung und für die pixelbasierte SR-Rekonstruktion.

4.2 EnhanceNet Super-Resolution

Das *ENet-PAT* ist ein vorwärts gerichtetes, vollständig verbundenes CNN zur SISR, das bei der SR-Rekonstruktion ein kontradiktorisches Training berücksichtigt und das SR-Ergebnis merkmalsbasiert optimiert.

Üblicherweise optimieren SR CNNs ihr Trainingsmodell pixelweise im Bildraum. Dies führt zu hohen Spitzen-Signal-Rausch-Verhältnissen (PSNR), wenngleich nur eine geringe Übereinstimmung mit der menschlichen Wahrnehmung von Bildqualität [73, 105, 106] besteht. Dem entgegen legt das SISR-Verfahren *ENet-PAT* nach Sajjadi et al. [82] seinen Schwerpunkt auf die Rekonstruktion realistischer Texturen, anstelle der pixelgenauen Reproduktion von GT-Bildern. In Abbildung 4.3 ist der Unterschied an einem Minimalbeispiel [82] verdeutlicht. Es zeigt a) ein hochauflösendes GT-Bild, b) sein heruntergerechnetes LR-Bild sowie dessen SR-Ergebnisse c) nach einer optimalen pixelgenauen MSE-Optimierung (vgl. auch Gleichung 4.2) und d) nach einer merkmalsbasierten Optimierung. Der optimale MSE-Schätzer trainiert die Minimierung des euklidischen Abstands und gibt den Mittelwert vieler möglicher Lösungen aus, was zu einem unnatürlich wirkenden und unplausiblen SR-Ergebnis führt. Dagegen optimiert der *ENet-PAT*-Algorithmus auf mehrere Verlustfunktionen im Merkmalsraum hin, woraus ein scharfes Bild resultiert, das in seiner menschlichen Wahrnehmung nur schwer vom Originalbild zu unterscheiden ist. Allerdings weichen die Orientierungen der Rechtecke im SR-Ergebnis von denen im GT-Bild ab und führen deshalb bei einer pixelgenauen Bewertung, wie der PSNR-Messung, zu größeren Abweichungen als bei dem unscharfen MSE-Ergebnis.

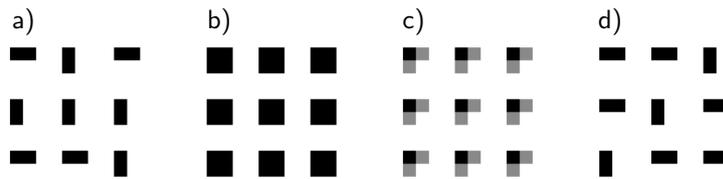


Abbildung 4.3: Das Minimalbeispiel nach [82] visualisiert a) ein hochauflösendes GT-Bild in Form verschiedenartig orientierter Rechtecke sowie b) sein heruntergerechnetes LR-Bild. Ferner veranschaulicht es den Unterschied einer c) konventionellen pixelbasierten SR-Rekonstruktion zu der d) merkmalsbasierten SR-Rekonstruktion nach dem ENet-PAT-Algorithmus.

Das ENet-PAT-CNN von Sajjadi et al. [82] gehört den vorwärts gerichteten, vollständig gefalteten Netzwerken (engl. feed-forward fully convolutional networks) an und realisiert ein so genanntes erzeugendes kontradiktorisches Netz (engl. generative adversarial network; GAN), das sich definitionsgemäß aus einem Generator-Netz G_{net} und seinem entgegen wirkenden Diskriminator-Netz D_{net} zusammensetzt [26, 27, 82, 107, 108]. Das Generator-Netz lernt Residuen-Bilder (Hochpass-Bilder) I^{res} aus bikubisch interpolierten LR-Bildern und ihren GT-Pendants, um damit ein SR-Bild I^{SR} abzuschätzen. Dabei zielt es auf eine maximale Übereinstimmung zwischen SR-Ergebnisbild I^{SR} und zugehörigem GT-Bild I^{GT} ab. Das SR-Ergebnis ergibt sich schließlich aus der Addition eines bikubisch interpolierten LR-Startbildes I^{kub} mit dem gelernten Residuen-Bild I^{res} [82]:

$$I^{SR} = I^{kub} + I^{res} \quad (4.4)$$

Das bewertende Diskriminator-Netz prüft, ob das ihm vorgelegte Bild vom Generator-Netz stammt oder ob es sich um ein GT-Bild aus dem Trainingsdatensatz handelt [82]. Es wird darauf trainiert das Generator-Ergebnisbild vom zugrundeliegenden GT-Bild zu unterscheiden. Sein Ziel ist es, selbst minimale Abweichungen zwischen dem SR-Ergebnisbild aus dem Generator-Netz und dem GT-Bild zu erkennen. Im Zuge der SR-Rekonstruktion optimiert das ENet-PAT-Verfahren auf die Verlustfunktionen der Gleichungen 4.5 bis 4.8 hin.

Die erste Verlustfunktion, der *Wahrnehmungsverlust im Merkmalsraum*, minimiert die mittlere quadratische Fehlerabweichung (MSE) im Merkmalsraum. Dazu bildet eine differenzierbare Funktion ϕ die beiden Intensitätsbilder I^{SR} und I^{GT} in den Merkmalsraum ab, wo anschließend die Fehlerabweichung nach der quadrierten l2-Norm berechnet wird [82]:

$$\mathcal{L}_P = \|\phi(I^{SR}) - \phi(I^{GT})\|_2^2 \quad (4.5)$$

Anstelle der klassischen pixelweisen Optimierung setzt Gleichung 4.5 eine merkmalsbasierte Optimierung um, wodurch das Netzwerk nicht auf möglichst identische Pixelwerte im Bildraum, sondern auf gleiche Merkmale im Merkmalsraum hin optimiert wird. Dies führt zu einem Ergebnisbild, das der menschlichen Wahrnehmung angenähert ist [73, 82]. Die alleinige Optimierung auf den Wahrnehmungsverlust im Merkmalsraum führt zu scharfen Ergebnisbildern, verursacht darin jedoch Artefakte, weil mit ihr Details in fein strukturierten Bildbereichen nicht reproduziert werden [82].

Um die Detailtreue im SR-Ergebnis zu erhöhen, erfolgt zusätzlich ein *kontradiktorisches Training*, bei dem Generator-Netz G_{net} und Diskriminator-Netz D_{net} wechselseitig gegeneinander optimiert werden. Für den Fall der SISR entspricht die Eingabe des Generators einem LR-

Bild I^{LR} , die des Diskriminators dagegen einem GT-Bild I^{GT} aus dem Trainingsdatensatz. Aus dem LR-Eingangsbild schätzt der Generator sein SR-Ausgabebild I^{SR} ($G_{net}(I^{LR}) = I^{SR}$), welches anschließend das Diskriminator-Netz mit den GT-Bildern aus dem Datensatz vergleicht. Während das Generator-Netz ein SR-Ergebnis anstrebt, das möglichst wenig von seinem GT-Pendant abweicht, hat das Diskriminator-Netz zum Ziel selbst geringste Abweichungen zwischen SR-Ergebnis und GT-Bild zu erkennen. Dazu minimiert das Training zwei Verlustfunktionen, eine für den Generator

$$\mathcal{L}_A = -\log(D_{net}(G_{net}(I^{LR}))) = -\log(D_{net}(I^{SR})) \quad (4.6)$$

und eine für den Diskriminator

$$\mathcal{L}_D = -\log(D_{net}(I^{GT})) - \log(1 - D_{net}(I^{SR})). \quad (4.7)$$

Die wechselweise Optimierung des kontradiktorischen Trainings trägt zu einer realistisch wirkenden SR-Rekonstruktion bei. Mit der Kombination aus den ersten drei Verlustfunktionen der Gleichungen 4.5 bis 4.7 werden hochfrequente Bilddetails rekonstruiert und die Bildschärfe in den SR-Ergebnissen nimmt zu. Allerdings tritt bei dem Versuch Teilbereiche zu glätten gelegentlich hochfrequentes Rauschen auf und das Ergebnis tendiert teilweise dazu an Kanten hohe Frequenzen hinzuzufügen. Dies führt in einigen Fällen zu Halos und scharfen Artefakten [82]. Die vierte Verlustfunktion reduziert den *Verlust der Texturanpassung*, indem die Struktur im SR-Ergebnisbild iterativ an die des GT-Zielbildes angeglichen wird. Das Netzwerk lernt durch den Trainingsprozess Bilder zu generieren, die die gleichen lokalen Texturen aufweisen wie die hochauflösenden Bilder. Texturabweichungen werden nach Gleichung 4.8 [82] mit einer Gram-Matrix $G(F) = F^T F \in \mathbb{R}^{n \times n}$ [82] unter Berücksichtigung von n Merkmalen bestimmt. Die Berechnungen erfolgen stückweise, um lokal ähnliche Strukturen zu verstärken. Die letzte Verlustfunktion hilft dem Netzwerk lokal sinnvolle Texturen zu erzeugen und reduziert die nach den ersten drei Verlustfunktionen verbliebenen Artefakte [82].

$$\mathcal{L}_T = \|G(\phi(I^{SR})) - G(\phi(I^{GT}))\|_2^2 \quad (4.8)$$

Die Autoren Sajjadi et al. stellen eine vor-trainierte Referenz-Implementierung von ihrem *ENet-PAT*-Algorithmus [109] zur Verfügung, die zur SR-Rekonstruktion aus einem einzelnen LR-Bild die oben beschriebenen Verlustfunktionen der Gleichungen 4.5 bis 4.8 berücksichtigt. Das Modell wurde mit allen Farbbildern des MSCOCO-Datensatzes [110] gelernt, deren kurze Bildseite mindestens 384 Pixel zählt [82]. Dies entspricht einer Gesamtanzahl von etwa 200k Bildern [82]. Zunächst skaliert der Algorithmus ein zur Verfügung gestelltes GT-Bild um den Faktor 4 herunter, um ein LR-Bild zu erhalten. Auf dieses LR-Eingangsbild wendet die Referenz-Implementierung den *ENet-PAT*-Algorithmus an und gibt ein SR-Ergebnisbild in der ursprünglichen Auflösung des GT-Bildes aus [109].

Kapitel 5

Multi-Frame Super-Resolution

In Kapitel 5 werden zwei Vertreter der Multi-Frame SR (MFSR) vorgestellt, welche der zweiten Hauptgruppe der geometrischen SR-Techniken aus der Übersicht in Abbildung 3.2 zugeordnet sind. Entgegen den SISR-Methoden aus Kapitel 4, werden bei der MFSR keine zusätzlichen Informationen aus einer Datenbank extrahiert, sondern komplementäre Bildinformationen [23] durch Mehrfachaufnahmen der Objektszene gewonnen. Bei der 1) *inkohärenten Fourier Ptychographie* in Abschnitt 5.1, als Vertreter der strukturierten Beleuchtung, wird die Objektszene mit einem konstanten Beleuchtungsmuster beleuchtet, dessen laterale Position für jedes LR-Bild variiert. Dagegen sind bei der 2) *Subpixel-Shift SR*, beschrieben in Abschnitt 5.2, die Einzelaufnahmen innerhalb einer LR-Bilderfolge im Bruchteil deren Pixelgröße zueinander verschoben. Während in den beiden Unterabschnitten 5.1.1 und 5.1.2 zunächst die Funktionsweise und anschließend das Simulationsmodell der inkohärenten Fourier Ptychographie beschrieben sind, werden in den Unterabschnitten 5.2.1 und 5.2.2 analog dazu die Funktionsweise sowie der in dieser Arbeit verwendete Algorithmus zur Umsetzung der Subpixel-Shift-Methode erläutert.

Bezogen auf das allgemeine Beobachtungsmodell aus Abschnitt 3.3 ist die MFSR-Methode immer dann möglich, wenn sich mindestens ein Parameter in Gleichung 3.1 beziehungsweise Abbildung 3.3 für aufeinanderfolgende LR-Bilder ändert [31]. Dazu zählen Bewegung, Unschärfe (optisch, atmosphärisch und/oder Bewegungsunschärfe), Zoom und Belichtung [31]. Wegen der erforderlichen Mehrfachaufnahmen erfolgt bei der MFSR die Steigerung der Bildauflösung auf Kosten der zeitlichen Auflösung.

5.1 Inkohärente Fourier Ptychographie

Die inkohärente Fourier ptychographische Bildgebung vereint Mehrfachaufnahmen unter variierender Beleuchtungssituation mit einem intensitätsmodulierten Muster, um hochfrequente Bildinformationen in den darstellbaren Bereich des optischen Systems zu verschieben und damit dessen numerische Apertur synthetisch zu erweitern.

Die inkohärente oder musterbeleuchtete Fourier Ptychographie verfolgt den Ansatz der strukturierten Beleuchtungsmikroskopie (engl. structured illumination microscopy; SIM) und hatte ursprünglich die Auflösungssteigerung beugungsbegrenzter Fluoreszenz-Mikroskope zum Ziel [49, 50, 52, 54, 111–113]. Dong et al. beschreiben in [112] erstmals den Einsatz der inkohärenten Fourier Ptychographie zur Auflösungssteigerung makroskopischer Kamerasysteme. Sie

steigert die Auflösung von Aufnahmen optisch rauer Oberflächen in großen Objektentfernungen (Größenordnung m). Als ein Vertreter der MFSR, resultiert der Informationsgewinn für die Auflösungssteigerung aus Mehrfachaufnahmen der Objektszene. Dabei unterscheidet sich die laterale Position eines statischen Musters für jede Einzelaufnahme innerhalb einer LR-Bilderfolge.

5.1.1 Funktionsweise der inkohärenten Fourier Ptychographie

Die inkohärente Fourier Ptychographie, nachfolgend mit FP abgekürzt, verschiebt hochfrequente Objektinformationen in den Bandpass der abbildenden Optik durch die Wechselwirkung der strukturierten Beleuchtung mit der Objektstruktur [54]. Dabei superponieren Strukturen des Beleuchtungsmusters mit den Objektstrukturen zu einem Überlagerungsmuster im Ortsraum multiplikativ [49]. Im Fourierraum entspricht dies einer Faltung vom Beleuchtungsmusterspektrum mit dem Objektspektrum [55], wodurch Objektfrequenzen, die über der Auflösungsgrenze des Objektivs liegen, in dessen Detektionsbereich verschoben werden. In Abbildung 5.1 ist die Wechselwirkung eines statischen Beleuchtungsmusters mit der Objektstruktur im Ortsraum visualisiert. Bei der Beleuchtung eines unbekanntes Musters in a) mit einem bekannten sinusförmigen Streifenmuster in b), resultiert dessen Überlagerungsmuster in c), welches schließlich von einer Kamera erfasst wird. Dabei verschiebt der Moiré-Effekt höherfrequente Bildinformationen in den Bandpass des optischen Systems [49, 50, 52]. Mit der Kenntnis über Beleuchtungs- und Überlagerungsmuster erfolgt die Rekonstruktion der Objektstruktur, wobei beliebige Beleuchtungsstrukturen verwendet werden können [54, 55].

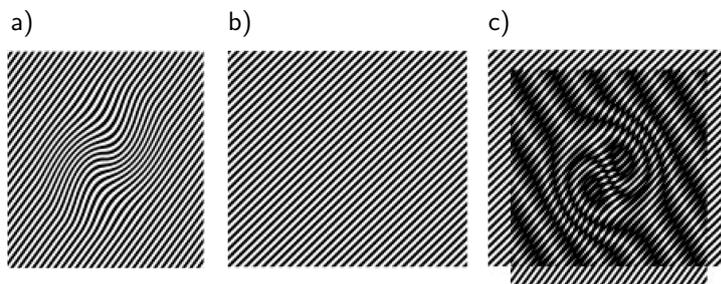


Abbildung 5.1: Aus der Beleuchtung einer unbekanntes Objektstruktur in a) mit einem bekannten Beleuchtungsmuster in b) resultiert das Überlagerungsmuster in c). Letzteres beinhaltet die Differenzfrequenzen aus der Beleuchtungsmusterfrequenz und den Ortsfrequenzen der unbekanntes Objektstruktur. Mit der Aufnahme des Überlagerungsmusters und der Kenntnis über das Beleuchtungsmuster lässt sich die Objektstruktur rekonstruieren, sofern die Differenzfrequenzen im Bandpass des optischen Systems liegen. Abbildung übernommen aus [50].

Prinzipiell ist die maximale Auflösungssteigerung der FP skalierbar [54, 114]. Um die Zielauflösung zu maximieren, ist es wünschenswert, dem Beleuchtungsmuster möglichst hohe Ortsfrequenzen aufzuprägen [50]. Ein definiertes Beleuchtungsmuster mit Ortsfrequenzen k_{Bel} beleuchtet die unbekanntes Objektszene k_{Obj} , wodurch ein Überlagerungsmuster entsteht. Liegen die Differenzfrequenzen k_{Diff} innerhalb der Grenzfrequenz des optischen Systems $|k_{Obj} - k_{Bel}| < k_{opt}$, sind die Überlagerungsmuster in der LR-Aufnahme sichtbar [50]. Für den Beleuchtungsfall mit sinusförmigen Streifen sind dies Moiré-Muster. Aus der Kenntnis über Beleuchtungsmuster k_{Bel} und Überlagerungsmuster k_{Diff} lässt sich die Objektszene k_{Obj} mit den

darin enthaltenen höheren Ortsfrequenzen rekonstruieren. Der Informationsgehalt nimmt mit der Anzahl an LR-Bildaufnahmen zu, wobei für jedes Einzelbild das Beleuchtungsmuster lateral verschoben ist. Dessen Schritt- oder Scangröße entspricht dabei der Größenordnung der feinsten Beleuchtungsstruktur [54]. Auf die Betrachtung im Fourierraum übertragen, führt jede zusätzliche Aufnahme zur virtuellen Erweiterung des Frequenzraumes.

Der FP-Algorithmus rekonstruiert die Objektszene und das Beleuchtungsmuster [54], indem er eine Phasenwiederherstellungstechnik (engl. phase retrieval) anwendet [115]. Verallgemeinert sucht er eine hochauflösende SR-Schätzung, die konsistent mit allen zur Verfügung gestellten LR-Aufnahmen ist [115]. Gleichung 5.1 [54, 115] beschreibt das Abbildungsmodell der FP im Fourierraum. Zunächst erfolgt die Multiplikation der hochauflösenden Objektszene I^{GT} mit dem Beleuchtungsmuster P_k im Ortsraum. Dessen Superpositionsergebnis wird mit $\mathcal{F}(I^{GT} \cdot P_k)$ in den Fourierraum transformiert, wo es mit der optischen Transferfunktion OTF der auflösungsbegrenzenden optischen Komponente multipliziert wird. Daraus resultiert die Fouriertransformierte des niedrig aufgelösten Intensitätsbildes $\mathcal{F}(I_k^{LR})$ für jedes k -te Einzelbild einer LR-Bilderfolge. Die FP zielt darauf ab aus den LR-Bildaufnahmen ein SR-Ergebnis I^{SR} zu rekonstruieren, das möglichst geringe Abweichungen zu seiner idealen GT-Objektszene I^{GT} aufweist.

$$\mathcal{F}(I_k^{LR}) = OTF \cdot \mathcal{F}(I^{GT} \cdot P_k) \quad (5.1)$$

Im Wesentlichen kehrt der Rekonstruktionsprozess die beiden Multiplikationen in Gleichung 5.1 um. Die nachfolgende Beschreibung stammt maßgeblich aus Zhengs Publikation in [115]. Als Startpunkt dient eine Erstabschätzung der GT-Objektszene I_{est}^{GT} aus der niedrig aufgelösten LR-Bilderfolge I_k^{LR} oder als zufällige Schätzung. Die Multiplikation von I_{est}^{GT} mit dem Beleuchtungsmuster P_k entspricht der Schätzung eines ersten hochauflösenden HR-Bildes $I_k^{HR} = I_{est}^{GT} \cdot P_k$ im Ortsraum. Nach dessen Transformation in den Ortsfrequenzraum, wird die Erstschtzung dort nach dem Bandpassverhalten des optischen Systems, in Form dessen OTF , degradiert. Nun aktualisiert jedes weitere LR-Bild I_k^{LR} den Informationsgehalt in der HR-Schätzung $I_{updated,k}^{HR}$, um deren zusätzlich enthaltene Objektinformationen aus der jeweiligen Beleuchtungssituationen P_k im Fourierraum nach folgendem Zusammenhang:

$$\mathcal{F}(I_{updated,k}^{HR}) = \mathcal{F}(I_k^{HR}) + OTF \cdot (\mathcal{F}(I_k^{LR}) - OTF \cdot \mathcal{F}(I_k^{HR})) \quad (5.2)$$

Dessen Rücktransformation in den Ortsraum führt zur aktualisierten HR-Schätzung $I_{updated,k}^{HR}$, mit der die Erstabschätzung der GT-Objektszene I_{est}^{GT} im Ortsraum wie folgt aktualisiert wird:

$$I^{SR} = I_{est}^{GT} + \frac{P_k}{(\max(P_k))^2} \cdot (I_{updated,k}^{HR} - I_{est}^{GT} \cdot P_k) \quad (5.3)$$

Der Aktualisierungsprozess wiederholt sich für alle Beleuchtungsmuster und LR-Bilder bis die Lösung konvergiert. Als Konvergenzbedingung dient entweder eine definierte Anzahl an Iterationen [54, 115] oder ein Schwellwert [112, 115], den es für Ergebnisse aufeinanderfolgender Iterationen zu unterschreiten gilt.

5.1.2 FP-Simulationsmodell für PMD-Kameras

Ein allgemeingültiges Simulationsmodell der muster-beleuchteten FP ist dem Anhang A.1 als auskommentiertes Octave-/Matlab-Skript angefügt und basiert maßgeblich auf Zhengs Code in [115]. Es berücksichtigt dabei die Eigenschaften der PMD-Kamera, wie sie in Abschnitt 2.2, Tabelle 2.1 angegeben sind. Abbildung 5.2 dient als Visualisierungshilfe für das Simulationsmodell und zeigt in Zeile 1 das Abbildungsmodell der FP im Ortsraum. Ein konstantes, jedoch lateral variables Speckle-Muster beleuchtet das GT-Objekt, dessen Bildaufnahme in den Fourierraum transformiert wird. Dort filtert die *OTF* des optischen Systems, in Form einer quadratischen Blende, die hochfrequenten Bildanteile heraus. Als Ausgabe resultiert eine muster-beleuchtete LR-Bilderfolge, dessen Beleuchtungssituation für jedes Einzelbild geringfügig voneinander abweicht. Das Tiefpassverhalten im Orts-Frequenzraum ist in Zeile 2 für einen Downsampling-Faktor von 2 dargestellt. Für höhere Downsampling-Faktoren würde sich der Bandpass und damit das weiße Quadrat weiter verkleinern. Zeile 3 zeigt exemplarisch ein bandbegrenztetes LR-Bildpaar der Bilderfolge im Orts- und Frequenzraum.

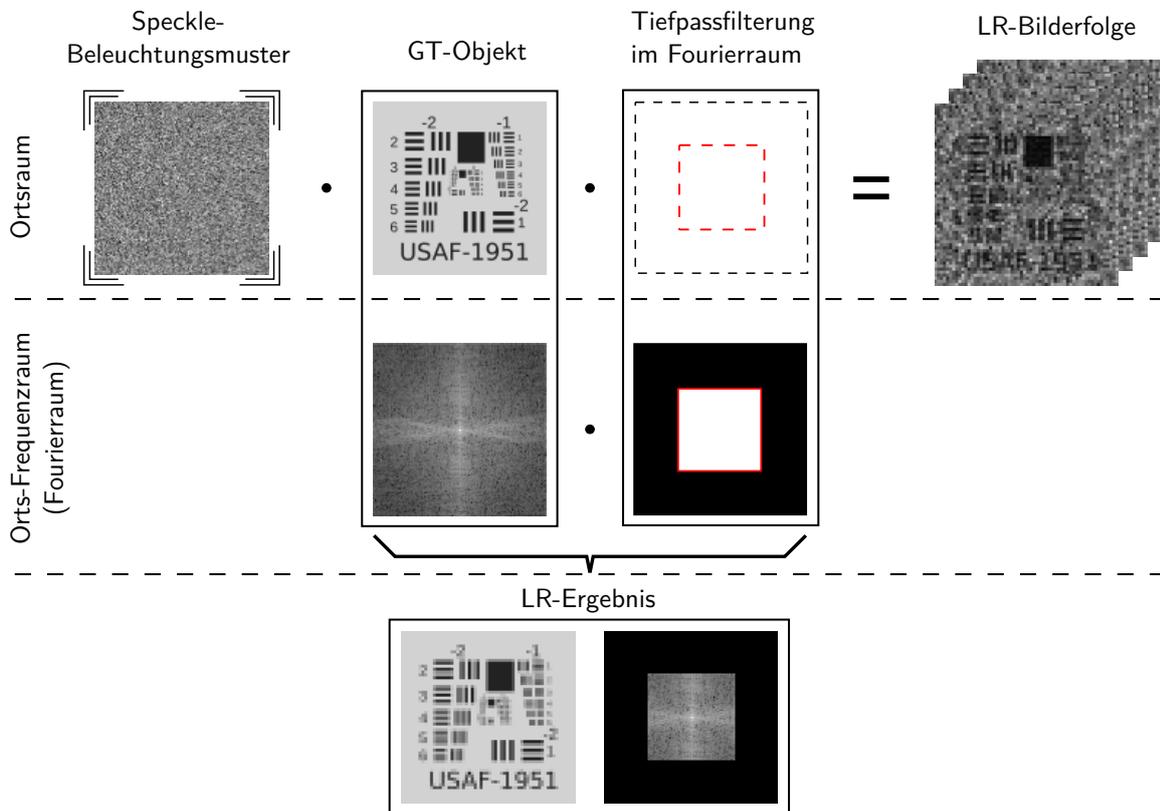


Abbildung 5.2: Das Simulationsmodell der inkohärenten Fourier Ptychographie wechselt bei seinen Berechnungen zwischen dem Orts- (Zeile 1) und Fourierraum (Zeile 2). Im Ortsraum beleuchtet ein konstantes Speckle-Muster die Objektszene, dessen laterale Position sich für jede Einzelaufnahme in der Größenordnung der angestrebten HR-Pixelgröße verschiebt. Das Downsampling erfolgt im Orts-Frequenzraum über eine Tiefpassfilterung, dessen Bandpass die OTF des begrenzenden optischen Elements bestimmt. Die Ausgabe ist eine bandbegrenzte LR-Bilderfolge, dessen Beleuchtungssituation für jedes Einzelbild variiert.

Zur besseren Übersichtlichkeit wird der Programmablauf des Skripts in Anhang A.1 in 4 Blöcke unterteilt. Im ersten Block wird ein hochauflösendes GT-Bild im Ortsraum eingelesen und anschließend mit Hilfe der schnellen Fourier Transformation (engl. fast Fourier transform; FFT) in den Orts-Frequenzraum umgewandelt. Dies stellt die ideale GT-Objektszene in Orts- und Frequenzraum dar, die es zu rekonstruieren gilt und in Abbildung 5.2 als schwarz umrahmtes GT-Objekt hervorgehoben ist. Daraufhin erfolgt im zweiten Block in den Zeilen 10 bis 15 die Modellierung des optischen Systems. Dabei werden zwei Annahmen für das optische Modell getroffen: Zum einen übersteigt das Auflösungsvermögen der abbildenden Optik die laterale Ortsauflösung des Sensors und zum anderen sind die Sensor- und Pixelflächen quadratisch. Demnach wirken Pixelgröße und -abstand noch vor dem Auflösungsvermögen der Optik als limitierende Größen, was sich in Form einer quadratischen Blende im Orts-Frequenzraum äußert und in Abbildung 5.2 durch das rote Quadrat gekennzeichnet ist. In Zeile 15 des Skripts wird das Frequenzspektrum des niedrig auflösenden LR-Bildes durch die Multiplikation des GT-Bildes mit der begrenzenden Sensor- OTF (rotes Quadrat) im Fourierraum definiert. In Abbildung 5.2 ist diese Tiefpassfilterung in Zeile 2 verbildlicht und das daraus resultierende LR-Ergebnis im Orts- und Fourierraum in Zeile 3 aufgezeigt. Zurück im Skript definiert der dritte Block das Beleuchtungsmuster und erzeugt eine Bilderfolge niedrig aufgelöster LR-Bilder, deren Anzahl an Einzelbildern in den Zeilen 18 und 19 definiert ist. Das Beleuchtungsmuster entspricht hier einem Speckle-Muster mit Strukturen in der Größenordnung der angestrebten HR-Bildpunkte und einer Varianz von 0,5 (Zeile 20), welche den Kontrast im Beleuchtungsmuster steuert. Das Verschieben des Beleuchtungsmusters transformiert höherfrequente Bildinformationen in den Bandpass (rotes Quadrat) des optischen Systems, wobei sich grundsätzlich beliebige Beleuchtungsmuster umsetzen lassen. Die Zeilen 25 bis 30 bestimmen den Fahrweg, mit dem das Beleuchtungsmuster über das GT-Intensitätsbild gerastert wird. Dieser ist Zheng's Publikation [115] entnommen und wird über die Funktion $gseq()$ aufgerufen. Damit erfolgt der Fahrweg in Schlangenlinien, wobei der Startpunkt im Ursprung des Frequenzraums liegt und schrittweise zum Randbereich hinaus läuft [115]. Das zugehörige Octave-/Matlab-Skript ist ebenfalls dem Anhang A.1 angefügt. In Zeile 40 wird die LR-Bilderfolge, unter Berücksichtigung des schrittweise versetzten Beleuchtungsmusters, dem idealen GT-Bild sowie der OTF nach dem Abbildungsmodell in Gleichung 5.1, erzeugt. Dieser Ablauf wird mit der ersten Zeile in Abbildung 5.2 veranschaulicht. Der letzte Block im Skript beinhaltet die Rekonstruktion der SR-Ergebnisse in Orts- und Orts-Frequenzraum anhand der zuvor erzeugten LR-Bilderfolge. In Zeile 47 führt zunächst eine einfache Mittelung aller LR-Bilder zu einer ersten Schätzung eines HR-Intensitätsbildes als Startwert. Diese Erstschätzung wird dann in Zeile 58 im Fourierraum nach Gleichung 5.2 und in Zeile 60 im Ortsraum nach Gleichung 5.3 aktualisiert. Der Wechsel zwischen den mathematischen Berechnungsräumen dient der Einschränkung des Lösungsraums. Im Ortsraum dienen die Intensitätsmessungen der LR-Aufnahmen I_k^{LR} als Objektbeschränkung, um Lösungskonvergenz sicherzustellen [115]. Im Fourierraum dagegen wirkt die begrenzende OTF des optischen Systems als Nebenbedingung [115] für die Suche nach einem SR-Ergebnis [112]. In vorliegendem Skript endet die Aktualisierung der SR-Schätzung nach 10 Iterationen.

5.2 Subpixel-Shift Super-Resolution

Die Subpixel-Shift SR nutzt mehrere LR-Aufnahmen, die im Bruchteil ihrer Pixelgröße zueinander verschoben sind, um damit die Abtastrate bei der digitalen Bildaufnahme einer ortskontinuierlichen Objektszene virtuell zu erhöhen.

Seinen Ursprung findet die Subpixel-Shift (SpS) SR im Orts-Frequenzraum, wo sie durch Tsais und Huangs Publikation [33] im Jahre 1984 erstmals einem größeren Fachpublikum vorgestellt wurde. Während sich SR-Verfahren im Orts-Frequenzraum durch hohe Recheneffizienz auszeichnen, reagieren sie empfindlich auf Fehler im Abbildungsmodell [23, 44] und scheitern an der Umsetzung komplexer Bewegungsmodelle [23, 44]. Wegen diesen Einschränkungen etablierten sich die Methoden im Ortsraum [23] und die SR-Klassifizierung in Abbildung 5.3 beschränkt sich auf diesen Berechnungsraum. Zu den verbreitetsten Methoden zählen die ungleichförmige Interpolation [22, 25, 44, 68, 116], die iterative Rückprojektion (engl. iterative back-projection; IBP) [22, 35, 36, 117], die Projektion auf konvexe Mengen (engl. projection on convex sets; POCS) [22, 36, 118–120] und die regularisierte SR [22, 23, 31, 32, 41–44, 121].

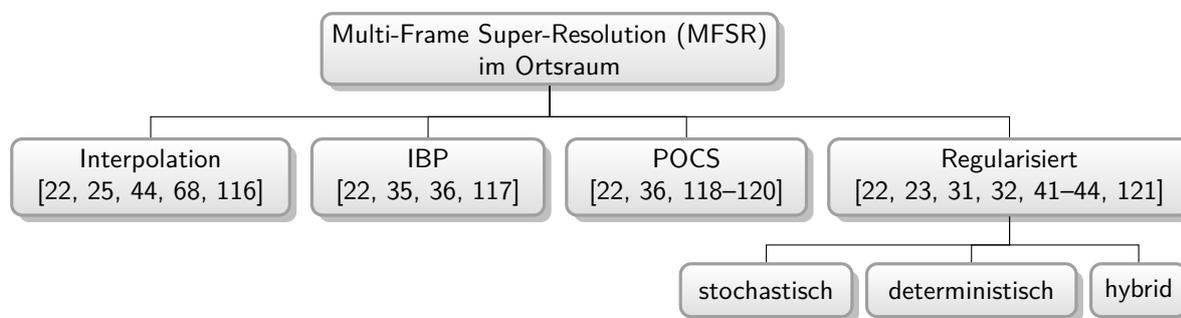


Abbildung 5.3: Klassifikation von MFSR-Methoden im Ortsraum nach ihrer Rekonstruktionstechnik.

Die ungleichförmige Interpolation überträgt Subpixel-verschobene LR-Bilder in ein gemeinsames, hochauflösendes HR-Bild und wirkt dort unter Verwendung von Deblurring-Methoden dem Unschärfeeffekt entgegen, um damit die Bildqualität im SR-Ergebnis zu verbessern [22, 25]. Dennoch führen diese Methoden an kontrastreichen Bildübergängen regelmäßig zu Artefakten wie Unschärfe, blockige Pixel und Aliasing [44]. Dieser Unzulänglichkeit nimmt sich die iterative Rückprojektion an. Dazu schätzt die IBP ein erstes HR-Bild aus Simulationsdaten und vergleicht anschließend dessen Ergebnis mit den LR-Messaufnahmen. Aus deren Differenz zueinander aktualisiert ein Algorithmus die HR-Schätzung iterativ bis eine Abbruchbedingung, beispielsweise in Form eines Fehlerminimums, erreicht ist [22, 36] und gibt das zugehörige Bild als SR-Ergebnis aus. Die IBP im Allgemeinen hat keine eindeutige Lösung und erschwert die Modellierung des optischen Systems sowie das Einbinden von a-priori-Wissen [22]. Einen anderen Ansatz verfolgt das POCS-Verfahren, indem es einen Satz konvexer und beschränkter Mengen definiert, die alle das SR-Ergebnis enthalten [120]. Dabei beschreibt jede Menge eine spezifische Eigenschaft durch einen Satz eindeutig bestimmter Vektoren [22]. Ein Algorithmus sucht den Schnittpunkt aller Mengen iterativ und findet damit das SR-Ergebnis [36]. POCS-Verfahren liefern keine eindeutige Lösung, konvergieren langsam und erfordern einen hohen Rechenaufwand [22]. Von den SR-Methoden im Ortsraum haben sich vor allem die regularisierten Methoden durchgesetzt. Dabei lösen Regularisierungsverfahren schlecht gestellte Probleme mit Hilfe zusätzlicher

Informationen oder Bedingungen (Vorinformationen, a-priori-Informationen), die für bestimmte, eingegrenzte Anwendungsfälle gültig sind [72]. Übertragen auf das SR-Problem beschränkt der Regularisierungsterm, in Form von a-priori-Wissen, die unendlich große Lösungsmenge auf ein eindeutiges Ergebnis, das am Besten zu den vorab definierten Bedingungen passt [31, 62]. Eine solche Bedingung ist beispielsweise der Kompromiss zwischen dem Weichzeichnen und Erhalten von Kanten [23, 31, 44]. Zudem beschleunigt a-priori-Wissen die Konvergenzrate beim Auffinden der SR-Lösung [31, 44]. Regularisierte SR-Methoden weisen eine hohe Flexibilität hinsichtlich der Modellierung des Beobachtungsmodells auf und liefern eine robuste, eindeutige Lösung [22, 31, 36]. Wie der Übersicht in Abbildung 5.3 zu entnehmen ist, unterscheidet die regularisierte SR stochastische, deterministische und hybride Methoden.

Stochastische Ansätze berechnen ein SR-Ergebnis, das mit höchster Wahrscheinlichkeit aus einer gemessenen LR-Bilderfolge resultiert. Zunächst werden die LR-Eingangsbilder, das Rauschen und das GT-Bild als statistische Signale angenommen [34, 36]. Anschließend schätzt ein Algorithmus das SR-Bild, indem er eine Wahrscheinlichkeitsdichtefunktion (engl. probability density function; PDF) auf das ideale GT-Bild hin maximiert [22, 34, 36]. Zu den verbreitetsten Schätzfunktionen zählen der Maximum-Likelihood-Schätzer (ML) als Methode der größten Plausibilität [122] und der Maximum-a-Posteriori-Schätzer (MAP) als Modalwert der a-posteriori-Verteilung bei einer Bayes-Schätzung [44, 122]. Entgegen den ML-Schätzern, fließt bei der MAP-Regularisierung a-priori-Wissen mit ein [22]. Beim Regularisierungsterm nutzen MAP-basierte Regularisierungstechniken in der Regel Markov-Netzwerke (engl. Markov random field; MRF) oder eine Variante der totalen Variation (engl. total variation; TV) [44].

Deterministische Regularisierungsmethoden erlauben, im Gegensatz zur Modellierung mit statistischen Signalen bei der stochastischen Regularisierung, eine eindeutige Beschreibung des Beobachtungsmodells [22]. Dazu schätzen sie die Registrierungsparameter aller LR-Eingangsbilder und lösen das inverse SR-Problem, indem sie dieses mit Hilfe von a-priori-Wissen in ein gut gestelltes Problem umwandeln [22, 44]. Zu den Regularisierungstechniken zählen unter anderem die Methode der iterativ gewichteten kleinsten Quadrate (engl. iteratively reweighted least squares; IRLS), ein gewichtetes Gaußmodell, die Huber-Norm sowie der Tikhonov-Schätzer der kleinsten Quadrate mit seinen L-Normen L_1 -, L_2 - und L_p -Norm [44].

Hybride Regularisierungstechniken sind Kombinationen aus der stochastischen und deterministischen Regularisierung [44]. Eine allgemeine Zusammenfassung regularisierter SR-Verfahren bieten Park et al. und Khattab et al. in ihren Publikationen [22, 44].

5.2.1 Funktionsweise der Subpixel-Shift Super-Resolution

Grundsätzlich bezieht die SpS SR seine Zusatzinformationen aus expliziter Redundanz [81] in Form eines sich ändernden Aliasing-Effekts innerhalb einer LR-Bilderfolge [22, 31, 36, 38, 57, 70, 123]. Aliasing tritt bei der Digitalisierung analoger, kontinuierlicher Signale auf [124], weil es das Whittaker-Kotelnikow-Shannon (WKS)-Abtasttheorem (auch als Shannon-Nyquist-Abtasttheorem bekannt) [125] verletzt. Das Theorem besagt, dass die höchste in dem abzutastenden Signal enthaltene Frequenz maximal der halben Abtastfrequenz entsprechen darf, um es exakt rekonstruieren zu können [126]. Unter Verwendung mehrerer Messungen mit abweichendem Aliasing-Effekt, lassen sich jedoch Ortsfrequenzen identifizieren und rekonstruieren [36], die oberhalb dieser halben Abtastfrequenz liegen. Das heißt, ein Algorithmus rekonstruiert hochfrequente Bildanteile aus den niederfrequenten Komponenten einer LR-Bilderfolge. Vorausset-

zung dafür ist, dass deren benachbarte Einzelbilder im Bruchteil der Pixelgröße relativ zueinander verschoben sind [23, 26, 32, 44, 57, 68, 70, 81, 117, 123]. Im allgemeingültigen Beobachtungsmodell nach Gleichung 3.1 und Abbildung 3.3 beinhaltet der Bewegungsoperator M_k die subpixelgenaue Bildposition sämtlicher LR-Eingangsbilder. Jedes einzelne davon tastet Objektbereiche ab, die in der vorhergehenden Aufnahme zwischen den Bildpunkten lag. Dabei spielt es grundsätzlich keine Rolle, ob die Verschiebung in der Objekt- oder Bildebene stattfindet [22, 42, 43, 46]. Klassischerweise beschreibt die Literatur jedoch den Versatz des Bildsensors, beispielsweise mit Linearstellenheiten [40]. Verallgemeinert erfolgt die SR-Rekonstruktion nach der SpS-Methode in 3 Schritten [22, 42–44, 127]:

1. Bildregistrierung: Sofern nicht durch aktive Aktoren bekannt, schätzt ein zusätzlicher Algorithmus den Versatz aufeinanderfolgender Bilder in Subpixel-Genauigkeit.
2. Fusion: Mit Kenntnis über die exakte Lage der einzelnen LR-Bilder werden deren Bildinformationen positionstreu in ein gemeinsames HR-Gitter überführt.
3. Rekonstruktion: Regularisierungstechniken optimieren das SR-Ergebnis hinsichtlich seiner Bildqualität. Dazu wird a-priori-Wissen genutzt, um Bilddegradationen zu korrigieren und das HR-Bild wiederherzustellen.

Die erfolgreiche Umsetzung am realen Aufbau erfordert, dass das Auflösungsvermögen der abbildenden Optik mindestens so hoch ist wie die angestrebte laterale Ortsauflösung im HR-Bild [40]. Andernfalls ist die Bedingung des Aliasing nicht erfüllt und die SR-Rekonstruktion nicht möglich.

5.2.2 Iterativley Re-weighted Minimization Algorithmus

Der IRWSR-Algorithmus schätzt ein SR-Ergebnis aus einer Folge niedrig auflösender, Subpixelverschobener Einzelbilder. Bei der Rekonstruktion wird eine Bayes-Schätzung mit einer robusten und kantenschonenden Regularisierung kombiniert. Zusätzlich zum SR-Bild schätzt der Algorithmus seine Modellparameter automatisch nach einem iterativen Optimierungsmodell.

Zur SR-Rekonstruktion nach der SpS-Methode wird die iterativ neu gewichtete SR (engl. iteratively re-weighted super-resolution; IRWSR) nach Köhler et al. [42] umgesetzt. Sie kombiniert ein räumlich adaptives Bayes Modell mit einem iterativ gewichteten Gauß'schen Rauschmodell und einer gewichteten BTV-Regularisierung (engl. weighted bilateral total variation; WBTV). Damit gehört es in der Übersicht zu den MFSR-Verfahren in Abbildung 5.3 den hybriden, regularisierten SR-Modellen an. Darüber hinaus schätzt der Algorithmus gleichzeitig die Modellparameter und das SR-Ergebnis in einem iterativen Grob-zu-Fein-Schema. Er kombiniert die robuste und kantenschonende Bildrekonstruktion mit einer automatisierten Parameterschätzung. Neben seiner Flexibilität bei der Modellierung von Bewegungsmodellen ist er rechnerisch effizient und einfach zu implementieren. Sofern nichts anderes angegeben ist, stützen sich sämtliche Beschreibungen in diesem Kapitel auf Köhlers Publikation in [42] sowie auf seine Dissertation in [43].

Das Beobachtungsmodell der IRWSR entspricht der mathematischen Beschreibung in Gleichung 3.1, wobei deren lineare, rauminvariante Unschärfe durch die Punktpreisfunktion der Kamera angenommen wird. Damit gilt für alle LR-Bilder derselbe Unschärfe-Operator und B_k wird mit $k = 1$ zu B . Der Rausch-Operator n_k ist als gewichtete Gauß-Verteilung mit räumlich

variierender Standardabweichung umgesetzt. Entgegen rauminvarianter Rauschmodelle - wie additives weißes Gauß'sches Rauschen (engl. additive white gaussian noise; AWGN), Laplace-Verteilungen oder M-Schätzern - entspricht die gewichtete Gauß-Verteilung mit räumlich variierender Standardabweichung eher dem gemischten Rauschverhalten in realen Bildaufnahmen. Durch Substitution der Degradations-Operatoren Unterabtastung D , Unschärfe B und Bewegung M_k mit H_k , vereinfacht sich Gleichung 3.1 zu:

$$I^{LR_k} = D \cdot B \cdot M_k \cdot I^{GT} + n_k = H_k \cdot I^{GT} + n_k \quad (5.4)$$

Ausgehend von diesem SR-Beobachtungsmodell beschreibt der Satz von Bayes in Gleichung 5.5 die bedingte a-posteriori-Wahrscheinlichkeit $p(I^{SR}|I_k^{LR})$, dass das latente SR-Ergebnis I^{SR} aus einem Satz niedrig auflösender LR-Bilder I_k^{LR} resultiert. Dabei ist $p(I^{SR})$ die a-priori-Wahrscheinlichkeit des latenten SR-Bildes, $p(I^{LR_k})$ die a-priori-Wahrscheinlichkeit der LR-Bilder und $p(I_k^{LR}|I^{SR})$ die bedingte Wahrscheinlichkeit, dass die LR-Bilder I_k^{LR} im SR-Ergebnis I^{SR} enthalten sind.

$$p(I^{SR}|I_k^{LR}) = \frac{p(I_k^{LR}|I^{SR}) \cdot p(I^{SR})}{p(I_k^{LR})} \propto p(I_k^{LR}|I^{SR}) \cdot p(I^{SR}) \quad (5.5)$$

Eine MAP-Schätzung (engl. maximum a-posteriori) rekonstruiert das SR-Bild, welches am besten zur niedrig aufgelösten Bilderfolge I_k^{LR} passt. Dabei berücksichtigt sie a-priori-Informationen über das Beobachtungsmodell, um a-posteriori die Wahrscheinlichkeit über den Erhalt eines einzelnen SR-Ergebnisses zu maximieren. Dies entspricht der Maximierung des Posteriors $p(I^{SR}|I_k^{LR})$:

$$I_t^{SR} = \underset{I^{SR}}{\operatorname{argmax}} \left\{ p(I_k^{LR}|I^{SR}, \beta_t) \cdot p(I^{SR}|\alpha_t) \right\} \quad (5.6)$$

Die hier berücksichtigten a-priori Informationen α_t und β_t beschreiben latente Gewichte, die mit jeder Iteration t automatisch geschätzt werden. Das räumlich adaptive Gewicht α_t optimiert die Rekonstruktion von Diskontinuitäten, indem eine Normgewichtungsfunktion einzelne Bildregionen unterschiedlich stark gewichtet. Um den Einfluss isolierter, verrauschter Pixel zu reduzieren, erfolgt eine kantenschonende 3×3 Medianfilterung der latenten SR-Schätzung. Das Beobachtungsgewicht β_t berücksichtigt Abweichungen zwischen der latenten SR-Schätzung und jeder LR-Bildaufnahme. Zur Maximierung des Posteriors $p(I^{SR}|I_k^{LR})$ minimiert der IRWSR-Algorithmus schließlich das Energieminimierungsproblem F_t in jeder Iteration t :

$$I_t^{SR} = \underset{I^{SR}}{\operatorname{argmin}} F_t(I^{SR}) \quad (5.7)$$

$$F_t(I^{SR}) = (I^{LR} - H I^{SR})^\top B_t (I^{LR} - H I^{SR}) + \lambda_t \|A_t S I^{SR}\|_1 \quad (5.8)$$

Dabei enthalten die Diagonalmatrizen A_t und B_t die beiden Gewichte nach $A_t = \operatorname{diag}(\alpha_t)$ und $B_t = \operatorname{diag}(\beta_t)$. Die lineare Skalierungsmatrix S passt die Bildgröße im Regularisierungsterm $\lambda_t \|A_t S I^{SR}\|_1$ an die des Beobachtungsmodells $(I^{LR} - H I^{SR})^\top B_t (I^{LR} - H I^{SR})$ an, wobei H die Degradationsmechanismen des optischen Systems berücksichtigt. Ein Regularisierungsparameter λ_t balanciert a-priori-Wissen zum Beobachtungsmodell aus und wird ebenfalls mit jeder Iteration t neu geschätzt.

Obwohl das Minimierungsproblem in Gleichung 5.7 konvex ist, führt die Berücksichtigung der latenten Gewichte dazu, dass das gesamte Optimierungsproblem zu unterschiedlichen lokalen Minima konvergieren kann. Aus diesem Grund sucht die iterativ neu gewichtete Minimierung

das globale Minimum nach einem Grob-zu-Fein-Schema. Dessen Lösung fällt mit dem robusten und spärlichen SR-Rekonstruktionsproblem [38, 39, 41, 69, 70, 121] zu folgender nicht-konvexen Energieminimierungsfunktion zusammen:

$$F(I^{SR}) = \sum_{i=1}^{kM} \mathcal{L}([I^{LR} - H I^{SR}]_i) + \lambda \Phi(S I^{SR}) \quad (5.9)$$

Diese beschreibt im ersten Term die Datentreue (engl. data fidelity term) mit einer vorgeschalteten Verlustfunktion \mathcal{L} und im zweiten Term die Regularisierung $\Phi(S I^{SR})$ mit dem vorgeschalteten Regularisierungsparameter λ . Analog zu obiger Beschreibung passt die lineare Skalierungsmatrix S die Bildgröße im zweiten Term an die des ersten Terms an. Die Laufzahlen k und M beschreiben die Anzahl der LR-Bilder innerhalb der Bilderfolge sowie die Pixelanzahl in einem einzelnen LR-Bild mit $M = M_{vert} \cdot M_{horiz}$ Bildpunkten. Die Huber-Verlustfunktion $\mathcal{L}([I^{LR} - H I^{SR}]_i)$ im ersten Term definiert den Begriff der Datentreue und verhält sich entsprechend dem Gewicht β_t . Sie gewichtet kleine Residuen zwischen I_k^{LR} und $H I^{SR}$ zu einem nicht-negativen Schwellwert quadratisch und große Abweichung dazu linear, um robust gegenüber Ausreißern im LR-Bildeingang zu sein.

Dagegen definiert eine gemischte $l1/lp$ -Norm, die sich analog dem Gewicht α_t verhält, den Begriff der Regularisierung im zweiten Term. Sie unterdrückt hier Rauschen und erhält scharfe Kanten bei der Bildwiederherstellung. Dazu steuert der Regularisierungsterm das a-priori-Wissen, um ein sinnvolles SR-Ergebnis zu erhalten [23]. Die Regularisierung senkt ihren Einfluss auf Diskontinuitäten und erhöht ihn zur Glättung homogener Bildregionen, indem sie einen mit α_t gewichteten BTV-Prior einsetzt. Der adaptive Prior folgt einer mittelwertfreien, gewichteten Laplace-Verteilung mit räumlich variierenden Gewichten α_t . Für Diskontinuitäten gilt $\alpha = 0$, um den Einfluss der Regularisierung zu verringern. Dort greift die lp -Norm, wobei $0 < p < 1$ gilt. Dem entgegen werden homogene Regionen mit $\alpha = 1$ gewichtet, um dort eine Glättung zu bewirken. Für diese greift die $l1$ -Norm.

Der Regularisierungsparameter λ balanciert die beiden Terme der Datentreue und Regularisierung zueinander aus. Wird dieser Parameter zu niedrig geschätzt, ist das SR-Ergebnis schlecht konditioniert und durch Restrauschen beeinträchtigt, während im Falle einer Überschätzung die SR-Ergebnisse unscharf sind.

Die automatische Parameterschätzung der Gewichte sowie des Regularisierungsparameters ersetzt die manuelle Parameterabstimmung konventioneller MAP-Schätzer und vereinfacht damit den Einsatz unter realen Bedingungen. Die Optimierungsschleife wird so lange durchgeführt bis eine maximale Anzahl Iterationen erreicht oder eine Konvergenzbedingung erfüllt ist.

Kapitel 6

Methoden zur Bewertung von Auflösungssteigerungen

In folgendem Kapitel werden drei Methoden zur Bewertung von Auflösungssteigerungen vorgestellt. In Abschnitt 6.1 ist die Schräge-Kante-Methode beschrieben, die mit Hilfe eines Kontrastüberganges an einer schrägen Kante den Ortsfrequenzgang im Testbild misst. Ergänzend dazu wird in Abschnitt 6.2 die USAF-1951 Auflösungstesttafel vorgestellt, welche eine einfache Methode zur Bestimmung der erreichten Grenzauflösung sowie zum Erkennen potenzieller Bildfehler im SR-Ergebnis darstellt. Darüber hinaus ist in Abschnitt 6.3 eine relativ bewertende Rauschanalyse nach dem Spitzen-Signal-Rausch-Verhältnis (PSNR) erläutert, das die erreichte Bildqualität im SR-Ergebnis in Relation zu seinem idealen GT-Bild setzt.

Neben dem Anwendungsbereich, der rechnerischen Effizienz und der Handhabbarkeit von SR-Algorithmen, ist allen voran die erreichte Bildqualität im SR-Ergebnis ein Entscheidungskriterium bei der Wahl einer geeigneten SR-Methode. Üblicherweise bewertet die Literatur ihre SR-Ergebnisse relativ zu seinem GT-Bild nach einer Rauschanalyse. Wegen der geringen Bildauflösung des PMD-Sensors der verwendeten Kamera (vgl. dazu Tabelle 2.1) existieren jedoch keine hochauflösenden GT-Bilder, weshalb diese Methode der Qualitätsbeurteilung ausscheidet. Um die Abhängigkeit von einem idealen Referenzbild zu umgehen, wird in der vorliegenden Arbeit die Bewertung der Bildqualität im SR-Ergebnis nach dem erreichten optischen Auflösungsvermögen in Linienpaaren pro Pixel (LP/px) beziehungsweise in Linienpaaren pro Millimeter (LP/mm) verfolgt. Die dazu benötigten Testbilder zeigt die nachfolgende Abbildung:

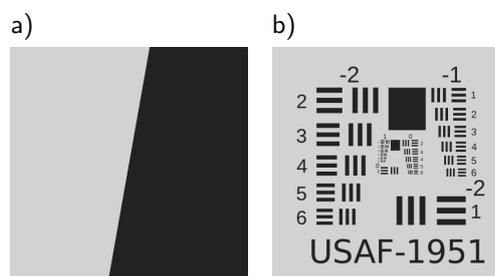


Abbildung 6.1: Das Schräge-Kante-Testbild in a) separiert zwei Bildbereiche unterschiedlicher Helligkeit, hier unter einem Winkel von 10° . Die USAF-1951 Auflösungstesttafel in b) dagegen besteht aus mehreren Gruppen schwarz-weißer, horizontal und vertikal orientierter Balkenpaare steigender Ortsfrequenz.

6.1 Ortsfrequenzganganalyse nach der Schräge-Kante-Methode

Zur Bestimmung des erreichten optischen Auflösungsvermögens im SR-Ergebnis misst das Schräge-Kante-Verfahren (engl. slanted edge; SE) dessen Ortsfrequenzgang (engl. spatial frequency response; SFR). Die SFR ist eine Annäherung an die Modulationstransferfunktion (engl. modulation transfer function; MTF) [128], die typischerweise in der Optik herangezogen wird, um die Abbildungsqualität von Objektiven zu quantifizieren [129]. Die MTF ist der Betrag der OTF [130] und beschreibt das Verhältnis von Bildkontrast zu Objektkontrast als Zahlenwert zwischen 0 und 1 oder als Prozentangabe zwischen 0 % und 100 % [129].

Das SE-Testbild separiert zwei Bildbereiche unterschiedlicher Helligkeit unter einer schrägen Kante, die idealerweise zwischen 3° und 30° gedreht ist [131]. In vorliegender Arbeit beträgt der Rotationswinkel der schrägen Kante sowohl für die Simulationen als auch für die Messungen 10° . Mit vertikaler Orientierung der schrägen Kante wird die horizontale Frequenzantwort gemessen, bei horizontaler Orientierung entsprechend die vertikale Frequenzantwort.

Die Schräge-Kante-Methode ist Bestandteil der ISO-Norm 12233 [132] und verfolgt das Ziel der SFR-Messung mit nur einer Aufnahme. Damit erweist sie sich als schnelle und robuste Methode einer standardisierten SFR-Messung, die eine hohe Vergleichbarkeit gewährleistet. Zur Auswertung dient das Open Source Tool *MTF mapper* in der Version 0.7.11 von van den Bergh [133]. Dessen Ausgabe beschreibt eine SFR-Kurve, die den Kontrast über der Ortsfrequenz aufträgt. Zum direkten Vergleich dient der MTF_{50} -Wert, welcher die übertragene Ortsfrequenz in LP/px bei einem Kontrastverhältnis von 50 % angibt. Die Standardabweichung beim relativen Fehler liegt für den MTF_{50} -Wert unterhalb 1 % [134].

Der SFR-Algorithmus identifiziert Kanten mit Hilfe des Sobel-Filters, einem diskreten Differenzierungsoperator. Die Filterung der schrägen Kante liefert für jede Zeile eine Kantenspreizfunktion (engl. edge spread function; ESF), dessen Maximum dem Umkehrpunkt des Kontrastüberganges im SE-Testbild entspricht. Bei vertikaler Kantenorientierung wandert die Phase der Kantenspreizfunktionen [135] - entsprechend der Hell-Dunkel-Grenze im SE-Testbild - zeilenweise entlang der horizontalen Richtung. Die Einzelergebnisse aller zeilenweisen Kantenspreizfunktionen werden in eine gemeinsame, hochauflösende Kantenspreizfunktion überführt [136]. Die numerische Differentiation dieser hochauflösenden ESF gibt die Linienspreizfunktion (engl. line spread function; LSF) aus. An dieser wird eine Fast-Fourier-Transformation durchgeführt, deren normierte Amplitude die MTF ergibt [137]. Die Ausgabe erfolgt als SFR-Kurve, welche den Kontrast über die Ortsfrequenz aufträgt. Für einen idealen Sensor mit rechteckförmiger Pixelgeometrie entspricht die Kurve einer sinc-Funktion [138, 139] nach Gleichung 6.1 mit der Pixelbreite s in horizontaler Richtung. Den zugehörigen Kurvenverlauf zeigt Abbildung 6.2.

$$MTF_{Sensor} = \|sinc(s)\| = \left\| \frac{\sin(\pi \cdot s)}{\pi \cdot s} \right\| \quad (6.1)$$

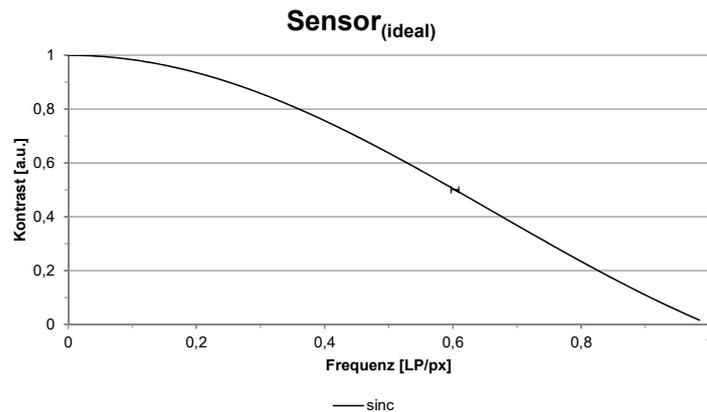


Abbildung 6.2: Der Ortsfrequenzgang (engl. spatial frequency response; SFR) eines idealen Bildsensors mit endlicher Pixelbreite folgt einer sinc-Funktion. Der Fehlerbalken visualisiert die Standardabweichung beim relativen Fehler bei der Bestimmung des MTF₅₀-Wertes von 1 %.

6.2 USAF-1951 Auflösungstesttafel

Im Jahre 1951 führte die United States Air Force die einfach zu interpretierende USAF-1951 Auflösungstesttafel [140] ein, mit der das Auflösungsvermögen von optischen System bewertet werden kann. Es besteht aus mehreren Gruppen mit sechs Elementen, die jeweils drei horizontale und vertikale schwarze Balken enthalten, wie bereits in Abbildung 6.1 b) gezeigt. Dabei beginnen die Gruppennumern bei -2 an zu zählen und geben gleichzeitig den Potenzexponenten auf Basis der Zweierpotenz ($2^{\text{Gruppennummer}}$) wie auch den Startwert der Ortsauflösung für eben diese Gruppe in LP/mm vor. Die Elementzahlen innerhalb einer Gruppe zählen von 1 bis 6. Ausgehend vom ersten Element in einer Gruppe, nimmt die Ortsfrequenz für jedes weitere Element mit der sechsten Wurzel von Zwei: $\sqrt[6]{2}$ zu [141]. Nachfolgende Gleichung fasst diesen Zusammenhang zusammen:

$$\text{Auflösungsvermögen} = 2^{\text{Gruppennummer} + \frac{(\text{Element} - 1)}{6}} \quad (6.2)$$

Die Gruppe mit den Elementen, die nicht mehr klar voneinander getrennt werden können, beschreibt die Grenzauflösung des optischen Systems. Als Beispiel ergibt sich für das dritte Element aus der Gruppe $+2$ ein Auflösungsvermögen von $2^{(2 + \frac{(3-1)}{6})} = 5,04 \text{ LP/mm}$.

Aufgrund seiner geometrisch komplexeren Struktur, verglichen mit dem Schräge-Kante-Testbild, ermöglicht die USAF-1951-Testtafel zudem das Erkennen potenzieller Bildfehler im SR-Ergebnis.

6.3 PSNR Rauschanalyse

Wie eingangs erwähnt, bewertet die Literatur SR-Techniken überwiegend unter Zuhilfenahme von Rauschanalysen. Die gängigsten Bewertungszahlen sind dabei der mittlere absolute Fehler (mean absolute error; MAE), das Signal-Rausch-Verhältnis (engl. signal-to-noise ratio; SNR) sowie der mittlere quadratische Fehler (engl. mean square error; MSE) mit seinen Abwandlungen als Spitzen-Signal-Rausch-Verhältnis (engl. peak signal-to-noise ratio; PSNR) und Wurzel der mittleren Fehlerquadratsumme (engl. root mean square error; RMSE). Allesamt quantifizieren sie die Abweichung vom SR-Ergebnis zu seinem fehlerfreien GT-Bild pixelweise.

Aufgrund ihrer herausragenden Stellung bei der Bewertung von SR-Methoden, erfolgt für die finalen SR-Simulationsergebnisse in Abschnitt 9.4 zusätzlich zur Bestimmung des optischen Auflösungsvermögens eine Rauschanalyse nach dem Spitzen-Signal-Rausch-Verhältnis (PSNR). Das PSNR ist eine Variante des mittleren quadratischen Fehlers (engl. mean square error; MSE), wie bereits in Gleichung 4.2 (exemplarisch für eine Folge niedrig aufgelöster Bildfolgen) aufgezeigt. Das PSNR und der MSE stehen wie folgt in Zusammenhang [142]:

$$PSNR = 10 \log_{10} \frac{I_{max}^2}{MSE} = 10 \log_{10} \frac{I_{max}^2}{\frac{1}{M} \sum_i (I_i^{SR} - I_i^{GT})^2} \quad (6.3)$$

Wobei M der Pixelanzahl in den Vergleichsbildern I_i^{SR} und I_i^{GT} entspricht und I_{max}^2 den maximal erreichbaren Helligkeitswert im Bild und damit den Dynamikbereich des Bildformats beschreibt. Die Einheit des PSNR ist das Dezibel (dB). Mit geringeren Bildfehlern nimmt der PSNR-Wert zu und damit im Falle der SR auch die Bildqualität im SR-Ergebnis [143].

Kapitel 7

Simulationen

In Kapitel 7 werden die Simulationsergebnisse der PCHIP-Interpolation sowie der drei vorgestellten SR-Verfahren beschrieben:

- PCHIP: Interpolationsmethode nach der Beschreibung in Abschnitt 3.1.
- ENet: Lernbasierte SISR nach dem EnhanceNet-Algorithmus aus Abschnitt 4.2.
- FP: Inkohärente Fourier Ptychographie als MFSR-Ansatz der strukturierten Beleuchtung nach dem Algorithmus aus Abschnitt 5.1.2.
- SpS: Subpixel-Shift MFSR nach dem IRWSR-Algorithmus aus Abschnitt 5.2.2.

Sämtliche Simulationen basieren auf synthetisch erzeugten SE- und USAF-1951-Testbildern, wie sie bereits in Abbildung 6.1 aufgezeigt sind. Während deren Erzeugung in Abschnitt 7.1 beschrieben ist, werden in dem daran anknüpfenden Abschnitt 7.2 die Implementierungen der einzelnen oben genannten Verfahren erläutert und wo erforderlich dessen Simulationsparameter definiert. Im Anschluss daran folgt in Abschnitt 7.3 die Darstellung der Simulationsergebnisse sowie die Quantifizierung der erreichten Bildqualität nach den Methoden der Ortsfrequenzganganalyse und dem USAF-1951-Testbild. In dem Unterabschnitt 7.3.1 werden die Simulationsergebnisse der einzelnen SR-Techniken diskutiert und anschließend deren Eignung hinsichtlich der Umsetzung einer generischen SR bewertet. Abschließend sind in dem Unterabschnitt 7.3.2 die wesentlichen Erkenntnisse zu einem Zwischenfazit zusammengefasst, woraus diejenige Methode abgeleitet wird, die sich am besten für die generische Auflösungssteigerung von PMD-Kameras eignet.

7.1 Erzeugung synthetischer Testbilder

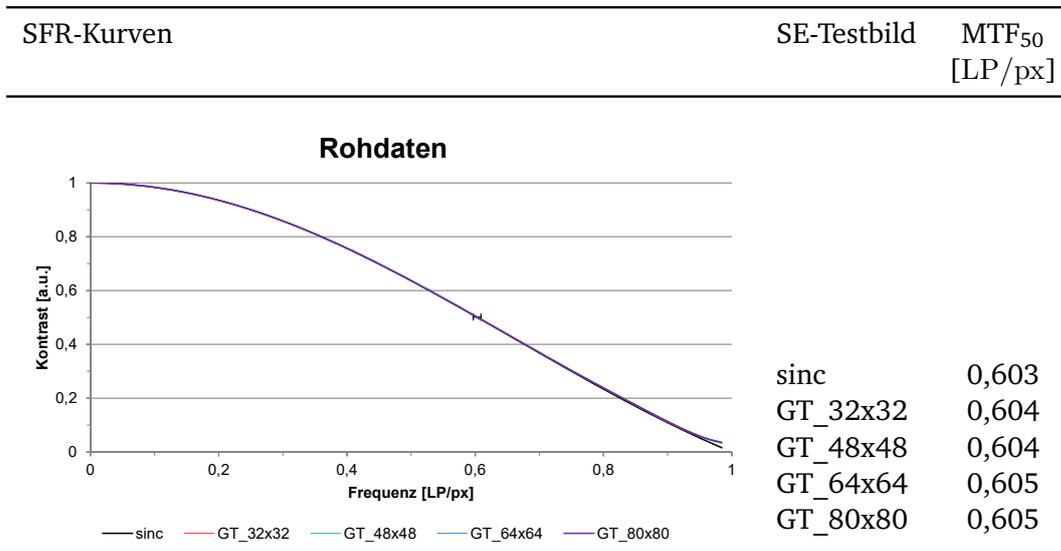
Sämtliche GT-Bilder entstammen synthetisch erzeugten Ursprungsbildern in einer Größe von 2400 x 2400 Bildpunkten mit den diskreten Helligkeitswerten 34 und 210. Ausgehend davon reduziert ein block-mittelndes Interpolationsverfahren das Ursprungsbild auf die in Tabelle 7.2 aufgeführten Bildauflösungen.

Für das SE-Testbild sind die Zielauflösungen 32 x 32, 48 x 48, 64 x 64 und 80 x 80 Bildpunkte groß. Damit entsprechen sie ganzzahligen Vielfachen der niedrig auflösenden LR-Bildgröße von 16 x 16 Bildpunkten. Diese Ausgangsgröße ergibt sich aus der zur Verfügung stehenden Zeilenanzahl der verwendeten PMD-Kamera bei einem Seitenverhältnis von 1:1. Diese Annahmen erfüllen die Voraussetzung, die in diesem Kapitel erlangten Simulationsergebnisse mit den

Messergebnissen aus dem experimentellen Aufbau im später folgenden Kapitel 9 vergleichen zu können.

Eine Ortsfrequenzganganalyse überprüft die Gültigkeit der als ideal angenommenen SE-GT-Bilder. Dazu werden in Tabelle 7.1 die Ortsfrequenzgänge der GT-Bilder aller vier Zielauflösungen mit jenem aus der Theorie zu erwartenden Kurvenverlauf für den auflösungsbegrenzenden PMD-Sensor gegenübergestellt. Dieser folgt nach Abschnitt 6.1, Gleichung 6.1 einer sinc-Funktion. Zusätzlich gibt die Tabelle den jeweils erreichten MTF_{50} -Wert für jedes Testbild an. Die maximale Abweichung bei den MTF_{50} -Werten von 0,002 LP/px zur sinc-Funktion liegt mit 0,3 % innerhalb der Messabweichung von 1 % und wird vernachlässigt.

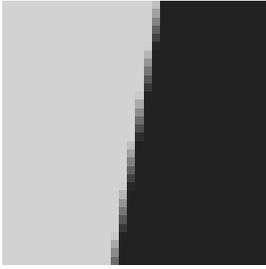
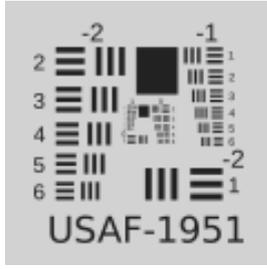
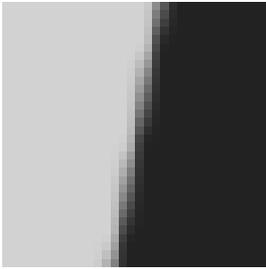
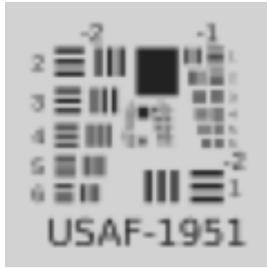
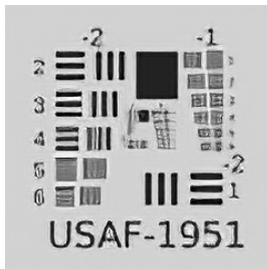
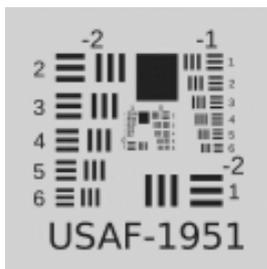
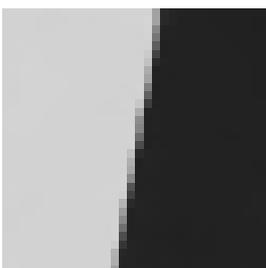
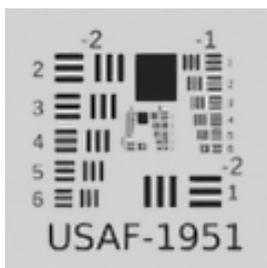
Tabelle 7.1: Der Ortsfrequenzgang aller SE-GT-Bilder stimmt mit dem aus der Theorie zu erwartenden Kurvenverlauf nach der sinc-Funktion überein.



Um die komplexeren Objektstrukturen im USAF-1951-Testbild sinnvoll bewerten zu können, liegt dort die Ausgangsgröße der LR-Bilder bei 60 x 60 Bildpunkten. Analog zum SE-Testbild ergeben sich aus den ganzzahligen Skalierungsfaktoren vier Zielauflösungen in den Größen 120 x 120, 180 x 180, 240 x 240 und 300 x 300 Bildpunkte.

In Tabelle 7.2 sind den idealen GT-Bildern beider Testvarianten in der ersten Zeile mit jeder weiteren Zeile die Simulationsergebnisse der einzelnen SR-Methoden gegenübergestellt. Dabei sind in Fettdruck jene Bildauflösungen hervorgehoben, die dem abgebildeten SR-Ergebnis entsprechen. Bei den SR-Methoden sind in Klammern dahinter der umgesetzte Skalierungsfaktor sowie gegebenenfalls die für das SR-Ergebnis berücksichtigte Anzahl der LR-Eingangsbilder angeben.

Tabelle 7.2: Die SE- und USAF-1951-Testbilder sind nach ihren GT-Bildern sowie den Simulationsergebnissen der einzelnen SR-Methoden (PCHIP, ENET, FP und SpS mit NN-Interpolation) aufgelistet. In Fettdruck sind der Skalierungsfaktor sowie die LR-Bildanzahl hervorgehoben.

SE	Bildauflösung	USAF-1951	Bildauflösung
	<u>GT</u> 32 x 32 48 x 48 64 x 64 80 x 80		<u>GT</u> 120 x 120 180 x 180 240 x 240 300 x 300
	<u>PCHIP</u> 32 x 32 (2x) 48 x 48 (3x) 64 x 64 (4x) 80 x 80 (5x)		<u>PCHIP</u> 120 x 120 (2x) 180 x 180 (3x) 240 x 240 (4x) 300 x 300 (5x)
	<u>ENet</u> - 64 x 64 (4x) -		<u>ENet</u> - 240 x 240 (4x) -
	<u>FP</u> 32 x 32 (2x_N49) 48 x 48 (3x) 64 x 64 (4x) -		<u>FP</u> 120 x 120 (2x_N49) 180 x 180 (3x) 240 x 240 (4x) -
	<u>SpS (NN)</u> 32 x 32 (2x_N4) 48 x 48 (3x) 64 x 64 (4x) 80 x 80 (5x)		<u>SpS (NN)</u> 120 x 120 (2x_N4) 180 x 180 (3x) 240 x 240 (4x) 300 x 300 (5x)

7.2 Umsetzung der Simulationen

Nachfolgend ist die Vorgehensweise bei der Simulation der SR-Methoden erläutert und für die beiden MFSR-Ansätze der inkohärenten FP sowie der Subpixel-Shift SR sind zusätzlich die berücksichtigten Simulationsparameter definiert.

PCHIP-Interpolation

Analog zur Erzeugung der GT-Bilder in Abschnitt 7.1 rechnet ein block-mittelndes Interpolationsverfahren das niedrig auflösende Einzelbild aus einem hochauflösenden Ursprungsbild herunter, welches anschließend als LR-Eingangsbild bei der Umsetzung der *PCHIP-Interpolation* dient. Für das SE-Testbild liegt die LR-Bildgröße bei 16 x 16 Bildpunkten, für das USAF-1951-Testbild dagegen bei 60 x 60 Bildpunkten. Die LR-Versionen beider Testbilder entsprechen einer idealisierten Betrachtung, deren optisches Auflösungsvermögen lediglich durch die Bildauflösung limitiert ist. Somit berücksichtigt die Simulation der PCHIP-Methode hier keine weiteren Bilddegradationen, wie beispielsweise Rauschen oder Unschärfe. Sie vergrößert ein einzelnes LR-Bild mit ganzzahligen Skalierungsfaktoren zwischen 2- und 5-facher Auflösungssteigerung, nach der Beschreibung in Abschnitt 3.1.

EnhanceNet SISR

Die Simulation der lernbasierten SISR nach dem *EnhanceNet* greift auf die vortrainierte Referenz-Implementierung *ENet-PAT* in der Version 1.01 von Sajjadi et al. [82, 109] zurück. Entsprechend der Beschreibung in Abschnitt 4.2 vergrößert sie ein einzelnes LR-Eingangsbild um ausschließlich den Skalierungsfaktor 4. Außerdem sind bei dieser SISR-Simulation keine weiteren Bilddegradationen berücksichtigt.

Inkohärente Fourier Ptychographie

Zur Simulation der inkohärenten Fourier Ptychographie wird das Octave-/Matlab-Skript aus Anhang A.1 ausgeführt, wie zuvor in Abschnitt 5.1.2 beschrieben. Für jede Simulation sind in dem Skript vorab der Skalierungsfaktor F_{Aufl} , das zugrundeliegende GT-Bild *objectIntensity* sowie die Anzahl der erwünschten LR-Eingangsbilder *patternNum* als Quadratzahl ungerader Ganzzahlen *arraySize* zu definieren.

In Tabelle 7.3 sind die Simulationsparameter zu den hier gezeigten Simulationen zusammengefasst. Unabhängig vom Vergrößerungsfaktor nimmt die Anzahl der LR-Eingangsbilder innerhalb einer Bildfolge quadratisch zu. Abgesehen von einer intensitätsmodulierten Beleuchtung sind den Eingangsbildern keine weiteren Bilddegradationen auferlegt.

Tabelle 7.3: Simulationsparameter zur inkohärenten FP: Neben den GT-Bildern erwartet das Skript Angaben zum Skalierungsfaktor F_{Aufl} sowie zur Anzahl der niedrig aufgelösten Eingangsbilder.

F_{Aufl}	2	3	4
SE-GT	GT_32x32	GT_48x48	GT_64x64
USAF-GT	GT_120x120	GT_180x180	GT_240x240
Anzahl LR-Bilder	9, 25, 49, 81	9, 25, 49, 81, 121, 169	9, 25, 49, 81, 121, 169, 225

Subpixel-Shift SR nach dem IRWSR-Algorithmus mit NN-Interpolation

Die Simulation der Subpixel-Shift SR erfolgt mit der Multi-Frame Super-Resolution Toolbox von Köhler et al. in der Version 1.6.1 [144]. Der darin umgesetzte IRWSR-Algorithmus verrechnet eine Folge niedrig auflösender Eingangsbilder, gemäß der Beschreibung in Abschnitt 5.2.2.

Die Simulation der IRWSR-Rekonstruktion erfordert jedoch zunächst eine LR-Bilderfolge, deren Einzelbilder im Bruchteil ihrer Pixelgröße zueinander verschoben sind. Diese wird mit dem kommentierten Octave-/Matlab-Skript, der dem Anhang A.2 exemplarisch für die 2-fache Auflösungssteigerung angefügt ist, erzeugt. In den Zeilen 18 bis 21 werden sämtliche Einzelbilder einer Bilderfolge in der angestrebten SR-Bildauflösung aus einem höher auflösenden GT-Bild extrahiert. Dabei sind die aufeinanderfolgenden Einzelbilder um ganzzahlige Pixel relativ zueinander verschoben. Der erwünschte Subpixel-Versatz in den LR-Bildern resultiert aus dem nachgelagerten Downsampling in Zeile 27. Nach Köhler [43] beschreiben Dirac-Impulse den diskreten Downsampling-Operator und sind als Nearest Neighbor (NN)-Interpolation zu implementieren. Zusätzlich zum Downsampling nach der NN-Interpolation werden mit dem Skript auch die anderen Bilddegradationen des allgemeinen Beobachtungsmodells aus Abschnitt 3.3 umgesetzt. In den Zeilen 26, 28 und 29 sind die Unschärfe- und Rauschfilter definiert, die übergreifend für alle Skalierungsfaktoren und LR-Bilderfolgen gleichermaßen gelten. Unschärfe bezeichnet hier den weichzeichnenden Einfluss der Optik und ist als isotroper Gauß'scher Glättungsfilter mit zweidimensionalem Filterkern realisiert. Dessen Standardabweichung von $\sigma_B = 0,5$ steuert den Grad der Unschärfe und übernimmt hier den von der Literatur [42, 43] vorgeschlagenen Wert. Außerdem steuert die Standardabweichung $\sigma_{AWGN} = 0,01$ das additive weiße Gauß'sche Rauschen (AWGN), welches dem LR-Bild Zufallszahlen innerhalb der Standardabweichung einer Normalverteilung überlagert. Der Wert ist so gewählt, dass der visuelle Rauscheindruck des synthetisch erzeugten LR-Bildes dem eines gemessenen Intensitätsbildes der PMD-Kamera gleicht. Als weiterer Rauscheinfluss ist den LR-Bildern Poisson-verteiltes Schrotrauschen aufgeprägt. Die zur Simulation berücksichtigten Parameter sind in Tabelle 7.4 zusammengefasst.

Zur SR-Rekonstruktion erwartet der IRWSR-Algorithmus, zusätzlich zur LR-Bilderfolge und dem Skalierungsfaktor, Angaben zur PSF-Breite des optischen Systems sowie zum Subpixel-Versatz der aufeinanderfolgenden LR-Bilder. Beide Angaben sind in Tabelle 7.5 den LR-Bilderfolgen und ihren Skalierungsfaktoren zugeordnet. Dabei wurden die PSF-Breiten für jede Simulation empirisch ermittelt: Es sind jeweils diejenigen PSF-Breiten eingetragen, bei deren SR-Ergebnis die SFR-Kurve am wenigsten von der des zugehörigen GT-Bildes abweicht. Damit beschreibt die PSF-Breite die Unschärfe des optischen Gesamtsystems unter Berücksichtigung all seiner Degradationen. Die Subpixel-Verschiebungen sind mit dem Erzeugen der synthetischen LR-Bilderfolge bekannt und in Bruchteilen der Pixelgröße angegeben.

Tabelle 7.4: Parametersatz zur Erstellung der LR-Bilderfolge für die Simulation der SpS SR mit NN-Interpolation.

Anzahl LR-Bilder	4	9	16	25
LR-Bildauflösung	16 x 16			
F_{Aufl}	2	3	4	5
σ_B	0,5			
σ_{AWGN}	0,01			
Poisson-Rauschen	ja			

Tabelle 7.5: Parametersatz zur SR-Rekonstruktion nach dem IRWSR-Algorithmus aus der synthetisch erzeugten LR-Bilderfolge mit NN-Interpolation.

Anzahl LR-Bilder	4	9	16	25
LR-Bildauflösung	16 x 16			
F_{Aufl}	2	3	4	5
PSF-Breite	0,30	0,16	0,15	0,12
Subpixel-Versatz [Px]	1/2	1/3	1/4	1/5

7.3 Simulationsergebnisse auf Basis der synthetischen Testbilder

In diesem Abschnitt werden die Ortsfrequenzganganalysen der einzelnen SR-Methoden für ihre jeweils umgesetzten Skalierungsfaktoren zusammengefasst. Dazu sind die Ortsfrequenzgänge der SR-Ergebnisse dem theoretisch zu erwartenden sinc-Kurvenverlauf in Diagrammen gegenübergestellt. Zudem geben die Tabellen die erreichten MTF_{50} -Werte der abgebildeten SFR-Kurven an.

PCHIP-Interpolation

Das Diagramm in Tabelle 7.6 zeigt die Ortsfrequenzgänge der Ergebnisse bei 2-, 3-, 4- und 5-facher Auflösungssteigerung nach der PCHIP-Interpolation. Mit steigendem Skalierungsfaktor nimmt die Abweichung zum theoretischen Auflösungsvermögen nach der sinc-Funktion zu. Analog dazu verhalten sich die MTF_{50} -Werte.

EnhanceNet SISR

In Tabelle 7.7 ist der Ortsfrequenzgang für das SR-Ergebnis bei 4-facher Auflösungssteigerung nach dem ENet-PAT-Algorithmus jenen der sinc-Funktion und PCHIP-Interpolation gegenübergestellt. Bis etwa 0,35 LP/px reicht der SFR-Kurvenverlauf des ENet-SR-Ergebnisses an den der sinc-Funktion heran, bevor der Bildkontrast für höhere Ortsfrequenzen einbricht.

Tabelle 7.6: Ortsfrequenzganganalyse für die PCHIP-Simulation bei 2-, 3-, 4- und 5-facher Auflösungssteigerung.

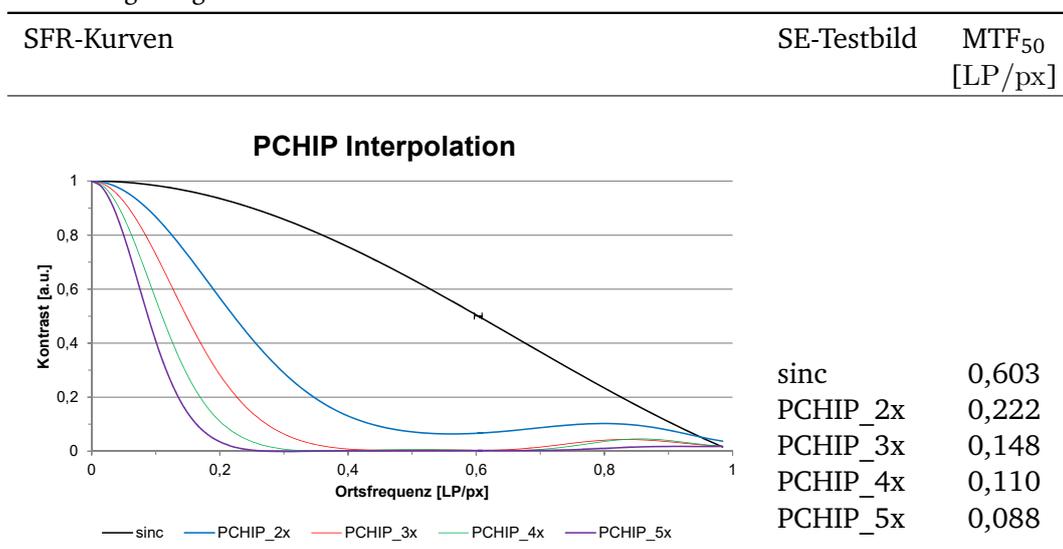
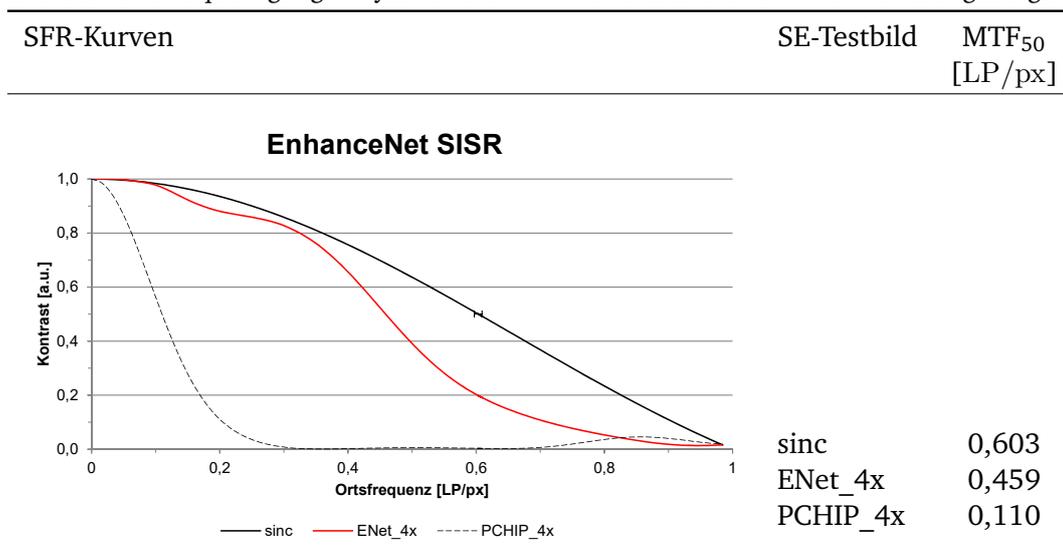


Tabelle 7.7: Ortsfrequenzganganalyse für die EnhanceNet SISR bei 4-facher Auflösungssteigerung.



Inkohärente Fourier Ptychographie

Bei der inkohärenten FP sind die Simulationsergebnisse bei 2-, 3- und 4-facher Auflösungssteigerung separat voneinander in den Tabellen 7.8 bis 7.10 zusammengetragen. Darin werden jeweils die Ortsfrequenzgänge und MTF₅₀-Werte bei steigender Anzahl an LR-Eingangsbildern mit denjenigen der zugehörigen GT-Bilder (sinc-Funktion) und PCHIP-Interpolationen verglichen.

Tabelle 7.8: Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 2-facher Auflösungssteigerung.

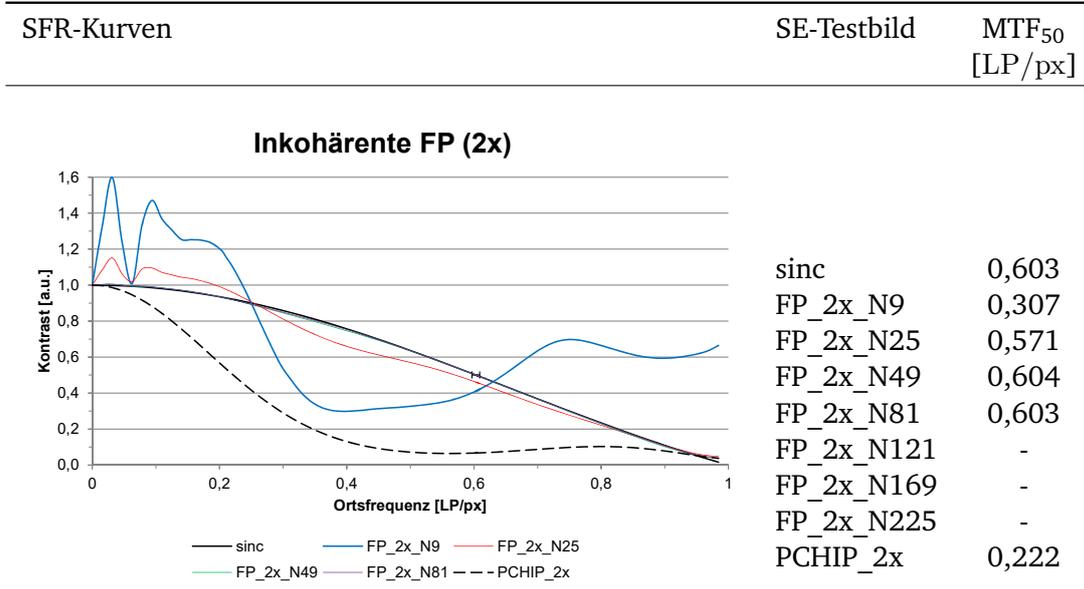


Tabelle 7.9: Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 3-facher Auflösungssteigerung.

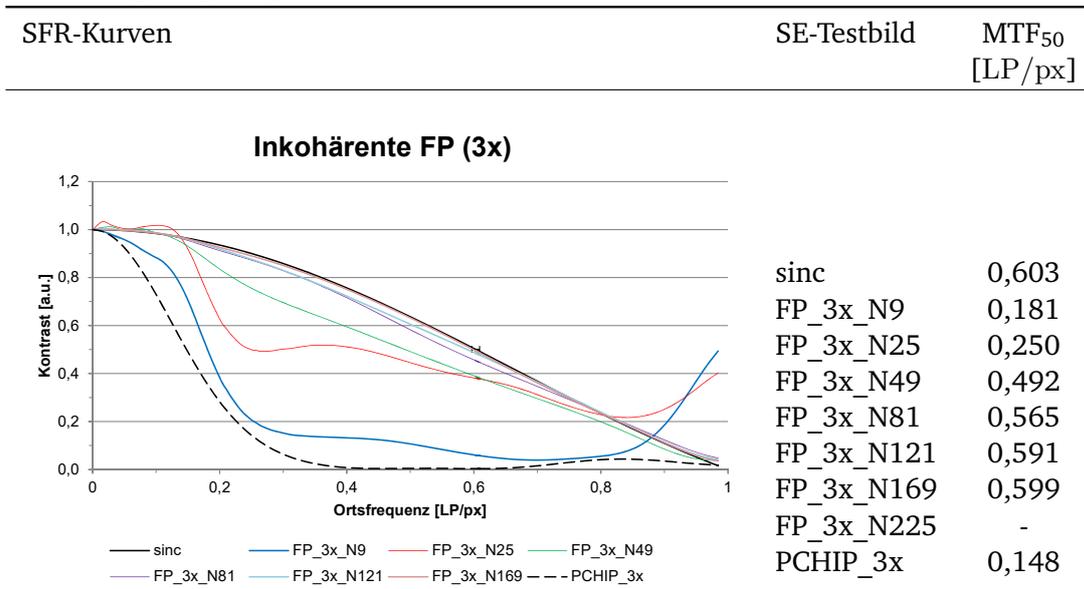
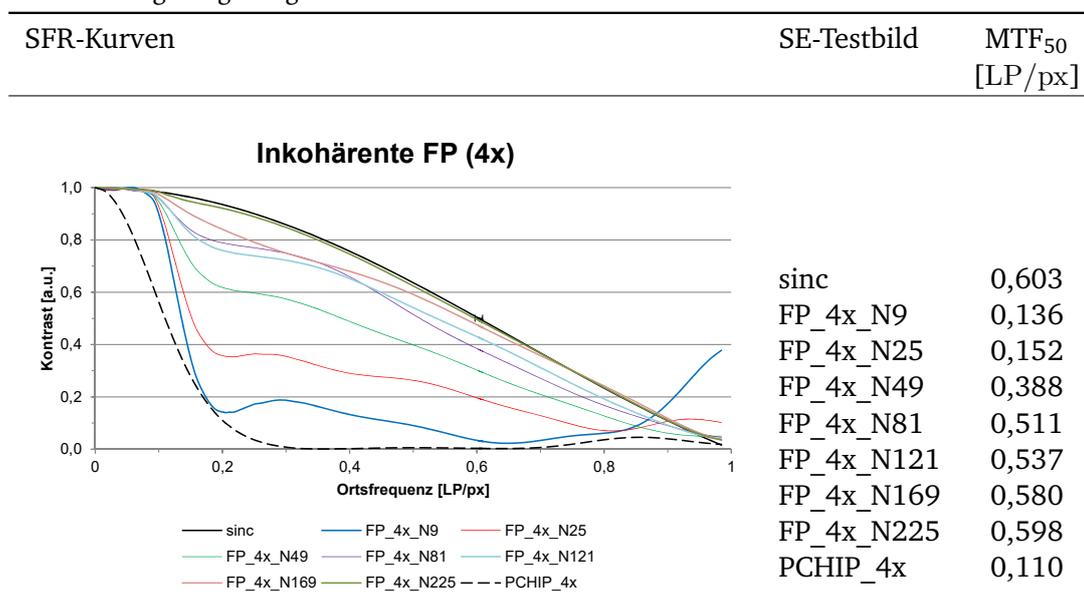


Tabelle 7.10: Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 4-facher Auflösungssteigerung.



Ergänzend zur Ortsfrequenzganganalyse wird in den Tabellen 7.11 und 7.12 die Abhängigkeit der Bildqualität von der Anzahl an LR-Eingangsbildern exemplarisch für die 2- und 4-fache Auflösungssteigerung visualisiert. Die oberen Reihen zeigen die SR-Rekonstruktionen des SE-Testbildes, die unteren diejenigen des USAF-1951-Testbildes, wobei die Anzahl der Eingangsbilder quadratisch zunimmt. Je weniger LR-Bilder zur Verfügung stehen, desto stärker treten Artefakte im SR-Ergebnisbild in Erscheinung.

Tabelle 7.11: Abhängigkeit der Bildqualität im SR-Ergebnis bei 2-facher Auflösungssteigerung von der Anzahl an LR-Eingangsbildern (v.l.n.r.): N=9, N=25, N=49 und N=81.

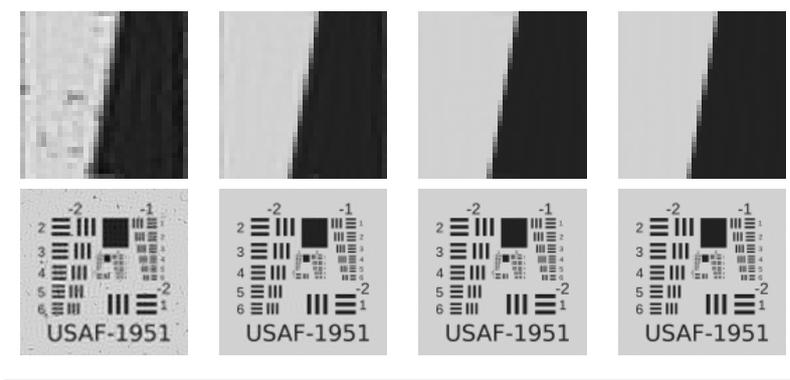
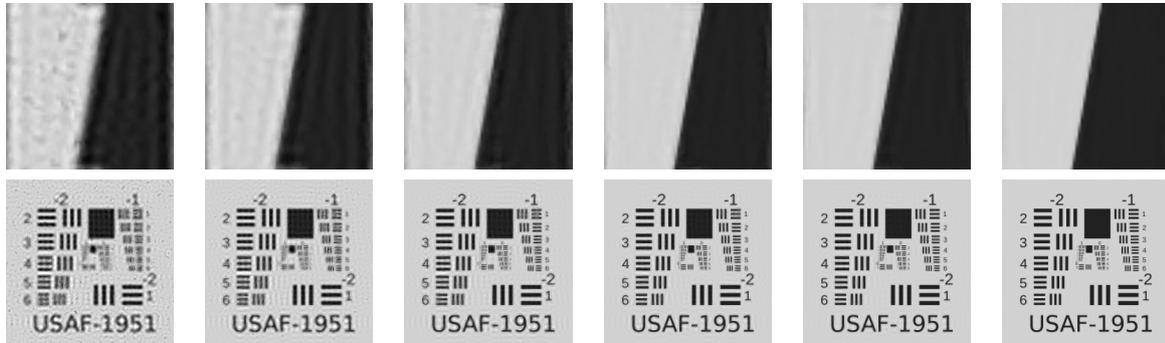


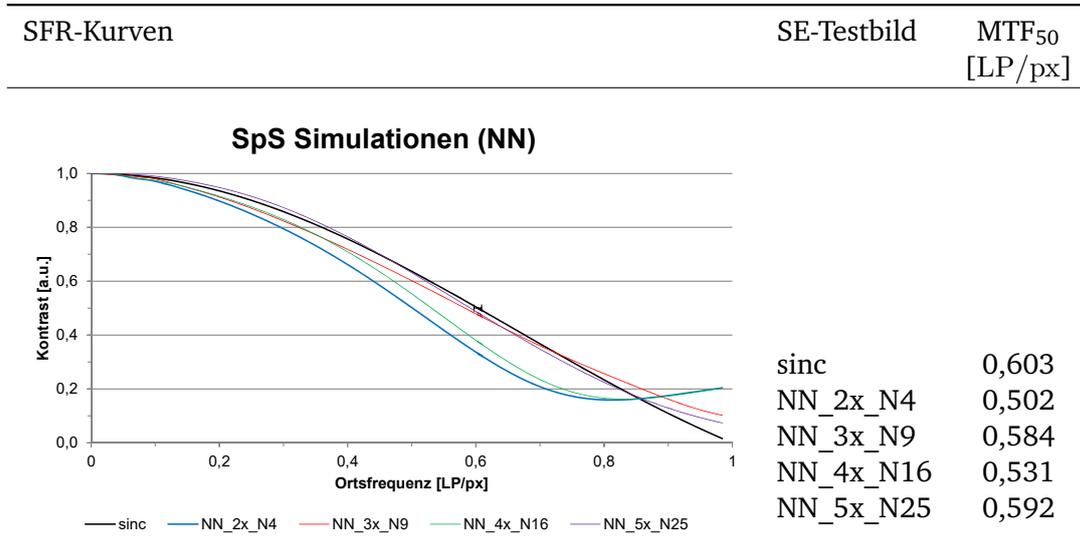
Tabelle 7.12: Abhängigkeit der Bildqualität im SR-Ergebnis bei 4-facher Auflösungssteigerung von der Anzahl an LR-Eingangsbildern (v.l.n.r.): N=9, N=25, N=49, N=81, N=121 und N=169.



Subpixel-Shift SR nach dem IRWSR-Algorithmus mit NN-Interpolation

In Tabelle 7.13 sind die Ortsfrequenzgänge der SR-Ergebnisse nach der SpS-Methode bei 2- bis 5-facher Auflösungssteigerung sowie der Ortsfrequenzgang der sinc-Funktion vereint. Dabei nimmt bei der SpS-Methode die Anzahl der LR-Eingangsbilder quadratisch mit dem Skalierungsfaktor zu. Wie in Abschnitt 7.2 beschrieben, ist hier der Downsampling-Operator zur Erzeugung der synthetischen LR-Einzelbilder mit einer NN-Interpolation umgesetzt.

Tabelle 7.13: Ortsfrequenzganganalyse für die Subpixel-Shift SR nach dem IRWSR-Algorithmus mit NN-Interpolation bei 2- bis 5-facher Auflösungssteigerung.



7.3.1 Interpretation der Simulationsergebnisse

Die Bewertung der Simulationsergebnisse erfolgt für jede SR-Technik separat und bezieht sich auf die SR-Rekonstruktionen der synthetischen SE- und USAF-1951-Testbilder. Es sei darauf hingewiesen, dass ein direkter Vergleich der erreichten Grenzauflösungen aus der Ortsfrequenzganganalyse und dem USAF-1951-Testbild unzulässig ist, weil sich die Bildauflösungen beider Testbilder unterscheiden.

Rohdaten

Zunächst sind in Tabelle 7.1 die Ortsfrequenzgänge der SE-Rohdaten gegenübergestellt. Darin decken sich die SFR-Kurvenverläufe aller vier GT-Bilder mit der theoretisch zu erwartenden sinc-Funktion. Der maximale absolute Fehler beim MTF_{50} -Wert ist für die zwei GT-Bilder mit den Bildauflösungen 64×64 sowie 80×80 Bildpunkten gleich groß und beträgt $0,002 \text{ LP/px}$. Dies entspricht einem maximalen relativen Fehler von $0,33 \%$ und kann vernachlässigt werden, weil er innerhalb der in Abschnitt 6.1 angegebenen Messabweichung von 1% der Auswertesoftware *MTF mapper* liegt. Folglich ist die Annahme der Rohdaten als ideale SE-GT-Bilder gültig.

Im USAF-1951-Testbild aus Tabelle 7.2 sind die Linienpaare in Block -1, Element 3 nicht mehr eindeutig voneinander getrennt wahrnehmbar. Nach Gleichung 6.2 resultiert daraus für das GT-USAF-1951-Testbild mit der Bildauflösung 120×120 Pixel eine Grenzauflösung von $0,63 \text{ LP/mm}$.

PCHIP-Interpolation

Bei der Auflösungssteigerung nach der PCHIP-Interpolation fällt die Bildqualität im Ergebnisbild mit steigendem Skalierungsfaktor rapide ab, wie aus Tabelle 7.6 ersichtlich ist. Bereits bei doppelter Pixelanzahl entlang einer Bildrichtung sinkt der MTF_{50} -Wert im SE-Testbild um 63% auf eine Ortsfrequenz von $0,222 \text{ LP/px}$. Mit zunehmendem Vergrößerungsfaktor reduziert sich der Kantenkontrast im Interpolationsergebnis weiter, was sich durch einen abnehmenden MTF_{50} -Wert äußert.

In Tabelle 7.2 zeigt sich ein ähnliches Verhalten für die ermittelte Grenzauflösung im USAF-1951-Testbild. Bei 2-facher Auflösungssteigerung fällt die Grenzauflösung um 37% auf $0,40 \text{ LP/mm}$, wonach die Linienpaare in Gruppe -2, Element 5 gerade nicht mehr voneinander getrennt wahrnehmbar sind.

Im Allgemeinen verschlechtert die Interpolation eines einzelnen LR-Bildes die Bildqualität im Ergebnisbild deutlich sicht- und messbar. Für das menschliche Auge äußern sich diese Qualitätseinbußen in unscharfen Ergebnisbildern mit flachen Kontrastübergängen an Kanten. Hohe Ortsfrequenzen, wie sie bei der USAF-1951 Auflösungstesttafel auftreten, verschwimmen bis hin zur Unkenntlichkeit.

EnhanceNet SISR

Anders verhält es sich bei der SISR-Methode nach dem ENet-PAT-Algorithmus. Dafür zeigt das SR-Ergebnis an der schrägen Kante in Tabelle 7.2 nicht nur augenscheinlich einen scharfen

Kontrastübergang, sondern auch die Ortsfrequenzanalyse in Tabelle 7.7 bestätigt diesen ersten visuellen Eindruck. Trotz der vergleichsweise hohen Auflösungssteigerung um den Faktor 4, liegt der SFR-Kurvenverlauf bei Ortsfrequenzen bis rund $0,35 \text{ LP/px}$ dicht an der theoretischen Auflösungsgrenze. Erst danach fällt der Kontrast ab, so dass der MTF_{50} -Wert mit $0,459 \text{ LP/px}$ rund 24 % unterhalb dem der sinc-Funktion liegt. Dennoch übersteigt er mit dieser erreichten Ortsfrequenz die der PCHIP-Interpolation um gut das Vierfache bei derselben Auflösungssteigerung. Allerdings zeigt der ENet-PAT-Algorithmus deutliche Schwächen bei der Rekonstruktion komplexer Geometrien, wie das SR-Ergebnis des USAF-1951-Testbildes in Tabelle 7.2 zeigt. Artefakte stören die Darstellung von Buchstaben und Ziffern sowie die der Linienpaare mit höherer Ortsfrequenz bis hin zur Unkenntlichkeit. Dies wird auf zwei Ursachen zurückgeführt: 1) Der Algorithmus optimiert merkmalsbasiert, um die menschliche Wahrnehmung nachzuempfinden. Daraus folgt jedoch eine ungenaue Zuordnung im Subpixel-Bereich [82] und die Rekonstruktion feiner Objektstrukturen kann sich vom tatsächlichen GT-Bild unterscheiden. Das Minimalbeispiel in Abbildung 4.3, Abschnitt 4.2 veranschaulicht diesen Zusammenhang. Diese Rekonstruktionsfehler fallen an den klaren Strukturen des USAF-1951-Testbildes schneller auf als bei den natürlichen Fotomotiven, wie sie Trainingsdatensätze typischerweise enthalten und anhand derer lernbasierte SISR-Verfahren üblicherweise bewertet werden. 2) Außerdem beinhaltet der MSCOCO-Trainingsdatensatz, mit dem der ENet-PAT-Algorithmus trainiert wurde, keine Bilder mit solchen speziellen und unnatürlichen Geometrien, weshalb deren Rekonstruktion gar nicht erst gelernt wurde. Aufgrund der starken Artefakte sind die Linienpaare in Gruppe -2, Element 5 nicht mehr eindeutig voneinander getrennt wahrnehmbar und die Grenzauflösung der SR-Rekonstruktion liegt bei $0,40 \text{ LP/mm}$. Damit entspricht sie der Grenzauflösung der PCHIP-Interpolation bei 2-facher Auflösungssteigerung.

Im Gegensatz zur PCHIP-Interpolation, erhöht die EnhanceNet SISR nicht nur die Bildauflösung, sondern auch die Bildqualität im SR-Ergebnis. Demnach steigert der ENet-PAT-Algorithmus das Auflösungsvermögen im SE-Testbild um gut das Vierfache, bezogen auf den MTF_{50} -Wert bei 4-facher Auflösungssteigerung. Die Grenzen zeigen sich dagegen am USAF-1951-Testbild. Bildinhalte, die der Algorithmus zuvor nicht gelernt hat sowie Objektszenen, bei denen es auf eine Subpixel-genaue Rekonstruktion ankommt, kann der Algorithmus nicht störungsfrei rekonstruieren. Resümierend lässt sich festhalten, dass die Auflösungssteigerung mit nur einem Eingangsbild möglich ist, sofern der spätere Anwendungsbereich bekannt ist und entsprechend im Datensatz für den Trainingsprozess berücksichtigt wird. Eine generische Auflösungssteigerung ist damit jedoch nicht möglich.

Inkohärente Fourier Ptychographie

In Tabelle 7.2 sind die SR-Rekonstruktionen der inkohärenten Fourier Ptychographie für beide Testbilder bei 2-facher Auflösungssteigerung und unter Berücksichtigung von 49 LR-Eingangsbildern dargestellt. Dabei ist deren erreichte Bildqualität augenscheinlich vom jeweiligen GT-Pendant nicht zu unterscheiden. Die objektive Betrachtung der SFR-Kurven in Tabelle 7.8 bestätigt diesen Eindruck für das SE-Testbild, da die SFR-Kurve für die SR-Rekonstruktion aus 49 LR-Eingangsbildern auf der sinc-Funktion verläuft. Zudem fällt in den Tabellen 7.8 bis 7.10 auf, dass für jeden Skalierungsfaktor das GT-Bild nahezu identisch rekonstruiert werden kann, solange nur hinreichend viele LR-Eingangsbilder zur Verfügung stehen. Während dafür bei 2-facher Skalierung noch 49 LR-Bilder ausreichen, benötigt die 4-fache Auflösungssteigerung bereits 225

LR-Eingangsbilder. Stehen dagegen nur wenige Eingangsbilder zur Verfügung, resultieren im SR-Bild deutlich wahrnehmbare Artefakte. Diese zeichnen sich durch fleckige Inhomogenitäten in sonst homogenen Bildbereichen sowie Doppelkanten an Kontrastübergängen aus, wie sie die SR-Ergebnisse in den Tabellen 7.11 und 7.12 aufweisen. Diese Bildfehler zeichnen sich für das SE-Testbild ebenfalls bei den Ortsfrequenzganganalysen in den Tabellen 7.8 bis 7.10 ab. Dabei tragen zwei Effekte maßgeblich zur verminderten Bildqualität bei:

1. Die inhomogene Beleuchtung der Objektszene mit Speckle-Mustern verursacht dunkle Flecken in den SR-Ergebnisbildern, wenn diese mit nur wenigen LR-Eingangsbildern rekonstruiert werden. Bei der 2-fachen Auflösungssteigerung sind diese Artefakte besonders prägnant, wenn die SR-Rekonstruktion aus 9 LR-Eingangsbildern erfolgt, bei der 4-fachen Auflösungssteigerung dagegen auch noch bei der SR-Rekonstruktion aus 25 LR-Eingangsbildern, wie den Tabellen 7.11 und 7.12 zu entnehmen ist. Grund dafür ist eine unvollständige Ausleuchtung der Objektszene aufgrund zu weniger Aufnahmen. Dieser Umstand kann ebenfalls den zugehörigen SFR-Kurven entnommen werden. Dort führen die Inhomogenitäten bei den untersuchten Auflösungssteigerungen zu mitunter starken Kontrastschwankungen im unteren Orts-Frequenzbereich um $0,15 \text{ LP/px}$ herum. Wegen der im Verhältnis größeren Speckle-Ausdehnung bei 2-facher Auflösungssteigerung, beeinflusst die inhomogene Beleuchtung dessen SR-Ergebnis stärker als bei höheren Skalierungsfaktoren. Bereits bei 3-facher Auflösungssteigerung fallen die Ausschläge deutlich geringer aus, bis sie bei 4-facher Vergrößerung nicht mehr nennenswert in Erscheinung treten. Außerdem führt die inhomogene Ausleuchtung in den Ergebnisbildern zu mitunter starken Kontrastanstiegen bei hohen Ortsfrequenzen $>0,80 \text{ LP/px}$. Je prägnanter die Fehlstellen an der Hell-Dunkel-Grenze ausgeprägt sind, desto stärker tritt dieser Kontrastanstieg in Erscheinung, wie ein Vergleich mit den erreichten Bildqualitäten in den Tabellen 7.11 und 7.12 zeigt. Generell gilt: Erst mit einer hinreichend hohen Anzahl an LR-Eingangsbildern überlagern sich die Speckle-Muster zu einer homogenen Objektbeleuchtung im SR-Ergebnisbild.
2. Die inverse Fourier-Transformation einer diskreten Rechteckblende verursacht Ringing-Artefakte [145], die sich durch Geisterkanten an horizontalen und vertikalen Kontrastübergängen in der SR-Rekonstruktion äußern. Das in Kapitel 5.1.2 beschriebene Simulationsmodell wendet im Fourierraum einen idealen Tiefpass an, der Frequenzen abrupt sperrt. Mit steigender Anzahl an Eingangsbildern vergrößert sich der Tiefpassbereich und höhere Frequenzen, die zuvor noch herausgefiltert wurden, werden nun übertragen. Dies vermindert das Auftreten von Ringing-Artefakten, wie die Abbildungen in Tabelle 7.12 zeigen. Enthält die LR-Bilderfolge hinreichend viele Einzelbilder, rekonstruieren diese den gesamten Fourierraum und die Grenzfrequenz des auflösungs-gesteigerten optischen Systems ist erreicht. Sobald dieser Zustand erreicht ist, geht dessen Tiefpass-Charakter mit diskreter Blende verloren und damit ebenso die Ringing-Artefakte im SR-Ergebnis. Die Ringing-Artefakte führen in den SFR-Kurven zu einem lokalen Minimum, das für die 2- bis 4-fachen Auflösungssteigerungen in den Ortsfrequenzbereichen um $0,4 \text{ LP/px}$, $0,3 \text{ LP/px}$ und $0,2 \text{ LP/px}$ herum liegt, wie den Tabellen 7.8 bis 7.10 zu entnehmen ist. Dabei zeigt sich das Verhalten bei der 4-fachen Auflösungssteigerung und für $N = 9$ bis $N = 49$ Eingangsbilder am deutlichsten.

Die inkohärente Fourier Ptychographie ist in der Lage ein synthetisches GT-Bild vollständig zu rekonstruieren, sofern hinreichend viele LR-Einzelbilder zur Rekonstruktion beitragen. Dabei

beeinflusst der gewählte Skalierungsfaktor die erforderliche Mindestanzahl an Eingangsbildern. Während bei 2-facher Auflösungssteigerung bereits 49 LR-Bilder ausreichen, um das GT-Bild vollständig zu rekonstruieren, werden bei 4-facher Skalierung dafür bereits 225 Eingangsbilder benötigt. Die erreichte Bildqualität im SR-Ergebnisbild hängt maßgeblich von seinem Rauschen ab, das bei den hier verwendeten synthetischen Testbildern auf eine inhomogene Ausleuchtung zurückzuführen ist. Für die Adaption der inkohärenten FP auf PMD-Kameras wirkt sich jedoch die Forderung eines bekannten, unveränderlichen Beleuchtungsmusters zur SR-Rekonstruktion einschränkend auf ausschließlich zweidimensionale Messanwendungen aus, weil tiefenausgedehnte Objekte das statische Beleuchtungsmuster verzerren würden. Mit den typischerweise dreidimensionalen Objektszenen scheidet damit die generische Auflösungssteigerung von PMD-Kameras mit Hilfe der inkohärenten FP aus.

Subpixel-Shift SR nach dem IRWSR-Algorithmus mit NN-Interpolation

Die SR-Rekonstruktionen beider Testbilder nach der Subpixel-Shift SR mit NN-Interpolation erreichen in den Ergebnisbildern augenscheinlich nahezu die Bildqualität der zugrunde liegenden GT-Bilder, wie die Gegenüberstellung in Tabelle 7.2 zeigt. Diesen visuellen Eindruck bestätigt die Ortsfrequenzganganalyse für das SE-Testbild in Tabelle 7.13. Demnach verläuft die SFR-Kurve für das SR-Ergebnis bei 2-facher Auflösungssteigerung aus 4 LR-Eingangsbildern größtenteils parallel versetzt zur sinc-Funktion und weicht beim MTF_{50} -Wert mit $0,502 \text{ LP/px}$ knapp 17 % von dem des GT-Bildes ab. Mit Zunahme der LR-Eingangsbilder reichen die SFR-Kurvenverläufe dichter an die theoretische Auflösungsgrenze heran, wobei bei 3- und 5-facher Auflösungssteigerung das GT-Bild nahezu vollständig rekonstruiert wird. Auffällig ist hierbei jedoch, dass die SFR-Kurven für ungerade Anzahlen an LR-Einzelbildern besser mit dem Verlauf der sinc-Funktion übereinstimmen als die der geraden Anzahlen an Eingangsbildern. Diese fallen im Vergleich dazu ab einer Ortsfrequenz von rund $0,4 \text{ LP/px}$ stärker ab. Außerdem übersteigen die SFR-Kurven aller 4 SR-Ergebnisse das theoretische Auflösungsvermögen für Ortsfrequenzen größer $0,85 \text{ LP/px}$. Diese Beobachtungen sind an dieser Stelle noch nicht geklärt und sollen vor einer tiefer gehenden Untersuchung mit später folgenden Messreihen am realisierten Aufbau bestätigt werden.

Die Linienpaare im USAF-1951-Testbild bei 2-facher Auflösungssteigerung sind in Gruppe -1, Element 2 nicht mehr eindeutig voneinander getrennt und bestimmen damit dessen Grenzauflösung mit $0,56 \text{ LP/mm}$. Demnach ist das SR-Ergebnis des USAF-1951-Testbildes um 11 % schlechter aufgelöst als das GT-Bild.

7.3.2 Zwischenfazit zu den Simulationsergebnissen

Resümierend zeigen die Simulationsergebnisse aller drei (echten) SR-Methoden Ergebnisbilder mit deutlich gesteigertem Auflösungsvermögen. Wie zuletzt mit den Interpretationen der SR-Ergebnisse erörtert, eignet sich von den drei Methoden letztlich nur die Subpixel-Shift SR zur generischen Auflösungssteigerung von PMD-Kameras.

Bei der *EnhanceNet SISR* hängt die Qualität der SR-Rekonstruktion maßgeblich vom Trainingsdatensatz mit seinen LR- und GT-Bildpaaren ab. Da es die PMD-Kamera für Anwendungen im

Außenbereich nicht mit hochauflösenden Sensoren gibt, existieren keine GT-Bilder für den Trainingsprozess. Zudem begrenzt der Datensatz, mit dem die lernbasierten SISR-Verfahren trainiert werden, das spätere Anwendungsfeld auf die damit gelernten Bildeigenschaften. Eine generische SR ist damit ausgeschlossen.

Die *inkohärente Fourier Ptychographie* erlaubt eine Skalierung der Auflösungssteigerung über die Anzahl der berücksichtigten LR-Eingangsbilder. Allerdings scheidet sie wegen ihrer Anforderung nach einem bekannten und unveränderlichen Beleuchtungsmuster für die Auflösungssteigerung dreidimensionaler Objektszenen aus. Somit ist auch mit diesem Verfahren eine generische SR für PMD-Kameras nicht uneingeschränkt realisierbar.

Die *Subpixel-Shift SR* überzeugt mit hochwertigen SR-Rekonstruktionen bei verhältnismäßig geringer Anzahl an LR-Eingangsbildern. Für die 2-fache Auflösungssteigerung reichen bereits 4 LR-Eingangsbilder aus, um beim MTF_{50} -Wert bis auf 17% an den des SE-GT-Bildes heranzukommen. Bereits mit 9 LR-Eingangsbildern hat die SR-Rekonstruktion des synthetischen SE-Testbildes die ursprüngliche Bildqualität seines GT-Pendants nahezu erreicht. Außerdem eignet sich die SR-Methode für beliebige Anwendungsfälle sowie zur Auflösungssteigerung dreidimensionaler Objektszenen. Aus diesen Gründen verfolgt diese Arbeit den Ansatz der SpS SR für die experimentellen Untersuchungen am Demonstrator.

Kapitel 8

Demonstratoren

In folgendem Kapitel werden zwei Demonstratoren vorgestellt, mit denen die Simulationsergebnisse der SpS-SR aus dem vorherigen Kapitel verifiziert werden. Beim Sensor-Shift-Ansatz, beschrieben in Abschnitt 8.1, wird der Bildsensor im Bruchteil seiner Pixelgröße bewegt, um Subpixel-verschobene Bildaufnahmen zu erhalten. Er ist bereits aus der Literatur bekannt [40] und dient als Referenz für den zweiten realisierten Aufbau in Abschnitt 8.2. Dieser neuartige Lens-Shift-Ansatz [146] verschiebt anstelle des Bildsensors eine Linse im optischen System, woraus ein Subpixel-genauer Bildversatz in der Bildebene erfolgt. Abschließend werden in Abschnitt 8.3 die für diese Arbeit geltenden Messbedingungen beschrieben.

Zur erfolgreichen Umsetzung der SpS-SR muss in den LR-Bildaufnahmen Aliasing vorhanden sein, wie bereits in Abschnitt 5.2.1 erläutert. Demzufolge muss die Grenzauflösung der abbildenden Optik betragsmäßig so viel höher sein als die Grenzauflösung des niedrig auflösenden Sensors, wie der angestrebte Skalierungsfaktor von LR- zu SR-Bild ist. Andernfalls limitiert die Abbildungsleistung des Objektivs die Auflösungssteigerung im Ergebnisbild. Die Grenzauflösung ist hier nach Burns und Williams [147] für einen Kontrast von 10 % definiert.

Um eine ausreichend hohe Abbildungsleistung der Optik sicherzustellen, werden die im jeweiligen Aufbau verwendeten Objektive vermessen. Dazu konnte auf die *ImageMaster* MTF-Messanlage der Firma Trioptics zurückgegriffen werden, welche die Arri AG, München dankenswerterweise für die hier gezeigten Messaufnahmen zur Verfügung gestellt hat. Dessen axiale und außeraxiale Messgenauigkeit gibt der Hersteller mit $\pm 2\%$ [148] an. Die Messergebnisse sind in Form von MTF-Kurven dem theoretisch erreichbaren Auflösungsvermögen des PMD-Sensors gegenübergestellt.

8.1 Sensor-Shift Verfahren

Beim Sensor-Shift-Demonstrator setzen Mikrometerschrauben den transversalen Bildversatz direkt in der Bildebene um, indem sie den Bildsensor lateral verschieben. Durch die aktive Sensorbewegung ist für jede LR-Aufnahme die Sensorposition in Subpixel-Genauigkeit bekannt, ohne diese über einen optionalen Algorithmus schätzen zu müssen. Als Bezugspunkt für die LR-Aufnahmen dient bei diesem Aufbau die optische Achse, die in Nominalstellung das Zentrum des Bildsensors durchstößt. Die Positionen der einzelnen LR-Bilder umlaufen die optische Achse konzentrisch und sind relativ zueinander verschoben. Generell hängt die erreichte Bildqualität

im SR-Ergebnis davon ab, wie exakt der Bildversatz aufeinanderfolgender LR-Aufnahmen bekannt ist [31, 32, 64]. Wegen der großen Pixelausdehnung des PMD-Sensors im Sub-Millimeter-Bereich, gewährleisten die eingesetzten Mikrometerschrauben eine hinreichend genaue Sensorpositionierung.

In Abbildung 8.1 ist die Abbildungsleistung der abbildenden Optik der des PMD-Sensors gegenübergestellt, wobei für das Objektiv sowohl die tangentielle als auch die sagittale MTF-Messkurve aufgetragen sind. Die SFR-Kurve des PMD-Sensors ergibt sich aus seiner Pixelbreite nach der sinc-Funktion in Gleichung 6.1. Dementsprechend zeigt die schwarze Kurve das theoretische Auflösungsvermögen des Sensors für die in Abschnitt 2.2 vorgestellte PMD-Kamera auf. Bei einem Kontrast von 10% überträgt das Objektiv noch 23 LP/mm, der PMD-Sensor dagegen 6 LP/mm. Somit übersteigt die Grenzauflösung der Optik die des Sensors um das 3,8-fache. Die MTF-Kurvenverläufe des Objektivs decken sich für die tangentialen und sagittalen Richtungen über den gesamten Messbereich.

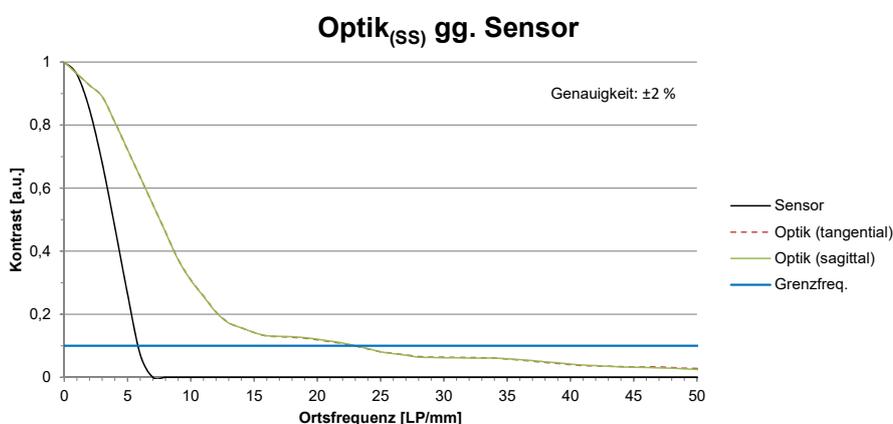


Abbildung 8.1: Die Grenzfrequenz, im Diagramm als blaue Linie hervorgehoben, ist erreicht, wenn der Bildkontrast auf 10% abgefallen ist. Nach dieser Definition übersteigt die Abbildungsleistung der Optik (6 LP/mm) die des Sensors (23 LP/mm) um den Faktor 3,8.

8.2 Lens-Shift Verfahren

Die Lens-Shift-Methode beschreibt einen neuartigen Ansatz zur Umsetzung der SpS-SR, bei dem Mikroaktoren eine Linse des optischen Systems im Sub-Millimeterbereich verschieben. Dabei umlaufen die LR-Aufnahmepositionen, analog zur Sensor-Shift-Methode, die optische Achse konzentrisch. Der Bildversatz in der Bildebene resultiert indirekt aus dem Transversalversatz der Linse und nicht aus einer Sensorbewegung. Plakativ beschrieben invertiert der Lens-Shift-Ansatz die klassische optische Bildstabilisierung (engl. optical image stabilizer, OIS), die Hersteller teilweise in modernen Kamera-Objektiven umsetzen. Damit ist der Aufbau dafür prädestiniert, auf etablierte opto-mechanische Hardware der Optik-Industrie zurückzugreifen.

Das optische System ist als dreilinsiges Petzval-Objektiv realisiert, wie in Abbildung 8.2 skizziert. Es besteht aus zwei plankonvexen Asphären mit je 50 mm Brennweite sowie einer Plankonkavlinse mit -25 mm Brennweite. Seine rechnerische Systembrennweite liegt bei 42,7 mm. Der maximale horizontale Feldwinkel erreicht $6,5^\circ$, die paraxiale Vergrößerung beträgt $\gamma = -0,01246$

und der Eintrittspupillendurchmesser liegt bei 20 mm. Dabei beziehen sich diese Angaben auf Simulationen mit dem sequentiellen Raytracer *Zemax* bei einer Auslegungs-Wellenlänge von 850 nm.

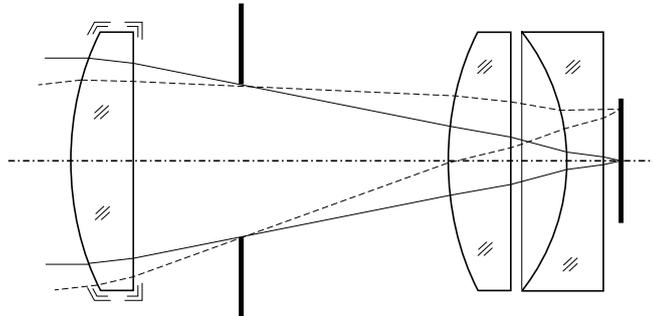


Abbildung 8.2: Der Lens-Shift-Aufbau ist als dreilinsiges Petzval-Objektiv realisiert, dessen Eintrittslinse mit motorgesteuerten Aktuatoren transversal verschoben wird.

Bei der Auslegung des optischen Systems ist darauf zu achten, dass sich das Verschieben der optischen Komponente auf ihre Abbildungseigenschaften auswirken kann. Darum wird in Abbildung 8.3 die Abbildungsleistung der Optik für die zentrierte und maximal dezentrierte Linsenposition mit der des PMD-Sensors verglichen. Für jede Linsenposition muss das Auflösungsvermögen der abbildenden Optik mindestens so hoch sein wie die angestrebte laterale Ortsauflösung im HR-Bild ist [40].

Unter Verwendung kostengünstiger Kataloglinsen liefert das ausgelegte optische System eine ausreichend hohe Abbildungsqualität, um für sämtliche Linsenpositionen das Auflösungsvermögen des Sensors deutlich zu übersteigen. Dies belegen die Messkurven in Abbildung 8.3, wonach das Auflösungsvermögen des Objektivs das des PMD-Sensors um einen Faktor >8 übertrifft. Für den senkrechten Lichteintritt und einem Kontrastabfall auf 10% löst das Objektiv sowohl bei zentrierter als auch bei maximal dezentrierter Linsenstellung mehr als 50 LP/mm auf, der PMD-Sensor dagegen nach wie vor 6 LP/mm. Auch hier decken sich die MTF-Kurvenverläufe des Objektivs für die tangentialen und sagittalen Richtungen über den gesamten Messbereich.

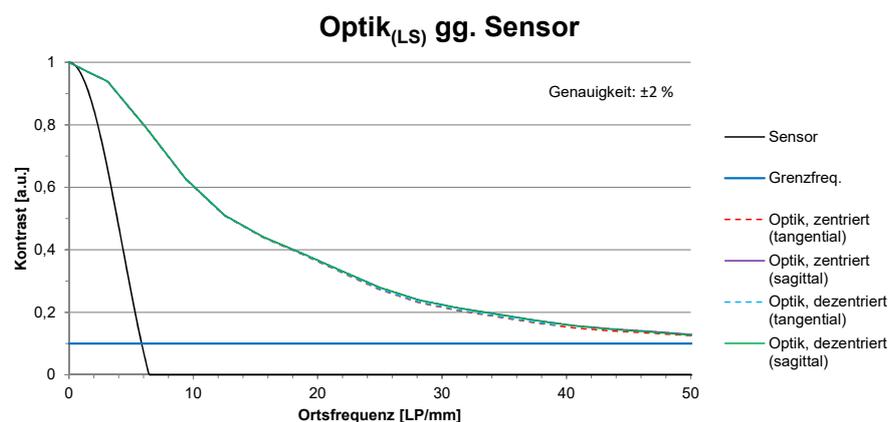


Abbildung 8.3: Sowohl bei zentrierter als auch maximal dezentrierter Linsenstellung übertrifft die Grenzauflösung der abbildenden Optik die des PMD-Sensors um einen Faktor >8 .

8.3 Messbedingungen

Unabhängig von der Demonstrator-Variante, erfolgen die LR-Aufnahmen mit der in Kapitel 2.2 vorgestellten PMD-Kamera. Während der SS-Aufbau auf das Original-Objektiv der PMD-Kamera zurückgreift, setzt der LS-Aufbau das in vorhergehendem Kapitel beschriebene Objektiv ein. Bei den aufgenommenen Bilderfolgen handelt es sich um die Amplitudenbilder des Schräge-Kante-Testbildes. Dieses ist mit einer Grundfläche von 0,6 m x 0,6 m in 3,5 m Entfernung zur Bildebene aufgestellt. Entsprechend dem synthetischen SE-Testbild ist seine schräge Kante um 10° gedreht. Der Kontrastunterschied von schwarzem zu weißem Bildbereich ergibt sich über unterschiedlich stark reflektierende Oberflächen. Eine 2 m x 4 m große, mit licht-absorbierendem Molton Stoff ausgekleidete Messzelle minimiert Störlicht sowie Multi-Pfad-Fehler [6, 149] und trägt damit zu konstanten Messbedingungen bei. Abbildung 8.4 zeigt das quadratische SE-Testbild auf einer Positionier-Plattform in der lichtabsorbierenden Messzelle.



Abbildung 8.4: Das SE-Testbild ist drehbar gehalten und lässt sich auf der Plattform frei im Raum positionieren. Um Einflüsse von Störlicht zu minimieren, befindet es sich in einer lichtabsorbierenden Messzelle.

Kapitel 9

Experimentelle Ergebnisse

Mit diesem Kapitel werden die Simulationsergebnisse der Subpixel-Shift-Methode aus Abschnitt 7.3 verifiziert. Dazu nehmen die zwei realisierten Demonstratoren aus dem vorherigen Kapitel, der Sensor-Shift- wie auch der Lens-Shift-Aufbau, Folgen niedrig auflösender Amplitudenbilder mit der PMD-Kamera aus Abschnitt 2.2 auf.

Bei der Auswertung in Abschnitt 9.1 werden die SR-Ergebnisse aus den Messungen denen aus der Simulation gegenübergestellt. Dabei deckt die Ortsfrequenzganganalyse Abweichungen auf, die auf ein unzureichendes Simulationsmodell bei der Erzeugung synthetischer LR-Bilderfolgen zurückzuführen sind. Dementsprechend erfolgen in Abschnitt 9.2 Anpassungen sowohl an den Simulations- als auch an den Rekonstruktionsparametern des allgemeinen Beobachtungsmodells. Deren Simulationsergebnisse werden in Abschnitt 9.3 erneut den SR-Ergebnissen aus den Messaufnahmen gegenübergestellt und zeigen damit, dass mit dem angepassten Beobachtungsmodell schlüssige SR-Rekonstruktionen vorhersagbar sind. In Abschnitt 9.3.1 wird der Einfluss der Parameterwahl auf das SR-Ergebnis synthetisch erzeugter Bilderfolgen erläutert, bevor in Abschnitt 9.3.2 eine finale Diskussion zu den Simulationsergebnissen erfolgt. Dabei werden die Genauigkeit des Rauschmodells, der limitierende Faktor bei der Erhöhung des Auflösungsvermögens nach der SpS-SR-Methode sowie dessen Leistungsfähigkeit bei der SR-Rekonstruktion behandelt. Daran anknüpfend ist in Abschnitt 9.3.3 das Optimierungspotenzial des vorgeschlagenen Simulationsmodells aufgezeigt, bevor in Abschnitt 9.4 eine PSNR-Rauschanalyse auf die SR-Rekonstruktionen der finalen Simulationsergebnisse beider Testbilder angewendet wird. Die letzten Untersuchungen zur SpS-SR erfolgen in Abschnitt 9.5 mit ihrer Anwendung auf die niedrig auflösenden Amplituden- und Entfernungsbilder einer dreidimensionalen Objektszene, welche dem Lens-Shift-Aufbau entstammen. Mit einer Diskussion der ausformulierten Forschungsfragen in Abschnitt 9.6 wird dieses Kapitel abgeschlossen.

9.1 Messergebnisse

Zur SR-Rekonstruktion erwartet der IRWSR-Algorithmus Angaben zum Subpixel-Versatz der aufeinanderfolgenden LR-Aufnahmen sowie zur PSF-Breite des verwendeten optischen Systems. Der Subpixel-Versatz aufeinanderfolgender Aufnahmen entspricht dem Fahrweg der Sensorbewegung und wird am Nonius der Mikrometerschrauben abgelesen. Die PSF-Breite des optischen Systems wird empirisch ermittelt, wie schon bei der SpS-Simulation mit NN-Interpolation in Abschnitt 7.2: Für jede LR-Bilderfolge erfolgen SR-Rekonstruktionen mit systematisch abnehmender PSF-Breite, wobei sein Startwert bei 0,50 beginnt. Das SR-Ergebnis, dessen SFR-Kurve am wenigsten von der sinc-Funktion abweicht, gibt die PSF-Breite des optischen Systems vor.

Über die gesamten Messreihen beider Demonstratoren hinweg führt eine PSF-Breite von 0,30 zu den besten SR-Ergebnissen. Dies erscheint schlüssig, weil die PSF-Breite der PMD-Kamera nicht von der Anzahl an LR-Eingangsbildern oder dem Skalierungsfaktor abhängt und deshalb für alle Skalierungsfaktoren identisch bleiben sollte. In Tabelle 9.1 sind die Rekonstruktionsparameter der aufgenommenen Messreihen zusammengefasst.

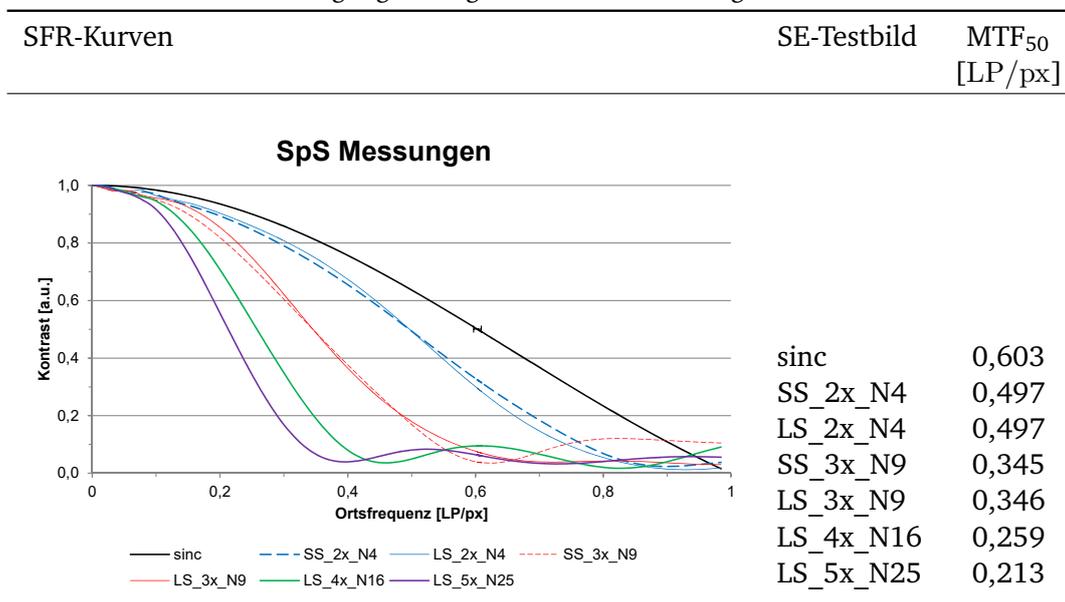
Tabelle 9.1: Parametersatz für den IRWSR-Algorithmus zur SR-Rekonstruktion aus den aufgenommenen LR-Bilderfolgen.

Anzahl LR-Bilder	4	9	16	25
LR-Bildauflösung	16 x 16			
F_{Aufl}	2	3	4	5
PSF-Breite	0,30			
Subpixel-Versatz [Px]	1/2	1/3	1/4	1/5

Vergleich von Sensor-Shift- und Lens-Shift-Demonstrator anhand deren SR-Ergebnisse

Die Ortsfrequenzganganalyse in Tabelle 9.2 attestiert den beiden Demonstrator-Varianten eine hohe Übereinstimmung bei den SR-Ergebnissen für die 2- und 3-fache Auflösungssteigerung. Unter Berücksichtigung der relativen Fehlertoleranz beim Bestimmen des MTF_{50} -Wertes (siehe Abschnitt 6.1), sind die Messergebnisse für beide Demonstratoren als gleich zu bewerten. Die geringfügigen Abweichungen im SFR-Kurvenverlauf zueinander sind auf die unterschiedliche Umsetzung beim Anfahren der Subpixel-Positionen der einzelnen LR-Bildaufnahmen zurückzuführen und können vernachlässigt werden. Folglich ist es für die SR-Rekonstruktion irrelevant, woher der Subpixel-Versatz in der Bildebene herrührt und die SR-Ergebnisse beider Demonstratoren werden als gleich angenommen. In der Konsequenz wird für die 4- und 5-fache Auflösungssteigerung auf zusätzliche Messungen mit dem Sensor-Shift-Demonstrator verzichtet.

Tabelle 9.2: Die Ortsfrequenzganganalyse der SR-Ergebnisse aus den Messreihen beider Demonstrator-Varianten zeigt eine hohe Übereinstimmung. Damit belegt sie die Gleichstellung beider Varianten bei der Erzeugung niedrig auflösender Bilderfolgen.



Gegenüberstellung von Simulations- und Messergebnissen

Mit den SR-Ergebnissen aus den LR-Aufnahmen der PMD-Kamera, erfolgt in diesem Abschnitt ein direkter Vergleich mit deren Simulations-Pendants. Dazu stellen die Tabellen 9.3 bis 9.6 die SFR-Kurven der gemessenen denen der simulierten SR-Ergebnisse gegenüber. Der besseren Übersichtlichkeit wegen, erfolgen die Vergleiche für jeden Vergrößerungsfaktor in einem separaten Diagramm. Dabei sind die einzelnen SFR-Kurven in gewohnter Weise farblich kodiert: Die schwarze, durchgezogene Kurve beschreibt die theoretische Auflösungsgrenze der PMD-Kamera nach einer sinc-Funktion. Die Untergrenze bei der Auflösungssteigerung visualisiert die schwarze, gestrichelte Kurve als Ergebnis der PCHIP-Interpolation bei dem jeweils vorliegenden Skalierungsfaktor. Dazwischen verlaufen die farbigen, durchgezogenen Kurven als SR-Ergebnisse aus Simulation und Messung. Dabei steht rot für das Simulationsergebnis unter Berücksichtigung der NN-Interpolation, blau für das Messergebnis aus dem Sensor-Shift-Aufbau und grün für das Messergebnis aus dem Lens-Shift-Aufbau.

Während bei der 2-fachen Auflösungssteigerung die SFR-Kurven aus den Messungen scheinbar gut mit jener aus der Simulation übereinstimmen, weichen sie bereits ab der 3-fachen Auflösungssteigerung deutlich von der Simulation ab. Diese Abweichungen nehmen mit steigendem Skalierungsfaktor zu, so dass sich die SR-Ergebnisse aus den Messaufnahmen mit dem Lens-Shift-Aufbau jenen Ergebnissen der PCHIP-Interpolation annähern. Mit Verweis auf Abschnitt 7.3.1 sei festgehalten, dass die Messreihen das Phänomen einer leistungsfähigeren SR-Rekonstruktion aus LR-Bilderfolgen mit ungerader Anzahl an Eingangsbildern nicht bestätigen.

Tabelle 9.3: Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 2-facher Auflösungssteigerung.

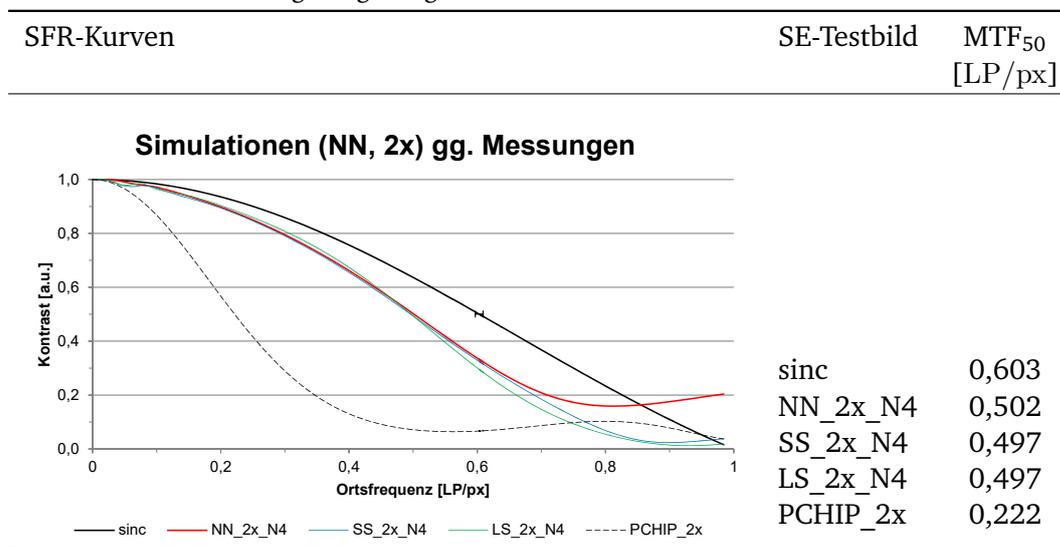


Tabelle 9.4: Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 3-facher Auflösungssteigerung.

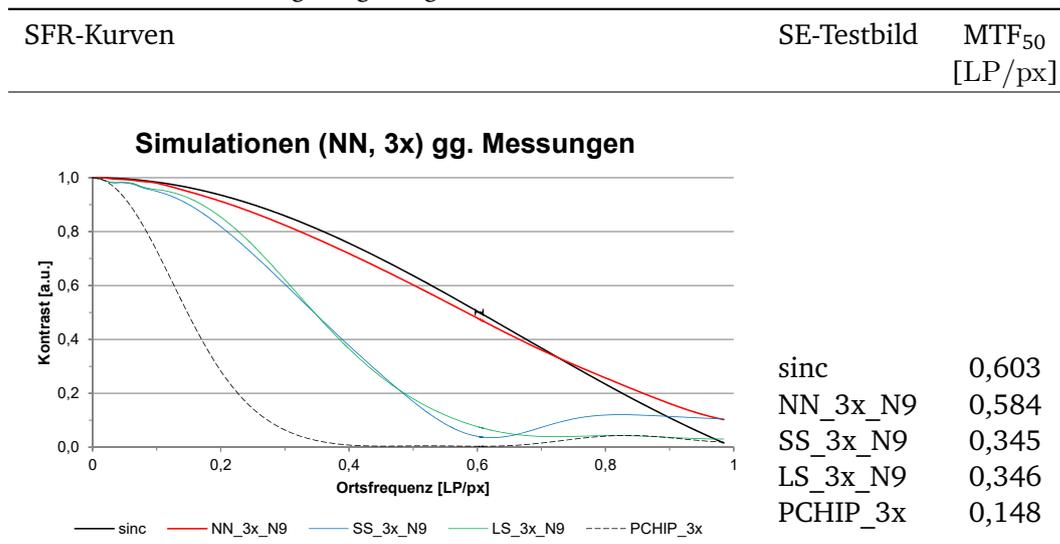


Tabelle 9.5: Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 4-facher Auflösungssteigerung.

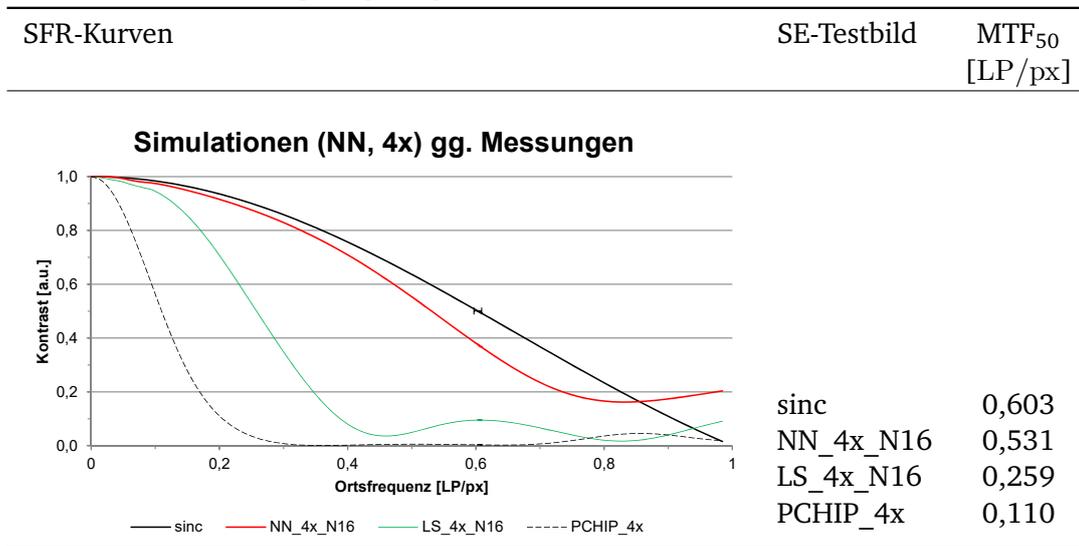
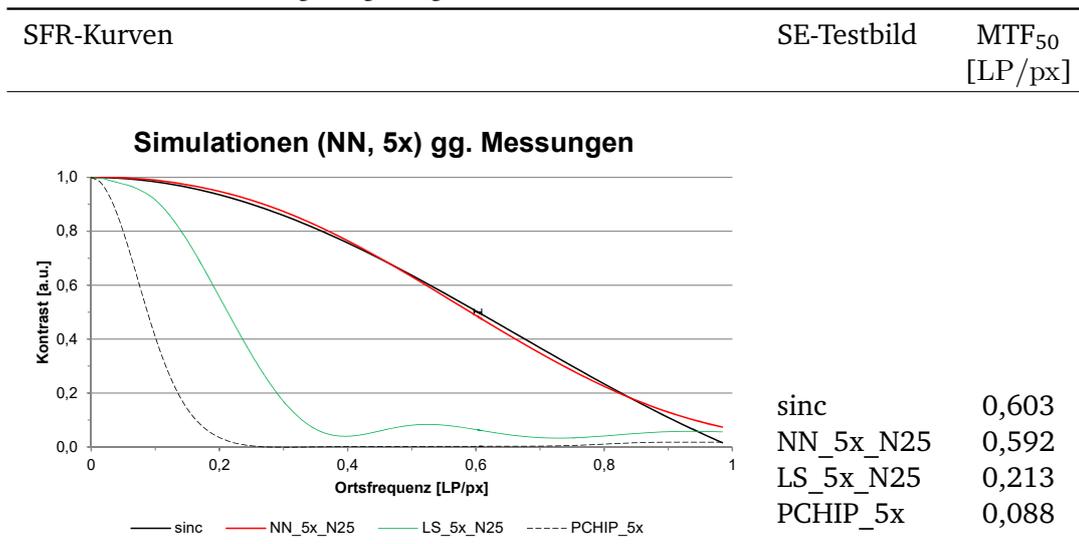


Tabelle 9.6: Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 5-facher Auflösungssteigerung.



9.2 Anpassung des Super-Resolution Beobachtungsmodells

Der Vergleich der Simulations- und Messergebnisse im vorhergehenden Abschnitt zeigt für die Auflösungssteigerungen ab einem Skalierungsfaktor von 3 signifikante Abweichungen in den SR-Ergebnissen. Dagegen zeigen die Messergebnisse aus den beiden Demonstratoren eine hohe Übereinstimmung, obwohl sich deren Umsetzung bei der Aufnahme von LR-Bilderfolgen unter-

scheidet. Dementsprechend wird die Datenerzeugung für beide Messaufbauten als gültig angenommen und in diesem Abschnitt folgt die Anpassung des allgemeinen Beobachtungsmodells zur Simulation von SR-Ergebnissen.

Anpassung der Simulationsparameter an die Kameraeigenschaften

Zur Simulation des allgemeinen Beobachtungsmodells aus Abschnitt 7.2 erzeugt das Skript aus Anhang A.2 eine synthetische LR-Bilderfolge, dessen Downsampling mit einer NN-Interpolation umgesetzt ist. Diese Vorgehensweise spiegelt jedoch nicht das Verhalten realer Bildaufnahmen wider, welches ein Gedankenspiel verdeutlicht: Eine kontinuierliche Objektszene wird näherungsweise als das hochauflösende GT-Bild in gewünschter Zielauflösung betrachtet. Dieses gilt es mit mehreren niedrig auflösenden Einzelbildern zu rekonstruieren. Dabei ist die Auflösung des abtastenden Sensors um ein ganzzahliges Vielfaches kleiner als die Zielauflösung im GT-Bild. Das bedeutet ein LR-Pixel deckt mehrere HR-Pixel gleichzeitig ab und sein Helligkeitswert entspricht dem arithmetischen Mittel aller von ihm abgedeckten Bildpunkte. Dagegen löscht der Algorithmus beim Downsampling nach der NN-Interpolation benachbarte HR-Pixel, um zu dem niedrig auflösenden Bild zu gelangen. In der Folge entspricht in diesem Beispiel der Helligkeitswert eines LR-Pixels nicht dem Mittelwert aus vier HR-Pixeln, sondern lediglich dem eines Einzelnen von ihnen. Insbesondere an Kontrastübergängen kann die NN-Interpolation somit steilere Gradienten verursachen als sie im Originalbild tatsächlich vorliegen. Um das Downsampling-Verhalten bei den Simulationen realistischer abzubilden, wird die NN-Interpolation durch eine mittelwertbildende (blockproc, BP) Interpolation ersetzt.

Zusätzlich zur inkorrekten Downsampling-Methode stimmen die angenommenen Simulationsparameter aus Abschnitt 7.2, Tabelle 7.4 und 7.5, nicht zwingend mit den Eigenschaften der verwendeten PMD-Kamera überein. Um diesbezüglich eine Übereinstimmung zu finden, werden die Simulationsparameter mit Hilfe empirischer Untersuchungen auf Basis der synthetisch erzeugten SE-Testbilder ermittelt. Ziel ist es, diejenigen Simulationsparameter zu finden, mit denen die vorhergesagten SR-Ergebnisse am wenigsten von jenen aus den Messaufnahmen abweichen. Dazu wird zunächst die PSF-Breite von 0,30 der PMD-Kamera übernommen, wie sie bereits in Abschnitt 9.1 mit Hilfe der Messreihen ermittelt wurde. Anschließend wird mit dem Salt-and-Pepper-Rauschen das Rauschverhalten im angepassten Simulationsmodell erweitert, um bei der Erzeugung synthetischer LR-Bilderfolgen näherungsweise den Einfluss ungültiger Pixel zu simulieren. Hier führt ein Rauschdichte-Parameter von $\rho_{SNP} = 0,002$ über alle Skalierungsfaktoren hinweg zu guten Übereinstimmungen in den SR-Ergebnissen aus Simulations- und Messdaten. Mit Kenntnis über die konstante PSF-Breite der PMD-Kamera und dem Rauschverhalten, wie es augenscheinlich am Besten zu den Amplitudenbildern passt, erfolgt eine Abschätzung des Unschärfebeitrags der Abbildungsoptik. Dieser ist bei der Generierung von LR-Bilderfolgen mit dem Octave-/Matlab-Skript in Anhang A.3 als isotroper Gauß'scher Glättungsfilter mit einer Standardabweichung σ_B realisiert. Ausgehend von einem Startwert von $\sigma_B = 0,5$, wird sein Wert beim Erzeugen von LR-Bilderfolgen mit einer Schrittweite von 0,01 sukzessive abwärts gezählt, während die Werte für PSF-Breite und Rauschparameter unverändert bleiben. Derjenige Wert, mit dem die simulierte SR-Rekonstruktion am wenigsten von jener aus den Messaufnahmen abweicht, entspricht dem gesuchten σ_B -Wert. Über alle vier Skalierungsfaktoren hinweg ist dies für $\sigma_B = 0,44$ der Fall.

Zusammengefasst setzt das angepasste Beobachtungsmodell drei Optimierungen um: 1) Die Interpolationsmethode beim Downsampling wechselt von der NN- zur mittelwertbildenden BP-Interpolation, um seinem Verhalten bei realen Bildaufnahmen gerecht zu werden. 2) Der Unschärfebeitrag σ_B der Abbildungsoptik wird über empirische Untersuchungen ermittelt und modelliert damit das Verhalten der PMD-Kamera genauer. 3) Das Rauschmodell wird um das Salt-and-Pepper-Rauschen ρ_{SnP} ergänzt, um das Rauschverhalten der PMD-Kamera genauer abzubilden. Das auf die PMD-Kamera angepasste Beobachtungsmodell zeigt Abbildung 9.1, wobei seine finalen Simulations- und Rekonstruktionsparameter in Tabelle 9.7 zusammengefasst sind. Das zugehörige Octave-/Matlab-Programm ist dem Anhang A.3 angefügt.

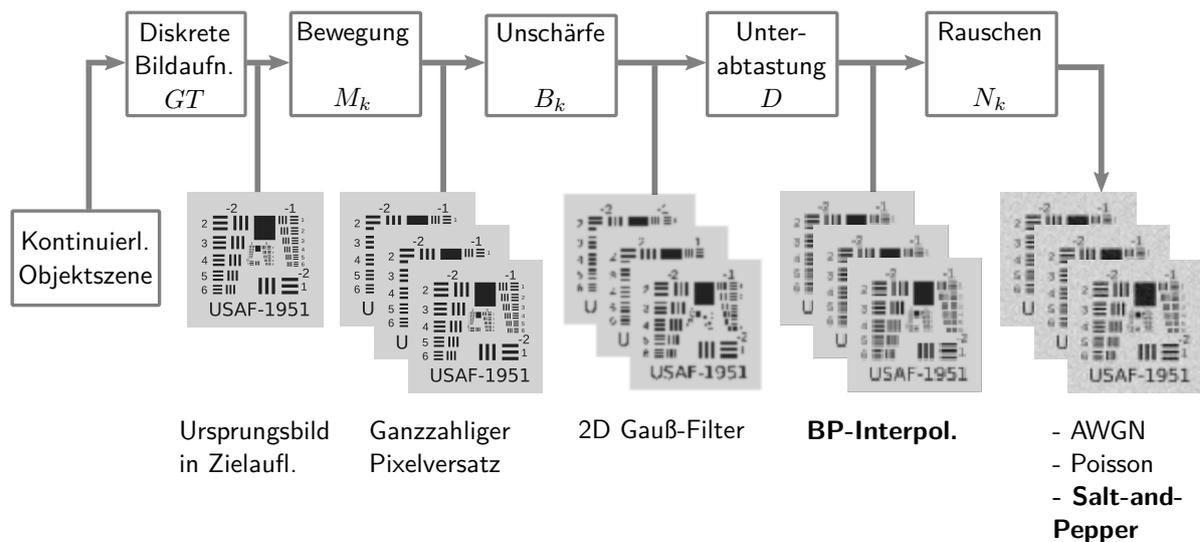


Abbildung 9.1: Das an die PMD-Kamera angepasste Beobachtungsmodell nutzt für das Downsampling eine mittelwertbildende BP-Interpolation und ergänzt das Rauschmodell um Salt-and-Pepper Rauschen, das näherungsweise fehlerhafte Pixelwerte simuliert.

Anpassung der Rekonstruktionsparameter an die Kameraeigenschaften

Bei den Rekonstruktionsparametern ändert sich lediglich die Angabe der PSF-Breite. Diese ergibt sich aus den Messreihen mit der PMD-Kamera in Abschnitt 9.1 zu 0,30 und gilt, entgegen den veränderlichen PSF-Breiten aus Tabelle 7.5, nun für alle Skalierungsfaktoren gleichermaßen. In Tabelle 9.7 sind diejenigen Simulations- und Rekonstruktionsparameter aufgeführt, mit denen die SR-Rekonstruktionen aus den Simulationsdaten über alle vier untersuchten Skalierungsfaktoren hinweg die geringsten Abweichungen zu jenen aus den Messdaten aufweisen.

Tabelle 9.7: Die für das angepasste Beobachtungsmodell modifizierten wie auch ergänzten Simulations- und Rekonstruktionsparameter sind in Fettdruck hervorgehoben.

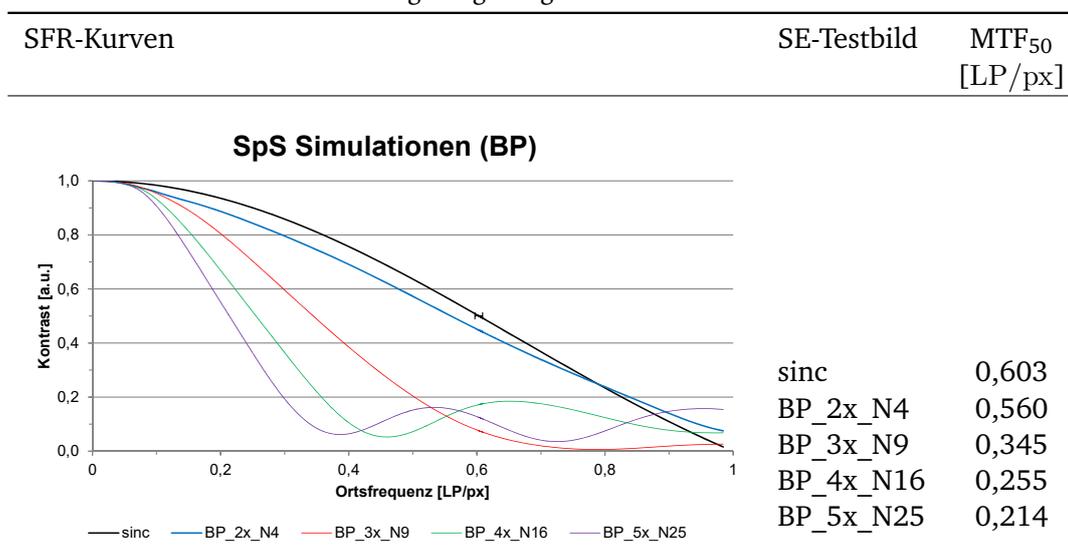
F_{Aufl}	2	3	4	5
LR-Bildauflösung		16 x 16		
Anzahl LR-Bilder	4	9	16	25
σ_B	0,44			
σ_{AWGN}	0,01			
Poisson-Rauschen	ja			
ρ_{SnP}	0,002			
PSF-Breite	0,30			
Subpixel-Versatz [Px]	1/2	1/3	1/4	1/5

9.3 Simulationsergebnisse der SpS-SR mit angepasstem Beobachtungsmodell

Nach Anpassung des Beobachtungsmodells, sind in Tabelle 9.8 die aktualisierten Simulationsergebnisse für die vier Skalierungsfaktoren in bekannter Weise gegenübergestellt. Demnach reicht einzig die Bildqualität der SR-Rekonstruktion aus 4 Eingangsbildern an die des GT-Bildes heran. Hier läuft die SFR-Kurve der Simulation (blaue Kurve) größtenteils dicht an der des theoretischen Auflösungsvermögens (schwarze Kurve) entlang. Bei einer Ortsfrequenz von knapp $0,8 \text{ LP/px}$ schneiden sich beide Kurven und die SFR-Kurve der Simulation übersteigt die der sinc-Funktion für darüber liegende Frequenzen. Der MTF_{50} -Wert für das SR-Ergebnis aus 4 LR-Bildern weicht mit $0,56 \text{ LP/px}$ rund 7 % von dem des GT-Bildes ab. Mit steigender Anzahl an LR-Eingangsbildern, und damit mit zunehmendem Skalierungsfaktor, nimmt diese Abweichung zu. Zur besseren Übersichtlichkeit erfolgt der Vergleich von Simulations- zu Messergebnissen in den Tabellen 9.9 bis 9.12 für jeden Skalierungsfaktor einzeln. Darstellungsart und Farbkodierung folgen jener in Abschnitt 9.1, wobei die rote Kurve jetzt das SR-Ergebnis nach dem angepassten Beobachtungsmodell - unter Berücksichtigung der BP-Interpolation - darstellt. Zudem wird mit der PCHIP-Interpolation für jeden Skalierungsfaktor die Untergrenze der Auflösungssteigerung beschrieben, in den Diagrammen als schwarze gestrichelte Linie dargestellt. Für die Interpolation dient jeweils das erste Bild der synthetischen LR-Bilderfolgen.

Die höchste Übereinstimmung im SFR-Kurvenverlauf zeigen die Simulations- und Messdaten bei der 5-fachen Auflösungssteigerung, dargestellt in Tabelle 9.9. Hier decken sich die SFR-Kurven aus Simulation und Messung über den gesamten relevanten Frequenzbereich bis zum ersten Nebenminimum bei knapp $0,4 \text{ LP/px}$. Darüber hinaus liegen zusätzlich das Nebenmaximum bei etwa $0,55 \text{ LP/px}$ sowie das zweite Nebenminimum bei rund $0,75 \text{ LP/px}$ deckungsgleich bei denselben Ortsfrequenzen. Lediglich die Kontrastwerte der Nebenmaxima fallen für die auf den Messdaten basierende SFR-Kurve kleiner aus. Unter Berücksichtigung der relativen Fehler-toleranz beim Bestimmen des MTF_{50} -Wertes (siehe Abschnitt 6.1), sind die MTF_{50} -Werte der Simulations- und Messergebnisse als gleich zu bewerten. Beträgsmäßig weichen die MTF_{50} -Werte der SR-Rekonstruktionen 65 % vom Idealwert der sinc-Funktion ab.

Tabelle 9.8: Ortsfrequenzganganalyse für die Subpixel-Shift SR nach Anpassung des Simulationsmodells bei 2- bis 5-facher Auflösungssteigerung.



Auch bei der 4-fachen Auflösungssteigerung, dargestellt in Tabelle 9.10, können die SFR-Kurvenverläufe aus den Simulations- und Messergebnissen über das relevante Frequenzspektrum hinweg als übereinstimmend angenommen werden. Erst über das erste Nebenminimum bei etwa 0,45 LP/px hinaus, weichen die Lagen der Nebenmaxima und -minima voneinander ab. Außerdem weist das Nebenmaximum in der Simulation einen höheren Kontrast auf als das der Messung. Beim MTF₅₀-Wert weichen Simulation und Messung mit 1,5 % geringfügig höher voneinander ab als die Fehlertoleranz groß ist. Wegen der geringen Diskrepanz werden die beiden Werte dennoch als näherungsweise gleich angenommen. Deren Abweichung zum theoretisch erreichbaren MTF₅₀-Wert beträgt rund 57 %.

Für die 3-fache Auflösungssteigerung in Tabelle 9.11 kommen zusätzlich zu den Messergebnissen aus dem Lens-Shift-Aufbau diejenigen aus dem Sensor-Shift-Aufbau (blaue Kurve) hinzu. Dort zeigt sich erstmals eine Abweichung zwischen Simulation und Messung noch vor Erreichen des ersten Nebenminimums. Zunächst laufen die SFR-Kurven aller drei SR-Ergebnisse dicht beieinander, bevor sie ab mittleren Ortsfrequenzen von etwa 0,5 LP/px auseinanderdriften. Dies führt dazu, dass die ersten Nebenminima der drei SFR-Kurven zwischen 0,6 LP/px und 0,8 LP/px gestreut liegen. Außerdem weist die SFR-Kurve aus dem Ergebnis des Lens-Shift-Aufbaus einen Kontrastanstieg im höheren Ortsfrequenzbereich um 0,8 LP/px herum auf. Der MTF₅₀-Wert deckt sich für die Simulations- und Messergebnisse und liegt somit innerhalb seiner Messtoleranz, allerdings weicht er 43 % vom Idealwert ab.

Zuletzt deckt die Ortsfrequenzganganalyse in Tabelle 9.12 für die 2-fache Auflösungssteigerung die größten Diskrepanzen zwischen Simulation und Messung auf. Dabei unterscheiden die Messkurven auch hier die SR-Ergebnisse aus Sensor-Shift- und Lens-Shift-Aufbau. Ab etwa 0,4 LP/px laufen die SFR-Kurvenverläufe aus Simulations- und Messdaten auseinander, wobei dabei das Simulationsergebnis für die höherliegenden Ortsfrequenzen stets einen höheren Kontrast aufweist. Im Ortsfrequenzbereich von 0,8 LP/px bis 1,0 LP/px übersteigt die SFR-Kurve der Simulation das theoretisch mögliche Auflösungsvermögen. Die MTF₅₀-Werte sind für die

Messungen gleich und weichen von dem der Simulation um 11 % ab. Vom Idealwert der sinc-Funktion weicht der MTF_{50} -Wert aus der Simulation um 7 % ab, der aus den Messungen dagegen um 18 %.

Tabelle 9.9: Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 5-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.

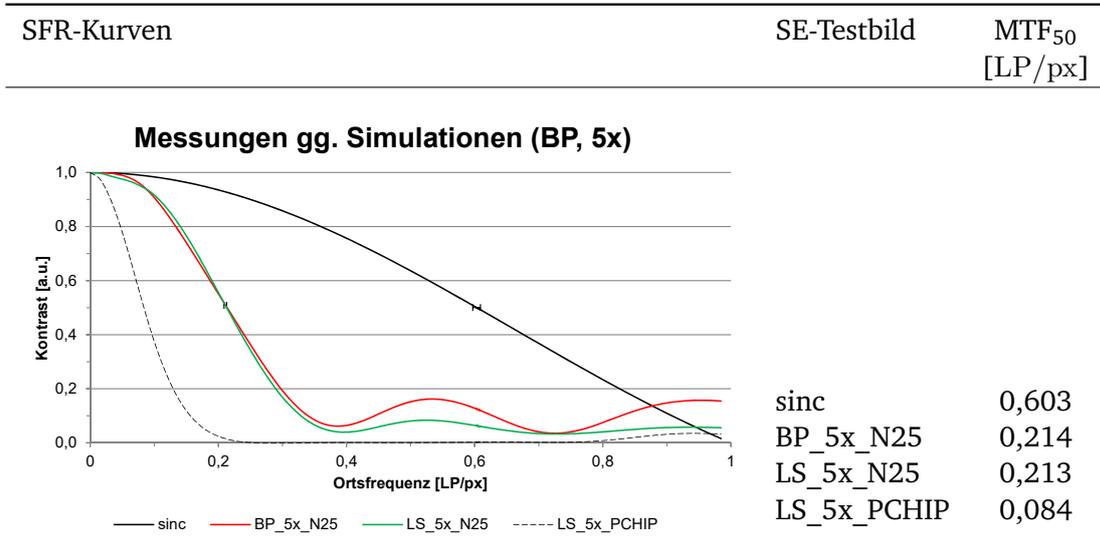


Tabelle 9.10: Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 4-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.

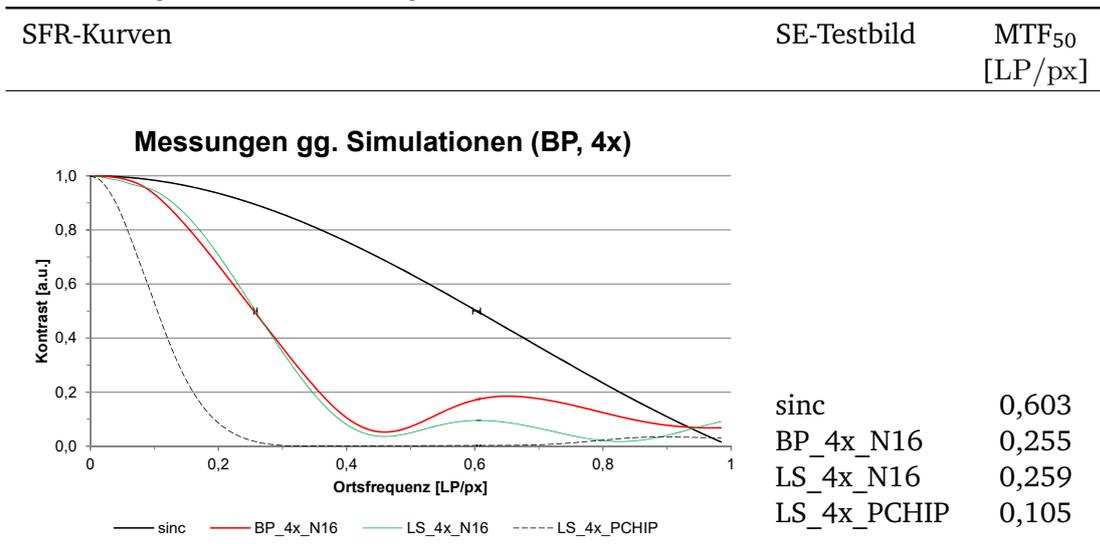


Tabelle 9.11: Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 3-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.

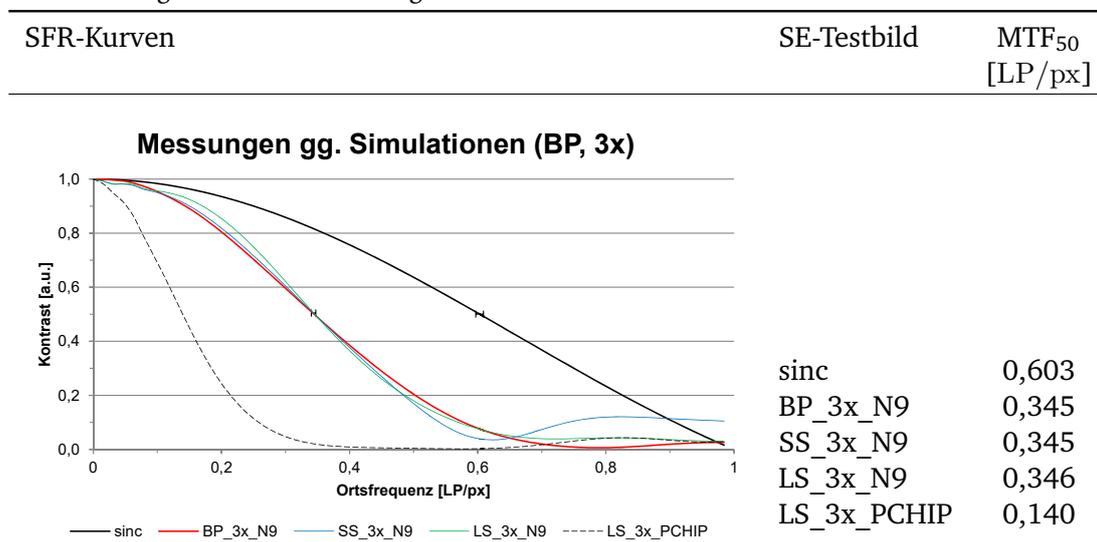
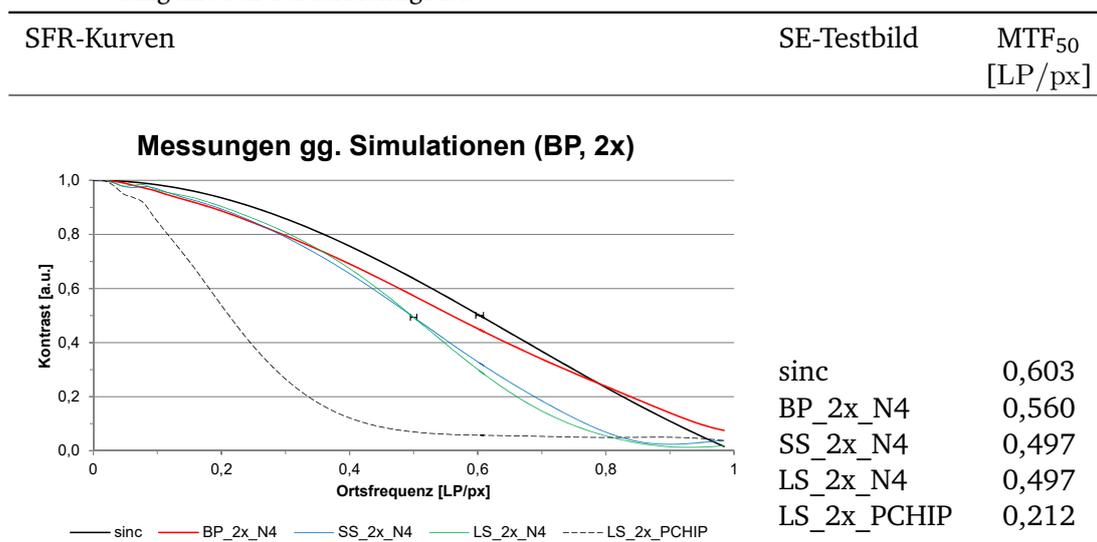


Tabelle 9.12: Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 2-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.



9.3.1 Einfluss der Parameterwahl auf das Simulationsergebnis

Die Anpassung des allgemeinen Beobachtungsmodells in Abschnitt 9.2 beschreibt die empirische Ermittlung von Simulations- und Rekonstruktionsparametern, welche in Tabelle 9.7 nochmals zusammengefasst sind. Zur besseren Nachvollziehbarkeit sind nachfolgend deren Einflüsse auf den SFR-Kurvenverlauf im SR-Ergebnis beschrieben. Die Beobachtungen beziehen sich auf das Simulationsergebnis bei 5-facher Auflösungssteigerung aus Tabelle 9.9, weil dessen Ortsfrequenzganganalyse am genauesten mit jener aus der Messung übereinstimmt. Das Verhalten gilt jedoch für jeden Skalierungsfaktor gleichermaßen.

- Das Gauß'sche Glättungsfilter mit seiner Standardabweichung σ_B verschiebt das erste Nebenminimum der SFR-Kurve. Mit zunehmendem σ_B wandert dieses nach links in den Bereich niedriger Ortsfrequenzen.
- Das additive weiße Gauß'sche Rauschen (AWGN) mit seiner Standardabweichung σ_{AWGN} verschiebt das erste Nebenminimum für große Rauschbeiträge in Richtung der kleinen Ortsfrequenzen. Zudem reduziert sich mit Erhöhung des Rauschbeitrags der Kontrast bei den hohen Ortsfrequenzen nahe dem ersten Nebenminimum.
- Die Berücksichtigung von Poisson-verteilterm Rauschen führt zu einer besseren Rekonstruktion hoher Ortsfrequenzen nahe dem ersten Nebenminimum, was sich wiederum durch dessen Wandern in Richtung hoher Ortsfrequenzen äußert.
- Das Salt-and-Pepper-Rauschen wirkt sich primär auf die Amplitude der Nebenmaxima aus, deren Kontrast mit steigender Rauschdichte ρ_{SNP} zunimmt. Zudem verschiebt sich das erste Nebenminimum mit steigendem Rauschbeitrag in Richtung der niedrigen Ortsfrequenzen.
- Die PSF-Breite beeinflusst die Position des ersten Nebenminimums sowie den SFR-Kurvenverlauf bis dorthin. Als einziger Vertreter der Rekonstruktionsparameter, dessen Angabe der IRWSR-Algorithmus erwartet, wirkt sich dieser Parameter nicht auf die Bilderzeugung sondern auf dessen Berücksichtigung bei der SR-Rekonstruktion aus. Wird die PSF-Breite zu hoch gewählt, gewichtet der Algorithmus den Unschärfebeitrag des optischen Gesamtsystems bei der SR-Rekonstruktion zu schwach und im SR-Ergebnis dominieren hochfrequente Bildanteile. Dies verursacht im SFR-Kurvenverlauf eine Kontrastüberhöhung im niederfrequenten Bereich und verschiebt das erste Nebenminimum in Richtung höherer Ortsfrequenzen. Wird die PSF-Breite dagegen zu niedrig gewählt, gewichtet der Algorithmus den Unschärfebeitrag des optischen Systems zu stark und im SR-Ergebnis dominieren niedrige Ortsfrequenzen. Durch die reduzierte Bildschärfe in der SR-Rekonstruktion fällt die SFR-Kurve bereits im Bereich niedriger Ortsfrequenzen steil ab und das erste Nebenminimum wandert in Richtung der niedrigen Ortsfrequenzen.

9.3.2 Interpretation der Simulationsergebnisse mit angepasstem Beobachtungsmodell

Nachfolgend erfolgt eine Diskussion der finalen Simulationsergebnisse aus Abschnitt 9.3. Dabei werden Unzulänglichkeiten bei der Modellierung des Rauschmodells aufgezeigt und die Leis-

tungsfähigkeit der SR-Rekonstruktion nach dem Subpixel-Shift-Ansatz bewertet. Abschließend wird der limitierende Einfluss bei der Auflösungssteigerung der hier eingesetzten niedrig auflösenden PMD-Kamera herausgearbeitet.

Abbildung des statistisch verteilten Rauschmodells

Eine Analyse der Ortsfrequenzgänge in den Tabellen 9.9 bis 9.12 zeigt, dass die Abweichungen bei den SFR-Kurvenverläufen zwischen Simulation und Messung mit abnehmendem Skalierungsfaktor zunehmen. Die Abweichungen liegen für den MTF_{50} -Wert bei 2-facher Auflösungssteigerung in Tabelle 9.12 bereits bei 11 %. Darüber hinaus weist die SFR-Kurve aus dem Simulationsergebnis für Ortsfrequenzen ab etwa $0,8 \text{ LP/px}$ einen höheren Kontrast auf als dies theoretisch möglich ist. Im Allgemeinen resultieren diese Diskrepanzen aus einem unzureichenden Rauschmodell bei der Erzeugung der synthetischen LR-Bilderfolgen. Sie fallen insbesondere dort ins Gewicht, wo eine geringe Anzahl an LR-Bildern die statistisch verteilten Rauschparameter abbilden sollen. So tragen bei 2-facher Auflösungssteigerung lediglich 4 LR-Bilder mit jeweils 16×16 Bildpunkten zur Rauschmodellierung bei, während es bei 5-facher Auflösungssteigerung mit $25 \times 16 \times 16$ Bildpunkten mehr als das Sechsfache sind. Es erscheint nachvollziehbar, dass sich ein Modell statistisch verteilter Rauschparameter mit zunehmender Anzahl an Eingangsbildern dem Rauschverhalten realer Sensoren annähert. Für die 3-fache Auflösungssteigerung entsprechen die Simulationsergebnisse in erster Näherung den Messergebnissen und erlauben eine bedingte Vorhersage der SR-Ergebnisse. Der SFR-Kurvenverlauf in Tabelle 9.11 liegt im relevanten Frequenzbereich bis zum ersten Nebenminima größtenteils auf dem aus den Messungen und die MTF_{50} -Werte sind gleich groß. Bei den darüber liegenden Skalierungsfaktoren wird die SR-Vorhersage als korrekt angenommen. Bei der 5-fachen Auflösungssteigerung in Tabelle 9.9 stimmen selbst die Positionen der Nebenextrema der Kurvenverläufe aus Simulation und Messung überein. Lediglich außerhalb des relevanten Frequenzbereichs übersteigen die Amplitudenwerte der Simulation die der Messung.

Nach der Zusammenfassung im vorherigen Abschnitt 9.3.1 beeinflussen die Rauschparameter im Wesentlichen die Positionen von Nebenminima und -maxima sowie die Amplituden der Nebenmaxima im SFR-Kurvenverlauf. Außerdem wandern mit einem zunehmendem Unschärfebeitrag die ersten Nebenminima nach links in Richtung der niedrigen Ortsfrequenzen. Für die Amplitudenausschläge der Nebenmaxima zeigt sich in erster Linie das Salt-and-Pepper-Rauschen verantwortlich, das fehlerhafte Pixelwerte modellieren soll. Streng genommen kommt dieser Rauschmechanismus nur einer groben Näherung gleich, weil er lediglich in schwarzen und weißen Bildpunkten unterscheidet. Bei realen Aufnahmen können ungültige Pixel jedoch beliebige Grauwerte annehmen. Zudem wirken sich auch das Gauß'sche Glättungsfilter und das additive weiße Gauß'sche Rauschen auf die Amplituden der Nebenmaxima aus. Die Parametrierung beschreibt letztlich das Ausloten von Unschärfeeinflüssen mit Amplitudenausschlägen in den Nebenmaxima, wobei der Unschärfe eine höhere Gewichtung beigemessen wird.

Mit Beschränkung auf den relevanten Frequenzbereich, ermöglicht die hier vorgestellte Simulation nach dem angepassten Beobachtungsmodell aussagekräftige Vorhersagen ab einem Skalierungsfaktor von 3. Die Genauigkeit der Simulationsergebnisse nimmt mit der Anzahl an LR-Bildern weiter zu, weil sich statistische Abweichungen weniger stark auswirken. Dagegen ist das hier umgesetzte Rauschmodell für die 2-fache Auflösungssteigerung zu ungenau und erlaubt dafür keine aussagekräftigen Vorhersagen hinsichtlich der SR-Rekonstruktionen.

Leistungsfähigkeit der SpS-SR bei der Auflösungssteigerung

In Tabelle 9.13 ist die Leistungsfähigkeit bei der Auflösungssteigerung nach der Subpixel-Shift SR in Form eines Performance-Faktors dargestellt. Dieser gibt das Verhältnis der MTF₅₀-Werte des SR-Ergebnisses aus dem Lens-Shift-Aufbau zu seinem PCHIP-Interpolations-Pendant an. Über alle Skalierungsfaktoren hinweg liegt er bei Werten zwischen 2,3 und 2,5. Zuletzt ist in der untersten Zeile der Tabelle die prozentuale Abweichung beim MTF₅₀-Wert vom jeweiligen SR-Ergebnis zur idealen sinc-Funktion aufgeführt.

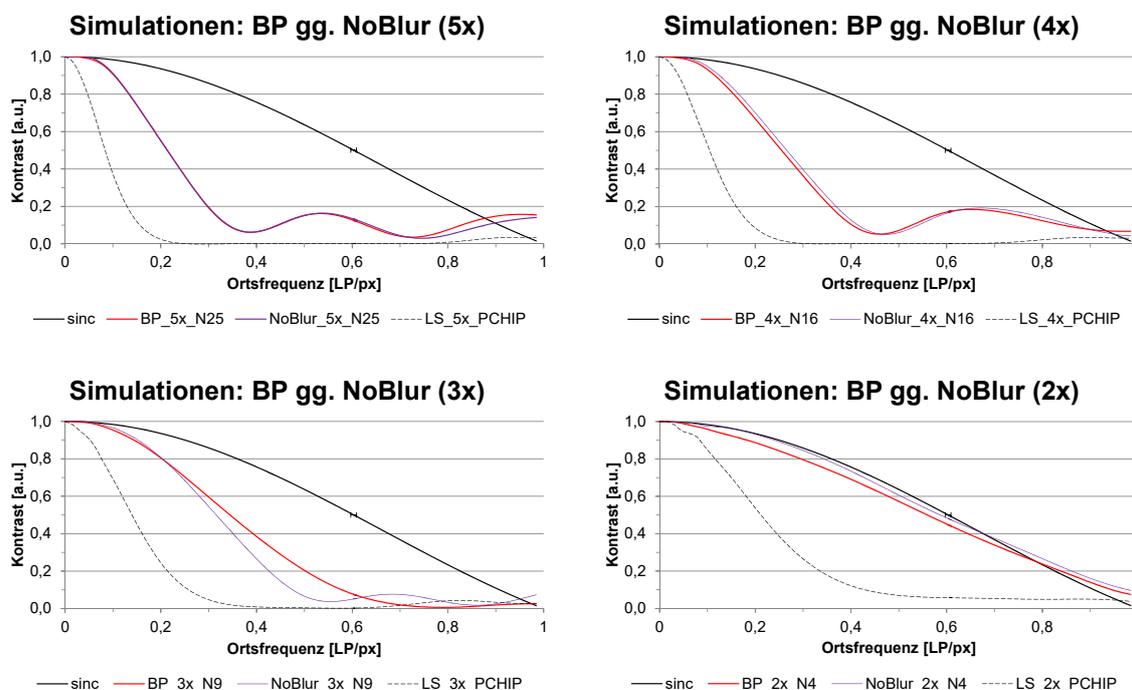
Tabelle 9.13: Leistungsfähigkeit der SpS-SR im Vergleich zur PCHIP-Interpolation, wobei sämtliche Bild-daten dem Lens-Shift-Aufbau entstammen.

Skalierung	$F_{Aufl} = 2$	$F_{Aufl} = 3$	$F_{Aufl} = 4$	$F_{Aufl} = 5$
MTF ₅₀ _LS_PCHIP [LP/px]	0,212	0,140	0,105	0,084
MTF ₅₀ _SpS_LS [LP/px]	0,497	0,346	0,259	0,213
Performance-Faktor	2,3	2,4	2,5	2,5
MTF ₅₀ _sinc [LP/px]			0,603	
Abweichung_sinc	18 %	43 %	57 %	65 %

Limitierender Faktor bei der Auflösungssteigerung

Abschließend wird in diesem Unterabschnitt geprüft, welcher Faktor letztendlich limitierend auf die Auflösungssteigerung mit der SpS-SR wirkt. Nach der Beschreibung in Kapitel 8, wird zunächst das Auflösungsvermögen der abbildenden Optik als begrenzende Größe bei der Auflösungssteigerung angenommen. Mit Blick auf den Gauß'schen Glättungsfilter, der bei den finalen Simulationen mit einer Standardabweichung von $\sigma_B = 0,44$ realisiert ist, kommen jedoch Zweifel auf. Als default-Wert gibt das IRWSR-Skript eine Standardabweichung von $\sigma_B = 0,5$ vor. Damit verglichen erscheint der umgesetzte Wert von 0,44 recht hoch. Vor allem unter Berücksichtigung der guten Abbildungsleistung beider Objektive im Vergleich zur lateralen Ortsauflösung des PMD-Sensors, wie in den Diagrammen der Abbildungen 8.1 und 8.3 in Kapitel 8 dargestellt. Eine Plausibilitätsprüfung schafft Klarheit: In einer letzten Simulation wird der Unschärfebeitrag der Optik ignoriert. Das heißt, es gelten die Simulationsparameter aus Tabelle 9.7, jedoch ohne Berücksichtigung der Gauß'schen Unschärfe. Erwartungsgemäß sollte sich nun die Bildqualität im SR-Ergebnis weiter verbessern, weil der limitierende Einfluss des Objektivs so nicht zum Tragen kommt. Die Ortsfrequenzganganalyse in Tabelle 9.14 vergleicht die Simulationsergebnisse unter Berücksichtigung der Standardabweichung von $\sigma_B = 0,44$ (BP, rote Kurve) mit den Simulationsergebnissen ohne Berücksichtigung der Standardabweichung (NoBlur, lilafarbene Kurve).

Tabelle 9.14: Die Diagramme veranschaulichen den limitierenden Einfluss der Sensorunschärfe auf die Auflösungssteigerung der PMD-Kamera: Auch für den Fall, dass keine Linsenunschärfe berücksichtigt ist, decken sich die SFR-Kurven mit jenen aus den vorhergehenden Simulation- und Messdaten.



Analog zu obiger Beschreibung, verlaufen auch bei diesem Vergleich die SFR-Kurven für höhere Skalierungsfaktoren dichter beieinander. Dennoch erlauben sie die Schlussfolgerung, dass sich die beiden Simulationsvarianten entsprechen. Die geringen Abweichungen im Kurvenverlauf bei der 2- und 3-fachen Auflösungssteigerung sind auf die wechselseitige Wirkung zweier Ursachen zurückzuführen. 1) Die randomisierten Rauschparameter unterscheiden sich für jede synthetisch erzeugte LR-Bilderfolge. 2) Wegen der wenigen Eingangsbilder sowie deren geringer Bildauflösung, wirken sich auch nur geringfügig voneinander abweichende Rauschparameter auf die jeweiligen SR-Ergebnisse aus. Bezogen auf den hier vorgestellten und untersuchten Lens-Shift-Aufbau bedeuten diese Erkenntnisse, dass der Unschärfebeitrag des Sensors die erreichte Bildqualität im SR-Ergebnis limitiert und nicht, wie ursprünglich angenommen, das Auflösungsvermögen der Abbildungsoptik. Bei der Simulation der Sensorunschärfe zeigt sich primär die BP-Interpolation mit seinem mittelwertbildenden Charakter verantwortlich.

9.3.3 Optimierungspotenzial am angepassten Simulationsmodell

Wie im vorherigen Abschnitt 9.3.2 ausgeführt, erlauben die Simulationen nach dem angepassten Beobachtungsmodell sinnvolle Vorhersagen für das SR-Ergebnis ab einem Skalierungsfaktor von 3. Liegen dagegen weniger Datenpunkte vor, versprechen Optimierungen am Rauschmodell präzisere Vorhersagen. Zum einen modelliert das aktuelle Rauschmodell rauminvariantes additives weißes Gauß'sches Rauschen und weicht somit vom raumvarianten Rauschverhalten in echten Bildaufnahmen ab [42]. Die Rausch-Modellierung nach einer gewichteten Gauß-Verteilung mit räumlich variierender Standardabweichung entspricht eher dem gemischten Rauschverhalten echter Bildaufnahmen [42] und zeigt damit Potenzial für genauere Simulationsergebnisse. Des Weiteren simuliert das umgesetzte Salt-and-Pepper-Rauschen fehlerhafte Pixelwerte nur in grober Näherung, weil es lediglich zwischen schwarzen und weißen Grauwerten unterscheidet. Bei realen Aufnahmen können diese Pixelfehler hingegen auch beliebige Zwischenwerte annehmen, deren Berücksichtigung eine Optimierung am Rauschmodell bedarf.

Speziell auf den Lens-Shift-Aufbau bezogen, könnten außerdem Änderungen beim Unschärfebeitrag durch variierende Linsenpositionen berücksichtigt werden, um ein noch allgemeiner gefasstes Simulationsmodell zu erhalten. Dies wird jedoch erst dann relevant, wenn die Abbildungsleistung der Optik die Auflösungssteigerung limitiert.

9.4 Rauschanalyse

In folgendem Abschnitt werden die Simulationsergebnisse aller drei SR-Methoden mit Hilfe einer Rauschanalyse verglichen. Diese berücksichtigt für die SpS-Methode ausschließlich die SR-Ergebnisse aus dem angepassten Beobachtungsmodell mit BP-Interpolation, weil dieses das Verhalten der realen PMD-Kamera besser simuliert als das allgemeine Beobachtungsmodell mit NN-Interpolation. Eine Rauschuntersuchung auf Basis von Messdaten ist nicht möglich, weil keine hochauflösenden GT-Aufnahmen der Objektszenen existieren und damit das Referenzbild für den Vergleich fehlt.

In Tabelle 9.15 sind die Spitzen-Signal-Rausch-Verhältnisse der SE- und USAF-Testbilder für verschiedene Skalierungsfaktoren als PSNR-Werte gegenübergestellt. Dabei entsprechen die untersuchten SR-Bilder den Rekonstruktionsergebnissen aus synthetisch erzeugten LR-Eingangsbildern, wie in Abschnitt 7.1 erläutert.

Qualitativ steht das Verhalten der PSNR-Werte im Einklang mit den Erkenntnissen aus den Ortsfrequenzganganalysen. Für das geometrisch einfach gehaltene SE-Testbild liefern alle drei SR-Methoden gute Rekonstruktionsergebnisse mit PSNR-Werten bis zu 50 dB. Unter Berücksichtigung der Anzahl an LR-Eingangsbildern, sticht insbesondere der KI-basierte ENet-PAT-Algorithmus hervor. Dieser erreicht für ein einzelnes LR-Eingangsbild und bei 4-facher Auflösungssteigerung einen PSNR-Wert von 35,6 dB. Damit liegt er gleichauf mit dem SpS-Ansatz bei nur 2-facher Vergrößerung und nahezu in der Größenordnung der FP mit 81 LR-Eingangsbildern bei ebenfalls 4-facher Auflösungssteigerung. Die FP zeigt für ansteigende Skalierungsfaktoren ein sich wiederholendes Muster beim Rauschverhalten. So erreichen die SR-Rekonstruktionen für jeden untersuchten Vergrößerungsfaktor einen PSNR-Wert von rund 50 dB, wenn nur hinreichend viele LR-Eingangsbilder zur Verfügung stehen. Wohingegen sie bei der Effizienz über alle Vergrößerungsfaktoren hinweg schlechter abschneiden als die SpS-Methode. Während die

Tabelle 9.15: Gegenüberstellung der PSNR-Werte für das Schräge-Kante- und USAF-1951-Testbild bei 2- bis 5-facher Auflösungssteigerung.

PSNR [dB]	Schräge Kante (SE)				USAF-1951			
	2x	3x	4x	5x	2x	3x	4x	5x
Ground Truth	∞	∞	∞	∞	∞	∞	∞	∞
PCHIP	27,2	24,6	23,6	22,2	18,3	16,6	16,0	15,5
ENet	-	-	35,6	-	-	-	18,4	-
FP_N9	26,2	26,1	24,9	-	23,6	19,2	18,1	-
FP_N25	40,2	31,9	28,1	-	36,3	23,7	20,3	-
FP_N49	49,3	38,3	32,3	-	44,8	29,8	23,7	-
FP_N81	49,7	45,6	38,1	-	49,6	37,4	28,2	-
FP_N121	-	47,4	38,1	-	-	46,1	31,4	-
FP_N169	-	49,5	42,5	-	-	-	40,3	-
FP_N225	-	-	49,2	-	-	-	45,9	-
SpS_BP	35,6	30,8	29,3	28,8	27,7	23,1	22,0	21,0

FP für die 2-, 3- und 4-fache Auflösungssteigerung jeweils 9, 25 und 49 Eingangsbilder benötigt, um näherungsweise die gleiche Größenordnung beim Rauschwert zu erreichen, reichen der SpS-Methode bereits 4, 9 und 16 Eingangsbilder aus. Folglich benötigt die SpS-Methode weniger Eingangsbilder als die FP, um im SR-Ergebnis annähernd die gleiche Bildqualität zu erreichen. Die Beobachtungen aus dem SE-Testbild sind größtenteils auf das USAF-Testbild übertragbar. Einzig die SR-Qualität beim ENet-PAT-Algorithmus bricht bei dem geometrisch komplexeren Testbild deutlich ein. Bei 4-facher Auflösungssteigerung liegt der PSNR-Wert mit 18,4 dB nur noch im Bereich der FP-Rekonstruktion aus 9 Eingangsbildern und unterhalb des Rauschwertes der SpS-Rekonstruktion aus 16 Eingangsbildern. Wie zuvor bei der Ortsfrequenzganganalyse in Abschnitt 7.3.1, schlägt sich die fehlerhafte Rekonstruktion der hohen Ortsfrequenzen auch im PSNR-Wert nieder.

Damit bestätigt die Rauschanalyse die Beobachtungen zur Bildqualität anhand der Ortsfrequenzganganalyse. Vorteilhaft für die Rauschanalyse ist deren Unabhängigkeit vom Testbild. So lange ein hochauflösendes Referenzbild vorliegt, lässt sich die Rauschanalyse auf nahezu beliebige Objektszenen anwenden. Für den direkten Vergleich unterschiedlicher SR-Methoden ist allerdings sicherzustellen, dass den einzelnen SR-Ergebnissen der identische Bildausschnitt zugrunde liegt.

9.5 Anwendung der Subpixel-Shift SR auf Entfernungsbilder

Abschließend erfolgt die Anwendung der SpS-SR-Methode auf die Entfernungsbilder der PMD-Kamera. Dabei soll geprüft werden, ob die lateralen Auflösungssteigerungen in Amplituden- und Entfernungsbild gleich sind. In Abbildung 9.2 ist die Aufnahme einer dreidimensionalen Objektszene gezeigt, die aus Boxen unterschiedlicher Orientierungen und Oberflächenbeschaffenheiten vor einem lichtabsorbierenden, schwarzen Hintergrund besteht. Mit dem roten Rahmen ist das Sichtfeld der PMD-Kamera visualisiert. Die zugehörigen Amplituden- und Entfernungsbilder sind in Abbildung 9.3 dargestellt, wobei die Teilbilder a) und b) dem Amplituden-

und Entfernungsbild nach Anwendung der PCHIP-Interpolation sowie c) und d) den jeweiligen SR-Ergebnissen nach Anwendung der SpS-SR-Methode auf die Messaufnahmen aus dem Lens-Shift-Aufbau entsprechen. Während die PCHIP-Interpolationen auf niedrig auflösenden Einzelaufnahmen basieren, werden zur Auflösungssteigerung mit der SpS-Methode je Bildtyp die Informationen aus 4 LR-Aufnahmen überlagert.

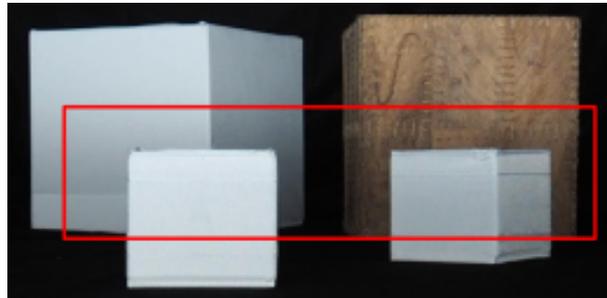


Abbildung 9.2: Objektszene mit Boxen unterschiedlicher Orientierung und Oberflächenbeschaffenheit vor einem lichtabsorbierenden Hintergrund. Das rot markierte Rechteck markiert das Sichtfeld der PMD-Kamera.

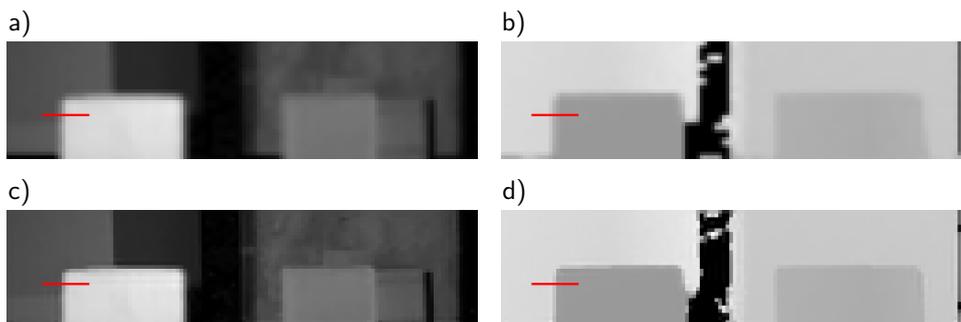


Abbildung 9.3: Die Amplitudenbilder der PMD-Kamera sind nach Anwendung der PCHIP-Interpolation in a) sowie der SpS-SR in c) abgebildet. Analog dazu sind die zugehörigen Entfernungsbilder in b) und d) dargestellt.

Rein visuell erscheinen die SR-Ergebnisse sowohl im Amplituden- als auch im Entfernungsbild schärfer als die der PCHIP-Interpolation. Ein Gradientenvergleich an der linken Kante der vordersten weißen Box (entsprechend den roten Markierungen in Abbildung 9.3) bestätigt diesen ersten Eindruck. In den Diagrammen der Abbildungen 9.4 und 9.5 sind die Grauwerte über die Bildpunkte entlang des Linienprofils aufgetragen. Je weniger Bildpunkte für den Übergang vom hellen in den dunklen Bildbereich eingenommen werden, desto schärfer wirkt seine Kante im Grauwertbild.

Beim Vergleich der beiden Bildtypen in Tabelle 9.16 zeigt sich, dass der Grauwertbereich Δ_{gl} in den Amplitudenbildern rund doppelt so groß ist wie der in den zugehörigen Entfernungsbildern. Der Grauwertbereich ergibt sich aus den minimalen und maximalen Helligkeitswerten entlang des untersuchten Linienprofils. Dabei sind die unteren bzw. oberen 10 % ausgenommen, um den Einfluss von Ausreißern zu minimieren. Die zugehörigen Pixelpositionen ergeben sich durch lineare Interpolation und werden auf eine Nachkommastelle angegeben. Das Δ_{px} beschreibt mit der Anzahl an Bildpunkten die Strecke, den der Kontrastübergang benötigt. Im interpolierten

Entfernungsbild verläuft der Kontrastübergang über 4,3 Bildpunkte, im interpolierten Amplitudenbild dagegen über 4,0. Nach Anwendung der SpS-Methode, reduziert sich die Anzahl der erforderlichen Bildpunkte um 1,6 Pixel im Entfernungsbild bzw. 1,0 Pixel im Amplitudenbild. Im Verhältnis zum Amplitudenbild wirkt sich die Auflösungssteigerung im Entfernungsbild noch deutlicher aus, wie der Vergleich der Gradientensteigungen q zeigt. Demnach nimmt die Steigung des Gradienten für das Amplitudenbild um den Faktor 1,4 und für das Entfernungsbild um den Faktor 1,6 zu.

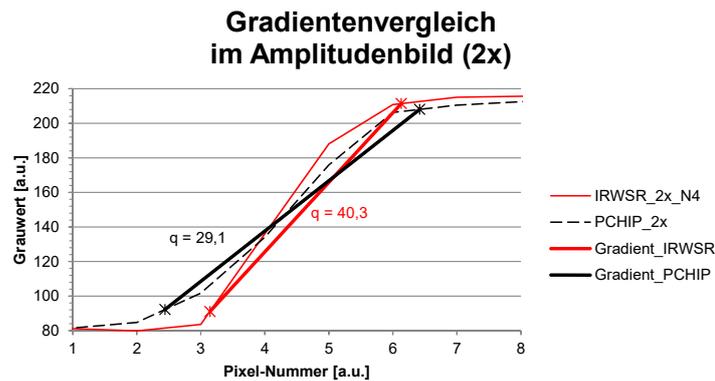


Abbildung 9.4: Gradientenvergleich im Amplitudenbild an der linken Kante der Box für die PCHIP-Interpolation und das IRWSR-Ergebnis.

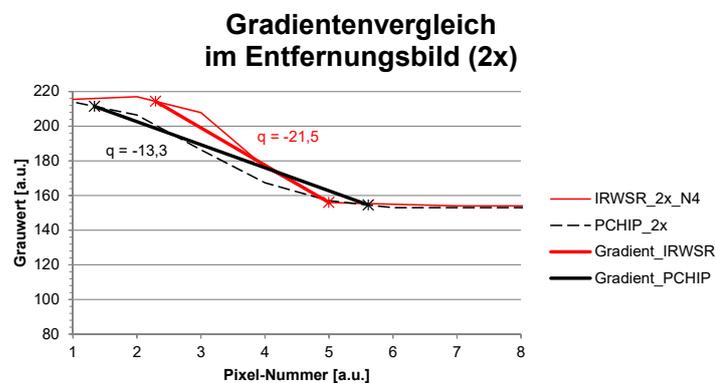


Abbildung 9.5: Gradientenvergleich im Entfernungsbild an der linken Kante der Box für die PCHIP-Interpolation und das IRWSR-Ergebnis.

Wie die Untersuchungen am Kontrastübergang einer diskreten Stufe zeigen, ist eine laterale Auflösungssteigerung im Entfernungsbild ebenfalls nachweisbar. Der Vergleich der Gradientensteigungen zeigt, dass das Entfernungsbild stärker von dem auflösungssteigernden SpS-Verfahren profitiert als das Amplitudenbild. Damit ist klar, dass sich die Auflösungssteigerung im Entfernungsbild nicht unmittelbar aus der des Amplitudenbildes ableiten lässt. Außerdem lassen die Untersuchungen keine Schlussfolgerung auf eine gesteigerte Tiefenauflösung im Entfernungsbild zu.

Tabelle 9.16: Eigenschaften der Gradientensteigungen am Kontrastübergang im Amplituden- und Entfernungsbild.

Bildtyp	Amplitudenbild		Entfernungsbild	
	PCHIP_2x	IRWSR_2x_N4	PCHIP_2x	IRWSR_2x_N4
Δgl	115,8	120,3	56,9	58,3
Δpx	4,0	3,0	4,3	2,7
Steigung q	29,1	40,3	13,3	21,5
Performance-Faktor		1,4		1,6

9.6 Abschließende Diskussion

Zur Einordnung des wissenschaftlichen Beitrags dieser Arbeit, werden die erlangten Forschungsergebnisse im Kontext der eingangs formulierten Hypothesen in Kapitel 1 diskutiert.

Hypothese 1: SR-Techniken sind in der Lage SR-Ergebnisse gesteigerter Bildqualität zu rekonstruieren, unabhängig von der bereitgestellten Bildauflösung der LR-Eingangsbilder.

Die erste Hypothese ist für alle drei untersuchten SR-Techniken gültig, wie mit den Simulations- und Messergebnissen in den Abschnitten 7.3 und 9.1 aufgezeigt. Die SISR, inkohärente FP und SpS-SR steigern das Auflösungsvermögen in deren SR-Rekonstruktionen mitunter erheblich, obwohl die LR-Eingangsbilder mit Bildauflösungen von 16 x 16 Bildpunkten für das Schräge-Kante-Testbild sowie von 60 x 60 Bildpunkten für die USAF-1951-Auflösetesttafel eine vergleichsweise geringe Datenmenge bereitstellen. Die wesentlichen Vor- und Nachteile der einzelnen SR-Methoden sind mit dem Zwischenfazit in Abschnitt 7.3.2 zusammengefasst.

Hypothese 2: Zur Bewertung der erreichten Bildqualität im SR-Ergebnis liefert die Ortsfrequenzganganalyse qualitativ dieselben Erkenntnisse wie die etablierte Rauschanalyse.

Auch diese Hypothese trifft zu, wie allen voran in Abschnitt 9.4 dargelegt ist. Ein Vergleich der einzelnen SR-Techniken nach der Ortsfrequenzganganalyse und der PSNR-Rauschanalyse attestiert den beiden Bewertungsverfahren qualitativ dieselben Erkenntnisse zur erreichten Bildqualität im SR-Ergebnis zu liefern. Die Vergleichsgrundlage hierzu bilden die Simulationsergebnisse beider Testbilder, dem Schräge-Kante-Testbild und der USAF-1951-Auflösetesttafel, bei 2- bis 5-fachem Skalierungsfaktor. Messdaten können für diese Vergleichszwecke nicht herangezogen werden, weil für die verwendete PMD-Kamera keine GT-Bilder existieren und somit kein Referenzbild für die PSNR-Bewertung zur Verfügung steht.

Hypothese 3: Für den SR-Rekonstruktionsprozess nach der Subpixel-Shift-Methode ist der Entstehungsort der Subpixel-Bewegung nicht relevant.

Die dritte Hypothese wird mit dem Vergleich der SR-Ergebnisse aus den beiden realisierten Demonstratoren in Abschnitt 9.1 bestätigt. Darin werden die Ortsfrequenzgänge der SR-Bilder bei 2-, und 3-facher Auflösungssteigerung gegenübergestellt, deren LR-Eingangsbilder einmal aus dem Sensor-Shift- und das andere mal aus dem Lens-Shift-Aufbau stammen. Der größtenteils deckungsgleiche Verlauf der SFR-Kurven belegt die Gleichstellung beider Demonstratoren und damit die Ortsunabhängigkeit der Subpixel-Bewegung.

Hypothese 4: Bei der Subpixel-Shift SR begrenzt das Auflösungsvermögen der abbildenden Optik die maximal erreichbare Auflösungssteigerung im SR-Ergebnis.

Diese Hypothese wird mit der Untersuchung zum limitierenden Faktor bei der Auflösungssteigerung nach der SpS-Methode in Abschnitt 9.3.2 widerlegt. Für die in dieser Arbeit verwendete PMD-Kamera wirkt sich der Unschärfebeitrag des PMD-Sensors einschränkender auf die erreichte Bildqualität in der SR-Rekonstruktion aus als die Auflösungsgrenze des Objektivs.

Hypothese 5: Für die Subpixel-Shift SR ist die Steigerung des lateralen Auflösungsvermögens in beiden Bildtypen, dem Amplituden- wie auch dem Entfernungsbild, betragsmäßig gleich.

Auch diese Hypothese kann widerlegt werden, wie in Abschnitt 9.5 gezeigt. Nach Anwendung der SpS-SR auf Subpixel-verschobene Amplituden- und Entfernungsbildpaare einer dreidimensionalen Objektszene, zeigt ein Gradientenvergleich abweichende Verbesserungen in den SR-Ergebnissen von Amplituden- und Entfernungsbild. Während das laterale Auflösungsvermögen im SR-Amplitudenbild um den Faktor 1,4 höher ist als bei seinem PCHIP-Vergleichsbild, ist er beim SR-Entfernungsbild um das 1,6-fache höher. Demnach scheint das Entfernungsbild stärker von der SR-Rekonstruktion zu profitieren als das Amplitudenbild.

Kapitel 10

Zusammenfassung und Ausblick

Bezugnehmend auf die in der Einleitung in Abschnitt 1 definierten Arbeitsziele, erfolgt nachfolgend eine Zusammenfassung der wesentlichen Erkenntnisse dieser Arbeit sowie ein kurzer Ausblick in potenzielle, weiterführende Forschungstätigkeiten.

Das übergeordnete Ziel zur Applikation einer generischen SR-Methode auf 3D-PMD-Kameras niedriger Bildauflösung wird mit Umsetzung der Subpixel-Shift Super-Resolution an (SpS-SR) einem neuartigen Lens-Shift-Aufbau demonstriert. Hierbei erhöht ein iterativ neu gewichteter SR-Algorithmus (IRWSR) das laterale Auflösungsvermögen im SR-Ergebnis durch die Überlagerung niedrig auflösender Mehrfachaufnahmen, ohne dabei die sensorspezifischen Eigenschaften der zugrunde liegenden Kamera zu verändern.

Die Auflösungssteigerung wird nach dem optischen Auflösungsvermögen in LP/px als absolut messbares Gütekriterium im SR-Amplitudenbild bewertet, womit die direkte Gegenüberstellung von Simulations- und Messergebnissen möglich ist. Dazu erfolgt eine Ortsfrequenzganganalyse an einem Schräge-Kante-Testbild, während die USAF-1951-Auflösungstesttafel vorrangig dem Erkennen potenzieller Bildfehler im SR-Ergebnis dient.

Im Allgemeinen sind die SR-Techniken auch bei sehr niedrig auflösenden Eingangsbildern mit Größen von lediglich 16 x 16 Bildpunkten erfolgreich. Erste Untersuchungen auf Grundlage synthetisch erzeugter und niedrig auflösender (low resolution, LR) Amplitudenbilder liefern nach ihrer SR-Methode folgende grundlegende Erkenntnisse:

- Die KI-lernbasierte *Single Image SR (SISR)* verspricht, unter Verwendung künstlicher Intelligenz, hohe Bildqualitäten bei großen Vergrößerungsfaktoren, so lange das spätere Anwendungsgebiet bekannt ist und deren wesentliche Bildmerkmale mit einem Datensatz trainiert wurden. Andernfalls können unplausible Artefakte die Bildqualität im SR-Ergebnis verringern.
- Die *inkohärente Fourier Ptychographie (FP)* ermöglicht gewissermaßen die Skalierung der Auflösungssteigerung, solange eine hinreichend hohe Anzahl an LR-Bildern zur Verfügung steht. Limitierend wirkt sich ihr Einsatzgebiet auf ausschließlich ebene Objektszenen aus, die eine Auflösungssteigerung dreidimensionaler Objekte ausschließt.
- Die *Subpixel-Shift SR (SpS-SR)* stellt eine flexible und zugleich robuste Methode zur Erhöhung der lateralen Ortsauflösung dar, deren Auflösungssteigerung aus einer Vielzahl Subpixel-verschobener LR-Bilder resultiert.

Ein Vergleich der Simulationsergebnisse zeigt, dass einzig die klassische SpS-SR imstande ist die geforderte generische Auflösungssteigerung von 3D-PMD-Kameras erfolgreich umzusetzen.

Die Gegenüberstellung der SpS-SR-Simulationsergebnisse nach dem allgemeinen Beobachtungsmodell mit Nearest Neighbor (NN)-Interpolation und den Messergebnissen deckt teils gravierende Abweichungen in den SR-Ergebnissen auf. Dabei erfährt der MTF_{50} -Wert seine größte Fehlerabweichung von rund 65 % bei einer 5-fachen Auflösungssteigerung, wie Tabelle 9.13 zu entnehmen ist. Diese Diskrepanz kann auf eine unrealistische Umsetzung des allgemeinen Beobachtungsmodells der SR zurückgeführt werden. Insbesondere dessen umgesetzte NN-Interpolation beim Downsampling, wie es von der Literatur vorgeschlagen wird, entspricht nicht dem Verhalten bei der realen Bildaufnahme mit digitalen Bildsensoren und wird durch eine mittelwertbildende Blockproc (BP)-Interpolation ersetzt. Außerdem werden beim angepassten Beobachtungsmodell der SR die Simulations- und Rekonstruktionsparameter an die Eigenschaften der verwendeten PMD-Kamera angepasst. Damit sind realistische SR-Vorhersagen unter Zuhilfenahme synthetisch erzeugter LR-Bilderfolgen ab einer 3-fachen Auflösungssteigerung möglich. Dann liegt die maximale Fehlerabweichung beim MTF_{50} -Wert in der Größenordnung seiner Messtoleranz mit einer Standardabweichung beim relativen Fehler von $< 1\%$. Bei 2-facher Auflösungssteigerung liegt die Abweichung zwischen Simulations- und Messergebnis dagegen bei verhältnismäßig großen 11 %. Dies resultiert aus einer zu geringen Datenmenge bei der SR-Rekonstruktion, womit das Simulationsmodell die statistisch verteilten Rauschparameter nicht hinreichend genau abbilden kann.

Zur Bestätigung der Simulationsergebnisse und Demonstration der Leistungsfähigkeit der SpS-SR dienen zwei Messaufbauten. Ein konventioneller Sensor-Shift-Aufbau verschiebt den Bildsensor, während der neuartige Lens-Shift-Aufbau einen transversalen Linsenversatz im Objektiv umsetzt. Dabei verspricht der Lens-Shift-Aufbau eine gegebenenfalls auch nachrüstbare Lösung zur kostengünstigen Umsetzung der SpS-SR für beliebige Digitalkameras zu sein, indem auf opto-mechanische Standardkomponenten zurückgegriffen werden kann. Ein Vergleich der SR-Ergebnisse aus den Messaufnahmen beider Aufbauten belegt, dass der Lens-Shift-Ansatz gleichwertige Bildaufnahmen generiert. Folglich ist es für die SR-Rekonstruktion nach der SpS-Methode irrelevant, wo der Subpixel-Versatz herrührt. Abgesehen von der 2-fachen Auflösungssteigerung, decken sich die SR-Ergebnisse aus den Messaufnahmen mit jenen aus den Simulationsdaten bei Berücksichtigung der mittelwertbildenden Blockproc-Interpolation.

Die Leistungsfähigkeit der SpS-SR wird mit den SR-Ergebnissen aus den Messreihen des Lens-Shift-Demonstrators ermittelt. Ein Performance-Faktor gibt für jeden Skalierungsfaktor das Verhältnis der MTF_{50} -Werte zwischen SpS-SR und PCHIP-Interpolation an. Demnach liegt die maximal erreichbare Auflösungssteigerung über alle untersuchten Messreihen hinweg zwischen einem Faktor von 2,3 und 2,5. In Bezug auf die theoretische Grenzauflösung, reichen die MTF_{50} -Werte der SR-Ergebnisse bei 2- bis 5-facher Auflösungssteigerung der Reihe nach bis auf 18 %, 43 %, 57 % und 65 % an den einer idealen sinc-Funktion heran.

Wie bereits bei der Auflistung zu den grundlegenden Erkenntnissen oben erwähnt, eignen sich die SISR und inkohärente FP nur für bestimmte Anwendungsgebiete. Dort können sie jedoch mit teilweise sehr hohen Bildqualitäten auftrumpfen. Die SpS-SR dagegen ist in ihrem Einsatzgebiet flexibler, allerdings ist ihre maximal erreichbare Auflösungssteigerung stärker von der bereitgestellten Hardware abhängig. In den hier realisierten Demonstratoren definiert nicht das Auflösungsvermögen der Objektive die Grenze bei der Auflösungssteigerung, sondern der Unschärfebeitrag des PMD-Sensors.

Abschließend wird die SpS-Methode auf gemessene, niedrig auflösende Amplituden- und Entfernungsbildpaare einer dreidimensionalen Objektszene angewandt. Ein Gradientenvergleich zeigt, dass die Auflösungssteigerungen in Amplituden- und Entfernungsbild geringfügig voneinander abweichen. Während für das Amplitudenbild das Auflösungsvermögen um das 1,4-fache über dem der PCHIP-Interpolation bei gleichem Skalierungsfaktor liegt, steigert sich das Auflösungsvermögen im Entfernungsbild um das 1,6-fache.

Ausblick

Abgesehen von dem in Abschnitt 9.3.3 angesprochenen Optimierungspotenzial beim Simulationsmodell, hält die rechnergestützte Auflösungssteigerung von PMD-Kameras im Allgemeinen weiterführende Forschungstätigkeiten bereit:

- Wie bereits die Rauschanalysen, bewertet auch der Gradientenvergleich aus Abschnitt 9.5 laterale Auflösungssteigerungen im Entfernungsbild lediglich anhand einer relativen Veränderung zwischen dem SR-Ergebnis und seinem Bezugsbild. Von Interesse kann dagegen auch ein absolutes Maß zur Bewertung des Auflösungsvermögens im Entfernungsbild sein, analog der in dieser Arbeit vorgestellten Ortsfrequenzanalyse bei den Amplitudenbildern.
- Zudem bewertet der Gradientenvergleich die Steigerung der lateralen Ortsauflösung im Amplituden- und Entfernungsbild, sprich die Auswirkung der SR-Methode auf die Bildauflösung bzw. effektive Pixelgröße im Grauwertbild. Weiterführende Forschungen könnten die Steigerung der Tiefenauflösung von PMD-Kameras mit Hilfe von SR-Methoden untersuchen.
- Außerdem könnte die Eignung von Phasenbildern, die eine PMD-Kamera ohnehin für jedes Amplituden- und Entfernungsbildpaar aufnimmt, zur Auflösungssteigerung geprüft werden. Dies könnte letztlich die Auflösungssteigerung in Echtzeit sowie von Videoaufnahmen ermöglichen.

Anhang A

Codeverzeichnis

A.1 Simulationsmodell zur inkohärenten Fourier Ptychographie

Quellcode A.1: Simulationsmodell zur inkohärenten Fourier Ptychographie.

```
1 %% Incoherent Fourier Ptychographie simulation
2 close all; clear all;
3 pkg load image;
4
5 %% Part 1: read ground truth image
6 objectIntensity = im2double(imread('GT_intensity_input.png'));
7 figure;imshow(objectIntensity,[]); title('GT_intensity_input');
8 objectFT = fftshift(fft2(objectIntensity));
9
10 %% Part 2: setup the incoherent imaging system to get an LR image
11 F_Aufl = 2; % along one direction
12 [m n] = size(objectIntensity);
13 lowpass = m/F_Aufl; % Downsampling
14 OTF = zeros(m,n);
15 OTF((end-lowpass)/2+1:(end+lowpass)/2,(end-lowpass)/2+1:(end+lowpass)
    /2) = ones; % low-pass filter limits OTF in Fourier domain
16
17 %% Part 3: Create illumination pattern and generate LR image sequence
18 arraySize = 5; % number of LR input images in one direction
19 patternNum = arraySize^2; % total number of LR input images
20 pattern = imnoise(ones(m,n),'speckle',0.5);
21 % speckle illumination pattern
22 step = 1; % scan step is equal to speckle size
23
24 % defining the scanning path:
25 for j=1:arraySize
26     xLocation(1,1+arraySize*(j-1):arraySize+arraySize*(j-1)) = -(
        arraySize-1)/2:1:(arraySize-1)/2)*step;
27     yLocation(1,1+arraySize*(j-1):arraySize+arraySize*(j-1)) = ((
        arraySize-1)/2-(j-1))*step;
28 end;
```

```
29 snakeSeq = gseq(arraySize); % call seperate function 'gseq.m'
30 patternSeq = zeros(m,n,patternNum);
31
32 % generating a speckle pattern sequence:
33 for i=1:patternNum
34     patternSeq(:,:,i) = circshift(pattern,[xLocation(snakeSeq(i))
35         yLocation(snakeSeq(i))]);
36
37 % generation of LR image sequence:
38 imSeqLowRes = zeros(m,n,patternNum);
39 for i=1:patternNum
40     lowResFT = OTF.*fftshift(fft2(objectIntensity.*patternSeq(:,:,i))
41         ); % applying low-pass filter in Fourier domain
42     imSeqLowRes(:,:,i) = abs(iff2(iff2shift(lowResFT))); % LR image
43         sequence in spatial domain
44     imSeqLowResFT(:,:,i) = fftshift(fft2(imSeqLowRes(:,:,i))); % LR
45         image sequence in Fourier domain
46 end
47 figure; imshow(imSeqLowRes(:,:,1),[]); title('Pattern_illuminated_
48     LR_input');
49
50 %% Part 4: pattern-illuminated FP recovery
51 objectRecover = sum(imSeqLowRes,3)/patternNum; % initial guess
52 for loopnum=1:10
53     for i=1:patternNum
54         object_pattern=objectRecover.*patternSeq(:,:,i); % update
55             initial guess with pattern sequence in spatial domain
56         object_pattern2=object_pattern;
57         object_patternFT=fftshift(fft2(object_pattern));
58         lowResFT1=OTF.*object_patternFT; % Downsample initial guess in
59             Fourier domain
60         im_lowRes=iff2(iff2shift(lowResFT1));
61
62         im_lowRes=imSeqLowRes(:,:,i).*exp(1i.*angle(im_lowRes)); %
63             update LR image sequence with phase information from
64             downsampled initial guess
65         lowResFT2=fftshift(fft2(im_lowRes));
66         object_patternFT=object_patternFT+conj(OTF)./(max(max((abs(OTF))
67             .^2))).*(lowResFT2-lowResFT1); % update SR estimation in
68             Fourier domain
69         object_pattern=iff2(iff2shift(object_patternFT)); % inverse
70             fast Fourier transformation
71         objectRecover=objectRecover+patternSeq(:,:,i).*(object_pattern-
72             object_pattern2)./(max(max(patternSeq(:,:,i))).^2); % update
73             SR estimate in spatial domain
```

```

61     end
62 end
63 figure;imshow(abs(objectRecover),[]); title('Pattern-illuminated_FP
        _recovery_in_spatial_domain');
64 imwrite(abs(objectRecover),['FP_SR_Recover','.png']);
65 objectRecoverFT=fftshift(fft2(objectRecover));
66 figure;imshow(log(abs(objectRecoverFT)),[]); title('Pattern-
        illuminated_FP_recovery_in_Fourier_domain');

```

Quellcode A.2: Reihenfolge, in welcher die LR-Bilder für den Rekonstruktionsprozess verwendet werden.
Aus [115] übernommen.

```

1 function seqf=gseq(arraysize)
2 n=(arraysize+1)/2;
3 arraysize=2*n-1;
4 sequence=zeros(2,arraysize^2);
5 sequence(1,1)=n;sequence(2,1)=n;
6 dx=+1;dy=-1;stepx=+1;stepy=-1;
7 direction=+1;counter=0;
8 for i=2:arraysize^2
9     counter=counter+1;
10    if (direction==+1)
11        sequence(1,i)=sequence(1,i-1)+dx;
12        sequence(2,i)=sequence(2,i-1);
13        if (counter==abs(stepx))
14            counter=0;direction=direction*-1;
15            dx=dx*-1;stepx=stepx*-1;
16            if stepx>0
17                stepx=stepx+1;
18            else
19                stepx=stepx-1;
20            end
21        end
22    else
23        sequence(1,i)=sequence(1,i-1);
24        sequence(2,i)=sequence(2,i-1)+dy;
25        if (counter==abs(stepy))
26            counter=0;direction=direction*-1;
27            dy=dy*-1;stepy=stepy*-1;
28            if (stepy>0)
29                stepy=stepy+1;
30            else
31                stepy=stepy-1;
32            end
33        end
34    end
35 end

```

```
36 seq=(sequence(1,:) -1)* arraysize+sequence(2,:);
37 seqf(1,1:arraysize ^2)=seq;
38 end
```

A.2 Erzeugen einer LR-Bilderfolge nach dem allgemeinen Beobachtungsmodell

Quellcode A.3: Extraktion einer subpixel-verschobenen LR-Bilderfolge aus einem idealen GT-Bild nach dem allgemeinen Beobachtungsmodell.

```
1 %% Generating a LR image sequence (general observation model)
2 %% Default parameters from literature with nearest-neighbor
  interpolation
3 clear all; close all;
4 pkg load image; % Octave's image package to handle matlab image
  processing commands
5
6 %% Read ground truth image
7 GT = im2double(imread('SE_ideal_48x48.png'));
8     figure(); imshow(GT,[]); title('Original_slanted_edge');
9
10 %% Parameter settings
11 F_Aufl = 2;
12 k = 4;
13 sigma_B = 0.50;
14 sigma_AWGN = 0.010;
15
16 %% Crop HR images from GT image and shift them relative to each other
  by integer pixel values
17 %% imcrop on larger image is applied to avoid disturbing edge effects
18 HR{1} = imcrop(GT, [7 14 31 31]); % [x_ini y_ini width height]
19 HR{2} = imcrop(GT, [6 14 31 31]);
20 HR{3} = imcrop(GT, [6 15 31 31]);
21 HR{4} = imcrop(GT, [7 15 31 31]);
22
23 %% Model image degradations
24 %% Downsampling + noise
25 for i = 1:k
26     s{i} = imsmooth(HR{i}, 'gaussian', sigma_B); % Isotropic Gaussian
  smoothing: Gaussian filter with standard deviation 'sigma_B'
27     s{i} = imresize(s{i},(1/F_Aufl), 'nearest'); % Downsampling with
  nearest-neighbor method
28     s{i} = imnoise(s{i}, 'gaussian', 0.0, sigma_AWGN^2); % Additive
  white Gaussian noise with zero-mean (0.0) and variance '
  sigma_AWGN^2'
29     s{i} = imnoise(s{i}, 'poisson'); % Poisson noise
```

```

30 LRImages(:,:,i) = s{i};
31 LRImages_write(:,:,i) = im2uint8(s{i}); % save LR sequence as 8bit
    grey scale images
32 imwrite(LRImages_write(:,:,i), strcat('SE_NN_LR0',num2str(i),'.
    png'));
33 end
34
35 %% Save LR image sequence as '.mat'
36 save LRseq_SE_NN_GaussBlur0p50_StdDev_mean0p0_sigma0p010_Poisson_n4.
    mat LRImages;

```

A.3 Erzeugen einer LR-Bilderfolge nach dem angepassten Beobachtungsmodell

Quellcode A.4: Extraktion einer subpixel-verschobenen LR-Bilderfolge aus einem idealen GT-Bild nach dem angepassten Beobachtungsmodell.

```

1 %% Generating a LR image sequence (adapted observation model)
2 %% Selected parameters according to the realized optical system with
    block averaging interpolation + Salt-and-pepper noise
3 clear all; close all;
4 pkg load image; % Octave's image package to handle matlab image
    processing commands
5
6 %% Read ground truth image
7 GT1 = im2double(imread('SE_ideal_48x48.png'));
8 figure(); imshow(GT1,[]); title('Original_slanted_edge');
9
10 %% Parameter settings
11 F_Aufl = 2;
12 k = 4;
13 sigma_B = 0.44;
14 sigma_AWGN = 0.010;
15 rho_SnP = 0.002;
16
17 %% Crop HR images from GT image and shift them relative to each other
    by integer pixel values
18 %% imcrop on larger image is applied to avoid disturbing edge effects
19 HR{1} = imcrop(GT, [8 11 31 31]); % [x_ini y_ini width height]
20 HR{2} = imcrop(GT, [7 11 31 31]);
21 HR{3} = imcrop(GT, [7 12 31 31]);
22 HR{4} = imcrop(GT, [8 12 31 31]);
23
24 %% Model image degradations
25 %% Downsampling + noise
26 for i = 1:k

```

```
27 s{i} = imsmooth(HR{i}, 'gaussian', sigma_B); % Isotropic Gaussian
    smoothing: Gaussian filter with standard deviation 'sigma_B'
28 s{i} = blockproc(s{i}, [F_Aufl, F_Aufl], 'mean2'); % Downsampling
    with block averaging method
29 s{i} = imnoise(s{i}, 'gaussian', 0.0, sigma_AWGN^2); % Additive
    white Gaussian noise with zero-mean (0.0) and variance '
    sigma_AWGN^2'
30 s{i} = imnoise(s{i}, 'salt_&_pepper', rho_SnP); % Salt-and-pepper
    noise with noise density 'rho_SnP'
31 s{i} = imnoise(s{i}, 'poisson'); % Poisson noise
32 LRImages(:, :, i) = s{i};
33 LRImages_write(:, :, i) = im2uint8(s{i}); % save LR sequence as 8bit
    grey scale images
34 imwrite(LRImages_write(:, :, i), strcat('SE_BP_LR0', num2str(i), '.
    png'));
35 end
36
37 %% Save LR image sequence as '.mat'
38 save LRseq_SE_BP_GaussBlur0p44_StdDev_mean0p0_sigma0p010_SnP0p002_
39 Poisson_n4.mat LRImages;
```

Literaturverzeichnis

- [1] R. Schwarte. “3D-Kamera nach Laufzeitverfahren”. Dt. Pat. DE4439298A1. 13. Juni 1996.
- [2] H. G. Heinol. “Untersuchung und Entwicklung von modulationslaufzeitbasierten 3D-Sichtsystemen”. Diss. Universität Siegen, 2001.
- [3] R. Lange. “3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology”. Diss. Universität Siegen, 2000.
- [4] R. Lange und P. Seitz. “Solid-state time-of-flight range camera”. In: *IEEE Journal of Quantum Electronics* 37.3 (März 2001), S. 390–397. ISSN: 0018-9197. DOI: 10.1109/3.910448.
- [5] R. Lange u. a. “Time-of-flight range imaging with a custom solid state image sensor”. In: Bd. 3823. 1999. DOI: 10.1117/12.360988. URL: <https://doi.org/10.1117/12.360988>.
- [6] M. Lindner und A. Kolb. “Calibration of the intensity-related distance error of the PMD TOF-camera”. In: *Intelligent Robots and Computer Vision XXV* 6764 (Sep. 2007). DOI: 10.1117/12.752808.
- [7] M. Lindner u. a. “Time-of-Flight sensor calibration for accurate range sensing”. In: *Computer Vision and Image Understanding* 114 (Dez. 2010), S. 1318–1328. DOI: 10.1016/j.cviu.2009.11.002.
- [8] M. Schmidt. “Analysis, Modeling and Dynamic Optimization of 3D Time-of-Flight Imaging Systems”. Diss. Universität Heidelberg, 2011.
- [9] S. Foix, G. Alenyà und C. Torras. “Lock-in Time-of-Flight (ToF) Cameras: A Survey”. In: *IEEE Sensors Journal* 11 (2011), S. 1917–1926.
- [10] F. Pedrotti u. a. *Optik für Ingenieure: Grundlagen*. Springer Berlin Heidelberg, 2002. ISBN: 978-3-540-22813-6.
- [11] M. Löffler-Mang. *Optische Sensorik*. Vieweg + Teubner Studium. Fertigung. Vieweg+ Teubner Verlag, 2011. ISBN: 9783834883087.
- [12] A. Rostásy und T. Sievers. *Handbuch Mediatektur: Medien, Raum und Interaktion als Einheit gestalten. Methoden und Instrumente*. transcript Verlag, 2018. ISBN: 9783839425176. URL: <https://books.google.de/books?id=0VJxDwAAQBAJ>.
- [13] Microsoft. *Azure Kinect DK hardware specifications*. <https://docs.microsoft.com/en-us/azure/kinect-dk/hardware-specification>. Original document from Microsoft. Feb. 2020.
- [14] Isrd Group. *Computer Graphics*. McGraw-Hill Education (India) Pvt Limited, 2005. ISBN: 9780070593763. URL: <https://books.google.de/books?id=WxzId7sBZiMC>.

- [15] T. Möller u. a. “Robust 3D Measurement with PMD Sensors”. In: *Proceedings of the 1st Range Imaging Research Day at ETH*. 2005, S. 3–906467.
- [16] R. Baer, M. Barfuß und D. Seifert. *Beleuchtungstechnik: Grundlagen*. HUSS-MEDIEN GmbH, Verlag Technik, 2016. ISBN: 9783341016343.
- [17] ifm electronic gmbh. *efector250 - O3D303*. <https://www.ifm.com/de/de/product/O3D303.html>. Original document from ifm electronic gmbh. Juli 2015.
- [18] ifm electronic gmbh. *efector250 - O3M151*. <https://www.ifm.com/products/ae/ds/O3M151.html>. Original document from ifm electronic gmbh. Okt. 2014.
- [19] ifm electronic gmbh. *Kamerasysteme für mobile Arbeitsmaschinen*. [https://www.ifm.com/download/files/ifm-kamerasysteme-mobile-arbeitsmaschinen-industrielle-bildverarbeitung-o3m-2018/\\$file/ifm-kamerasysteme-mobile-arbeitsmaschinen-industrielle-bildverarbeitung-o3m-2018.pdf](https://www.ifm.com/download/files/ifm-kamerasysteme-mobile-arbeitsmaschinen-industrielle-bildverarbeitung-o3m-2018/$file/ifm-kamerasysteme-mobile-arbeitsmaschinen-industrielle-bildverarbeitung-o3m-2018.pdf). Original document from ifm electronic gmbh. Jan. 2018.
- [20] ifm electronic gmbh. *3D-Sensorik für mobile Arbeitsmaschinen*. <https://www.ifm.com/binaries/content/assets/pdf-files/de/de/produktneuheiten/2016/sps/updates/ifm-o3m150-3d-sensorik-fuer-mobile-arbeitsmaschinen-d-16.pdf>. Original document from ifm electronic gmbh. Apr. 2016.
- [21] ifm electronic gmbh. *Software-Handbuch - Mobiler 3D Smart-Sensor - O3M151 - Basic Function*. <https://www.ifm.com/ch/de/product/O3M151?tab=documents>. Original document from ifm electronic gmbh. Dez. 2016.
- [22] S.C. Park, M.K. Park und M.G. Kang. “Super-resolution image reconstruction: A technical overview”. English. In: *IEEE Signal Processing Magazine* 20.3 (Mai 2003), S. 21–36. ISSN: 1053-5888. DOI: 10.1109/MSP.2003.1203207.
- [23] L. Yue u. a. “Image super-resolution: The techniques, applications, and future”. In: *Signal Processing* 128 (Mai 2016). DOI: 10.1016/j.sigpro.2016.05.002.
- [24] William T. Freeman, Thouis R. Jones und Egon C Pasztor. “Example-Based Super-Resolution”. In: *IEEE Comput. Graph. Appl.* 22.2 (2002), S. 56–65. DOI: 10.1109/38.988747.
- [25] J. Tian und K. K. Ma. “A survey on super-resolution imaging”. In: *Signal, Image and Video Processing* 5 (Sep. 2011), S. 329–342. DOI: 10.1007/s11760-010-0204-6.
- [26] K. Hayat. “Super-Resolution via Deep Learning”. In: *CoRR* abs/1706.09077 (2017). arXiv: 1706.09077. URL: <http://arxiv.org/abs/1706.09077>.
- [27] Z. Wang, J. Chen und S. C. H. Hoi. “Deep Learning for Image Super-resolution: A Survey”. In: *CoRR* abs/1902.06068 (2019). arXiv: 1902.06068. URL: <http://arxiv.org/abs/1902.06068>.
- [28] P. Shirley, M. Ashikhmin und S. Marschner. *Fundamentals of Computer Graphics*. Taylor & Francis, 2009. ISBN: 9781568814698. URL: <https://books.google.de/books?id=hJPRgRrkXvUC>.
- [29] N. Birbaumer und R.F. Schmidt. *Biologische Psychologie*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013. ISBN: 9783662007891. URL: <https://books.google.de/books?id=w9A1BgAAQBAJ>.

-
- [30] N.S. Kaisare. *Computational Techniques for Process Simulation and Analysis Using MATLAB®*. CRC Press, 2017. ISBN: 9781498762120. URL: <https://books.google.de/books?id=Y3JQDwAAQBAJ>.
- [31] K. Nasrollahi und T. Moeslund. “Super-resolution: A comprehensive survey”. In: *Machine Vision and Applications* 25 (Aug. 2014), S. 1423–1468. DOI: 10.1007/s00138-014-0623-4.
- [32] N. Nguyen, P. Milanfar und G. Golub. “A computationally efficient superresolution image reconstruction algorithm”. In: *IEEE Transactions on Image Processing* 10.4 (2001), S. 573–583.
- [33] R. Y. Tsai und T. S. Huang. “Multiframe image restoration and registration”. In: *Advances in Computer Vision and Image Processing*. 1984, S. 317–339.
- [34] S. Borman und R. L. Stevenson. “Super-resolution from image sequences-a review”. In: *1998 Midwest Symposium on Circuits and Systems (Cat. No. 98CB36268)*. 1998, S. 374–378.
- [35] M. Irani und S. Peleg. “Improving Resolution by Image Registration”. In: *CVGIP: Graph. Models Image Process.* 53.3 (Apr. 1991), S. 231–239. ISSN: 1049-9652. DOI: 10.1016/1049-9652(91)90045-L. URL: [https://doi.org/10.1016/1049-9652\(91\)90045-L](https://doi.org/10.1016/1049-9652(91)90045-L).
- [36] M. Elad und A. Feuer. “Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images”. In: *Trans. Img. Proc.* 6.12 (Dez. 1997), S. 1646–1658. ISSN: 1057-7149. DOI: 10.1109/83.650118. URL: <https://doi.org/10.1109/83.650118>.
- [37] A. Zomet und S. Peleg. “Efficient super-resolution and applications to mosaics”. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. Bd. 1. 2000, 579–583 vol.1.
- [38] A. Zomet, A. Rav-Acha und S. Peleg. “Robust super-resolution”. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. Bd. 1. 2001.
- [39] S. Farsiu u. a. “Fast and robust super-resolution”. In: *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429) 2* (2003).
- [40] M. Ben-Ezra, A. Zomet und S. K. Nayar. “Jitter Camera: High Resolution Video from a Low Resolution Detector”. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR’04*. Washington, D.C., USA: IEEE Computer Society, 2004, S. 135–142.
- [41] S. Farsiu u. a. “Fast and Robust Multiframe Super Resolution”. In: *Trans. Img. Proc.* 13.10 (Okt. 2004), S. 1327–1344. ISSN: 1057-7149. DOI: 10.1109/TIP.2004.834669. URL: <https://doi.org/10.1109/TIP.2004.834669>.
- [42] T. Köhler u. a. “Robust Multiframe Super-Resolution Employing Iteratively Re-Weighted Minimization”. In: *IEEE Transactions on Computational Imaging* 2 (März 2016), S. 42–58. DOI: 10.1109/TCI.2016.2516909.
- [43] T. Köhler. “Multi-Frame Super-Resolution Reconstruction with Applications to Medical Imaging”. Diss. Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, 2018.

- [44] M. M. Khattab u. a. “Regularization-based multi-frame super-resolution: A systematic review”. In: *Journal of King Saud University - Computer and Information Sciences* (Nov. 2018). DOI: 10.1016/j.jksuci.2018.11.010.
- [45] R. Nair u. a. “A Survey on Time-of-Flight Stereo Fusion”. In: Jan. 2013. DOI: 10.1007/978-3-642-44964-2_6.
- [46] T. Köhler u. a. “A Unified Bayesian Approach to Multi-Frame Super-Resolution and Single-Image Upsampling in Multi-Sensor Imaging”. In: *Proceedings of the British Machine Vision Conference (BMVC)*. Hrsg. von Mark W. Jones Xianghua Xie und Gary K. L. Tam. BMVA Press, Sep. 2015, S. 143.1–143.12. ISBN: 1-901725-53-7. DOI: 10.5244/C.29.143. URL: <https://dx.doi.org/10.5244/C.29.143>.
- [47] I. Eichhardt, D. Chetverikov und Z. Jankó. “Image-guided ToF depth upsampling: a survey”. In: *Machine Vision and Applications* 28 (Feb. 2017), S. 1–16. DOI: 10.1007/s00138-017-0831-9.
- [48] M. González-Audicana u. a. “Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition”. In: *Geoscience and Remote Sensing, IEEE Transactions on* 42 (Juli 2004), S. 1291–1299. DOI: 10.1109/TGRS.2004.825593.
- [49] M. G. L. Gustafsson. “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy”. In: *Journal of Microscopy* 198.2 (2000), S. 82–87. DOI: 10.1046/j.1365-2818.2000.00710.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2818.2000.00710.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2818.2000.00710.x>.
- [50] Mats G. L. Gustafsson. “Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution”. In: *Proceedings of the National Academy of Sciences* 102.37 (2005), S. 13081–13086. ISSN: 0027-8424. DOI: 10.1073/pnas.0406877102. eprint: <https://www.pnas.org/content/102/37/13081.full.pdf>. URL: <https://www.pnas.org/content/102/37/13081>.
- [51] J. García, Z. Zalevsky und D. Fixler. “Synthetic aperture superresolution by speckle pattern projection”. In: *Opt. Express* 13.16 (Aug. 2005), S. 6073–6078. DOI: 10.1364/OPEX.13.006073. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-13-16-6073>.
- [52] R. Heintzmann und M. Gustafsson. “Subdiffraction resolution in continuous samples”. In: *Nature Photonics* 3 (Juli 2009), S. 362–364. DOI: 10.1038/nphoton.2009.102.
- [53] E. Mudry u. a. “Structured illumination microscopy using unknown speckle patterns”. In: *Nature Photonics* 6 (Mai 2012), S. 312–315. DOI: 10.1038/NPHOTON.2012.83.
- [54] S. Dong u. a. “High-resolution fluorescence imaging via pattern-illuminated Fourier ptychography”. In: *Opt. Express* 22.17 (Aug. 2014), S. 20856–20870. DOI: 10.1364/OE.22.020856. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-17-20856>.
- [55] N. Chakrova u. a. “Studying different illumination patterns for resolution improvement in fluorescence microscopy”. In: *Opt. Express* 23.24 (Nov. 2015), S. 31367–31383. DOI: 10.1364/OE.23.031367. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-23-24-31367>.

-
- [56] M. Irani und S. Peleg. "Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency". In: *Journal of Visual Communication and Image Representation* 4.4 (1993), S. 324–335. ISSN: 1047-3203. DOI: <https://doi.org/10.1006/jvci.1993.1030>. URL: <http://www.sciencedirect.com/science/article/pii/S1047320383710308>.
- [57] R. C. Hardie, K. J. Barnard und E. E. Armstrong. "Joint MAP Registration and High Resolution Image Estimation Using a Sequence of Undersampled Images". In: (1996).
- [58] Z. Lin und H. Y. Shum. "Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 26.1 (Jan. 2004), S. 83–97. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2004.1261081. URL: <https://doi.org/10.1109/TPAMI.2004.1261081>.
- [59] M. Ben-Ezra u. a. "Penrose Pixels for Super-Resolution". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7 (2011), S. 1370–1383.
- [60] R. Fattal. "Image Upsampling via Imposed Edge Statistics". In: *ACM Trans. Graph.* 26.3 (Juli 2007), 95–es. ISSN: 0730-0301. DOI: 10.1145/1276377.1276496. URL: <https://doi.org/10.1145/1276377.1276496>.
- [61] J. Sun, Z. Xu und H. Y. Shum. "Image super-resolution using gradient profile prior". In: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. 2008, S. 1–8.
- [62] H. Aly und E. Dubois. "Image up-sampling using total-variation regularization with a new observation model". In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 14 (Nov. 2005), S. 1647–59. DOI: 10.1109/TIP.2005.851684.
- [63] K. Kim und Y. Kwon. "Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32 (Juli 2010), S. 1127–1133. DOI: 10.1109/TPAMI.2010.25.
- [64] Q. Shan u. a. "Fast Image/Video Upsampling". In: *ACM Trans. Graph.* 27 (Dez. 2008), S. 153. DOI: 10.1145/1457515.1409106.
- [65] H. Zhang u. a. "Non-Local Kernel Regression for Image and Video Restoration". In: Sep. 2010, S. 566–579. DOI: 10.1007/978-3-642-15558-1_41.
- [66] R. Timofte, V. De Smet und L. Van Gool. "Anchored Neighborhood Regression for Fast Example-Based Super-Resolution". In: *The IEEE International Conference on Computer Vision (ICCV)*. Dez. 2013.
- [67] C. Y. Yang, C. Ma und M. H. Yang. "Single-Image Super-Resolution: A Benchmark". In: Sep. 2014, S. 372–386. DOI: 10.1007/978-3-319-10593-2_25.
- [68] J. Yang und T. Huang. "Image Super-Resolution: Historical Overview and Future Challenges". In: Dez. 2017, S. 1–34. ISBN: 9781315217994. DOI: 10.1201/9781439819319-1.
- [69] V. Patanavijit und S. Jitapunkul. "A Lorentzian Stochastic Estimation for a Robust Iterative Multiframe Super-Resolution Reconstruction with Lorentzian-Tikhonov Regularization". In: *EURASIP Journal on Advances in Signal Processing* 2007 (Jan. 2007), S. 21–21. DOI: 10.1155/2007/34821.

- [70] S. Farsiu u. a. “Advances and challenges in Super-Resolution”. In: *International Journal of Imaging Systems and Technology* 14 (Jan. 2004). DOI: 10.1002/ima.20007.
- [71] A. Rieder. “Einführung: Was ist ein inverses Problem?” In: *Keine Probleme mit Inversen Problemen: Eine Einführung in ihre stabile Lösung*. Wiesbaden: Vieweg+Teubner Verlag, 2003, S. 1–19. ISBN: 978-3-322-80234-7. DOI: 10.1007/978-3-322-80234-7_1. URL: https://doi.org/10.1007/978-3-322-80234-7_1.
- [72] J. Kaschenz. “Regularisierung unter Berücksichtigung von Residuentoleranzen”. Doctoral Thesis. Berlin: Technische Universität Berlin, Fakultät VI - Planen Bauen Umwelt, 2006. DOI: 10.14279/depositonce-1303. URL: <http://dx.doi.org/10.14279/depositonce-1303>.
- [73] J. Johnson, A. Alahi und L. Fei-Fei. “Perceptual Losses for Real-Time Style Transfer and Super-Resolution”. In: Bd. 9906. Okt. 2016, S. 694–711. ISBN: 978-3-319-46474-9. DOI: 10.1007/978-3-319-46475-6_43.
- [74] X. Jia, H. Chang und T. Tuytelaars. “Super-Resolution with Deep Adaptive Image Resampling”. In: (Dez. 2017).
- [75] G. Lin u. a. “Image Super-Resolution Using a Dilated Convolutional Neural Network”. In: *Neurocomputing* (Sep. 2017). DOI: 10.1016/j.neucom.2017.09.062.
- [76] P. Liu, Y. Hong und Y. Liu. “Multi-Branch Deep Residual Network for Single Image Super-Resolution”. In: *Algorithms* 11 (Sep. 2018), S. 144. DOI: 10.3390/a11100144.
- [77] Z. Tang u. a. “A Joint Residual Network with Paired ReLUs activation for Image Super-Resolution”. In: *Neurocomputing* 273 (Aug. 2017). DOI: 10.1016/j.neucom.2017.07.061.
- [78] J. Yang u. a. “Image Super-Resolution Via Sparse Representation”. In: *Image Processing, IEEE Transactions on* 19 (Dez. 2010), S. 2861–2873. DOI: 10.1109/TIP.2010.2050625.
- [79] X. Zeng und L. Yang. “A robust multiframe super-resolution algorithm based on half-quadratic estimation with modified BTV regularization”. In: *Digital Signal Processing* 23 (Jan. 2013), S. 98–109. DOI: 10.1016/j.dsp.2012.06.013.
- [80] X. Song, Y. Dai und X. Qin. “Deep Depth Super-Resolution : Learning Depth Super-Resolution using Deep Convolutional Neural Network”. In: (Juli 2016), S. 13.
- [81] W. Shi u. a. “Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network”. In: (Sep. 2016).
- [82] M. S. M. Sajjadi, B. Schölkopf und M. Hirsch. “EnhanceNet: Single Image Super-Resolution through Automated Texture Synthesis”. In: *Computing Research Repository (CoRR)* abs/1612.07919 (2016). arXiv: 1612.07919. URL: <http://arxiv.org/abs/1612.07919>.
- [83] R. Timofte, V. De Smet und L. Van Gool. “A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution”. In: *ACCV*. 2014.
- [84] C. Dong u. a. “Image Super-Resolution Using Deep Convolutional Networks”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (Dez. 2014). DOI: 10.1109/TPAMI.2015.2439281.
- [85] Z. Wang u. a. “Deep Networks for Image Super-Resolution with Sparse Prior”. In: Dez. 2015, S. 370–378. DOI: 10.1109/ICCV.2015.50.

-
- [86] G. Riegler u. a. “Conditioned Regression Models for Non-blind Single Image Super-Resolution”. In: Dez. 2015, S. 522–530. DOI: 10.1109/ICCV.2015.67.
- [87] J. Kim, J. Lee und K. Lee. “Deeply-Recursive Convolutional Network for Image Super-Resolution”. In: (Nov. 2015).
- [88] Y. Wang u. a. “End-to-End Image Super-Resolution via Deep and Shallow Convolutional Networks”. In: *IEEE Access* PP (Juli 2016). DOI: 10.1109/ACCESS.2019.2903582.
- [89] J. Kim, J. Lee und K. Lee. “Accurate Image Super-Resolution Using Very Deep Convolutional Networks”. In: Juni 2016, S. 1646–1654. DOI: 10.1109/CVPR.2016.182.
- [90] Jiwon Kim, Jung Lee und Kyoung Lee. “Deeply-Recursive Convolutional Network for Image Super-Resolution”. In: Juni 2016, S. 1637–1645. DOI: 10.1109/CVPR.2016.181.
- [91] J. Yamanaka, S. Kuwashima und T. Kurita. “Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network”. In: (Juli 2017).
- [92] W. Yang u. a. “Deep Edge Guided Recurrent Residual Learning for Image Super-Resolution”. In: *IEEE Transactions on Image Processing* PP (Apr. 2016). DOI: 10.1109/TIP.2017.2750403.
- [93] Y. Tai, J. Yang und X. Liu. “Image Super-Resolution via Deep Recursive Residual Network”. In: Juli 2017. DOI: 10.1109/CVPR.2017.298.
- [94] D. Glasner, S. Bagon und M. Irani. “Super-Resolution from a Single Image”. In: *ICCV*. 2009. URL: <http://www.wisdom.weizmann.ac.il/~vision/SingleImageSR.html>.
- [95] C.Y. Yang, J. B. Huang und M. H. Yang. “Exploiting self-similarities for single frame super-resolution”. In: *Proc. IEEE Asian Conf. Comput. Vis.* (Jan. 2010), S. 497–510.
- [96] G. Freedman und R. Fattal. “Image and video upscaling from local self-examples”. In: *ACM Trans. Graph.* 28 (Jan. 2011), S. 1–10.
- [97] M. Hornacek u. a. “Depth Super Resolution by Rigid Body Self-Similarity in 3D”. In: Juni 2013, S. 1123–1130. DOI: 10.1109/CVPR.2013.149.
- [98] J. B. Huang, A. Singh und N. Ahuja. “Single Image Super-resolution from Transformed Self-Exemplars”. In: Juni 2015. DOI: 10.1109/CVPR.2015.7299156.
- [99] R. Zeyde, M. Elad und M. Protter. “On Single Image Scale-Up Using Sparse-Representations”. In: Bd. 6920. Juni 2010, S. 711–730. DOI: 10.1007/978-3-642-27413-8_47.
- [100] J. Cong und B. Xiao. “Minimizing Computation in Convolutional Neural Networks”. In: 8681 (Sep. 2014), S. 281–290. DOI: 10.1007/978-3-319-11179-7-36.
- [101] C. Gershenson. “Artificial Neural Networks for Beginners”. In: (Sep. 2003).
- [102] A. Veit, M. Wilber und S. Belongie. “Residual Networks Behave Like Ensembles of Relatively Shallow Networks”. In: *Advances in Neural Information Processing Systems* (Mai 2016).
- [103] S. Li u. a. “A Two-channel Convolutional Neural Network for Image Super-Resolution”. In: *Neurocomputing* 275 (Sep. 2017). DOI: 10.1016/j.neucom.2017.08.041.
- [104] M. Bevilacqua u. a. “Low-Complexity Single Image Super-Resolution Based on Nonnegative Neighbor Embedding”. In: (Sep. 2012). DOI: 10.5244/C.26.135.

- [105] V. Laparra u. a. “Perceptual image quality assessment using a normalized Laplacian pyramid”. In: *Electronic Imaging 2016* (Feb. 2016), S. 1–6. DOI: 10.2352/ISSN.2470-1173.2016.16.HVEI-103.
- [106] Z. Wang u. a. “Image Quality Assessment: From Error Visibility to Structural Similarity”. In: *Image Processing, IEEE Transactions on* 13 (Mai 2004), S. 600–612. DOI: 10.1109/TIP.2003.819861.
- [107] I. Goodfellow u. a. “Generative Adversarial Nets”. In: *ArXiv* (Juni 2014).
- [108] M. McCann, K. Jin und M. Unser. “A Review of Convolutional Neural Networks for Inverse Problems in Imaging”. In: *IEEE Signal Processing Magazine* 34 (Okt. 2017). DOI: 10.1109/MSP.2017.2739299.
- [109] *EnhanceNet-Code*. Eingesehen am 20.06.2020. URL: <https://github.com/msmsajjadi/EnhanceNet-Code>.
- [110] T. Y. Lin u. a. “Microsoft COCO: Common Objects in Context”. In: (Mai 2014).
- [111] R. Heintzmann und C. Cremer. “Laterally modulated excitation microscopy: Improvement of resolution by using a diffraction grating”. In: *Proc. SPIE* 3568 (Jan. 1999). DOI: 10.1117/12.336833.
- [112] S. Dong u. a. “Incoherent Fourier ptychographic photography using structured light”. In: *Photonics Research* 3 (Feb. 2015). DOI: 10.1364/PRJ.3.000019.
- [113] S. Dong u. a. “Resolution doubling with a reduced number of image acquisitions”. In: *Biomedical optics express* 6 (Aug. 2015), S. 2946–52. DOI: 10.1364/BOE.6.002946.
- [114] G. Zheng, R. Horstmeyer und C. Yang. “Wide-field, high-resolution Fourier ptychographic microscopy”. In: *Nature Photonics* 7 (Sep. 2013), S. 739–745. DOI: 10.1038/NPHOTON.2013.187.
- [115] G. Zheng. *Fourier ptychographic imaging: A MATLAB® tutorial*. Morgan & Claypool Publishers, 2016, S. 1–95. ISBN: 978-1-6817-4273-1.
- [116] H. Ur und D. Gross. “Improved resolution from subpixel shifted pictures”. In: *CVGIP: Graphical Model and Image Processing* 54 (März 1992), S. 181–186. DOI: 10.1016/1049-9652(92)90065-6.
- [117] S. Peleg, D. Keren und L. Schweitzer. “Improving image resolution using subpixel motion”. In: *Pattern Recognition Letters* 5 (März 1987), S. 223–226. DOI: 10.1016/0167-8655(87)90067-5.
- [118] H. Stark und P. Oskoui. “High-resolution image recovery from image-plane arrays, using convex projections”. In: *Journal of the Optical Society of America. A, Optics and image science* 6 (Dez. 1989), S. 1715–26. DOI: 10.1364/JOSAA.6.001715.
- [119] A. Tekalp, M. Özkan und I. Sezan. “High-resolution image reconstruction from lower-resolution imagesequences and space-varying image restoration”. In: Bd. 3. Apr. 1992, 169–172 vol.3. ISBN: 0-7803-0532-9. DOI: 10.1109/ICASSP.1992.226249.
- [120] Andrew Patti, Ibrahim Sezan und A. Tekalp. “Super Resolution Video Reconstruction with Arbitrary Sampling Lattices and Non-zero Aperture Time”. In: *IEEE Transactions on Image Processing* vol. 6 (Aug. 1997), pp. 1064–1076.

-
- [121] S. Farsiu u. a. "Robust Shift and Add Approach to Super-Resolution". In: *Proceedings of SPIE - The International Society for Optical Engineering* 5203 (Okt. 2003). DOI: 10.1117/12.507194.
- [122] H. Holling und G. Gediga. *Statistik - Wahrscheinlichkeitstheorie und Schätzverfahren*. Bachelorstudium Psychologie. Hogrefe Verlag, 2013. ISBN: 9783840921353. URL: https://books.google.de/books?id=0WeuYUaE%5C_uYC.
- [123] P. Vandewalle, S. Süssstrunk und M. Vetterli. "A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution". In: *Eurasip Journal on Applied Signal Processing* 2006 (Jan. 2006). DOI: 10.1155/ASP/2006/71459.
- [124] B. Jähne. *Digitale Bildverarbeitung*. Springer Berlin Heidelberg, 2013. ISBN: 978-3-662-06735-2. URL: <https://books.google.de/books?id=10-0BgAAQBAJ>.
- [125] P.A. Höher. *Grundlagen der digitalen Informationsübertragung: Von der Theorie zu Mobilfunkwendungen*. Springer Fachmedien Wiesbaden, 2013. ISBN: 9783834822147. URL: <https://books.google.de/books?id=rzggBAAAQBAJ>.
- [126] N. Bauer und Fraunhofer-Gesellschaft zur Förderung der Angewandten Forschung. *Handbuch zur industriellen Bildverarbeitung: Qualitätssicherung in der Praxis*. Fraunhofer-IRB-Verlag, 2007. ISBN: 9783816773863. URL: <https://books.google.de/books?id=ZmBvhgYUgOAC>.
- [127] T. Pham, L. Van Vliet und K. Schutte. "Robust Fusion of Irregularly Sampled Data Using Adaptive Normalized Convolution". In: *EURASIP Journal on Applied Signal Processing, 2006* 2006 (Jan. 2006). DOI: 10.1155/ASP/2006/83268.
- [128] J. Roland. "A study of slanted-edge MTF stability and repeatability". In: *Proceedings of SPIE - The International Society for Optical Engineering* 9396 (Feb. 2015). DOI: 10.1117/12.2077755.
- [129] H. H. Nasse. *Wie liest man MTF- Kurven?* Zeiss. Geschäftsbereich Photo-Objektive. Dez. 2008. URL: [http://www.zeiss.de/C12567A8003B8B6F/EmbedTitelIntern/CLN_30_MTF_de/\\$File/CLN_MTF_Kurven_DE.pdf](http://www.zeiss.de/C12567A8003B8B6F/EmbedTitelIntern/CLN_30_MTF_de/$File/CLN_MTF_Kurven_DE.pdf).
- [130] H. Tiziani. *Beurteilung der Bildqualität mit Hilfe der optischen Übertragungsfunktion*. 1975, S. 13.
- [131] D. Williams. "Benchmarking of the ISO 12233 Slanted-edge Spatial Frequency Response Plug-in." In: Jan. 1998, S. 133–136.
- [132] ISO. *ISO 12233:2014(en) — Photography — Electronic still picture imaging — Resolution and spatial frequency responses*. Beuth Verlag, Berlin, 2014.
- [133] F. van den Bergh. *MTF mapper v. 0.7.11*. <https://sourceforge.net/projects/mtfmapper/files/windows/>. Feb. 2019.
- [134] F. van den Bergh. *MTF Mapper user documentation*. Mai 2012.
- [135] M. Estriebeau und P. Magnan. "Fast MTF measurement of CMOS imagers using ISO 12233 slanted-edge methodology". In: *Proc. SPIE* 5251 (Jan. 2003). DOI: 10.1117/12.513320.
- [136] D. A. Kerr. *Determining MTF with a Slant Edge Target*. http://dougkerr.net/Pumpkin/articles/MTF_Slant_Edge.pdf. Okt. 2010.

- [137] K. Kohm. “Modulation transfer function measurement method and results for the Orbview-3 high resolution imaging satellite”. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 35 (Jan. 2004).
- [138] G. D. Boreman. *Modulation Transfer Function in Optical and Electro-optical Systems*. SPIE tutorial texts. SPIE Press, 2001. ISBN: 0819441430.
- [139] K. Fliegel. “Modeling and Measurement of Image Sensor Characteristics”. In: *Radioengineering* 13 (Dez. 2004).
- [140] ISO. *MIL-Standard-150A - Military Standard: Photographic Lenses*. United States Military Standards - US Air Force, Mai 1959.
- [141] E. Messina und J. M. Evans. “Standards for visual acuity”. In: *National Institute for Standards and Technology* (2006).
- [142] P. Ndajah u. a. “An investigation on the quality of denoised images”. In: *International Journal of Circuits Systems and Signal Processing* 5 (Juli 2011), S. 423–.
- [143] J. Ouwerkerk. “Image super-resolution survey”. In: *Image Vision Comput.* 24 (Okt. 2006), S. 1039–1052. DOI: 10.1016/j.imavis.2006.02.026.
- [144] *Multi-Frame Super-Resolution Toolbox*. Eingesehen am 27.07.2020. URL: <https://www5.cs.fau.de/research/software/multi-frame-super-resolution-toolbox/>.
- [145] E. Süße H. and Rodner. *Bildverarbeitung und Objekterkennung - Computer Vision in Industrie und Medizin*. SpringerVieweg, 2014. ISBN: 978-3-8348-2605-3. DOI: 10.1007/978-3-8348-2606-0.
- [146] H. Lietz, M. M. Hassan und J. Eberhardt. “Optical lens-shift design for increasing spatial resolution of 3D ToF cameras”. In: *Advanced Optical Technologies* 6.1 (2017), S. 39–46. DOI: <https://doi.org/10.1515/aot-2016-0063>. URL: <https://www.degruyter.com/view/journals/aot/6/1/article-p39.xml>.
- [147] P. Burns und D. Williams. “Sampling efficiency in digital camera performance standards”. In: *Proceedings of SPIE - The International Society for Optical Engineering* (Jan. 2008). DOI: 10.1117/12.766359.
- [148] *Produktübersicht ImageMaster® HR*. <https://trioptics.com/de/produkte/imagemaster-hr-universal-abbildungsqualitaet-mtf-pruefen/>. Eingesehen am 28.08.2020.
- [149] S. A. Guomundsson, H. Aanæs und R. Larsen. “Environmental Effects on Measurement Uncertainties of Time-of-Flight Cameras”. In: Aug. 2007, S. 1–4. ISBN: 1-4244-0969-1. DOI: 10.1109/ISSCS.2007.4292664.

Abbildungsverzeichnis

2.1	Die 4-Phasen-Messung rekonstruiert das reflektierte, sinusförmige Nutzsignal (blaue Kurve) aus vier äquidistanten Einzelmessungen $Q_1(n, m)$ bis $Q_4(n, m)$ innerhalb einer Schwingungsperiode f_{mod}	8
2.2	Der 3D-Sensor <i>O3M151</i> [18] der Firma <i>ifm electronic gmbh</i> , bestehend aus dem Empfänger (links) und Emitter (rechts), arbeitet nach dem PMD-Prinzip. Die Fotoaufnahme entstammt der Werbeschrift in [20].	11
3.1	Unter den Interpolationsergebnissen mit kontinuierlichem Signalverlauf liefert das PCHIP-Verfahren den steilsten Gradienten für sprunghafte Änderungen im Wertebereich, hier in Form ganzzahliger Helligkeitswerte.	14
3.2	Die Übersicht untergliedert sich in eine Auswahl der gängigsten Techniken (in Ebene 4) zur geometrischen SR (Ebene 1) nach ihrem Berechnungsraum (Ebene 2) sowie der benötigten Anzahl an niedrigauflösenden Eingangsbildern (Ebene 3). Ebene 5 deutet den Einsatzzweck nach einer generischen oder spezifischen SR am Beispiel der lernbasierten SISR an. In Fettdruck sind die in dieser Arbeit untersuchten Verfahren hervorgehoben.	16
3.3	Das allgemeine Beobachtungsmodell beschreibt den Prozess von einem hochauflösenden Ursprungsbild zu mehreren degradierten Bildern geringerer Bildauflösung. Zu den berücksichtigten Degradationsmechanismen zählen in sequentieller Reihenfolge Bewegung, Unschärfe, Unterabtastung und Rauschen. Letzteres in Form von additivem weißem Gauß'schen Rauschen (AWGN) und Poissonverteiltem Schrotrauschen.	17
4.1	In einem SR CNN gewichten Filter W die eingelesenen Bildinformationen, bevor sie diese an die Neuronen übergeben und dort eine Aktivierungsfunktion f_{akt} über dessen Aktivierung entscheidet. In einem Trainingsprozess werden sämtliche Filter und Funktionsparameter so lange angepasst, bis die gewünschte Bildausgabe erreicht ist.	22
4.2	Die Gegenüberstellung trägt die erreichte Bildqualität im SR-Ergebnis als PSNR über der Netzwerktiefe als Anzahl der enthaltenen Schichten für einige bekannte Vertreter der SR CNNs auf. Alle Angaben beziehen sich auf den Set5-Datensatz nach Bevilacqua et al. [104] bei 4-facher Vergrößerung und für die pixelbasierte SR-Rekonstruktion.	25

4.3	Das Minimalbeispiel nach [82] visualisiert a) ein hochauflösendes GT-Bild in Form verschiedenartig orientierter Rechtecke sowie b) sein heruntergerechnetes LR-Bild. Ferner veranschaulicht es den Unterschied einer c) konventionellen pixelbasierten SR-Rekonstruktion zu der d) merkmalsbasierten SR-Rekonstruktion nach dem ENet-PAT-Algorithmus.	26
5.1	Aus der Beleuchtung einer unbekanntnen Objektstruktur in a) mit einem bekannten Beleuchtungsmuster in b) resultiert das Überlagerungsmuster in c). Letzteres beinhaltet die Differenzfrequenzen aus der Beleuchtungsmusterfrequenz und den Ortsfrequenzen der unbekanntnen Objektstruktur. Mit der Aufnahme des Überlagerungsmusters und der Kenntnis über das Beleuchtungsmuster lässt sich die Objektstruktur rekonstruieren, sofern die Differenzfrequenzen im Bandpass des optischen Systems liegen. Abbildung übernommen aus [50].	30
5.2	Das Simulationsmodell der inkohärenten Fourier Ptychographie wechselt bei seinen Berechnungen zwischen dem Orts- (Zeile 1) und Fourierraum (Zeile 2). Im Ortsraum beleuchtet ein konstantes Speckle-Muster die Objektszene, dessen laterale Position sich für jede Einzelaufnahme in der Größenordnung der angestrebten HR-Pixelgröße verschiebt. Das Downsampling erfolgt im Orts-Frequenzraum über eine Tiefpassfilterung, dessen Bandpass die OTF des begrenzenden optischen Elements bestimmt. Die Ausgabe ist eine bandbegrenzte LR-Bilderfolge, dessen Beleuchtungssituation für jedes Einzelbild variiert.	32
5.3	Klassifikation von MFSR-Methoden im Ortsraum nach ihrer Rekonstruktionstechnik.	34
6.1	Das Schräge-Kante-Testbild in a) separiert zwei Bildbereiche unterschiedlicher Helligkeit, hier unter einem Winkel von 10° . Die USAF-1951 Auflösungsstesttafel in b) dagegen besteht aus mehreren Gruppen schwarz-weißer, horizontal und vertikal orientierter Balkenpaare steigender Ortsfrequenz.	39
6.2	Der Ortsfrequenzgang (engl. spatial frequency response; SFR) eines idealen Bildsensors mit endlicher Pixelbreite folgt einer sinc-Funktion. Der Fehlerbalken visualisiert die Standardabweichung beim relativen Fehler bei der Bestimmung des MTF_{50} -Wertes von 1 %.	41
8.1	Die Grenzfrequenz, im Diagramm als blaue Linie hervorgehoben, ist erreicht, wenn der Bildkontrast auf 10 % abgefallen ist. Nach dieser Definition übersteigt die Abbildungsleistung der Optik (6 LP/mm) die des Sensors (23 LP/mm) um den Faktor 3,8.	60
8.2	Der Lens-Shift-Aufbau ist als dreilinsiges Petzval-Objektiv realisiert, dessen Eintrittslinse mit motorgesteuerten Aktuatoren transversal verschoben wird.	61
8.3	Sowohl bei zentrierter als auch maximal dezentrierter Linsenstellung übertrifft die Grenzauflösung der abbildenden Optik die des PMD-Sensors um einen Faktor >8	61

8.4	Das SE-Testbild ist drehbar gehalten und lässt sich auf der Plattform frei im Raum positionieren. Um Einflüsse von Störlicht zu minimieren, befindet es sich in einer licht-absorbierenden Messzelle.	62
9.1	Das an die PMD-Kamera angepasste Beobachtungsmodell nutzt für das Downsampling eine mittelwertbildende BP-Interpolation und ergänzt das Rauschmodell um Salt-and-Pepper Rauschen, das näherungsweise fehlerhafte Pixelwerte simuliert.	69
9.2	Objektszene mit Boxen unterschiedlicher Orientierung und Oberflächenbeschaffenheit vor einem lichtabsorbierenden Hintergrund. Das rot markierte Rechteck markiert das Sichtfeld der PMD-Kamera.	80
9.3	Die Amplitudenbilder der PMD-Kamera sind nach Anwendung der PCHIP-Interpolation in a) sowie der SpS-SR in c) abgebildet. Analog dazu sind die zugehörigen Entfernungsbilder in b) und d) dargestellt.	80
9.4	Gradientenvergleich im Amplitudenbild an der linken Kante der Box für die PCHIP-Interpolation und das IRWSR-Ergebnis.	81
9.5	Gradientenvergleich im Entfernungsbild an der linken Kante der Box für die PCHIP-Interpolation und das IRWSR-Ergebnis.	81

Tabellenverzeichnis

2.1	Eigenschaften der PMD-Kamera <i>O3M151</i>	11
7.1	Der Ortsfrequenzgang aller SE-GT-Bilder stimmt mit dem aus der Theorie zu erwartenden Kurvenverlauf nach der sinc-Funktion überein.	44
7.2	Die SE- und USAF-1951-Testbilder sind nach ihren GT-Bildern sowie den Simulationsergebnissen der einzelnen SR-Methoden (PCHIP, ENET, FP und SpS mit NN-Interpolation) aufgelistet. In Fettdruck sind der Skalierungsfaktor sowie die LR-Bildanzahl hervorgehoben.	45
7.3	Simulationsparameter zur inkohärenten FP: Neben den GT-Bildern erwartet das Skript Angaben zum Skalierungsfaktor F_{Aufl} sowie zur Anzahl der niedrig aufgelösten Eingangsbilder.	47
7.4	Parametersatz zur Erstellung der LR-Bilderfolge für die Simulation der SpS SR mit NN-Interpolation.	48
7.5	Parametersatz zur SR-Rekonstruktion nach dem IRWSR-Algorithmus aus der synthetisch erzeugten LR-Bilderfolge mit NN-Interpolation.	48
7.6	Ortsfrequenzganganalyse für die PCHIP-Simulation bei 2-, 3-, 4- und 5-facher Auflösungssteigerung.	49
7.7	Ortsfrequenzganganalyse für die EnhanceNet SISR bei 4-facher Auflösungssteigerung.	49
7.8	Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 2-facher Auflösungssteigerung.	50
7.9	Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 3-facher Auflösungssteigerung.	50
7.10	Ortsfrequenzganganalyse für die inkohärente Fourier Ptychographie bei 4-facher Auflösungssteigerung.	51
7.11	Abhängigkeit der Bildqualität im SR-Ergebnis bei 2-facher Auflösungssteigerung von der Anzahl an LR-Eingangsbildern (v.l.n.r.): $N=9$, $N=25$, $N=49$ und $N=81$	51
7.12	Abhängigkeit der Bildqualität im SR-Ergebnis bei 4-facher Auflösungssteigerung von der Anzahl an LR-Eingangsbildern (v.l.n.r.): $N=9$, $N=25$, $N=49$, $N=81$, $N=121$ und $N=169$	52

7.13 Ortsfrequenzganganalyse für die Subpixel-Shift SR nach dem IRWSR-Algorithmus mit NN-Interpolation bei 2- bis 5-facher Auflösungssteigerung.	52
9.1 Parametersatz für den IRWSR-Algorithmus zur SR-Rekonstruktion aus den aufgenommenen LR-Bilderfolgen.	64
9.2 Die Ortsfrequenzganganalyse der SR-Ergebnisse aus den Messreihen beider Demonstrator-Varianten zeigt eine hohe Übereinstimmung. Damit belegt sie die Gleichstellung beider Varianten bei der Erzeugung niedrig auflösender Bilderfolgen.	65
9.3 Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 2-facher Auflösungssteigerung.	66
9.4 Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 3-facher Auflösungssteigerung.	66
9.5 Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 4-facher Auflösungssteigerung.	67
9.6 Gegenüberstellung von Erwartungs- und Messwerten mit der Ortsfrequenzganganalyse bei 5-facher Auflösungssteigerung.	67
9.7 Die für das angepasste Beobachtungsmodell modifizierten wie auch ergänzten Simulations- und Rekonstruktionsparameter sind in Fettdruck hervorgehoben. . .	70
9.8 Ortsfrequenzganganalyse für die Subpixel-Shift SR nach Anpassung des Simulationsmodells bei 2- bis 5-facher Auflösungssteigerung.	71
9.9 Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 5-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.	72
9.10 Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 4-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.	72
9.11 Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 3-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.	73
9.12 Gegenüberstellung von Simulations- und Messergebnissen nach der Ortsfrequenzganganalyse bei 2-facher Auflösungssteigerung, wobei für die Simulation das angepasste Beobachtungsmodell berücksichtigt ist.	73
9.13 Leistungsfähigkeit der SpS-SR im Vergleich zur PCHIP-Interpolation, wobei sämtliche Bilddaten dem Lens-Shift-Aufbau entstammen.	76
9.14 Die Diagramme veranschaulichen den limitierenden Einfluss der Sensorunschärfe auf die Auflösungssteigerung der PMD-Kamera: Auch für den Fall, dass keine Linsenunschärfe berücksichtigt ist, decken sich die SFR-Kurven mit jenen aus den vorhergehenden Simulation- und Messdaten.	77

9.15 Gegenüberstellung der PSNR-Werte für das Schräge-Kante- und USAF-1951-Testbild bei 2- bis 5-facher Auflösungssteigerung.	79
9.16 Eigenschaften der Gradientensteigungen am Kontrastübergang im Amplituden- und Entfernungsbild.	82

Abkürzungs- und Symbolverzeichnis

Abkürzungen

2D	Zweidimensional
3D	Dreidimensional
BP	Blockproc, hier als Funktion zur Umsetzung einer mittelwertbildenden Interpolation gemeint
BTV	Bilateral total variation (zur Regularisierung)
ANN	Artificial neural network, dt. künstliches neuronales Netz
AWGN	Additive white gaussian noise, dt. additives weißes Gauß'sches Rauschen
CNN	Convolutional neural network, dt. faltendes neuronales Netz
ENet	EnhanceNet SISR-Algorithmus
ESF	Edge spread function, dt. Kantenspreizfunktion
GAN	Generative adversarial network, dt. erzeugendes kontradiktorisches Netz
GT	Ground truth, dt. Grundwahrheit
FFT	Fast Fourier transformation, dt. schnelle Fourier Transformation
FP	Fourier Ptychographie
HR	High resolution, dt. hochauflösend
I	Image, dt. Bild
IBP	Iterative back-projection SR-Algorithmus
IR	Infrared, dt. Infrarot
IRLS	Iteratively reweighted least squares, dt. iterativ gewichtete kleinste Quadrate (zur Regularisierung)
LED	Light-emitting diode, dt. lichtemittierende Diode
LiDAR	Light detection and ranging
LP	Line pair, dt. Linienpaar
LR	Low resolution, dt. niedrigauflösend
LS	Lens-Shift, dt. Linsenversatz

LSF	Line spread function, dt. Linienspreizfunktion
IRWSR	Iterativley re-weighted super-resolution SR-Algorithmus
MAE	Mean absolute error, dt. mittlerer absoluter Fehler
MAP	Maximum-a-Posteriori (probability)
MFSR	Multi-Frame SR, dt. SR aus Mehrfachaufnahmen
ML	Maximum likelihood, dt. höchste Wahrscheinlichkeit
MRF	Markov random field, dt. Markow-Netzwerk (zur Regularisierung)
MSE	Mean squared error, dt. mittlerer quadratischer Fehler
MTF	Modulation transfer function, dt. Modulationsübertragungsfunktion
MTF₅₀	MTF-Wert bei einem Kontrastabfall auf 50%
NIR	Near-infrared, dt. nahes Infrarot
NN	Nearest Neighbor, dt. nächster Nachbar
OIS	Optical image stabilizer, dt. optische Bildstabilisierung
OTF	Optical transfer function, dt. optische Übertragungsfunktion
PCHIP	Piecewise cubic Hermite interpolating polynomial, dt. stückweise kubisch interpolierendes Hermite-Polynom
PDF	Probability density function, dt. Wahrscheinlichkeitsdichtefunktion
PMD	Photonic mixing device, dt. Photomischdetektor
POCS	Projection on convex sets SR-Algorithmus
PSF	Point spread function, dt. Punktspreizfunktion
PSNR	Peak signal-to-noise ratio, dt. Spitzen-Signal-Rausch-Verhältnis
px	Pixel, dt. Bildpunkt
RMSE	Root mean square error, dt. Wurzel der mittleren Fehlerquadratsumme
SE	Slanted edge, dt. schräge Kante (hier das Schräge-Kante-Testbild)
SFR	Spatial frequency response, dt. Ortsfrequenzgang
SIM	Structured illumination microscopy, dt. Mikroskopie mit strukturierter Beleuchtung
SISR	Single Image SR, dt. SR aus Einzelbildern
SNR	Signal-to-noise ratio, dt. Signal-Rausch-Verhältnis
SpS	Subpixel-Shift
SR	Super-Resolution
SS	Sensor-Shift, dt. Sensorversatz

ToF	Time-of-Flight, dt. Laufzeit
TV	Total variation (zur Regularisierung)
USAF	United States Air Force (hier das USAF-1951-Testbild)
VGA	Video Graphics Array
WBTv	Weighted bilateral total variation (zur Regularisierung)
WKS	Whittaker-Kotelnikow-Shannon Abtasttheorem

Symbolzeichen

α	Räumlich adaptive Gewichtung
α_t	Räumlich adaptive Gewichtung zur Iteration t
A_t	Diagonalmatrix mit Gewichtung α_t
β_t	Beobachtungsgewicht
β_t	Beobachtungsgewicht zur Iteration t
B_t	Diagonalmatrix mit Gewichtung β_t
B_k	Blurring-Operator zur Beschreibung der Unschärfe im Bild k
c	Lichtgeschwindigkeit = $2,998 \cdot 10^8$ m/s
γ	Paraxiale Vergrößerung
Δ_{gl}	Grauwertbereich beim Gradientevergleich
Δ_{px}	Pixelabstand beim Gradientevergleich
Δt	Äquidistante Zeitkonstante bei der 4-Phasen-Messung
$d(n, m)$	Entfernungswerte von Pixel (n,m)
D	Downsampling-Operator zur Beschreibung der Unterabtastung
D_{net}	Diskriminator-Netz
$\mathcal{F}\{\cdot\}$	Fouriertransformation
f_{act}	Aktivierungsfunktion der Neuronen
f_{mod}	Modulationsfrequenz des modulierten Lichts
F_{Aufl}	Auflösungssteigerungsfaktor
$F(I^{SR})$	Energieminimierungsfunktion
$F(I_t^{SR})$	Energieminimierungsfunktion zur Iteration t
G	Gram-Matrix

G_{net}	Generator-Netz
H_k	Zusammenfassung einzelner Bilddegradation zur Systemdegradation
I_{max}	Maximaler Helligkeitswert im Bild
I_{min}	Minimaler Helligkeitswert im Bild
I^{kub}	Bikubisch interpoliertes LR-Bild
I^{GT}	GT-Bild
I_i^{GT}	i-ter Bildpunkt im GT-Bild
I_{est}^{GT}	Erstabschätzung eines GT-Bildes bei der Fourier Ptychographie
I_k^{HR}	Erstabschätzung eines hochauflösenden Bildes bei der Fourier Ptychographie
$I_{updated,k}^{HR}$	Mit k-tem LR-Bild aktualisiertes hochauflösendes Bild bei der Fourier Ptychographie
I_k^{LR}	k-tes LR-Bild einer Bilderfolge
I^{SR}	SR-Bild
I_t^{SR}	SR-Bild zur Iteration t
I_i^{SR}	i-ter Bildpunkt im SR-Bild
I^{res}	Residuen-Bild/Differenz-Bild
k	Index zur Beschreibung der Anzahl an Einzelbildern innerhalb einer LR-Bilderfolge
k_{Bel}	Dem Beleuchtungsmuster aufgeprägte Ortsfrequenzen
k_{Diff}	Differenzfrequenzen im Überlagerungsmuster
k_{Obj}	In der Objektszene enthaltene Ortsfrequenzen
k_{Opt}	Grenzfrequenz, die ein optisches System übertragen kann
λ	Regularisierungsparameter/Kompromissparameter
λ_t	Regularisierungsparameter/Kompromissparameter der Iteration t
$\mathcal{L}_{\mathcal{A}}$	Kontradiktorische Verlustfunktion im ENet-PAT-Algorithmus
$\mathcal{L}_{\mathcal{P}}$	Wahrnehmungsverlust im ENet-PAT-Algorithmus
$\mathcal{L}_{\mathcal{T}}$	Verlustfunktion zur Texturanpassung im ENet-PAT-Algorithmus
$\mathcal{L}\{\cdot\}$	Verlustfunktion
l_1	l_1 -Norm = Taxicab Norm
l_2	l_2 -Norm = euklidische Norm
l_p	l_p -Norm
M	Anzahl an Bildpunkten innerhalb eines Bildes

M_{horiz}	Anzahl horizontaler Bildpunkte
M_{vert}	Anzahl vertikaler Bildpunkte
M_k	Motion-Operator zur Beschreibung von Relativbewegungen aufeinanderfolgender Einzelbilder
n_k	Noise-Operator zur Beschreibung von Rauschen im Bild k
φ_{mod}	Phasendifferenz des modulierten Lichts zwischen Sende- und Empfangssignal
$p(I_k^{LR})$	A-priori-Wahrscheinlichkeit der LR-Bilder
$p(I_k^{LR} I^{SR})$	Die bedingte Wahrscheinlichkeit, dass die LR-Bilder in einem SR-Ergebnis enthalten sind.
$p(I^{SR})$	A-priori-Wahrscheinlichkeit des SR-Bildes
$p(I^{SR} I_k^{LR})$	Die bedingte a-posteriori-Wahrscheinlichkeit, dass das latente SR-Ergebnis aus einem Satz niedrig auflösender LR-Bilder resultiert.
P_n	Beleuchtungsmuster im Ortsraum
m	Steigung m beim Gradientevergleich
$Q(n, m)$	Ladungsträgermengen von Pixel (n, m)
$\hat{\theta}$	Optimierungsfunktion/Zielfunktion
ρ_{SnP}	Rauschdichteparameter des Salt-and-Pepper-Rauschen
σ_B	Standardabweichung des Gauß'schen Glättungsfilters
σ_{AWGN}	Standardabweichung des additiven weißen Gauß'schen Rauschen
S	Skalierungsmatrix
s	Pixelbreite in horizontaler Richtung
t	Anzahl Iterationen bei der SR-Rekonstruktion
$\phi(I^{GT})$	Betrachtung des GT-Bildes im Merkmalsraum
$\phi(I^{SR})$	Betrachtung des SR-Bildes im Merkmalsraum
$\Phi\{\cdot\}$	Regularisierungsterm

Publikationsliste

Zeitschriften

H. Lietz, M. M. Hassan und J. Eberhardt. „Optical lens-shift design for increasing spatial resolution of 3D ToF cameras“. In: *Advanced Optical Technologies* vol. 6., no. 1, 2017, pp. 39-46. <https://doi.org/10.1515/aot-2016-0063>

Papers und Proceedings

- H. Lietz und J. Eberhardt. „Introduction to Fourier Ptychographic Imaging for 3D ToF Cameras“. In: *Engineering for a Changing World: 59th IWK, Ilmenau Scientific Colloquium, Technische Universität Ilmenau*, 2017. urn:nbn:de:gbv:ilm1-2017iwk:2
- M. M. Hassan, S. Malorodov, H. Lietz und J. Eberhardt. „Evaluation of consumer grade Time of Flight cameras for marker-less motion capture“. In: *TAR 2017 – Technically Assisted Rehabilitation Current Directions in Biomedical Engineering*, vol. 3, no. s1, 2017, pp. s1-s18. <https://doi.org/10.1515/cdbme-2017-1001>
- H. Lietz und J. Eberhardt. „Super-resolution PMD camera for applied metrology“. In: *Proc. SPIE 9751, Smart Photonic and Optoelectronic Integrated Circuits XVIII*, 975115, SPIE OPTO, 2016. <https://doi.org/10.1117/12.2209269>
- H. Lietz und J. Eberhardt. „Contribution to the standardization of 3D measurements using a high-resolution PMD camera“. In: *Proc. SPIE 9628, Optical Systems Design 2015: Optical Fabrication, Testing, and Metrology V*, 96281H SPIE Optical Systems Design, 2015. <https://doi.org/10.1117/12.2191042>
- H. Lietz, M. Kreh, P. Albrecht und J. Eberhardt. „Resolution enhancement of large pixel PMD sensors“. In: *DGaO-Proceedings 2015* (online). Deutsche Gesellschaft für angewandte Optik, 2015. ISSN: 1614-8436 - urn:nbn:de:0287-2015-B021-0
- H. Lietz, J. Baumgart und J. Eberhardt. „Entwicklung einer polychromatischen Beleuchtungseinheit für PMD-Kameras“. In: *DGaO-Proceedings 2014* (online). Deutsche Gesellschaft für angewandte Optik, 2014. ISSN: 1614-8436 - urn:nbn:de:0287-2014-B004-1

Poster und Konferenzbeiträge

- H. Lietz, M. Franzen, J. Eberhardt und S. Sinzinger. „Optimized deep learning algorithms for application with data from PMD cameras“. In: *DGaO-Proceedings 2018* (online, Poster). Deutsche Gesellschaft für angewandte Optik, 2018. ISSN: 1614-8436 -urn:nbn:de:0287-2018-P028-5
- H. Lietz und J. Eberhardt. „3D-PMD-Kameras in der Kraftfahrzeugtechnik – Potenzial und Limitierung“. In: *Anwenderforum Optische Sensorik im Automobilbereich*, Optence e.V., 2016.
- H. Lietz und J. Eberhardt. „Hochauflösende 3D-PMD-Kamera in der Kraftfahrzeugtechnik“. In: *Optische Technologien in der Fahrzeugtechnik*, 7. VDI-Tagung, VDI-Berichte Band Nummer 2278, 2016. ISBN 978-3-18-092278-2. ISSN 0083-5560

Fachvorträge

- H. Lietz, M. Franzen, J. Eberhardt und S. Sinzinger. „Optimized deep learning algorithms for application with data from PMD cameras“. *119. Jahrestagung der Deutschen Gesellschaft für angewandte Optik*, Aalen, 2018.
- H. Lietz und J. Eberhardt. „Introduction to Fourier ptychographic imaging for 3D ToF cameras“. *Engineering for a Changing World: 59th IWK, Ilmenau Scientific Colloquium, Technische Universität Ilmenau*, 2017.
- H. Lietz und J. Eberhardt. „Hochauflösende 3D-PMD-Kamera in der Kraftfahrzeugtechnik“. *Optische Technologien in der Fahrzeugtechnik*, 7. VDI-Tagung, Karlsruhe, 2016.
- H. Lietz und J. Eberhardt. „3D-PMD-Kameras in der Kraftfahrzeugtechnik – Potenzial und Limitierung“. *Anwenderforum Optische Sensorik im Automobilbereich*, Optence e.V., Bad Kreuznach, 2016.
- H. Lietz und J. Eberhardt. „Developing a PMD Jitter Camera“. *10th International Conference on Optics-Photonics Design & Fabrication (ODF '16)*, Weingarten, 2016.
- H. Lietz und J. Eberhardt. „Super-resolution PMD camera for applied metrology“. *Smart Photonic and Optoelectronic Integrated Circuits XVIII, 975115, SPIE OPTO*, San Francisco, USA, 2016.
- H. Lietz und J. Eberhardt. „Contribution to the standardization of 3D measurements using a high-resolution PMD camera“. *Optical Systems Design 2015: Optical Fabrication, Testing, and Metrology V, 96281H SPIE Optical Systems Design*, Jena, Germany, 2015.
- H. Lietz, M. Kreh, P. Albrecht und J. Eberhardt. „Resolution enhancement of large pixel PMD sensors“. *116. Jahrestagung der Deutschen Gesellschaft für angewandte Optik*, Brno, 2015.
- H. Lietz, J. Baumgart und J. Eberhardt. „Entwicklung einer polychromatischen Beleuchtungseinheit für PMD-Kameras“. *115. Jahrestagung der Deutschen Gesellschaft für angewandte Optik*, Karlsruhe, 2014.
- H. Lietz, J. Eberhardt und J. Baumgart. „An approach to generate colored images with PMD-cameras“. *European Optical Society Annual Meeting (EOSAM) 2014*, Berlin, 2014.
- H. Lietz, J. Baumgart und J. Eberhardt. „Entwicklung und Realisierung eines optischen Designs zur Kollimation von Multichip-LEDs“. *IEEE Workshop 2014 Industrielle Messtechnik & Kraftfahrzeugsensorik*, Mühlheim an der Ruhr, 2014.

Betreuung studentischer Abschlussarbeiten

- S. Peterfaj. „Aufbau und Programmierung einer mobilen Roboterplattform mit einer Microsoft Kinect zur Hinderniserkennung“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2014.
- D. Stopper. „Radiometrische Farbmessung texturierter Oberflächen“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2014.
- P. Albrecht. „Multiframe-Superresolution basierend auf einer Jitter-Kamera“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2015.
- M. Kreh. „Simulation optomechanischer Systeme zur Umsetzung eines Multiframe-Superresolution-Verfahrens für 3D-PMD-Kameras“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2015.
- F. Bürker. „Planung, Konstruktion, Bau, Programmierung einer flugfähigen Trägerplattform für Bildverarbeitungssysteme“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2017.
- M. Sträßle. „Programmierung einer flugfähigen Trägerplattform für Bildverarbeitungssysteme“. Hochschule Ravensburg-Weingarten, Bachelorthesis, 2017.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise ~~entgeltlich~~/unentgeltlich geholfen:

1. Herr Sebastian Braner bei der Vermessung der beiden untersuchten optischen Systeme mit einem MTF-Messtand bei der Firma *Arri AG*, München.
2. Herr Sebastian Ritter bei der Ansteuerung des motorgesteuerten Linsenhalters des Lens-Shift-Demonstrators.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt.

Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 10 der Promotionsordnung den Abbruch des Promotionsverfahrens zur Folge hat.

Metzingen, 14.03.2021

Ort, Datum

Henrik Lietz

Unterschrift