

# Super-Resolution Algorithmen für eine leistungsfähige Videosignalverarbeitung

von der Fakultät

Elektrotechnik und Informationstechnik

der Technischen Universität Dortmund

genehmigte

**Dissertation**

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

von

**Martin Max Richter**

Dortmund, 2011

Tag der mündlichen Prüfung: 08.02.2012

Hauptreferent: Prof. Dr.-Ing. Hartmut Schröder

Korreferent: Prof. Dr.-Ing. Holger Blume



# Super-Resolution Algorithmen für eine leistungsfähige Videosignalverarbeitung

von der Fakultät

Elektrotechnik und Informationstechnik

der Technischen Universität Dortmund

genehmigte

**Dissertation**

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

von

**Martin Max Richter**

Dortmund, 2011

Tag der mündlichen Prüfung: 08.02.2012

Hauptreferent: Prof. Dr.-Ing. Hartmut Schröder

Korreferent: Prof. Dr.-Ing. Holger Blume



# Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Arbeitsgebiet Schaltungen der Informationsverarbeitung der Technischen Universität Dortmund.

Dem Leiter des Arbeitsgebietes, Herrn Prof. Dr.-Ing. Hartmut Schröder, möchte ich an dieser Stelle meinen besonderen Dank für die Anregungen, Betreuung und Begutachtung dieser Arbeit aussprechen. Viele fruchtbare Diskussionen, seine Hilfestellung und das von ihm geschaffene hervorragende Arbeitsklima haben erheblich zum Gelingen dieser Arbeit beigetragen.

Herrn Prof. Dr.-Ing. Holger Blume, vom Institut für Mikroelektronische Systeme der Leibniz Universität Hannover, danke ich sehr für das Interesse an meiner Arbeit und die Übernahme des Korreferats.

Weiterhin möchte ich den Mitarbeitern der Firma Sony in Stuttgart, insbesondere den Herren Oliver Erdler, Paul Springer, Piergiorgio Sartor und Toru Nishi für die tolle Zusammenarbeit und die vielen Anregungen in Bezug auf Realisierungsmöglichkeiten danken.

Darüber hinaus möchte ich mich bei allen Kollegen des Lehrstuhls, studentischen Hilfskräften und Studien- und Diplomarbeitern bedanken, die durch ihren Einsatz und durch anregende Diskussionen zum Gelingen dieser Arbeit beigetragen haben.

Ein sehr herzlicher Dank geht auch an meine Familie, insbesondere meine Eltern und meine Schwester, die mir durch ihre Unterstützung und Fürsorge das Studium und die daran anschließende Tätigkeit als wissenschaftlicher Mitarbeiter überhaupt erst ermöglicht haben.

Zuletzt möchte ich mich noch bei allen Freunden bedanken, die mich immer unterstützt haben und mir immer mit Rat und Tat zur Seite gestanden haben.

Martin Richter



# Kurzfassung

Heutige Endgeräte müssen durch die große Formatvielfalt und verschiedenen Zuspielduellen Eingangsmaterial mit stark variierendem Störungsgrad und unterschiedlicher Auflösung verarbeiten können. Um dem Betrachter für all diese Quellen eine möglichst gute Bildqualität zu liefern, wird eine hochwertige, an das Bildmaterial gut anpassbare Videosignalverarbeitung benötigt. Sehr vielversprechend für diese Aufgabe sind iterative Rekonstruktionsalgorithmen, die bereits in anderen Bereichen wie dem maschinellen Sehen erfolgreich eingesetzt werden. Diese Algorithmen zeichnen sich durch eine zur Laufzeit ausgeführte Optimierung der Bildqualität unter der Verwendung von Bildmodellen aus. Jedoch sind diese Verfahren bisher in ihrer Anwendung begrenzt, da für eine Verarbeitung ohne Seiteneffekte die internen Modelle erfüllt sein müssen und durch den meist hohen Speicher- und Rechenaufwand relativ hohe Hardwarekosten entstehen.

In der vorliegenden Arbeit wird die Eignung dieser Verfahren für eine leistungsfähige Videosignalverarbeitung untersucht. Basierend auf den dort geltenden Anforderungen wird ein Entwurfskonzept für iterative Algorithmen erarbeitet, das die oben genannten Nachteile weitestgehend vermeidet sowie starke Verbesserungen der Algorithmen selbst ermöglicht. Dieses Konzept erlaubt die Beschreibung der Algorithmen als rekursive Filter, was zu einer Speicher- und Rechenaufwandsreduktion und darüber hinaus zu einer anschaulichen Interpretation führt. Die (interne) Verwendung einer Merkmalsanalyse und die Verbesserung einzelner Bausteine in den Algorithmen ermöglicht weiterhin die Anpassung an reales Videomaterial. Schwerpunkt der Untersuchungen bildet die Auflösungserhöhung durch Super-Resolution, welche ein aktuelles Forschungsthema darstellt und inzwischen auch in ersten Endgeräten eingesetzt wird. Mit Hilfe des Entwicklungskonzepts wurden zwei Algorithmen entwickelt, wobei der eine sehr aufwandgünstig ist und der andere auf die bestmögliche Qualität abzielt, dafür allerdings aufwändiger ist. Damit stehen je nach verfügbaren Ressourcen zwei Verfahren bereit, die durch eine Anpassung der internen Parameter bezüglich des Aufwandes noch gut skaliert werden können. Darüber hinaus wird gezeigt, wie die so entstandenen Algorithmen mit einer bildinhaltsabhängigen Signalverarbeitung kombiniert werden können, um die Flexibilität und ihr Einsatzgebiet weiter zu erhöhen. Als Erweiterung lässt sich daraus dann eine bildqualitätsgesteuerte Verarbeitung realisieren.

Den Abschluss bildet eine Bewertung der Verfahren in Bezug auf die erreichbare Bildqualität unter objektiven und subjektiven Gesichtspunkten, welche einen Rückschluss auf ihre Einsatzmöglichkeiten erlauben.



# Inhaltsverzeichnis

<b>Vorwort</b>	<b>I</b>
<b>Kurzfassung</b>	<b>I</b>
<b>Inhaltsverzeichnis</b>	<b>V</b>
<b>Abbildungsverzeichnis</b>	<b>IX</b>
<b>Tabellenverzeichnis</b>	<b>XIII</b>
<b>Notation</b>	<b>XV</b>
<b>1 Einleitung und Motivation</b>	<b>1</b>
1.1 Beiträge dieser Arbeit . . . . .	3
1.2 Aufbau der Arbeit . . . . .	5
<b>2 Ausgewählte Grundlagen</b>	<b>7</b>
2.1 Grundlagen der Interpolation und Abtastung . . . . .	7
2.2 Einführung der Super-Resolution . . . . .	10
2.3 Entwurfsmöglichkeiten für Algorithmen zur Super-Resolution . . . . .	11
2.3.1 Mathematische Ansätze für Rekonstruktionsmethoden . . . . .	13
2.3.2 Modelle zur Beschreibung des Super-Resolution Prozesses . . . . .	15
2.3.3 Lösungsstrategien . . . . .	22
2.3.4 FIR- und IIR-Filterstrukturen zur Lösung von Energiefunktionalen und Gleichungssystemen . . . . .	24
2.4 Bewegungsschätzung . . . . .	26
2.4.1 Blockmatching Verfahren . . . . .	27
2.4.2 Mehrreferenzbildschätzung . . . . .	31
2.4.3 Weitere Verfahren . . . . .	31
<b>3 Super-Resolution Verfahren</b>	<b>33</b>
3.1 Konzeptuelle Ansätze für die Super-Resolution . . . . .	33
3.1.1 Direktes und iteratives Verarbeitungskonzept . . . . .	34

3.1.2	Multiple Algorithm Blending (Multiple Algorithm Blending (MABLE)) . . . . .	36
3.1.3	Grundaufbau iterativer Super-Resolution Algorithmen . . . . .	38
3.2	Super-Resolution Algorithmen . . . . .	38
3.2.1	Auflösungserhöhung durch Detailsignaladdition . . . . .	40
3.2.2	Bildinhaltsabhängige Interpolation . . . . .	42
3.2.3	Nonuniform Interpolation . . . . .	44
3.2.4	Verfahren im Transformationsbereich . . . . .	44
3.2.5	Back-Projection Super-Resolution . . . . .	46
3.2.6	Statistische Verfahren . . . . .	48
<b>4</b>	<b>Entwurfskonzept für iterative Rekonstruktionsalgorithmen</b>	<b>53</b>
4.1	Formulierung des Entwurfskonzepts für iterative Rekonstruktionsalgorithmen	53
4.2	Anwendung des Entwurfskonzepts auf Regularisierungsalgorithmen . . . . .	55
4.2.1	Örtliche Regularisierung . . . . .	56
4.2.2	Örtlich-zeitliche Regularisierung . . . . .	60
<b>5</b>	<b>Entwurf von Super-Resolution Algorithmen für die Videosignalverarbeitung</b>	<b>65</b>
5.1	Vielseitig parametrisierbare Back-Projection Super-Resolution (VESBA) . . . . .	66
5.1.1	Modifikationen bei der Interpolation und Dezimation . . . . .	66
5.1.2	Veränderung der Unschärfefilter . . . . .	67
5.1.3	Alternativen bei der Bewegungskompensation . . . . .	67
5.1.4	Filterstrategien für die Detailsignale . . . . .	68
5.1.5	Mehrreferenzbildbewegungsschätzung . . . . .	68
5.2	Doppelt rekursive Mehrreferenzbild Super-Resolution (MUSIDORA) . . . . .	72
5.3	Aufstellen der Entwurfsgleichung . . . . .	73
5.3.1	Herleitung der Filtergleichung . . . . .	74
5.3.2	Verbesserung des entstandenen Verfahrens . . . . .	79
5.4	Zeitlich rein rekursive Super-Resolution (ESTERA) . . . . .	81
5.4.1	Entwurf des rein rekursiven Algorithmus durch Modifikation der Filtergleichung . . . . .	82
5.4.2	Verwendung einer Merkmalsanalyse zur Verbesserung der Ablaufsteuerung . . . . .	83
5.4.3	Vereinfachte Darstellung des rekursiven Super-Resolution Algorithmus	83
5.5	Beschreibung der Super-Resolution als bildinhaltsabhängige Signalverarbeitung . . . . .	83
5.5.1	Einsatz der Analysemaße innerhalb der Super-Resolution . . . . .	96
5.6	Konvergenzbetrachtungen . . . . .	98
5.6.1	Einzelbild Back-Projection Super-Resolution . . . . .	98
5.6.2	Mehrreferenzbild Back-Projection Super-Resolution . . . . .	101

5.6.3	Rekursive Super-Resolution . . . . .	102
<b>6</b>	<b>Auswertung</b>	<b>107</b>
6.1	Potential der Super-Resolution . . . . .	108
6.1.1	Simulationsergebnisse für eine aliasbehaftete Sequenz . . . . .	108
6.1.2	Simulationsergebnisse für eine Sequenz ohne Alias . . . . .	111
6.1.3	Systemtheoretisches Modell für den Super-Resolution Prozess . . . . .	112
6.2	Beurteilung der Mehrreferenzbild Back-Projection Super-Resolution . . . . .	117
6.2.1	Untersuchung des Datenmodells . . . . .	118
6.3	Beurteilung der rein rekursiven Super-Resolution (ESTERA) . . . . .	124
6.3.1	Ergebnisse für unterschiedliche zeitliche Gewichtungsfaktoren . . . . .	125
6.3.2	Verbesserungen durch den Ersatz einzelner Komponenten im Algorithmus . . . . .	125
6.3.3	Resultate unter Verwendung unterschiedlicher Vektorfeldauflösungen	126
6.3.4	Instabilität des Verfahrens und mögliche Gegenmaßnahmen . . . . .	127
6.4	Ergebnisse bei der Verwendung der Bildanalyse innerhalb der SR-Algorithmen	128
6.4.1	Beispiele für die Detail- und Kantenkarte . . . . .	128
6.4.2	Ergebnisse der Schätzung der minimalen gauß'schen Varianz für eine niedrig und eine hoch aufgelöste Sequenz . . . . .	128
6.4.3	Bewegungsvektorkängenabhängige Gewichtung der Detailsignale . . . . .	129
6.5	Vergleich mit anderen Interpolationsalgorithmen . . . . .	131
6.5.1	Generelle Bemerkungen zur Verwendung der objektiven Maße . . . . .	131
6.5.2	Bildbeispiele und objektive Messwerte für unterschiedliche Algorithmen . . . . .	132
6.6	Zusammenfassung . . . . .	136
<b>7</b>	<b>Zusammenfassung und Schlußfolgerungen</b>	<b>141</b>
7.1	Ausblick . . . . .	147
	<b>Literaturverzeichnis</b>	<b>149</b>
<b>A</b>	<b>Standardeinstellungen der Super-Resolution Algorithmen</b>	<b>157</b>
<b>B</b>	<b>Verwendete Testsequenzen</b>	<b>161</b>
	<b>Lebenslauf</b>	<b>165</b>



# Abbildungsverzeichnis

1.1	Charakterisierung der implementierten Verfahren nach ihrer Leistungsfähigkeit bezüglich Artefaktreduktion und Detailerzeugung bzw. -wiederherstellung. . . . .	5
2.1	Interpolation durch Einfügen von Nullen und anschließender Filterung. Blockschaltbild (oben) und Signalverläufe nach den einzelnen Blöcken (unten)	8
2.2	Spektrale Darstellung der Interpolation. Spektrum der Eingangssequenz, Zwischenergebnis nach Einfügen von Nullen mit eingezeichnetem Tiefpass, Interpolationsergebnis nach Standardinterpolation und nach Superresolution im Vergleich. . . . .	8
2.3	Mathematische Ansätze zur Beschreibung von Bildverarbeitungsaufgaben (vergleiche [CS05]) . . . . .	14
2.4	Abbildungsprozess zwischen hoch und niedrig aufgelöster Sequenz . . . . .	16
2.5	Wirkung der Total Variation Regularisierung: Eingangsbild (links) und mit Total Variation Regularisierung verarbeitetes Bild (rechts) . . . . .	20
2.6	Richtung der vier Ableitungen beim Huber Markov Zufallsfeld Bildmodell .	22
2.7	Bestrafungsfunktion für das Huber Markov Zufallsfeld Bildmodell . . . . .	22
2.8	Explite und implizite Lösungsstrategie für iterative Algorithmen . . . . .	24
2.9	Bestimmung des Bewegungsvektors mit dem 'Fullsearch' Verfahren (vergl. [Haa00]) . . . . .	28
2.10	Verfügbare örtliche und zeitliche Prädiktorkandidaten [Blu97] . . . . .	29
2.11	Relative Position der örtlichen und zeitlichen Prädiktoren beim Parallel Prädiktiven Blockmatching [Blu97] . . . . .	30
2.12	Korrekturstern für das prädiktive Blockmatching [Blu97] . . . . .	30
2.13	Bewegungsvektoren einer Mehrreferenzbildbewegungsschätzung . . . . .	31
3.1	Schematische Darstellung der direkten (nicht-iterativen) Bildverarbeitung .	34
3.2	Schematische Darstellung der iterativen Bildverarbeitung . . . . .	34
3.3	Prinzipschaltbild für das 'Multiple Algorithm Blending' . . . . .	37
3.4	Generalisierte Beschreibung iterativer Super-Resolution Algorithmen . . . .	38
3.5	Übersicht über Verfahren zur Steigerung der Auflösung in jedem Einzelbild	39
3.6	Beispielhafter Frequenzgang vor und nach der Bildschärfekorrektur . . . . .	41
3.7	Nonuniform Interpolation Super-Resolution . . . . .	45

---

3.8	Registrierung und Interpolation für die nonuniform Interpolation Super-Resolution . . . . .	45
3.9	Blockdiagramm zur Realisierung der Back-Projektion Super-Resolution . .	48
3.10	Realisierung der Modellierungen des Abbildungsprozesses (oben) und des Back-Projection Operators (unten) . . . . .	49
3.11	Kalman Filter basierte Super-Resolution . . . . .	52
4.1	Entwurfskonzept für iterative Rekonstruktionsalgorithmen . . . . .	55
4.2	Schematischer Ablauf des Regularisierungsalgorithmus . . . . .	57
4.3	Position der verwendeten Filtergewichte für die Berechnung des aktuellen Ausgangspixels. Die Lage ergibt sich direkt aus der mathematischen Berechnungsvorschrift . . . . .	58
4.4	Position der Filtergewichte für eine verbesserte Filterung innerhalb der Regularisierung . . . . .	59
4.5	Lage der Filterkoeffizienten und -Gewichte bei der örtlich-zeitlichen Regularisierung . . . . .	63
5.1	Ursprüngliches Schema der Back-Projektion mit Bewegungskompensation und der Inversen. . . . .	68
5.2	Back-Projektion Schema mit nur einer Bewegungskompensation. . . . .	69
5.3	Iterationsschema der Mehrreferenzbildbewegungsschätzung . . . . .	70
5.4	Wideblock-Splitting zur Verfeinerung des Vektorfelds . . . . .	71
5.5	Prediktorgeneration mit dem optischen Fluss . . . . .	72
5.6	Beziehung zwischen mathematischem Energiefunktional und bildverarbeitenden Operatoren . . . . .	74
5.7	Blockschaltbild des Mehrreferenzbild Super-Resolution-Systems . . . . .	78
5.8	Verbessertes Mehrreferenzbildverfahren mit doppelt rekursiver Verarbeitung (MUSIDORA) . . . . .	79
5.9	Strategie zur Erkennung der Kantenrichtung gemäß dem Huber Markov Zufallsfeld Bildmodell . . . . .	80
5.10	Verbesserte Kantenrichtungserkennung für das Bildmodell . . . . .	80
5.11	Vereinfachte Darstellung des rein rekursiven Super-Resolution Algorithmus mit integrierten Verbesserungen . . . . .	84
5.12	Frequenzselektive Gewichtung des Detailsignals. $N$ bezeichnet hier ausnahmsweise einmal abweichend die Länge des Filters in x- bzw. y-Richtung. . . . .	86
5.13	Vereinfachte schematische Darstellung des Super-Resolution Algorithmus als bildinhaltsabhängige Signaverarbeitung . . . . .	87
5.14	Blockschaltbild des finalen Systems (ESTERA) mit allen Blöcken . . . . .	88
5.15	Ablauf der Segmentierung zur Erkennung homogener Flächen und Detailbereichen . . . . .	91

---

5.16	Filtermasken für die Segmentierung. Tiefpassfilter zur Vorverarbeitung (links) und Hochpassmaske zur Detailerkennung (rechts) . . . . .	92
5.17	Lineare und quadratische Gewichtungsfunktion . . . . .	92
5.18	Beispielhafter Verlauf der Varianzschätzung mittels Minimal Description Length (vergl. [Brü09]) . . . . .	95
5.19	Ergebnis der Varianzschätzung mittels Minimal Description Length für unterschiedliche Bildbereiche . . . . .	95
5.20	Blockschaltbild für die Kontrastmessung . . . . .	96
5.21	Filterübertragungsfunktion bei der Super-Resolution pro Iteration (oben) und resultierende, direkte Gesamtübertragungsfunktion (unten) . . . . .	100
5.22	Darstellung der rekursiven Super-Resolution als direkte Verarbeitung unter Verwendung aller bisherigen Eingangsbilder. Auf die Gewichtung des zweiten Filters mit $(1 - \alpha_m)$ wurde verzichtet. . . . .	104
6.1	Rekursive Super-Resolution mit unterschiedlichen Parametern für eine kontrollierte Sequenz mit Alias. . . . .	109
6.2	Rekursive Super-Resolution mit unterschiedlichen Parametern für eine kontrollierte Sequenz ohne Alias. . . . .	110
6.3	Bildverarbeitungskette bei Verwendung von Super-Resolution . . . . .	112
6.4	Spektrum für das Ergebnis der Super-Resolution für eine Sequenz mit Subpixelbewegung und korrekt geschätzter Bewegung . . . . .	114
6.5	Spektrum für das Ergebnis der Super-Resolution für eine Sequenz mit Pixelbewegung und korrekt geschätzter Bewegung . . . . .	115
6.6	Alias im Falle von unbewegten Sequenzen . . . . .	116
6.7	Erzeugung von falschen Details durch eine fehlerhafte Bewegungsschätzung	117
6.8	Interpolations- und Dezimationsstrategien für die Mehrreferenzbild-SR . .	119
6.9	Einfluss unterschiedlicher zeitlicher Filterstrategien zur Bestimmung des Detailsignals . . . . .	120
6.10	Einfluss unterschiedlicher Filterstrategien bei inkorrekt Bewegungskompensation . . . . .	120
6.11	Alternative zeitliche Filtertechniken. Die rechte Teilabbildung entspricht den Einstellungen von MUSIDORA. . . . .	122
6.12	Eigenschaften des Huber MRF Bildmodells und des modifizierten HMRF auf das Ergebnis bei unterschiedlichen Schwellwerten $\alpha$ . . . . .	123
6.13	Feste oder adaptive Gewichtung des zeitlich vorigen Ergebnisbildes . . . .	124
6.14	Verbesserungen durch bildinhaltsabhängige Signalverarbeitung und adaptive Gewichtung der Komponenten . . . . .	126
6.15	Einfluss der Vektorfeldauflösung auf das Ergebnis der Super-Resolution . .	127
6.16	Gegenmaßnahmen gegen Divergenz und Kontrolle des Algorithmus durch Bildanalysemaße . . . . .	128

---

6.17	Detail- und Kantenkarte zur Steuerung des Algorithmus . . . . .	129
6.18	Bestimmung der minimalen lokalen Varianz für eine niedrig und eine hoch aufgelöste Sequenz. Bildbeispiel (oben) und Karte (unten) im Vergleich. . .	130
6.19	Anwendung der minimalen lokalen Varianzschätzung zur Kontrolle der Schärfe im Super-Resolution Algorithmus . . . . .	130
6.20	Vermeidung von Bewegungsvektorfehlern bei hohen Geschwindigkeiten. . .	131
6.21	Vergleich von hoch aufgelöster Referenz und zwei SR-Ergebnissen mit unterschiedlichem Schärfe und Detailgrad . . . . .	132
6.22	Ergebnisse der Sequenz 'stockholm' für unterschiedliche Interpolationsstrategien . . . . .	133
6.23	Ergebnisse der Sequenz 'wheel' für unterschiedliche Interpolationsverfahren	134
6.24	Deblocking und Interpolation niederqualitiver Videos . . . . .	138
6.25	Allgemeine Bildverbesserung durch den Super-Resolution Algorithmus . . .	139
7.1	Qualitative Beurteilung der Leistungsfähigkeit der entwickelten Algorithmen im Vergleich zu einem bestehenden SR-Ansatz . . . . .	146
A.1	Parametereinstellmöglichkeiten der Mehrreferenzbild Super-Resolution . . .	158
A.2	Parametereinstellmöglichkeiten der rein rekursiven Super-Resolution . . . . .	160
B.1	Sequenz 'football' . . . . .	162
B.2	Sequenz 'parkrun' . . . . .	162
B.3	Sequenz 'stockholm' . . . . .	163
B.4	Sequenz 'wheel' . . . . .	164
B.5	Sequenz 'matfight' . . . . .	164
B.6	Sequenz 'seeking' . . . . .	165
B.7	Sequenz 'foreman' . . . . .	166

# Tabellenverzeichnis

3.1	Eigenschaften der direkten und rekursiven Verarbeitung . . . . .	35
5.1	Parameter der Super-Resolution und ihre Steuerung durch Analysemaße . .	97
6.1	PSNR-Werte für unterschiedliche Interpolationsalgorithmen . . . . .	136
6.2	SSIM-Werte für unterschiedliche Interpolationsalgorithmen . . . . .	137



# Notation

## Abkürzungen

<b>3DRS</b>	3-Dimensional Recursive Search
<b>4K</b>	Kino- und Fernsehformat mit 4096x2160 Bildpunkten
<b>ADRC</b>	Adaptive Dynamic Range Coding
<b>AVC</b>	Advanced Video Coding
<b>DL</b>	Description Length
<b>HDTV</b>	High Definition Television
<b>HRS</b>	High Resolution Spline
<b>LCD</b>	Liquid Crystal Display
<b>LTl</b>	Luminance Transient Improvement
<b>MABLE</b>	Multiple Algorithm Blending
<b>MAP</b>	Maximum-A-Posteriori
<b>MDL</b>	Minimal Description Length
<b>NEDI</b>	New Edge Directed Interpolation
<b>PDE</b>	Partielle Differentialgleichung
<b>PPBM</b>	Parallel prädiktives Blockmatching
<b>PSNR</b>	Peak Signal to Noise Ratio
<b>SAD</b>	Summed Absolute Differences
<b>SDTV</b>	Standard Definition Television
<b>SR</b>	Super-Resolution
<b>SSIM</b>	Structural Similarity Index Measure
<b>TV</b>	Total Variation
<b>WS</b>	Weighted Selection

## Bezeichnung der neu entwickelten Algorithmen

<b>VESBA</b>	Voll parametrisierbarer Back-Projection Algorithmus (Mehrreferenzbild)
<b>ESTERA</b>	Rein rekursiver Back-Projection Algorithmus
<b>MUSIDORA</b>	Doppelt rekursiver Mehrbild Super-Resolution Algorithmus

## Symbole und mathematische Operatoren

### Allgemein

$\mathbf{A}$	Matrix
$\mathbf{A}^T$	Transponierte der Matrix $A$
$\mathbf{a}$	Vektor
$a$	Skalar oder skalare Funktion
$A$	Fouriertransformierte von $a$
$\mathbf{a}_i$	$i$ -tes Element des Vektors $\mathbf{a}$
$\mathbf{A}_{i,j}$	Matrizeintrag an der Stelle $i, j$
$a(x)$	Funktionswert an der Stelle $x$ (alternativ auch $a_x$ )
$\hat{a}$	Schätzung von $a$ (Kalman Filter)

### Operatoren und Funktionen

$\max$	Maximum
$\min$	Minimum
$\text{div}$	Divergenz
$\nabla$	Nabla Operator
$\Delta$	Laplace Operator
$*$	Faltung (diskret oder kontinuierlich)
$\frac{\partial}{\partial x}$	partielle Ableitung nach $x$
$p(\cdot)$	Wahrscheinlichkeit eines Ereignisses
$p(\cdot \cdot)$	bedingte Wahrscheinlichkeit
$F\{\}$	Fouriertransformation von einem Ausdruck
$\rho()$	Gewichtungsfunktion für das Huber Markov Zufallsfeld Bildmodell

---

$\Gamma ()$	Ersatzausdruck für das Huber Markov Zufallsfeld Bildmodell
$sign ()$	Signum Funktion
$\delta ()$	Dirac-Impuls

## Bezeichner

$\mathbf{y}$	Eingangsbild bzw. -sequenz
$\mathbf{y}_{int}$	interpolierte(s) Eingangsbild bzw. -sequenz
$\mathbf{z}$	Ausgangsbild bzw. -sequenz
$\mathbf{z}^{ws}$	Ergebnis der gewichteten Auswahl aus Eingangs- und Ergebnisbild
$\mathbf{z}^{mc}$	bewegungskompensiertes Ergebnisbild
$\mathbf{e}_{sys}$	Systemrauschen des Kalman Filter Prozesses
$\mathbf{e}_{mess}$	Messrauschen des Kalman Filter Prozesses
$\mathbf{e}_r$	Fehlervektor (Rauschen)
$\mathbf{b}$	lokaler blur Kernel, der aus $\mathbf{B}$ abgeleitet ist
$\mathbf{K}$	Kalman Matrix
$\mathbf{B}$	Blur Kernel
$\mathbf{D}$	Unterabtastungsmatrix
$\mathbf{F}$	Bewegungskompensationsmatrix
$\mathbf{H}_{BP}$	Back-Projection Operator
$E$	Energiefunktional
$\Omega$	Definitionsbereich des Bildes bzw. der Sequenz
$\lambda$	Lagrange'scher Multiplikator
$\lambda_{spat}, \lambda_{temp}$	örtlich und zeitlich definierte Lagrange'sche Multiplikatoren
$\beta$	Gewichtungsparameter für die MAP Super-Resolution
$D$	Datenterm der Regularisierung

$S_{spat}, S_{temp}$	örtliche und zeitliche Nebenbedingung (Regularisierung)
$\mathbf{b}_s, \mathbf{b}_t$	Filterkoeffizientenfelder (Regularisierung)
$diff\_t$	zeitliche Differenz (Absolutwert)
$\epsilon$	kleiner Wert (größer Null) zur Vermeidung der Division durch Null
$D_1, D_2$	Menge aller Pixel, die bei der Regularisierung zur Filterung genutzt werden
$\tau$	Iterationsschrittweite
$t_\alpha$	Schwellwert für das Huber Markov Zufallsfeld Bildmodell
$I$	Grauwertbild
$f$	Frequenz
$h()$	Impulsantwort
$H()$	Fouriertransformierte von $h$ , Übertragungsfunktion
$x, y$	kontinuierliche horizontale/vertikale Ortskoordinate
$k, l$	diskrete horizontale/vertikale Ortskoordinate
$t$	kontinuierliche Zeitvariable
$n$	diskrete Zeitvariable
$i, j$	diskrete relative, örtliche Verschiebung
$m$	diskrete relative, zeitliche Verschiebung
$v_x, v_y$	Bewegungsvektoren in x- und y-Richtung
$v_{x,m}, v_{y,m}$	Bewegungsvektoren des Mehrreferenzbildschätzers in x- und y-Richtung zum Zeitschritt $n + m$
$p, q$	Zahlvariable bei der Iteration
$r$	Zählindex bei der Konvergenzprüfung
$Y_*$	Ergebnis einer zeitlich bewegungskompensierten Filterung über alle Eingangsbilder $Y$
$Y_{fix,n}$	Informationsgehalt, der in jedem Eingangsbild $Y$ vorhanden ist

---

$Y_{var,n}$	Einzigartiger Informationsgehalt, nur in Eingangsbild mit Zeitindex $n$ vorhanden
$H_*$	multiplikative Gewichtung eines Degradationsoperators $H$ über die Eingangsbildanzahl $N$
$N$	Anzahl verwendeter Eingangsbilder
$d_i$	zweite Ableitung in eine der vier Hauptrichtungen
$\alpha, \alpha_1, \alpha_2$	fixe und adaptive Gewichtungparameter für die gewichtete Auswahl
$\kappa, \kappa_{fix}, \kappa_{adapt}$	adaptive und fixe Parameter zur Einstellung der gewichteten Auswahl
$c_1, c_2$	Gewichtungsfaktoren der Detailsignale innerhalb der Super-Resolution
$g$	Gewichtungsfaktor zur Berechnung des adaptiven Systemrauschens (Kalman-SR)
$g_v$	Ergebnis des geschwindigkeitsadaptiven Analysemaßes
$a, b$	lokale, diskrete Ausdehnung der Analysemasken in x- und y-Richtung
$\sigma$	Varianz des gauss'schen Filters
$\lambda_{DL}$	Gewichtungsfaktor zur Berechnung der Description Length
$w_{a,b}$	(gewichtete) Fenster mit Ausdehnung $a, b$



# Kapitel 1

## Einleitung und Motivation

Die Einführung von hochauflösendem Fernsehen (HDTV) und Codierverfahren wie H.264 AVC für die digitalen TV-Verteilnetze und Bildschirme bilden einen Meilenstein für die Bildqualität. In Zukunft wird dieser Trend durch 4K Videos und neue Codierstandards weiter fortgesetzt werden. Allerdings liegen viele Inhalte noch in Standardauflösung vor, sind sehr stark komprimiert oder besitzen eine noch niedrige Auflösung als SDTV. Die Großzahl der Fernseher verfügt aber inzwischen über eine Full-HD Auflösung von 1920x1080 Bildpunkten und ermöglicht damit die Darstellung von Videos mit hoher Auflösung und Qualität, was die Erwartungen des Zuschauers erheblich steigert und für niedrig aufgelöste und mit Codierartefakten behaftete Sequenzen hochwertige Algorithmen zur Videosignalverarbeitung erforderlich macht. Da das Eingangsmaterial sehr stark in seiner Qualität und Auflösung variieren kann, wird weiterhin eine sehr gut an die jeweilige Situation anpassbare Verarbeitung benötigt.

Diese Anforderungen können durch zwei unterschiedliche Strategien bewältigt werden. Auf der einen Seite bieten sich iterative Rekonstruktionsalgorithmen an (siehe z.B. [AK06] oder [CS05]), die zur Laufzeit eine Optimierung der Bildqualität unter Einbeziehung von Bildmodellen durchführen, so dass generell für ein breites Spektrum an Eingangsmaterial gute Ergebnisse hinsichtlich der (objektiven) Bildqualität erzielt werden können. Auf der anderen Seite kann eine gute Anpassung der Algorithmen an unterschiedliches Material durch eine bildinhaltsabhängige Signalverarbeitung [Kon01], [ZW09], [SMD09], [RDS<sup>+</sup>08] erreicht werden, indem eine Merkmalsanalyse auf dem Eingangsmaterial durchgeführt wird und die Verarbeitung dann an das jeweilige Material oder die jeweilige Bildsituation durch Parametervariation oder das Umschalten zwischen verschiedenen Algorithmen angepasst wird.

Gerade durch die große Formatvielfalt sind zwei wichtige Anwendungsmöglichkeiten dieser Algorithmen entscheidend, durch die eine erhebliche Verbesserung des Videomaterials sinnvoll und möglich ist. Durch eine Bekämpfung von Codierartefakten kann die Qualität von sehr niederratigen Videos z.B. von Web-Portalen deutlich gesteigert werden. Weiterhin ist eine Anpassung der Eingangsauflösung an die native Auflösung der Panels nötig. Hierzu können neben Standardverfahren, z.B. einer bilinearen Interpolation oder High Resolution

Spline Interpolation, Super-Resolution (SR) Algorithmen [NSLZ07], [FREM04], [SVW02], [BS99] angewendet werden. Bei diesen findet zusätzlich zur Anpassung der Pixelanzahl eine integrierte Störungsreduktion bei gleichzeitigem Detailerhalt oder sogar der Erzeugung neuer, hochfrequenter Details statt.

Im Bereich der Codierartefaktreduktion finden heute im Endgerät vor allem adaptive Filter [Pia08] Anwendung, die eine Analyse nach Blockgrenzen durchführen und die Stärke der Filterung an die erkannte Störung anpassen. Aus anderen Anwendungsgebieten, beispielsweise der Medizintechnik, sind Regularisierungsalgorithmen [ROF92] bekannt, die vor allem für die Reduktion von weißem Rauschen eingesetzt werden können. Diese Algorithmen gehören zur Gruppe der iterativen Rekonstruktionsalgorithmen, die während der Laufzeit eine an den Bildinhalt angepasste lokale Optimierung der Bildqualität durchführen. Allerdings benötigen diese Verfahren einen nicht unerheblichen Speicher- und Rechenaufwand, der für den Consumer-Bereich nachteilig ist, da hier neben der Bildqualität auch Aspekte wie Kosten und damit Hardwareaufwand berücksichtigt werden müssen. Einen guten Kompromiss bieten hier trainierte Filter [Kon01], [SHH08], die für jede Bildsituation optimale Filterkoeffizienten bereit stellen. Diese werden mit Hilfe von Trainingsmaterial angelehrt und zur Laufzeit ist deshalb nur noch eine einfache Klassifikation der Bildsituation und die Filterung selbst nötig.

Zur Interpolation finden meist Standardverfahren Anwendung [HA78], da diese sehr recheneffizient sind. Auch hier können trainierte Filter [Kon01], [Len09] eine deutliche Verbesserung der Bildqualität erzielen und teilweise sogar neue Details generieren. Super-Resolution Algorithmen zur Erhöhung der örtlichen Auflösung werden zwar seit den 1980er Jahren umfangreich erforscht [TH84], [NSLZ07], bedingt durch ihren höheren Rechenaufwand werden sie aber erst seit wenigen Jahren für den Consumer-Bereich attraktiv. Die meisten Verfahren basieren auch in diesem Fall auf iterativen Rekonstruktionsalgorithmen. Zielsetzung aller SR-Verfahren ist es, Frequenzanteile oberhalb der Nyquistfrequenz des Eingangsmaterials wiederherzustellen oder neu (als bestehende Schätzung) zu erzeugen. Das kann entweder durch die Addition synthetischer Detailsignale [FJP02], [SEF86], Bildmodelle in Kombination mit on-line oder off-line Optimierungsstrategien [Kon01], [BS99] oder die Ausnutzung von (örtlichem) Alias und Subpixelbewegungen zwischen benachbarten Bildern [SVW02] erreicht werden. Besonders nachteilig bei den meisten dieser Verfahren ist die hohe Anzahl von Bildspeichern, die für gute Ergebnisse benötigt wird, um die Informationen eines Ergebnisbildes aus mehreren Eingangsbildern zu kombinieren. Alternativ können auch hier Bildmodelle verwendet werden, um das Aussehen der Ergebnisbilder gezielt durch eine örtliche Verarbeitung zu beeinflussen und damit die Menge an Eingangsbildern zu reduzieren. Wichtige Beiträge, um den Speicheraufwand zu verringern oder die Robustheit gegenüber Störeinflüssen wie beispielsweise Rauschen oder fehlerhaften Bewegungsvektoren zu verbessern, wurden von Farsiu et al. [FREM04] und Stevenson et al. [BS99] erbracht. Im Consumer-Bereich werden durch die Industrie entweder rein örtliche, auf Bildmodellen basierende SR-Verfahren eingesetzt (z.B. das Toshiba Resolution+,

[YST<sup>+</sup>09]) oder zeitlich rein rekursive Verfahren [INXO10] bevorzugt. Mehrreferenzbildverfahren werden z.B. von Intel als besonderes Merkmal für Hochleistungsprozessoren oder für nicht echtzeitfähige Anwendungen bereitgestellt [Int07].

## 1.1 Beiträge dieser Arbeit

In dieser Arbeit wird untersucht, inwieweit sich die aus anderen Bereichen bekannten iterativen Rekonstruktionsalgorithmen für die Videosignalverarbeitung eignen. Schwerpunktmäßig wurden dazu Super-Resolution Algorithmen ausgewählt. Die Hauptbeiträge dieser Arbeit lassen sich wie folgt beschreiben. Um den Rahmenbedingungen der Videosignalverarbeitung gerecht zu werden und die iterativen Rekonstruktionsalgorithmen in diesem Bereich einsetzen zu können, wird ein aus vier Schritten bestehendes Entwurfskonzept vorgeschlagen. Die ersten zwei Schritte, die die Problemformulierung und die Ableitung des Algorithmus umfassen, werden bereits in ähnlicher Form z.B. in [AK06], [CS05], [Kau00], [DRS09a] durchgeführt. In der vorliegenden Arbeit wird eine erweiterte und umfassend formalisierte Beschreibung für diese Vorgehensweise geliefert. Ein weiterer Schritt in diesem Konzept besteht im Einbringen nachträglicher Verbesserungen in diese Algorithmen, wie es bereits in den (eigenen) Publikationen [DRS09a], [RBSK11] und [RDS<sup>+</sup>11] dargestellt ist. Diese Vorgehensweise ermöglicht zum einen die Ableitung speicher- und recheneffizienter Algorithmen durch eine geschickte Beschreibung der Algorithmen als rekursive und adaptive Filter und zum anderen die Integration zusätzlicher Komponenten in die Algorithmen, um konzeptionsbedingte Schwächen dieser zu beseitigen oder die Algorithmen weiter zu verbessern. Für zahlreiche Ansätze zur Regularisierung und Super-Resolution wird gezeigt, dass Verbesserungen bezogen auf den Aufwand sowie die Bildqualität in unterschiedlichen Schritten des Entwurfskonzeptes möglich sind. Diese können entweder direkt bei der Aufstellung des Ansatzes, der Ableitung des Algorithmus oder auch bei der nachträglichen Verbesserung erfolgen.

Im Rahmen der Dissertation wurde weiterhin ein Regularisierungsansatz entwickelt, der das Entwurfskonzept für iterative Algorithmen nutzt und auf einem neuen Ansatz mit zusätzlicher Nebenbedingung basiert. Im Ergebnis ergibt sich so ein örtlich-zeitlicher Regularisierungsalgorithmus, der nur eine sehr begrenzte Anzahl an Eingangsbildern für die Berechnung eines Ausgangsbildes benötigt, mit Bewegungsvektoren von Blockmatching-Verfahren [HBHO93] betrieben werden kann und robust gegenüber einer fehlerhaften Bewegungsschätzung ist. Diese Eigenschaften können bekannte Algorithmen, beschrieben z.B. in [AK06], [MTM<sup>+</sup>04], [YFL<sup>+</sup>03], nicht erfüllen.

Ein weiterer Beitrag dieser Arbeit ist die Implementierung mehrerer Algorithmen zur Super-Resolution mit unterschiedlichem Aufwand bei bestmöglicher Qualität der Ergebnisbilder. Durch eine variabel einstellbare Parametrisierung können diese Algorithmen für sehr unterschiedliche Anwendungszwecke genutzt werden, da sich in den Algorithmen selbst der

Aufwand gezielt beeinflussen lässt (z.B. durch verwendete Bildanzahl, Iterationen, usw.). Alle Algorithmen wurden mit Hilfe des Entwurfskonzeptes erarbeitet. Der erste Algorithmus basiert auf der Back-Projection [IP91], für den die Ableitung aus der Literatur bekannt ist aber nachträglich Verbesserungen eingefügt wurden, die zu einem geringeren Rechenaufwand führen. Dieser Algorithmus wird nachfolgend VESBA genannt. Der zweite Algorithmus (MUSIDORA) ist ebenfalls ein Verfahren, welches mehrere Referenzbilder zur Berechnung eines Ergebnisses nutzt und auf einem modifizierten MAP-Ansatz [BS99] aufbaut. Spätere konzeptionsbedingte Verbesserungen durch eine Verschiebung der zeitlichen Filterung aus der Rekursionsschleife in eine zusätzliche zeitliche Rekursionsschleife führen zu einem Algorithmus, der für ein Mehrreferenzbildverfahren einen vergleichsweise geringen Rechenaufwand benötigt und gleichzeitig eine hervorragende Bildqualität liefert. Als eine sehr rechenaufwandsgünstige Alternative wurde weiterhin ein zeitlich rein rekursives Verfahren (ESTERA) entwickelt, welches nur eine einzige Iteration pro Bild ausführt. Dieses besitzt damit eine feste Laufzeit und ist so für die Videosignalverarbeitung besonders attraktiv. Ein ähnliches Verfahren wird aber bereits in der Patentliteratur erwähnt [INXO10]. Es ist zeitgleich zu dem hier veröffentlichten Verfahren entstanden.

Die weitere Verbesserung der SR-Algorithmen durch eine Steuerung mit Hilfe einer Merkmalsanalyse ist ein weiterer Beitrag dieser Arbeit. Durch die Verwendung zeitlich adaptiver Gewichtungen der einzelnen Eingangsbilder kann die Robustheit erheblich gesteigert werden. Weiterhin ist mit anderen Maßen (z.B. der Schätzung der minimalen gauß'schen Varianz [Est00]) eine Anpassung an sehr unterschiedliches Eingangsmaterial möglich. Eine Messung des Kontrastverhältnisses zwischen Eingangs- und Ausgangssequenz macht eine Steuerung der Stärke der Detailwiederherstellung und Bildschärfe in den Algorithmen mit nur einem einzigen Parameter möglich.

Durch eine Umstellung der Systemgleichung ergibt sich die Struktur einer bildinhaltsabhängigen Signalverarbeitung [SMD09], [ZW09], [RDS<sup>+</sup>08]. Bei diesem Konzept werden Algorithmen mit unterschiedlichen Vorteilen in verschiedenen Situationen parallel auf das Eingangsmaterial angewendet und dann – basierend auf einer Merkmalsanalyse – der am besten passende Algorithmus ausgewählt. In dieser Arbeit konnte gezeigt werden, dass die rein rekursive Super-Resolution (ESTERA) als solch ein Konzept beschrieben werden kann. Durch einen Austausch von Einzelkomponenten oder zusätzlichen Steuerinformationen sind so weitere Verbesserungen möglich. Auch hier ist in [INXO10] ein ähnlicher Ansatz angedeutet, der aber nicht auf einer mathematischen Herleitung beruht.

Ein weiterer Beitrag dieser Arbeit besteht in der Implementierung zahlreicher Verfahren zur Beurteilung der Leistungsfähigkeit der einzelnen Verarbeitungskonzepte. Eine Übersicht über diese Algorithmen ist in Abbildung 1.1 gezeigt. In dieser Abbildung ist zu sehen, dass die Algorithmen sich über einen breiten und sehr unterschiedlichen Anwendungsbereich erstrecken. Die neu entstandenen Algorithmen zur Super-Resolution zeichnen sich dabei durch eine sehr gute Detailwiederherstellung bei gleichzeitig guter Störungsreduktion aus.

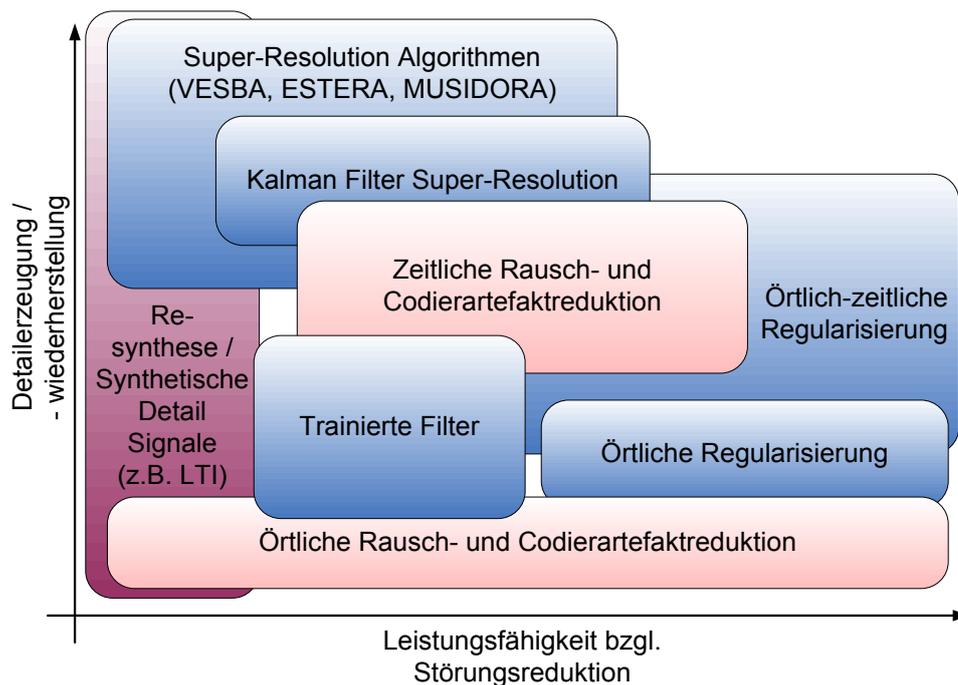


Abbildung 1.1: Charakterisierung der implementierten Verfahren nach ihrer Leistungsfähigkeit bezüglich Artefaktreduktion und Detailerzeugung bzw. -wiederherstellung.

Schließlich wird in dieser Arbeit eine systemtheoretische Beschreibung der Super-Resolution zur Veranschaulichung der nötigen Rahmenbedingungen (z.B. Alias und Subpixelwegung) entwickelt. Hierzu wurde eine Bildverarbeitungskette aus linearen Operatoren aufgestellt, die sich an dem Vorgehen für die Modellierung von Bewegungsunschärfe in [Sch00] und [Dol10] orientiert. Damit kann das Zusammenwirken von örtlicher und zeitlicher Filterung innerhalb der Super-Resolution dargestellt und dieses Modell für weitere Verbesserungen genutzt werden.

## 1.2 Aufbau der Arbeit

Die vorliegende Arbeit ist in folgende Bereiche unterteilt. Im nachfolgenden Kapitel 2 werden Grundlagen beschrieben, auf denen die betrachteten bzw. entwickelten bildinhaltsabhängigen und iterativen Rekonstruktionsalgorithmen zur Bildqualitätsverbesserung aufbauen. Zu diesen Grundlagen zählt u.a. die Vorstellung mathematischer Ansätze zur Beschreibung der Algorithmen, (iterative) Lösungsstrategien zur Minimumsuche und Verfahren zur Bewegungsschätzung, die für die zeitlichen Super-Resolution Algorithmen be-

nötigt werden. Kapitel 3 befasst sich mit Verfahren zur Super-Resolution. Um die Wirkungsweise der Algorithmen besser verstehen zu können, werden zuerst Verarbeitungskonzepte vorgestellt, an denen sich bereits die wichtigsten Vor- und Nachteile ableiten lassen. Bei den konkreten Verfahren wird hauptsächlich auf die iterativen Rekonstruktionsalgorithmen eingegangen, da diese später weiter verwendet werden. In Kapitel 4 wird dann das aus vier Schritten bestehende Entwurfskonzept für iterative Rekonstruktionsalgorithmen für die Videosignalverarbeitung vorgestellt. Dieses wird anschließend exemplarisch auf zwei Algorithmen zur Regularisierung angewendet. Während ersterer bereits aus der Literatur bekannt ist und erst nach der Ableitung des Algorithmus (nachträglich) Verbesserungen eingefügt werden, basiert letzterer auf einer Formulierung eines neuen Ansatzes mit einer zusätzlichen zeitlichen Nebenbedingung. Nachdem damit der Nachweis für die Einsatzmöglichkeiten des Entwurfskonzeptes erbracht ist, werden in Kapitel 5 die Super-Resolution Algorithmen selbst entworfen. Dazu wird zuerst ein bereits bekannter Back-Projection Algorithmus abgeleitet und dieser nachträglich verbessert. Die Analyse und Verbesserungen dieses Algorithmus bilden die Grundlage für einen modifizierten Ansatz für ein weiteres auf mehreren Referenzbildern basierendes Verfahren, welches durch eine doppelt rekursive Struktur zu einer für Mehrreferenzbildverfahren sehr guten Recheneffizienz führt und gleichzeitig eine sehr hohe Bildqualität garantieren kann. Als sehr rechenaufwandgünstige Alternative wird weiterhin ein rein rekursiver Algorithmus vorgeschlagen. Die variable Parametrisierbarkeit aller Algorithmen ermöglicht ein sehr breites Anwendungsfeld sowie die Auswahl eines optimalen Algorithmus für die jeweils vorliegenden Rahmenbedingungen (Aufwand, Bildqualität). Im Anschluss daran erfolgt die Beschreibung von Steuerungsmöglichkeiten dieser Algorithmen sowie die Beschreibung von Super-Resolution als bildinhaltsabhängige Signalverarbeitung. Das Kapitel endet mit Konvergenzbetrachtungen. Kapitel 6 enthält die Auswertung der SR-Algorithmen. Zuerst wird in diesem Kapitel die Fähigkeit der Detailwiederherstellung beispielhaft für den rein rekursiven Algorithmus gezeigt. Danach werden in einer Untersuchung die Einflussmöglichkeiten der wichtigsten Parameter aufgezeigt. Zum Schluss werden die Algorithmen mit in der Literatur beschriebenen Verfahren verglichen. Die Arbeit endet mit einer Zusammenfassung und Schlussfolgerungen in Kapitel 7.

# Kapitel 2

## Ausgewählte Grundlagen

Dieses Kapitel beinhaltet Grundlagen, die für die Erläuterung der iterativen Rekonstruktionsalgorithmen im Allgemeinen und die Algorithmen zur Super-Resolution im Speziellen benötigt werden. Da Super-Resolution immer mit einer Interpolation verbunden ist, werden zuerst in Abschnitt 2.1 die Grundlagen dieser vermittelt. Darauf aufbauend wird in Abschnitt 2.2 eine Einführung in die Thematik der Super-Resolution sowie deren geschichtliche Entwicklung gegeben. Für die Entwicklung iterativer Rekonstruktionsalgorithmen werden geeignete mathematische Werkzeuge benötigt. Deshalb werden im Abschnitt 2.3 passende mathematische Ansätze inklusive der wichtigsten Bildmodelle sowie iterative Lösungsstrategien vorgestellt. Da die meisten Super-Resolution Algorithmen Informationen aus mehreren zeitlich benachbarten Bildern zur Berechnung eines Ausgangsbildes verwenden, wird eine Bewegungsschätzung benötigt. Aus diesem Grund endet dieses Kapitel mit einer Erläuterung der wichtigsten Algorithmen zur Bewegungsschätzung in Abschnitt 2.4.

### 2.1 Grundlagen der Interpolation und Abtastung

Super-Resolution Algorithmen sind eine besondere Teilgruppe der Interpolationsalgorithmen. Im Allgemeinen ist es das Ziel einer Interpolation, Zwischenwerte einer Funktion mit Hilfe vorhandener Zahlenwerte zu bestimmen. Die generelle Vorgehensweise ist in Abbildung 2.1 gezeigt. In eine Folge aus Abtastwerten werden zuerst Nullen eingefügt, wodurch sich die Anzahl an Funktionswerten um den gewählten Interpolationsfaktor erhöht und die finale Anzahl an Stützstellen erreicht ist. Danach werden die Ergebnisse für die neu eingefügten Stützstellen durch die Filterung mit einem Interpolationstiefpass bestimmt. Die Form des Interpolationsfilters hat einen erheblichen Einfluss auf das Interpolationsergebnis. Während bei einer einfachen Pixelwiederholung (Filter mit rechteckförmiger Impulsantwort) Treppenstrukturen entstehen können, zeichnet sich das Ergebnis einer Spline-Interpolation [HA78] durch einen glatteren Kurvenverlauf aus. Eine Erklärung der Interpolation und die Eigenschaften der Interpolationsfilter im Frequenzbereich ist z.B. in [SB00] und [Len09] zu finden.

Wird ein kontinuierliches Signal abgetastet, so entstehen Wiederholspektren bei Vielfa-

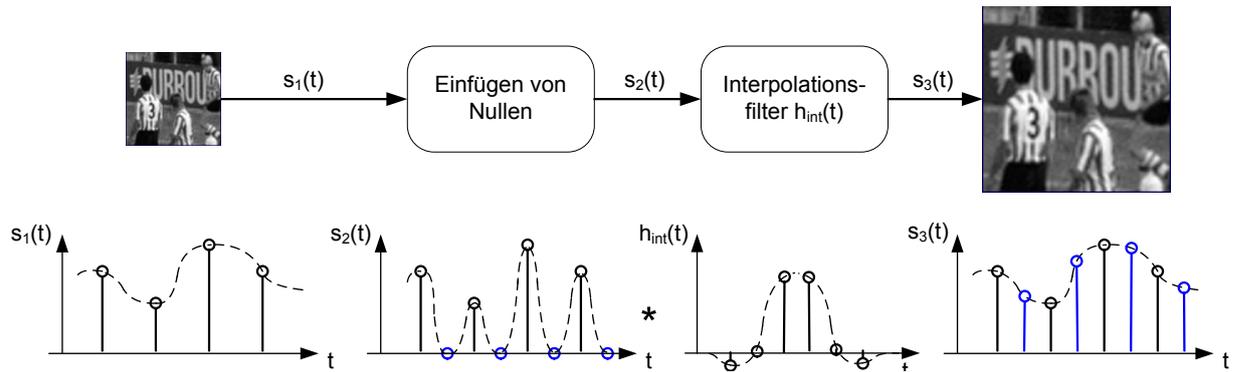


Abbildung 2.1: Interpolation durch Einfügen von Nullen und anschließender Filterung. Blockschaltbild (oben) und Signalverläufe nach den einzelnen Blöcken (unten)

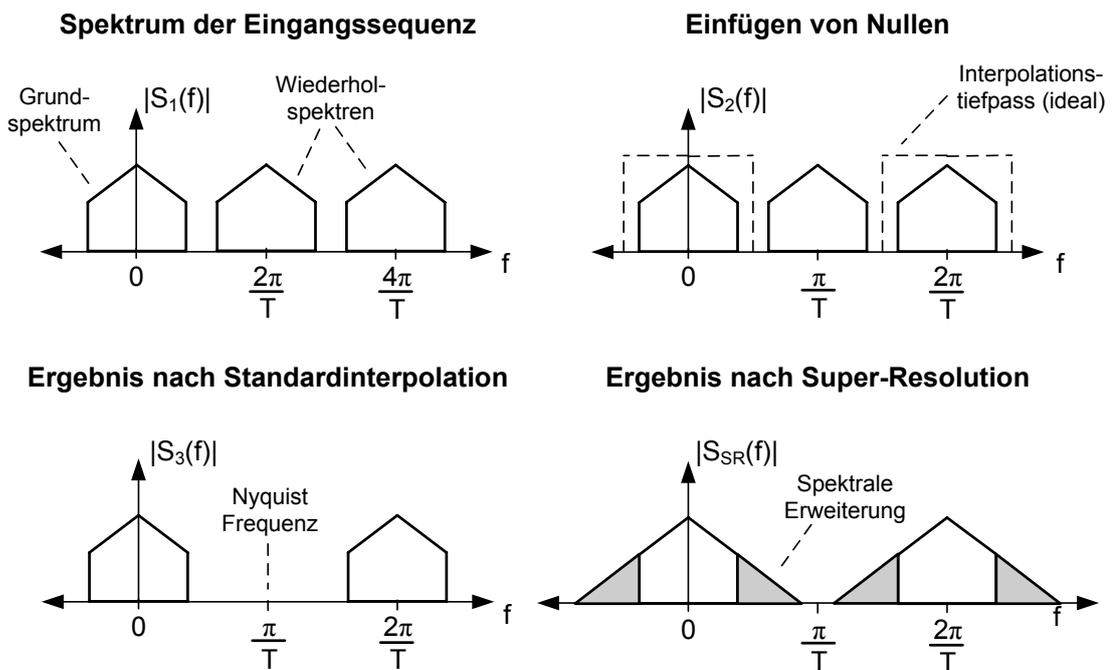


Abbildung 2.2: Spektrale Darstellung der Interpolation. Spektrum der Eingangssequenz, Zwischenergebnis nach Einfügen von Nullen mit eingezeichnetem Tiefpass, Interpolationsergebnis nach Standardinterpolation und nach Superresolution im Vergleich.

chen der Abtastfrequenz (siehe Abbildung 2.2). Soll die Abtastrate durch Interpolation erhöht werden, so müssen die Wiederholspektren zu Vielfachen der neuen Abtastfrequenz verschoben werden. Fügt man in die Originalsequenz Nullen ein, so führt dies lediglich zu einer Umskalierung der Frequenzachse, so dass (störende) Wiederholspektren entstehen. Diese Situation ist im Ausschnitt oben rechts von Abbildung 2.2 gezeigt. Anschaulich entstehen durch das Einfügen von Nullen mittel- und hochfrequente Störanteile, die im Frequenzbereich durch das Zwischenspektrum beschrieben werden können. Durch die sich anschließende Interpolationsfilterung können diese schließlich herausgefiltert werden. Diese mittel- und hochfrequenten Störanteile sind im Originalbereich ebenfalls beobachtbar, da die entstehende (kontinuierliche) Funktion nun durch die Nullstellen verläuft (siehe Signal  $s_2(t)$  in Abbildung 2.1). Um die Störspektren nun herauszufiltern, muss ein Tiefpass verwendet werden, dessen Grenzfrequenz auf die Spiegelfrequenz des ursprünglichen Abtastrasters ausgelegt ist. Da ein idealer Tiefpass auf Grund unendlicher Ausdehnung im Originalbereich in der Praxis nicht realisierbar ist, werden einfachere Approximationen dafür verwendet. Beispiele hierfür sind Filter mit rechteckförmiger, dreieckförmiger oder durch einen Spline beschriebene Filterimpulsantwort, die zu einer Pixelwiederholung bzw. lineare Interpolation oder High Resolution Spline (HRS) Interpolation führen. Je nach Wahl des Filters ergibt sich eine unterschiedliche Flankensteilheit und Komplexität. Dadurch werden die Störspektren unterschiedlich stark gedämpft sowie Anteile des Originalspektrum gefiltert. Dies macht sich durch die Sichtbarkeit von Störungen im Interpolationsergebnis einerseits als auch durch Unschärfe andererseits bemerkbar. Diese Unschärfe kann zum Teil durch eine nachträgliche Bildschärfekorrektur korrigiert werden (siehe Abschnitt 3.2.1). All diese beschriebenen Interpolationsverfahren verfolgen das Ziel, die Originalanteile bestmöglich zu erhalten und gleichzeitig die Störspektren möglichst effektiv herauszufiltern. Dabei bleibt die neu entstehende Lücke auf der Frequenzachse, welche auf dem höheren Abtastraster (neue hochfrequente) Information tragen könnte, ungenutzt. Das Auffüllen dieser Lücken mit sinnvollen Informationen ist das Ziel von Super-Resolution, wie es in Abbildung 2.2 unten rechts dargestellt ist. Diese Informationen können durch unterschiedliche Strategien erzeugt werden, welche im nachfolgenden Abschnitt 2.2 genauer beschrieben werden. Da diese Lücken im Spektrum erst durch eine Interpolation entstehen, kann eine Anwendung von Super-Resolution Algorithmen ohne vorherige Interpolation zu Störungen wie Alias führen.

Generell ist eine Übertragung der hier beschriebenen eindimensionalen Interpolation auf eine zwei- oder mehrdimensionale Verarbeitung sehr einfach möglich. Entweder wird eine separierte Verarbeitung in x- und y-Richtung durchgeführt oder aber eine echte 2-D Interpolation, wie sie beispielsweise in Abschnitt 3.2.2 beschrieben wird. Weitere Informationen finden sich z.B. in [SB00] und [Len09].

## 2.2 Einführung der Super-Resolution

Die historische Entwicklung von Super-Resolution wird bereits in zahlreichen Quellen behandelt (siehe z.B. [PPK03], [NSLZ07], [Mil05], [BS98]), so dass hier nur eine kurze Zusammenfassung basierend auf dieser Literatur gegeben wird. Im Bereich der Super-Resolution wird das Werk von Huang und Tsai aus dem Jahr 1984 [TH84] als die Pionierarbeit angesehen. In dieser Quelle wird das erste Mal die Strategie erwähnt, unter Ausnutzung von Alias aus mehreren niedrig aufgelösten Eingangsbildern ein hoch aufgelöstes Bild zu berechnen, wobei für die Berechnung der Ergebnispixel Eigenschaften der Fouriertransformation ausgenutzt werden. Auflösungsgewinne durch (globale) Subpixelbewegungen wurden von Ur und Gross [UG92] zur Super-Resolution ausgenutzt, so dass diese Arbeiten als Erweiterung von Huang und Tsai angesehen werden können. Zomet [ZP01] sieht in seiner Übersicht zum Thema Super-Resolution das erweiterte Abtasttheorem als Grundlage für diese Arbeit an, das auch in [VSVV07] als Basis für Super-Resolution verwendet wird. Wichtige Meilensteine sind die beispielbasierte (engl. Example Based) Super-Resolution [FJP02], bei der aus anderem Bildmaterial erzeugte hochfrequente Korrektursignale zum Eingangsbild hinzuaddiert werden, die iterative Back-Projection [IP91], welche Super-Resolution als ein Optimierungsproblem auffasst und die Formulierung der Super-Resolution in einem wahrscheinlichkeitsbasierten Framework unter Einbeziehung von Bildmodellen [BS99].

Obwohl das Thema Super-Resolution in der Literatur intensiv behandelt wird, finden sich dort keine präzisen Definitionen zu diesem Begriff. Die meisten Artikel (u.a. [BS98], [NSLZ07], [FREM04]) verstehen sinngemäß unter Super-Resolution die Berechnung eines hochfrequenten Bildes durch die Kombination von Informationen aus mehreren Eingangsbildern. In [Len09] wurde diese Definition erweitert, indem auch rein örtlich arbeitende Verfahren wie trainierte Filter [Kon01] als Super-Resolution Algorithmen aufgefasst werden, die durch Ausnutzung von Vorwissen hochfrequente Informationen erzeugen. Alternativ können diese hochfrequenten Information auch durch die Addition synthetischer Detailsignale generiert werden [Wu93].

In der hier vorliegenden Arbeit werden alle Algorithmen als Super-Resolution Algorithmen aufgefasst, die in jedem Ergebnisbild hochfrequente Anteile oberhalb der ursprünglichen Spiegelfrequenz (z.B. durch synthetische Korrektursignale) neu erzeugen oder aus vorhandenen Informationen wie Alias und Subpixelbewegungen wiederherstellen. Die unterschiedlichen Möglichkeiten zur Erzeugung oder Wiederherstellung werden später noch genauer betrachtet. Diese Erzeugung bzw. Wiederherstellung hochfrequenter Informationen erlaubt eine klare Abgrenzung zu herkömmlichen Interpolationsverfahren, die auf den Erhalt der Spektren der Originalsequenz abzielen. Insgesamt läßt sich eine Vielzahl von Verfahren unterscheiden. Die wichtigsten Ansätze werden in Kapitel 3 noch genauer vorgestellt. Die Berechnung der hochfrequenten Information kann durch die Addition synthetischer, nicht aus dem Bildinhalt selbst generierten Detailsignale [SEF86], [FJP02], on-line und off-line Optimierungsstrategien mit oder ohne Einbeziehung von Bildmodellen [Kon01], [TFM07]

und der Kombination von nicht redundanten Informationen aus mehreren Einzelbildern zu einem Ergebnisbild [NSLZ07] erreicht werden. Bei der Erzeugung dieser Information sind die größten Herausforderungen die präzise Schätzung der Subpixelbewegung, die Unterscheidung von neuer Information und Störungen sowie eine zeitlich stabile Positionierung der neu hinzukommenden Details in der Ergebnissequenz. Da die Unterscheidung von Details und Störungen sehr komplex ist, wird in den meisten Fällen auf eine hochpräzise und robuste Bewegungsschätzung gesetzt. Diese kann sowohl die Subpixelbewegung genau schätzen als auch eine exakte Positionierung der Details gewährleisten. Globale Bewegungsschätzungen gelten dabei als besonders robust, weshalb diese bevorzugt für die Super-Resolution eingesetzt werden [SVW02]. Die Robustheit dieser Schätzungen lässt sich dadurch erklären, dass nur sehr wenige Bewegungsparameter (je nach Bewegungsmodell zwischen zwei und zehn Parametern) aus einer sehr großen Datenmenge bestehend aus allen Pixeln eines Bildes geschätzt werden müssen.

Zusätzlich zur Erhöhung der Auflösung im Ergebnisbild weisen Super-Resolution noch weitere Eigenschaften auf. Verwendet ein Algorithmus ein Bildmodell oder nutzt die Kombination mehrerer Eingangsbilder zur Berechnung eines Ergebnisbildes, so kann neben der Auflösungserhöhung noch eine Störungsreduktion beobachtet werden. Dies führt zu einer besseren Wahrnehmung insbesondere bei Details mit sehr niedriger Amplitude, da für diese eine Maskierung durch starke Störungen sehr wahrscheinlich ist. Da ein Super-Resolution Algorithmus meist die Pixelzahl durch eine interne Interpolation erhöht und zusätzlich eine Störungsreduktion und Bildschärfeverbesserung umfasst, kann dieser mehrere Bildverarbeitungsaufgaben gleichzeitig lösen, die sonst von mehreren unterschiedlichen Algorithmen übernommen werden müssten. Dies ist eine weitere wichtige Eigenschaft von Super-Resolution Algorithmen.

## 2.3 Entwurfsmöglichkeiten für Algorithmen zur Super-Resolution

Kernziele beim Entwurf von Algorithmen für die Videosignalverarbeitung sind eine möglichst hohe Bildqualität der Ergebnisse bei einem gleichzeitig niedrigem Rechenaufwand. Zumeist widersprechen sich diese beiden Kriterien, so dass immer der beste Kompromiss zwischen beiden gefunden werden muss. Diese beiden Kriterien, erzielbare Ergebnisqualität und Rechenaufwand, müssen bereits bei der Entwicklung der Algorithmen berücksichtigt werden. Bei der Entwicklung von Algorithmen können im Allgemeinen unterschiedliche Strategien verfolgt werden. Eine sehr weit verbreitete und erfolgreiche Vorgehensweise ist der Entwurf eines Algorithmus basierend auf Modellannahmen. Kann dieses Modell direkt durch ein Filter oder eine Schaltung beschrieben werden, so ist der Entwurfsprozess bis auf eine gegebenenfalls noch anstehende Einstellung der (Modell-)Parameter bereits abge-

schlossen. Beispiele für diese Vorgehensweise finden sich in unterschiedlichen Bereichen der Videosignalverarbeitung, etwa der bewegungskompensierten inversen Filterung zur Reduktion von Bewegungsunschärfe [KV04a] oder adaptive Deblocking Algorithmen [KYKR98], [Pia08]. Obwohl diese Vorgehensweise nur sehr wenige generelle Schritte umfasst, kann der Entwurf sehr komplex sein, da die Findung eines geeigneten Modells nicht trivial ist und die Modelle teilweise eine sehr große Parameteranzahl benötigen, deren Einstellung eine große Herausforderung darstellt. Eine sich ggfs. anschließende Parameteroptimierung kann mittels Testsequenzen und objektiven Maßen automatisch erfolgen.

Eine weitere Möglichkeit, Algorithmen zur Videosignalverarbeitung zu entwerfen, ist die Auffassung des Algorithmus bzw. dessen Ziel (z.B. Störungsreduktion) als mathematischen Problemstellung. Diese Problemstellung kann z.B. ein Optimierungsproblem (Minimierung eines Energiefunktional), die Lösung einer partiellen Differentialgleichung oder eine Waveletanalyse bzw. -synthese mit integrierter Manipulation der Waveletbänder sein. Die Lösung dieser Problemstellung führt dann zu einem (meist iterativen) Algorithmus, der für den verwendeten Ansatz das optimale Ergebnis liefert. In [AK06], [CS05] und [WS04] werden diese Algorithmen intensiv betrachtet. Diese Algorithmen zeichnen sich durch eine sehr geringe Parameterzahl (meist weniger als 10) aus, so dass eine Parameteroptimierung oft sehr einfach möglich ist. Die iterative Lösung führt im Algorithmus zu einer Rückführungsschleife, die als eine Qualitätskontrolle und nachträgliche Korrektur des Ausgangsmaterials interpretiert werden kann, wie sie bei Algorithmen aus der Regelungstechnik bekannt ist. Auf Grund dieser Eigenschaften wird diese Entwurfsstrategie für die Entwicklung der Super-Resolution Algorithmen in dieser Arbeit verwendet. Zwar besitzt auch diese Entwurfsstrategie bzw. die daraus entwickelten Algorithmen Nachteile, allerdings überwiegen – wie später noch gezeigt wird – zum einen die Vorteile der Algorithmen den Nachteilen und zum anderen lassen sich diese Nachteile durch neue, in den Entwurfsprozess zu integrierende Schritte, bekämpfen.

Der gesamte Entwurf iterativer Rekonstruktionsalgorithmen lässt sich in mehrere Schritte unterteilen, die hier zu einem besseren Verständnis der späteren Kapitel erläutert werden sollen. Die Schritte zum Auffinden der Lösung sind im Einzelnen:

1. Aufstellung eines mathematischen Ansatzes
2. Transformationsschritte
3. (iterative) Lösung des Ansatzes

Diese drei Schritte sollen nun kurz erläutert werden. In den nachfolgenden Abschnitten werden diese dann getrennt voneinander betrachtet und die jeweils zur Verfügung stehenden Möglichkeiten vorgestellt. Die Aufstellung eines mathematischen Ansatzes hängt von der zu lösenden Bildverarbeitungsaufgabe (z.B. Artefaktreduktion oder Interpolation) ab. Mögliche Ansätze sind ein Energiefunktional oder eine partielle Differentialgleichung. In

den meisten Fällen enthält der mathematische Ansatz noch ein oder mehrere Bildmodelle, die das Aussehen der Lösung durch eine Einschränkung des Lösungsraumes beeinflussen können und zu einer mathematisch eindeutigen Lösung führen. Nachdem die Bildverarbeitungsaufgabe nun vollständig beschrieben ist, muss der Ansatz gelöst werden. In der Numerik [BSMM01] sind zahlreiche Gleichungslöser für lineare und nichtlineare Gleichungen bekannt. Um diese Verfahren einsetzen zu können, sind aber oft mehrere Transformationsschritte nötig, beispielsweise eine Diskretisierung oder die Überführung des gewählten Ansatzes in eine andere Beschreibungsform, die mit der Startvorschrift des Gleichungslösers übereinstimmt. Sind all diese Schritte abgeschlossen, ergeben sich konkrete Rechenvorschriften für den Algorithmus, der sich oft als rekursiver, adaptiver Filter realisieren lässt. Iterative Rekonstruktionsalgorithmen finden auch bereits in anderen Bereichen der Bild- und Videosignalverarbeitung Anwendung, u.a. in der Super-Resolution [FREM04], der Proscan-Konversion [KLF08], der Rauschreduktion durch Regularisierung [ROF92], in der Bewegungsschatzung [LK81] sowie in der Codierartefaktreduktion [Kau00], [DRS09a].

### 2.3.1 Mathematische Ansätze für Rekonstruktionsmethoden

In [AK06], [CS05], [WS04] werden die Grundlagen zur Formulierung von iterativen Rekonstruktionsalgorithmen beschrieben. Abbildung 2.3 fasst mögliche Ansätze inklusive wichtiger Eigenschaften und Vor- und Nachteile zusammen. Alle Ansätze besitzen die Gemeinsamkeit, dass sie aus zwei grundsätzlichen Arten von Termen bestehen. Terme, welche sowohl von der Eingangssequenz  $\mathbf{y}$  als auch von der Ausgangssequenz  $\mathbf{z}$  abhängig sind, werden als Datenterm bzw. Datenmodell bezeichnet. Diese gewährleisten eine Ähnlichkeit zwischen Eingangs- und Ausgangssequenz. Als Bildmodelle werden Terme bezeichnet, die nur von einer Sequenz (zumeist  $\mathbf{z}$ ) abhängig sind. Bildmodelle können bestimmte Strukturen oder Bildmerkmale erkennen bzw. beschreiben, so dass diese in der Ergebnissequenz hervorgehoben werden können. Mit  $\lambda$  kann eine Gewichtung zwischen beiden Arten von Termen vorgenommen werden. Mit diesem Faktor kann also die gewünschte Ähnlichkeit zwischen Eingangs- und Ausgangssequenz sowie gewünschter Strukturen im Ergebnisbild kontrolliert werden. Beide Arten von Termen werden im folgenden Abschnitt noch genauer erläutert.

Die Ansätze können in Energiefunktionale, partielle Differentialgleichungen (engl. Partial Differential Equation, PDE), stochastische Modellierung und Waveletanalyse bzw. -verarbeitung unterteilt werden. Mittels geeigneter Transformationen ist eine Überführung von einem in den anderen Ansatz möglich [CS05]. So kann beispielsweise die Wavelettheorie zur Lösung von Energiefunktionalen oder Ansätzen der Variationsrechnung verwendet werden [CDLL98].

Die Wahl des Ansatzes hängt deshalb von der Problemstellung bzw. der Problembeschreibung und der zur Verfügung stehenden Lösungsmöglichkeiten (Hardware) ab [AK06]. Während mittels Energiefunktionalen eine anschauliche Beschreibung inklusive interpre-

<p style="text-align: center;"><b>Energiefunktional</b></p> $\min_z \left\{ \int_{\Omega} (z-y)^2 d\Omega + \lambda \int_{\Omega}  \nabla z ^2 d\Omega \right\}$ <ul style="list-style-type: none"> <li>• Gute Interpretierbarkeit</li> <li>• Speicher- und Rechenintensive Lösungsalgorithmen</li> </ul>	<p style="text-align: center;"><b>Partielle Differentialgleichungen</b></p> $z_t = \Delta z + \lambda(y-z)$ <ul style="list-style-type: none"> <li>• Oft keine direkte Interpretierung möglich</li> <li>• Zahlreiche Lösungsalgorithmen vorhanden</li> <li>• Anwendung und Verhalten bestimmter partieller Differentialgleichung bekannt</li> </ul>
<p style="text-align: center;"><b>Stochastische Modellierung</b></p> $\max \left\{ p(z y) = \frac{p(z)p(y z)}{p(y)} \right\}$ <ul style="list-style-type: none"> <li>• Beschreibung stochastischer Merkmale (z.B. Texturen)</li> <li>• Suche nach der wahrscheinlichsten Lösung für das empfangene Signal</li> </ul>	<p style="text-align: center;"><b>Wavelets</b></p> $y = z + \frac{-\Delta z}{\lambda} = z + w$ <ul style="list-style-type: none"> <li>• Bandaufspaltung in niedrig und hochfrequente Bänder</li> <li>• Trennung und Optimierung bestimmter Eigenschaften gut möglich</li> </ul>

Abbildung 2.3: Mathematische Ansätze zur Beschreibung von Bildverarbeitungsaufgaben (vergleiche [CS05])

tierbarer Terme im Funktional entsteht, ist die direkte (numerische) Bestimmung des Minimums eines Funktionals meist sehr speicher- und rechenaufwändig, da die Matrizen und Vektoren innerhalb der Terme sehr groß werden können und eine Lokalisierung nicht immer möglich ist. Im Gegensatz dazu stehen für partielle Differentialgleichungen sehr viele schnelle Lösungsstrategien bereit, die insbesondere für die theoretische Physik entwickelt wurden [DRS09a]. Für einige Aufgaben wie beispielsweise eine Bildschärfverbesserung können auch direkt Differentialgleichungsansätze verwendet werden, da die Eigenschaften bekannter Differentialgleichungen genutzt werden können. Die Verwendung der Wärmeleitungsgleichung, der Laplace-Gleichung oder Shock-Filter [AK06] sind nur einige Beispiele dafür.

Die Beschreibung der Bildverarbeitung durch stochastische Prozesse aus der Wahrscheinlichkeitstheorie erleichtert die Beschreibung zufälliger Bildmerkmale bzw. -artefakte wie Texturen oder Rauschen. Weiterhin können dadurch Unsicherheiten im Modell gut ausgeglichen werden. Als Beispiele sind die MAP-Superresolution [BS99] oder der Kalman-Filter SR-Ansatz [FREM04] zu nennen, bei denen Störungen in der Eingangssequenz oder fehlerhafte Bewegungsvektoren im Modell berücksichtigt und damit unterdrückt werden können. Lediglich der Grad der Störung muss bekannt sein oder durch geeignete Strategien geschätzt werden.

Die Verarbeitung im Waveletbereich findet sowohl in der iterativen Rekonstruktionstechnik als auch bei direkten, nicht iterativen Verfahren Anwendung. Vorteil dieser Strategie ist die Aufteilung der Verarbeitung auf Teilbänder unterschiedlicher Frequenzen, die zum einen die Trennung von Störungen und Bildinhalten erleichtert und zum anderen eine gezielte Verarbeitung bestimmter Frequenzanteile [Mal99] möglich macht.

### 2.3.2 Modelle zur Beschreibung des Super-Resolution Prozesses

Für die Beschreibung des Super-Resolution Prozesses wird in den meisten Fällen ein Datenterm und ein Bildmodell benötigt. Der Datenterm hat hierbei die Aufgabe, die Ähnlichkeit zwischen niedrig aufgelöster und hoch aufgelöster Sequenz zu gewährleisten. Da jedes niedrig aufgelöste Bild (in Bezug auf die hoch aufgelöste Sequenz) nur über niedrige und mittlere Frequenzanteile verfügt, kann durch diesen Term nur eine Ähnlichkeit für diese Frequenzen garantiert werden. Hohe Frequenzanteile, die durch die Super-Resolution generiert bzw. wiederhergestellt werden sollen, können durch diesen Term nicht kontrolliert werden. Für diesen Zweck wird – wie bereits angesprochen – ein Bildmodell zur Kontrolle hochfrequenter Anteile benötigt. Da im Allgemeinen aber keine Informationen über die Eigenschaften der hohen Frequenzen in der Ergebnissequenz vorliegen, müssen für das Bildmodell Annahmen über die Charakteristik der hochfrequenten Anteile getroffen werden.

Das für die Super-Resolution am weitesten verbreitete Abbildungsmodell wird im nach-



Abbildung 2.4: Abbildungsprozess zwischen hoch und niedrig aufgelöster Sequenz

folgenden Abschnitt erläutert. Darauf folgt die Beschreibung von drei Bildmodellen, die unterschiedliche Annahmen über die hohen Frequenzanteile treffen. Die Auswahl eines dieser Bildmodelle hängt deshalb sehr stark von dem gewünschten Ergebnis ab.

### Datenmodell für die Super-Resolution

Als Datenmodell wird für die Super-Resolution das (lineare) Beobachtungsmodell verwendet, welches z.B. in [PPK03], [FREM04] für verschiedene Algorithmen eingesetzt wird und in Abbildung 2.4 dargestellt ist. Dieses Modell beschreibt die Beziehung zwischen hoch aufgelöster Sequenz  $\mathbf{z}$  und niedrig aufgelöster Sequenz  $\mathbf{y}$ . Basierend auf der Annahme, dass die niedrig aufgelöste Sequenz durch eine Aufnahme mittels Kamera aus einer generell deutlich höher aufgelösten kontinuierlichen Szenerie entstanden ist, kann jedes Bild durch eine Transformation bestehend aus Verschiebung bzw. Bewegung  $\mathbf{F}$ , Unschärfe durch schlechte Fokussierung und / oder Dezimationstiefpass  $\mathbf{B}$  und Unterabtastung  $\mathbf{D}$

ermittelt werden, so dass sich, unter Berücksichtigung von Rauschen  $\mathbf{e}_r$  für jedes niedrig aufgelöste Eingangsbild  $\mathbf{y}_n$  folgender Zusammenhang ergibt:

$$\mathbf{y}_n = \mathbf{DBF}_n \mathbf{z} + \mathbf{e}_r \quad (2.1)$$

Um nun ein Datenmodell aus diesem Zusammenhang zu formulieren, existieren in der Literatur verschiedene Schreibweisen. So ist die Formulierung eines Integrals über kontinuierliche Funktionen [ROF92], die Verwendung von Vektornormen [FREM04] oder reine Matrix-Vektorformulierungen [BS99] üblich, wobei über mathematische Umformungen ein direkter Zusammenhang zwischen den Formulierungen hergestellt werden kann. Für ein einheitliches Erscheinungsbild wird bei den hier vorgestellten Ansätzen immer erstere Schreibweise verwendet, da sich in dieser die Total Variation am anschaulichsten darstellen lässt. In den späteren Kapiteln werden die Ansätze zum Teil anders dargestellt, weil sich dadurch einfachere Lösungswege ergeben. Unter der Verwendung obiger Transformationsvorschrift ergibt sich so für das Energiefunktional:

$$\int_{\Omega} \sum_{n \in N} (\mathbf{y}_n - \mathbf{DBF}_n \mathbf{z})^2 \quad (2.2)$$

In diesem Funktional ist  $N$  die Anzahl aller am Super-Resolution Prozess beteiligter niedrig aufgelöster Eingangsbilder.  $\Omega$  bezeichnet den Bereich, über den das Integral berechnet wird. Dieser Integrationsraum umfasst in den meisten Fällen das gesamte Bild bzw. alle beteiligten Eingangsbilder. Setzt man geometrische Konvergenz voraus, so können Summe und Integral in der Berechnungsreihenfolge auch vertauscht werden, was je nach Lösung vorteilhaft sein kann. Bei diesem Funktional wird der mittlere quadratische Fehler zwischen Eingangssequenz  $\mathbf{y}$  und der der Transformation unterzogenen Ausgangssequenz  $\mathbf{z}$  berechnet. Mathematisch ergibt sich dieser Zusammenhang für die Verwendung der geometrischen Norm als Vergleichskriterium. In der Literatur wird alternativ auch die Maximumsnorm verwendet, so dass dort der maximale Fehler zwischen beiden Bildtypen kontrolliert werden kann [FREM04]. Die Wahl der Norm hängt vom gewünschten Ergebnis ab. Die Verwendung der geometrischen Norm, erlaubt eine Kontrolle des mittleren quadratischen Fehlers, weshalb große Ausreißer, hervorgerufen z.B. durch eine fehlerhafte Bewegungsschätzung, nur teilweise kontrolliert werden können. Diese können zwar durch die Maximumsnorm sehr gut kontrolliert werden, allerdings ergibt sich durch die Verwendung dieser Norm die Schwierigkeit, kleine Fehler nicht mehr oder nur noch ungenügend kontrollieren zu können. Durch Parameterangleichungen können aber für beide Normen sehr ähnliche Ergebnisse erzielt werden. In der vorliegenden Arbeit wurde ausschließlich die geometrische Norm verwendet und statt dessen das Ergebnis adaptiv angepasst. Gerade bei dem später vorgestellten rein rekursiven Algorithmus ist diese Vorgehensweise der Maximumsnorm überlegen, weil diese die Amplituden des Detailsignals auf die festen Werte -1, 0 und 1 einschränken würde, was Steuermöglichkeiten erheblich erschwert.

Das nur niedrige und mittlere Frequenzanteile mit diesem Funktional kontrolliert werden können, lässt sich durch den Unschärfefilter  $\mathbf{B}$  erklären. Je nach Wahl der Grenzfrequenz entfernt dieser alle hochfrequenten Anteile aus der hoch aufgelösten Sequenz  $\mathbf{z}$ . Dadurch kann beim Vergleich des tiefpassgefilterten Ergebnisbildes mit den niedrig aufgelösten Eingangsbildern das entstehende Korrektursignal in den hohen Frequenzbereichen nicht kontrolliert werden. Die dazu nötigen Bildmodelle werden in ihren Eigenschaften nachfolgend erklärt.

## Bildmodelle

Bildmodelle haben einen erheblichen Einfluss auf das Ergebnis iterativer Rekonstruktionsalgorithmen und die damit erreichbare Bildqualität. Durch sie ist es möglich, auch die neu entstehenden, hochfrequenten Anteile zu kontrollieren bzw. diesen eine gewünschte Form zu geben. In einem Optimierungsprozess werden sie als Nebenbedingung formuliert. Da nun durch die Kontrolle der hochfrequenten Anteile in Kombination mit dem bereits beschriebenen Datenmodell alle Frequenzanteile des Ergebnisses kontrolliert werden können, entsteht eine mathematisch eindeutige Lösung.

Generell dienen Bildmodelle zur Beschreibung wichtiger Merkmale in einem Ergebnisbild. Durch die Einbeziehung des Bildmodells in den Lösungsprozess weist das entstehende Ergebnis die Merkmale auf, die im Bildmodell formuliert wurden. Bildmodelle können für sehr unterschiedliche Zwecke verwendet und formuliert werden. In [CS05] findet sich eine ausführliche Beschreibung über Eigenschaften und Möglichkeiten von Bildmodellen. So können diese z.B. in Form einer deterministischen Funktion, einer Level-Set Repräsentation (z.B. für Segmentierungsaufgaben) oder eines Energiefunktionals formuliert werden. Da in der vorliegenden Arbeit auf Grund seiner sehr anschaulichen Beschreibungsform Energiefunktionale zur Formulierung der Ansätze gewählt wurden, sollen im folgenden nur Bildmodelle erläutert werden, die für die Super-Resolution relevant sind und die als Energiefunktional formuliert sind bzw. werden können. Die Tikhonov Regularisierung ermöglicht eine optimale Glättung des Rauschens, allerdings bestraft dieses Bildmodell Kanten sehr stark, was zu einer Unschärfe in diesen Bereichen führen kann. Diese Unschärfe kann durch ein adaptives Bildmodell, z.B. die Total Variation (TV) Regularisierung vermieden werden. Das Huber-Markov Zufallsfeld Bildmodell ist ebenfalls ein adaptives Bildmodell, wobei es zu seiner Berechnung nur Nachbarschaftsbeziehungen zu direkten Nachbarpixeln verwendet. Die Wahl des Bildmodells hängt also vor allem von zwei Aspekten ab, der zur Verfügung stehenden Hardware und dem gewünschten Aussehen des Ergebnisses. Trotz alledem muss berücksichtigt werden, dass alle Bildmodelle nur Hauptmerkmale von Bildern beschreiben können, die sich mathematisch formulieren lassen und wichtige Rahmenbedingungen erfüllen, die für die spätere Lösung nötig sind (z.B. Differenzierbarkeit, Konvexität, usw.). Aus diesen Gründen können komplexe Strukturen wie Texturen nur sehr bedingt abgebildet werden. Außerdem ist eine Unterscheidung von Bildmerkmalen und Störungen

auf Grund dieser Rahmenbedingungen sehr schwierig. So sind steile Kanten von Blocking-Artefakten nur durch komplexere, lokale Analysewerkzeuge unterscheidbar und nicht durch allgemein bekannte globale Bildmodelle iterativer Rekonstruktionstechniken. Komplexere, lokale Modelle wie z.B. ein Blocking-Modell [KYKR98] oder die Modellierung lokal variierender Bewegungsunschärfe [KV04b] finden deshalb in iterativen Rekonstruktionsalgorithmen bisher keine Anwendung. In der Artefaktreduktion und Super-Resolution wird in den meisten Fällen ein einfacher Regularisierungsansatz verwendet, der meist glättende Eigenschaften hat, um fehlerhafte Störstellen zu interpolieren oder Störungen wie Rauschen zu reduzieren.

### Tikhonov Regularisierung

Bei der Tikhonov Regularisierung handelt es sich um ein in [Tik43] vorgestelltes nicht adaptives Bildmodell, welches unabhängig von den Eingangspixelwerten dieselbe Filteroperation durchführt. Zur Beschreibung dieser Regularisierungsart kann folgendes Energiefunktional verwendet werden:

$$\int_{\Omega} |\nabla \mathbf{z}|^2 d\Omega \quad (2.3)$$

Dabei wird das Bild  $\mathbf{z}$  einer Hochpassfilterung mit dem Nabla-Operator unterzogen und das Ergebnis über den Bereich  $\Omega$  aufintegriert.  $\Omega$  erstreckt sich wie bei den Datenmodellen meist über das gesamte Bild, der Integrationsbereich kann aber auf einen Bildausschnitt verkleinert oder eine ganze Sequenz ausgedehnt werden. Durch die Hochpassfilterung werden hochfrequente Anteile extrahiert, so dass diese gesondert beurteilt werden können. Durch die quadratische Gewichtung der Signale werden insbesondere Anteile mit großer Amplitude wie beispielsweise steile Kanten sehr stark bestraft. Wird das Funktional nun minimiert, so ergibt sich als Ergebnis eine Tiefpassfilterung des Ausgangsbildes. Diese Filterung ist nicht adaptiv und kann dadurch realisiert werden, dass in jeder Iteration ein Hochpassanteil abgezogen wird. Das führt zu einer deutlichen Störungsreduktion. Jedoch werden auch die Kanten verschliffen, was Unschärfe zur Folge hat. Ein Hauptvorteil dieses Bildmodells ist seine einfache Realisierung durch feste Filterkoeffizienten. Der Nachteil, die Entstehung von Unschärfe, ist bei der Verwendung dieses Bildmodells in der Artefaktreduktion evtl. noch zu akzeptieren. Bei Super-Resolution, deren Ziel die Auflösungserhöhung ist, wäre diese Eigenschaft aber ungeeignet. Aus diesem Grund wird die Tikhonov Regularisierung im weiteren Verlauf der Arbeit nicht mehr betrachtet.

### Total Variation (TV) Regularisierung

Im Vergleich zur Tikhonov Regularisierung ist die Total Variation ein adaptives Bildmodell, welche von Rudin et al. in [ROF92] für einen Algorithmus zur kantenerhaltenden Rauschreduktion eingeführt wurde. Die Eigenschaft des Kantenerhalts ergibt sich dadurch,



Abbildung 2.5: Wirkung der Total Variation Regularisierung: Eingangsbild (links) und mit Total Variation Regularisierung verarbeitetes Bild (rechts)

dass auf die Quadrierung des Hochpassanteils verzichtet wird. Obwohl dieser Verzicht sehr simpel erscheint, hat dieser in der mathematischen Beschreibung weitreichende Folgen, die in [AK06] ausführlich diskutiert werden, hier aber der Vollständigkeit halber nur kurz erwähnt werden. Für den Kantenerhalt müssen ideale Sprünge beschrieben werden können, welche im Allgemeinen nicht differenzierbar sind, so dass eine Minimierung eigentlich nicht möglich ist. Erst durch eine Grenzwertbetrachtung und die Distributionentheorie kann dieses Problem gelöst werden.

Neben den stufenförmigen Kanten kann dieses Modell homogene Bereiche beschreiben, so dass dieses dazu geeignet ist, zwei Hauptmerkmale eines Bildes abzubilden: Steile, ideale (Stufen-)Kanten und homogene Flächen. Deshalb ist dieses Modell sehr gut für Comics und andere Bildinhalte ohne Texturen geeignet. Bei Bildern mit Texturen führt die Verwendung dieses Bildmodells zu einer Unschärfe in texturierten Bereichen oder sogar zu einer vollständigen Entfernung dieser Bildmerkmale (Comic-Transformation), da diese Strukturen nicht kontrolliert werden können und deshalb durch das Bildmodell entfernt werden (vergleiche Abbildung 2.5). Das Modell kann durch das folgende Funktional ausgedrückt werden:

$$\int_{\Omega} |\nabla \mathbf{z}| \, d\Omega \quad (2.4)$$

Im Vergleich zur Tikhonov Regularisierung ergibt sich bei einer Minimierung des entsprechenden Energiefunktional eine Filterstruktur, die Filterkoeffizienten beinhaltet, die proportional zum inversen des lokalen Gradienten sind [AK06], [DRS09a]. Dies gewährleistet den Erhalt von Kanten. Kleine und mittlere Variationen, die mit verhältnismäßig kleinen Gradienten korrespondieren, führen hingegen zu größeren Filterkoeffizienten. Infolgedessen werden neben Rauschen auch Bilddetails entfernt. In der Literatur werden mehrere Möglichkeiten vorgeschlagen, um dieses Problem zu bekämpfen. Dazu zählen eine adaptive Gewichtung des Total Variation Bildmodells in Abhängigkeit vom Bildinhalt [ZS06] oder die nicht lokalen (engl. non-local) Methoden, welche bei der Gradientenberechnung eine aus globalen Parametern bestimmte Gewichtungsfunktion einführen, um Unschärfe

in Texturbereichen zu vermeiden [ZC10]. Eine weitere Möglichkeit, die im Rahmen dieser Arbeit entwickelt wurde, wird in Kapitel 4 noch umfangreich beschrieben werden.

### Huber Markov Zufallsfeld (HMRF)

Das Huber Markov Zufallsfeld Bildmodell besitzt ebenfalls kantenerhaltende Eigenschaften, unterscheidet sich aber in mehreren Eigenschaften deutlich vom TV-Bildmodell. Zum einen wird zur Bestimmung des Hochpassanteils ein Laplace-Operator anstelle des Nabla-Operators verwendet. Zum anderen wird das Ergebnis einer (nichtlinearen) Gewichtungsfunktion  $\rho$  unterzogen, so dass die Bestrafung der Hochpassanteile je nach Amplitude unterschiedlich stark ausfällt. Im Vergleich zu den beiden anderen Bildmodellen kommt dieses Maß ursprünglich aus der Statistik. Die wichtigste Eigenschaft dieses Modells ist, dass nur direkte Nachbarschaftseigenschaften zur Berechnung der aktuellen Stelle benötigt werden. Für das Energiefunktional resultiert daraus folgender Term.

$$\int_{\Omega} \sum_i \rho(d_i(\mathbf{z})) d\Omega \quad (2.5)$$

Die Nachbarschaftseigenschaften führen bei diskreten Pixelanordnungen in einem rechteckigen Raster je nach Aufwand zu einer 4-er oder 8-er Nachbarschaft. Dadurch ergeben sich zwei bzw. vier Richtungen, für die die zweite Ableitung zu bilden ist. Die Summe über  $i$  beschreibt diesen Sachverhalt. Abbildung 2.6 zeigt für den Fall von vier Richtungen die zu berechnenden Ableitungen. Diese Ableitungen können als Kantendetektor interpretiert werden und formelmäßig folgendermaßen ausgedrückt werden:

$$d_i(\mathbf{z}) = \mathbf{z}_{k+di,l+dj} - 2\mathbf{z}_{k,l} + \mathbf{z}_{k-di,l-dj} \quad (2.6)$$

Wird die diskrete zweite Ableitung  $d$  in die gezeigten vier Hauptrichtungen gebildet, so nehmen die Verschiebungen  $di, dj$  die Werte -1, 1 und 0 an, je nachdem welche Richtung zu beschreiben ist.  $k, l$  bezeichnet die aktuelle Pixelposition. Anstelle nun diese Ableitungen direkt im Bildmodell zu verwenden, wird das Ergebnis einer Gewichtung unterzogen. Dadurch können hohe zweite Ableitungen, wie sie in Kantenbereichen oder texturierten Bereichen mit hohem Kontrast vorkommen, herunter gewichtet werden, so dass diese Bereiche während des Minimierungsprozesses einen kleineren Beitrag liefern und so letztendlich geschützt werden können. Kleinere Werte der zweiten Ableitungen, die beispielsweise im Rauschen ihre Ursache haben, können hingegen stärker bestraft werden, so dass für diesen Fall eine stärkere Glättung erzielt werden kann. Dies beruht auf der Annahme, dass Störungen im Vergleich zum Bildsignal bei akzeptablen PSNR-Werten eine relativ gesehen kleine Amplitude aufweisen. Die im HMRF-Bildmodell verwendete Gewichtungsfunktion zeigt Abbildung 2.7. Sie kann durch folgende Formel beschrieben werden:

$$\rho(d_i(\mathbf{z})) = \begin{cases} d_i(\mathbf{z})^2 & d_i(\mathbf{z}) \leq t_\alpha \\ 2t_\alpha |d_i(\mathbf{z})| - t_\alpha^2 & \text{sonst} \end{cases} \quad (2.7)$$

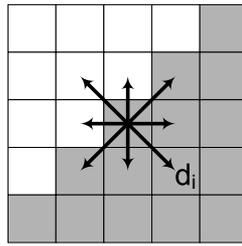


Abbildung 2.6: Richtung der vier Ableitungen beim Huber Markov Zufallsfeld Bildmodell

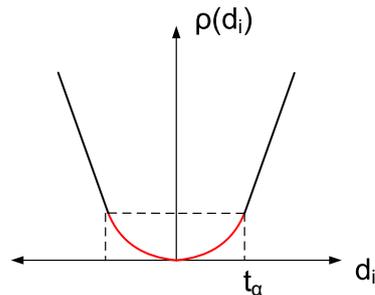


Abbildung 2.7: Bestrafungsfunktion für das Huber Markov Zufallsfeld Bildmodell

Der Schwellwert  $t_\alpha$  steuert dabei den Kantenerhalt oberhalb des Schwellwertes und die Glättung unterhalb von diesem. In Kapitel 5 wird genauer verdeutlicht, wie sich dieses Bildmodell nach einer Anpassung durch (lokale) Standardoperatoren der Bildverarbeitung realisieren lässt. Durch den Kantenerhalt bei gleichzeitiger Filterung der Bereiche mit niedriger Aktivität entspricht dieses Bildmodell den Eigenschaften einer kantensensitiven Rauschreduktion. Weitere Details zu diesem Bildmodell werden in [Bor96] erläutert.

### 2.3.3 Lösungsstrategien

Um den aufgestellten Ansatz nun in eine geeignete Lösungsform, z.B. eine Filterstruktur, zu überführen, muss eine Lösungsstrategie angewendet werden. Für diesen Zweck stehen durch die Mathematik und theoretische Physik eine Vielzahl von Lösungsstrategien bereit. Da dieses Thema bereits sehr umfangreich in der Literatur behandelt wurde (siehe u.a. [BSMM01], [AK06], [Cai06]), soll hier nur ein Einblick in die Verfahren gegeben werden, die für den weiteren Verlauf der Arbeit von Bedeutung sind. Während dieses Einblicks werden die einzelnen Verfahren klassifiziert und ein Bezug zur digitalen Filterung hergestellt.

### Transformationsschritte zur Vereinfachung der Lösungsstrategie

Da die meisten Ansätze für den kontinuierlichen Fall formuliert sind, muss der Ansatz diskretisiert werden, um maschinell gelöst werden zu können. Weiterhin ist es in manchen Fällen sinnvoll, den Ansatz in eine andere Beschreibungsform zu überführen, die inhaltlich genau dasselbe beschreibt, sich aber deutlich einfacher lösen lässt. Für beide Transformationsarten sollen nun Beispiele gegeben werden. Die Diskretisierung kann über eine Taylorentwicklung für beliebig hohe Ordnungen realisiert werden, wobei die Genauigkeit mit höherer Ordnung zunimmt [Bru01], [Bru06]. Wählt man eine sehr pragmatische Lösung mit Abbruch nach dem ersten Term für die Diskretisierung, ergeben sich sehr einfache Realisierungsstrukturen. So wird ein Integral in eine diskrete Summe und ein Gradient in ein Differenzschema überführt. Diese beiden Beschreibungen können beide auch als Filter interpretiert werden. Drei mögliche Strategien, einen Gradient zu approximieren, sind die Vorwärts-, Rückwärts und zentralen Differenzen:

$$(\mathbf{z}_x)_{k,l} \approx 0,5 \cdot (\mathbf{z}_{k,l} - \mathbf{z}_{k+1,l}) \quad (2.8)$$

$$(\mathbf{z}_x)_{k,l} \approx 0,5 \cdot (\mathbf{z}_{k,l-1} - \mathbf{z}_{k,l}) \quad (2.9)$$

$$(\mathbf{z}_x)_{k,l} \approx 0,5 \cdot (\mathbf{z}_{k-1,l} - \mathbf{z}_{k+1,l}) \quad (2.10)$$

Das gewählte Differenzschema kann einen erheblichen Einfluss auf das Aussehen der Lösung haben, da dadurch die Stärke der Glättung beeinflusst werden kann, wie es in [DRS09a] gezeigt wird. Dies kann dadurch erklärt werden, dass die beiden ersten Schemata einem Hochpass entsprechen, hingegen die zentralen Differenzen zu einem Bandpassfilter führen. Ebenso können auch Ableitungen höherer Ordnung durch Filter approximiert werden (siehe auch Abschnitt 4.2).

Eine Transformation des mathematischen Problems in eine andere Darstellung ist beispielsweise dann nötig, wenn für einen Ansatz nur sehr ineffiziente Lösungsstrategien bereit stehen, wie es für das Energiefunktional der Fall ist. Dies lässt sich durch die auf Matrizen basierende Formulierung erklären, so dass die Lösungsstrategie ebenfalls in Matrix-Vektordarstellung gegeben ist. Müssen nun während der Lösung Matrixmultiplikationen vorgenommen werden oder gelingt eine getrennte pixelweise Lösung des Problems nicht, so entsteht ein hoher Speicher- und Rechenaufwand. In [Cai06] werden zahlreiche Lösungsstrategien mit unterschiedlicher Komplexität und Konvergenzgeschwindigkeit vorgestellt, die für die Lösung eines MAP-Super-Resolution Verfahrens verwendet wurden. Der hohe Speicherbedarf und Rechenaufwand dieser Strategien erschwert aber die Akzeptanz der Super-Resolution in der Videosignalverarbeitung. Dieses Problem kann mit einer einfachen Transformationsvorschrift umgangen werden, welche als Euler-Lagrange Formalismus bezeichnet wird [AK06]:

$$\min \left\{ \int_{\Omega} F(\mathbf{z}, x, y, \mathbf{z}_x, \mathbf{z}_y) d\Omega \right\} \rightarrow \frac{\partial F}{\partial \mathbf{z}} - \frac{\partial}{\partial x} \frac{\partial F}{\partial \mathbf{z}_x} - \frac{\partial}{\partial y} \frac{\partial F}{\partial \mathbf{z}_y} = 0 \quad (2.11)$$

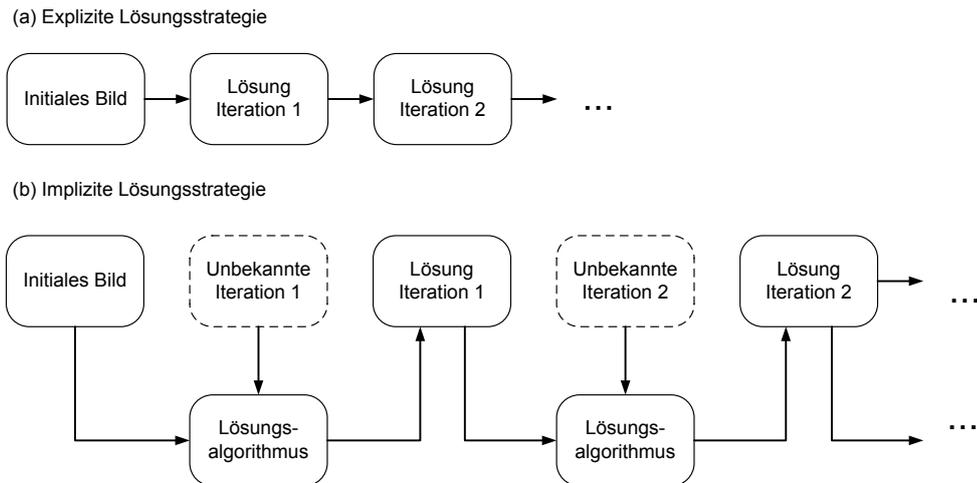


Abbildung 2.8: Explite und implizite Lösungsstrategie für iterative Algorithmen

$F$  entspricht einer Funktion, die vom (kontinuierlichen) Bild  $\mathbf{z}$ , den räumlichen Koordinaten  $x, y$  sowie den Ableitungen des Bildes in beide Raumrichtungen  $\mathbf{z}_x, \mathbf{z}_y$  abhängig ist. Das Ergebnis dieser Transformation ist eine partielle Differentialgleichung, für die es zahlreiche effiziente Lösungsstrategien gibt. Eine davon, die Kombination aus Gauß-Seidel Algorithmus und Lagged-Diffusivity, wird im nächsten Abschnitt beschrieben.

### 2.3.4 FIR- und IIR-Filterstrukturen zur Lösung von Energiefunktionalen und Gleichungssystemen

Möglichkeiten von elementweisen Lösungsstrategien werden in [AK06], [CS05], [FREM04] und [Kau00] beschrieben. Dazu können verschiedene Strategien angewendet werden. Zum einen kann durch die Variationsrechnung ein Energiefunktional in eine partielle Differentialgleichung überführt werden, die dann mit Standardverfahren wie Gauß-Seidel Algorithmus und/oder Lagged Diffusivity [VO96] gelöst werden kann. Durch ein geschicktes Umschreiben dieser Verfahren kann der Lösungsalgorithmus als nicht rekursive (FIR) oder rekursive Filterstruktur (IIR) beschrieben werden [Kau00], [DRS09a]. Nach [AK06] kann dabei zwischen der expliziten und impliziten Lösungsstrategie unterschieden werden (vergl. auch [DRS09b]), welche in Abbildung 2.8 dargestellt sind. Bei der impliziten Lösung werden neben bereits bekannten Elementwerten (Pixeln) auch (in der aktuellen Iteration noch) unbekannte Pixelwerte aus der Nachbarschaft benötigt, was zu einer IIR-Filterbeschreibung führt. Im Gegensatz dazu kommt die explizite Lösungsstrategie mit

Pixelwerten aus der vorigen Iteration aus, so dass eine Beschreibung durch einen FIR-Filter möglich ist. Beispiel für ein explizites Verfahren ist die Gradientenmethode:

$$\mathbf{z}_{q+1} = \mathbf{z}_q + \tau \nabla f(\mathbf{z}_q) \quad (2.12)$$

Die FIR-Filterstruktur ist deshalb gegeben, da die linke Seite der Gleichung nur von der Iteration  $q$  abhängt und nicht vom Schritt  $q + 1$  des aktuellen Ergebnisses. Die Gradientenmethode bestimmt iterativ (Iterationsvariable  $q$ ) die Lösung  $\mathbf{z}$ , welche die Funktion  $f$  minimiert. Bei Konvergenz des Verfahrens ist das Ergebnis der aktuellen und der vorherigen Iteration identisch, so dass sich  $\nabla f(\mathbf{z}_q) = 0$  ergibt. Dies ist genau die notwendige Bedingung für ein Minimum. Die hinreichende Bedingung, dass es sich tatsächlich um ein Minimum handelt und das dieses das globale Minimum ist, kann nur für konvexe Funktionen gewährleistet werden. Die Konvergenzgeschwindigkeit des Verfahrens hängt von der Iterationsschrittweite  $\tau$  ab. Wird diese zu klein gewählt, so kann die Konvergenz bei diesem Verfahren sehr langsam sein (siehe [Cai06]), eine zu große Schrittweite führt aber zu Divergenz.

Ein Beispiel für eine implizite Lösungsstrategie ist der Gauß-Seidel Algorithmus. Dieser löst ein allgemeines (lineares) Gleichungssystem der Form:

$$\mathbf{A} \cdot \mathbf{z} = \mathbf{y} \quad (2.13)$$

Anstelle nun das Ergebnis  $\mathbf{z}$  durch Inversion der Matrix  $\mathbf{A}$  und anschließender Multiplikation mit dem Eingangsbild  $\mathbf{y}$  direkt zu bestimmen, wird ein iteratives Lösungsverfahren gewählt, welches bei großen Matrizen den Rechenaufwand im Vergleich zur Inversion erheblich reduzieren kann. Das Lösungsverfahren wird dabei durch das elementweise Ausschreiben der Gleichung 2.13 und eine Umsortierung hergeleitet:

$$\left. \begin{array}{l} a_{1,1}\mathbf{z}_1 + a_{1,2}\mathbf{z}_2 + \dots + a_{1,N}\mathbf{z}_N = \mathbf{y}_1 \\ \vdots \\ a_{N,1}\mathbf{z}_1 + a_{N,2}\mathbf{z}_2 + \dots + a_{N,N}\mathbf{z}_N = \mathbf{y}_N \end{array} \right\} \sum_{j=1}^N a_{i,j}\mathbf{z}_j = \mathbf{y}_i \quad (2.14)$$

Auf der rechten Seite der Klammer befindet sich der Ausdruck für eine Zeile  $j$ . Löst man diese Gleichung nach dem aktuellen Element  $\mathbf{z}_i$  auf, so enthält man die Lösungsvorschrift für den Gauß-Seidel Algorithmus:

$$\mathbf{z}_i = \mathbf{y}_i - \sum_{j=1}^{i-1} a_{i,j}\mathbf{z}_j - \sum_{j=i+1}^N a_{i,j}\mathbf{z}_j \quad (2.15)$$

Durchläuft man die Matrix  $\mathbf{A}$  von oben links nach unten rechts und repräsentiert diese Matrix ein Bild, so sind die Pixelwerte der ersten Summe für die aktuelle Iteration bereits bekannt. Die anderen Pixel müssen aus einer weiter zurückliegenden Iteration verwendet

werden. Dadurch ergibt sich eine IIR-Filterstruktur. Diese wird noch deutlicher, wenn die Filterung auf den unmittelbaren Bereich um das aktuelle Pixel begrenzt ist, wie es in [Kau00] und [DRS09a] für Regularisierungsalgorithmen beschrieben ist. Sollen mit dem Gauß-Seidel Algorithmus auch nichtlineare Gleichungssysteme gelöst werden, so kann eine Linearisierung durch die so genannte Lagged-Diffusivity Methode [VO96] vorgenommen werden. Bei dieser Methode werden alle nichtlinearen Terme mit den bereits bekannten Werten berechnet, so dass sie im Gauß-Seidel Algorithmus als konstante Werte behandelt werden können. Nach der Konvergenz des Algorithmus können dann die nichtlinearen Terme erneut mit den aktuellen Werten berechnet werden. Dieses Vorgehen kann so lange wiederholt werden, bis auch die Neuberechnung konvergiert ist. Im Kapitel 4 wird dieses Vorgehen für Algorithmen der Regularisierung an einem Beispiel weiter verdeutlicht.

## 2.4 Bewegungsschätzung

Super-Resolution Algorithmen oder andere zeitlich arbeitende Verfahren berechnen ihre Ergebnisbilder durch geschickte Kombination von Informationen aus mehreren zeitlich benachbarten Bildern. Bewegen sich Objekte innerhalb der Sequenz oder werden Bewegungen durch die Kamera verursacht, muss diese Bewegung bestimmt und kompensiert werden, damit miteinander korrelierter Bildinhalt kombiniert werden kann. Ansonsten führt eine Kombination nicht korrelierter Bildinhalte zu sichtbaren Artefakten wie Unschärfe oder Kammstrukturen. Für Super-Resolution Algorithmen ist Bewegung in einer Sequenz sehr wichtig, da diese die Wahrscheinlichkeit nicht redundanter Informationen in den einzelnen Bildern erhöht, und zwar besonders, wenn es sich um eine Subpixelbewegung handelt. Die Bewegung selbst muss aber unbedingt korrekt geschätzt werden, da sonst eine Auflösungs-erhöhung durch eine zeitliche Filterung mehrerer Eingangsbilder nicht möglich ist. Da aber die Bestimmung eines Gütemaßes für die Korrektheit einer Bewegungsschätzung ein komplexes Problem darstellt, werden statt dessen sehr robuste und zuverlässige Algorithmen zur Bewegungsschätzung eingesetzt, so dass die kompensierten Bilder mit hoher Wahrscheinlichkeit korrekt sind und deshalb kein oder nur ein sehr einfaches Gütemaß benötigt wird, um die Qualität der kompensierten Bildinformation zu bestimmen. In der Literatur werden zu diesem Zweck oft robuste globale Bewegungsschätzer angewendet [SVW02], [FREM04], [PPK03]. Vorteile dieser Verfahren sind, dass nur wenige Parameter (z.B. von Kameraschwenks oder Zooms) aus einer großen Datenmenge geschätzt werden müssen und damit ein sehr zuverlässiges Ergebnis berechnet werden kann. Jedoch beschränken diese Verfahren den Einsatz der Super-Resolution auf Sequenzen, die nur eine überlagerte globale Bewegung der Kamera aufweisen. Nur sehr wenige Super-Resolution Verfahren greifen auf Blockmatching Verfahren zur Bewegungsschätzung zurück, die auch bei (lokalen) Objektbewegungen exakte Vektoren liefern können. Allerdings sind bei dieser Schätzung die Vektoren mit höherer Wahrscheinlichkeit fehlerhaft, weil nur eine kleinere Datenmenge für

jeden Block ausgewertet wird und diese auch noch von lokalen Störungen überlagert sein können. Da diese Verfahren aber vergleichsweise oft in der Endgeräte-Signalverarbeitung für Aufgaben wie Bildratenkonversion oder De-Interlacing eingesetzt werden [HBHO93], [Haa00], [Blu97] würde eine Verwendung dieser Verfahren zur Bewegungsschätzung für Super-Resolution Algorithmen die Einsatzchancen der Super-Resolution in der Signalverarbeitung für Endgeräte höchstwahrscheinlich erhöhen. Weiterhin kann so der Einsatzbereich der SR erheblich erweitert werden, weil nun auch in Bildbereichen mit lokalen Objektbewegungen die Auflösung potenziell erhöht werden kann. Die nächsten Abschnitte stellen nun die wichtigsten Grundlagen der Blockmatching Verfahren vor.

### 2.4.1 Blockmatching Verfahren

Blockmatching Verfahren können in zwei große Klassen unterteilt werden, den Algorithmen zur Suche nach der größten Ähnlichkeit zum aktuellen Block und den Algorithmen zur Bestimmung der *wahren Bewegung*. Beide Algorithmen wurden für unterschiedliche Anwendungsfälle entwickelt. In der Codierung ist es das Ziel, die Datenrate durch eine Ausnutzen zeitlicher Korrelationen zu reduzieren, so dass in benachbarten Bildern nach Blöcken mit größtmöglicher Ähnlichkeit (dem so genannten Residuum) gesucht wird. Dabei ist es nur von untergeordneter Bedeutung, ob dieser Block tatsächlich der (Objekt-) Bewegung entspricht. Im Gegensatz dazu wird in Anwendungen wie bewegungskompensierter Zwischenbildinterpolation oder Proscan-Konversion die *wahre Bewegung* benötigt [HBHO93], da nur so die Objektstrukturen erhalten bleiben. Da die tatsächliche Bewegung oft nicht mit dem Residuum übereinstimmt [BH04], unterscheiden sich auch die Blockmatching Algorithmen für beide Zwecke. Zur Bestimmung des Residuums wird häufig die erschöpfende Suche (engl. Fullsearch) eingesetzt (siehe u.a. [Haa00]). Das Grundprinzip soll mit Abbildung 2.9 verdeutlicht werden. Bei diesem Verfahren wird das aktuell vorliegende Bild in Blöcke unterteilt, wobei eine typische Blockgröße von 8x8 oder 16x8 beträgt. Aktuelle Bewegungsschätzer arbeiten dabei zumeist hierarchisch und erlauben eine adaptive Einteilung mit unterschiedlichen Blockgrößen in Abhängigkeit vom Bildinhalt [WSBL03]. Dadurch kann die Schätzgenauigkeit an Objektgrenzen erhöht werden, weil pro Block nur ein Vektor bestimmt wird und somit bei Grenzen zwischen sich unterschiedlich bewegendem Objekten nur für eine der beiden Bewegungen korrekt geschätzt werden kann. Nachdem das aktuelle Bild in Blöcke unterteilt ist, wird dann dieser Block im Referenzbild gesucht. Dazu wird ein Ähnlichkeitskriterium benötigt. Ein sehr oft verwendetes Kriterium ist die Summe absoluter Differenzen (SAD):

$$SAD = \sum_{i,j} |I(k+i, l+j, n) - I(k+i+v_x, l+j+v_y, n-1)| \quad (2.16)$$

Dabei wird die absolute Differenz der beiden Luminanzbilder  $I$  zwischen der Position im aktuellen Bild zum Zeitschritt  $n$  und der bewegungskompensierten Position zum Zeit-

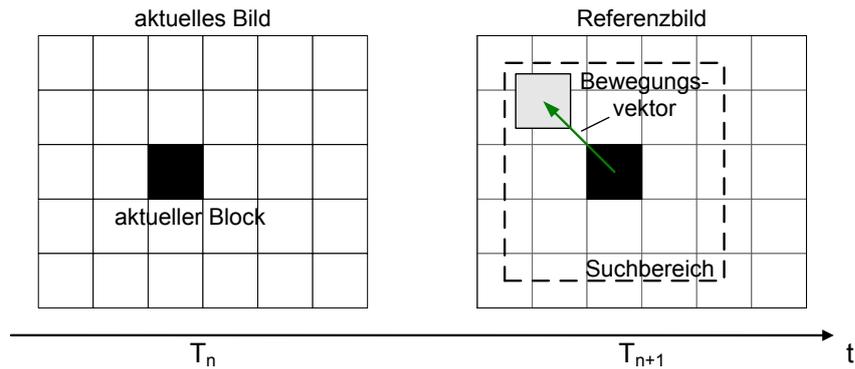


Abbildung 2.9: Bestimmung des Bewegungsvektors mit dem 'Fullsearch' Verfahren (vergl. [Haa00])

schritt  $n-1$  (verschoben um  $v_x$  und  $v_y$ ) berechnet und aufsummiert. Die Größe des Blockes, über den die absolute Differenz aufsummiert wird, beträgt in der Regel zwischen  $4 \times 4$  und  $32 \times 32$  Pixeln. Dieses Kriterium muss für alle möglichen Verschiebungen  $(v_x, v_y)$  bestimmt werden. Die Verschiebung, welche zum minimalen SAD führt, entspricht schließlich dem Bewegungsvektor. Um den Rechenaufwand zu begrenzen werden die Verschiebungen auf einen festen Suchbereich begrenzt. Dies beruht außerdem auf der Annahme, dass die Bewegung zwischen zwei Bildern nicht sehr groß ist (vergl. 3-D rekursive Suche [HBHO93]). Da die Überprüfung aller Blockpositionen sehr rechenaufwändig ist, wurden Varianten wie die hierarchische oder logarithmische Suche entwickelt, die allerdings zu der Gefahr führen können, nur ein lokales und nicht das absolute Minimum zu finden.

Im Gegensatz zur erschöpfenden Suche haben prädiktive Verfahren das Ziel, die wahre Bewegung zu schätzen [HBHO93], [Blu97]. Die 3-Dimensionale rekursive Suche (3DRS) war das erste dieser prädiktiven Verfahren, was später durch das parallel Prädiktive Blockmatching PPBM noch verbessert wurde. Die wahre Bewegung kann dadurch ermittelt werden, dass örtliche und zeitliche Korrelation zwischen benachbarten Blöcken benutzt werden, die Zahl der zu untersuchenden Vektoren auf sehr wenige aber dafür sehr zuverlässige Kandidaten eingeschränkt wird und zusätzliche zum SAD noch eine von den Kandidaten abhängige Bestrafung bei der Auswahl des finalen Bewegungsvektors verwendet wird. Diese Vorgehensweise bei der prädiktiven Bewegungsschätzung beruht auf zwei wichtigen Annahmen [HBHO93]:

- Objekte im Bild sind größer als die Blockgröße der Bewegungsschätzung
- (sich bewegende) Objekte besitzen eine Trägheit und ändern ihre Bewegung nur sehr langsam

Aus diesen beiden Annahmen lassen sich wichtige Grundzüge des Algorithmus ableiten. Dadurch, dass die Objekte meist größer als ein Block sind, besteht eine hohe Wahrschein-

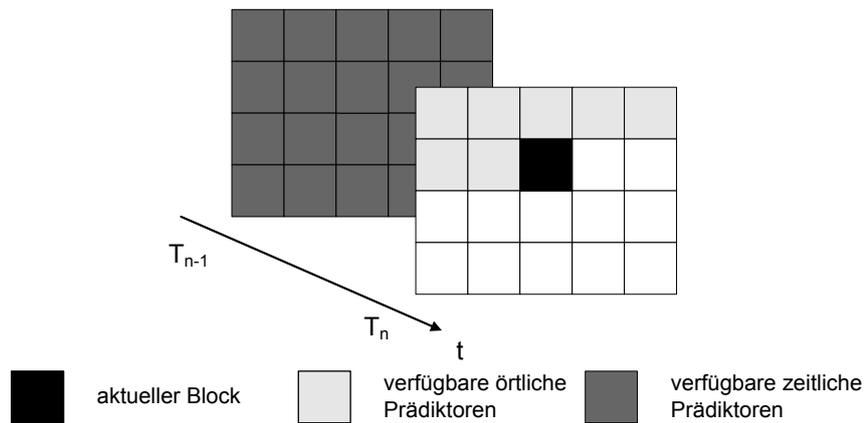


Abbildung 2.10: Verfügbare örtliche und zeitliche Prädiktorkandidaten [Blu97]

lichkeit, dass der aktuelle Block und seine Nachbarn identische Bewegungsvektoren haben werden. Gleiches gilt für die Blöcke aus dem aktuellen Bild und den Vektoren aus dem vorherigen Bild. Hierbei spielt die Trägheit der Objekte eine große Rolle, da die Bewegung nicht zu groß ist und es so eine große Überlappung der Bewegungsvektoren eines Objektes im aktuellen und vorherigen Vektorfeld gibt. Die Verwendung der Vektoren bereits abgearbeiteter Blöcke für die aktuelle Schätzung werden als Prädiktoren bezeichnet. Abbildung 2.10 verdeutlicht die zur Verfügung stehende Kandidatenmenge. Um den Rechenaufwand gering zu halten, wird die Kandidatenzahl stark eingeschränkt. Dazu werden für die Schätzung nur jeweils zwei örtliche und zwei zeitliche Prädiktoren herangezogen (siehe Abbildung 2.11). Die Anordnung der Prädiktoren beruht auf zahlreichen Simulationen [HBHO93], [Blu97]. Zusätzlich können Realisierungsaspekte berücksichtigt werden, so dass Ergebnisse parallel oder mit Pipelining Strukturen berechnet werden können [HBHO93]. So wird beispielsweise bei der 3-D rekursiven Suche (3DRS) nur auf örtliche Prädiktoren in der Blockreihe darüber zugegriffen, um bei der Berechnung Pipelining besser anwenden zu können, weil so das Ergebnis des direkt davor berechneten Blockes nicht vorliegen muss. Zusätzlich zu den bereits beschriebenen Prädiktoren werden noch weitere Kandidatenvektoren überprüft, um die wahre Bewegung bestimmen zu können. Da es viele unbewegte Bereiche in Bildern gibt, wird immer der Nullvektor überprüft. Weiterhin muss eine Geschwindigkeitsänderung der Vektoren innerhalb eines Bildes bzw. der Sequenz berücksichtigt werden. Aus diesem Grund wird ein Korrekturstern eingeführt, der in Abbildung 2.12 dargestellt ist. Dieser Korrekturstern wird auf die Spitze eines Prädiktors aufgesetzt und der zu prüfende Vektor um einen möglichen Kandidaten aus dem Korrekturstern korrigiert. Dies ermöglicht die Berücksichtigung von Änderungen im Vektorfeld wie Richtungsänderungen oder Beschleunigungen von Objekten. Rechenaufwand kann dadurch gering gehalten werden, dass pro Block wechselnd nur eine mögliche

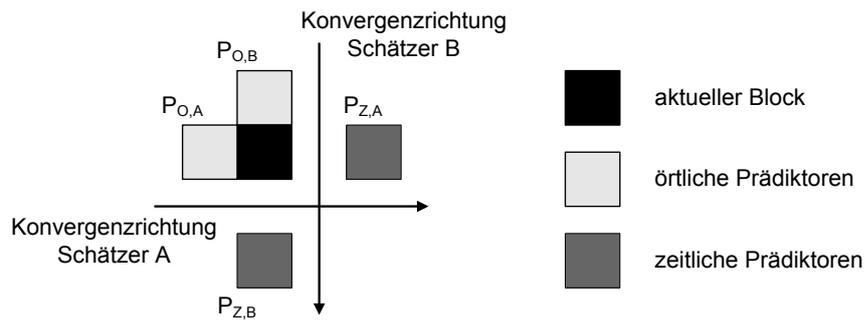


Abbildung 2.11: Relative Position der örtlichen und zeitlichen Prädiktoren beim Parallel Prädiktiven Blockmatching [Blu97]

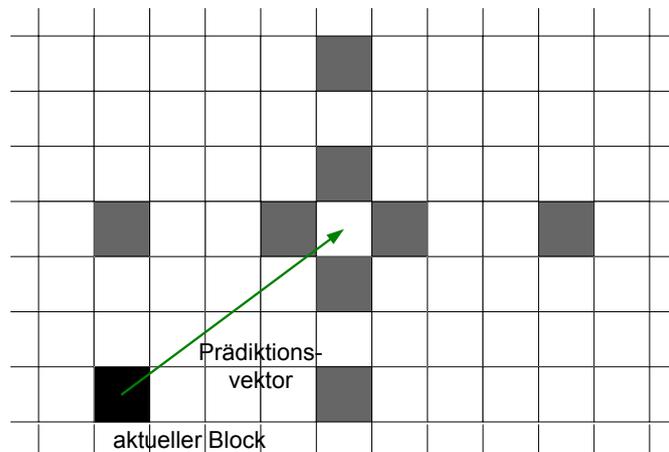


Abbildung 2.12: Korrekturstern für das prädiktive Blockmatching [Blu97]

Korrektur geprüft wird und andere Änderungen zyklisch für die nächsten Blöcke geprüft werden. Weitere Details zu dieser Strategie werden in [HBHO93] und [Blu97] beschrieben. Die Menge aller Kandidaten wird als Kandidatensatz bezeichnet. Durch die relativ kurzen Vektoren im Korrekturstern benötigt diese Schätzstrategie eine Konvergenzphase, um das finale Vektorfeld zu bestimmen. Um diese Einlaufphase zu verkürzen, kann die Schätzung pro Bild iterativ wiederholt werden, indem das bereits bestimmte aktuelle Vektorfeld aus einer früheren Iteration als zeitliche Prädiktion für den nächsten Iterationsdurchlauf verwendet wird. Dabei kann alternativ auch die Abarbeitungsreihenfolge, die normalerweise von links oben nach rechts unten erfolgt, umgekehrt werden [PK04]. Weiterhin werden zwei unabhängig voneinander arbeitende Schätzer verwendet, deren Prädiktoren orthogonal zueinander angeordnet sind, wie es Abbildung 2.11 verdeutlicht. Dies ist bei orthogonal zur Bewegungsrichtung verlaufenden Objektgrenzen von großem Vorteil, da vertikal zur Bewegungsrichtung die örtlichen Prädiktoren nur eine sehr geringe Änderung erfahren.

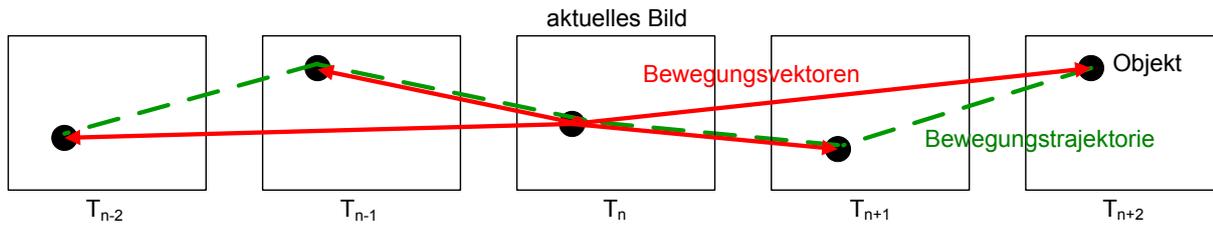


Abbildung 2.13: Bewegungsvektoren einer Mehrreferenzbildbewegungsschätzung

## 2.4.2 Mehrreferenzbildschätzung

Die bisher beschriebene Bewegungsschätzung bestimmt jeweils nur die Bewegung zwischen dem aktuellen und einem direkten Nachbarbild. Eine bidirektionale Bewegungsschätzung bestimmt Bewegungsvektoren sowohl zum vorigen als auch zum nachfolgenden Bild. In der Codierung wurde mit dem H.264 Standard eine Multirahmen Bewegungsschätzung (engl. Multi-Reference Frame Motion Estimation) eingeführt, die zusätzlich Bewegungsvektoren zu weiter entfernten Referenzbildern bestimmt (siehe Abbildung 2.13). In [WSBL03] wird diese Strategie näher beschrieben. Dieser Algorithmus ist eine Weiterentwicklung der erschöpfenden Suche, wobei der Suchbereich für die anderen Referenzbilder in Abhängigkeit der bereits bestimmten Vektoren für die anderen Referenzbilder angepasst wird. Speichergrößen und Recheneffizienz machen diese Korrektur erforderlich. Im Rahmen der vorliegenden Dissertation wurde dieses Verfahren erheblich erweitert, um auch die wahre Bewegung schätzen zu können. Diese Erweiterungen werden in Kapitel 5 beschrieben.

## 2.4.3 Weitere Verfahren

Neben Blockmatching Verfahren existieren beispielsweise noch Gradientenverfahren (auch optischer Fluss genannt) [LK81], [HS81] und in der Fourier-Domäne arbeitende Verfahren. Beim Gradientenverfahren wird die Bewegung mittels örtlicher und zeitlicher Gradienten berechnet. Basis für diese Berechnung ist die Identitätsgleichung:

$$I(x, y, t) = I(x + v_x t, y + v_y t, t - 1) \quad (2.17)$$

Diese Gleichung beruht auf der Annahme, dass der Bildinhalt des aktuellen und vorigen Bildes identisch ist, sofern das vorige Bild bewegungskompensiert wird. Aus dieser Gleichung lässt sich nach einer Taylorentwicklung folgende Gleichung herleiten:

$$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0 \quad (2.18)$$

Dies führt zu einem unterbestimmten Gleichungssystem mit den zwei Unbekannten  $v_x$  und  $v_y$  für den Bewegungsvektor. Zur Lösung dieses Problems existieren sowohl lokale Ansätze

über eine Minimierung des mittleren quadratischen Fehlers [LK81] als auch ein Energiefunktionalansatz unter der Einbeziehung von Nebenbedingungen [HS81]. Details können in der entsprechenden Literatur gefunden werden.

Phasenkorrelation [KH75] ist ein auf der Fouriertransformation basierendes Verfahren zur Bewegungsschätzung. Bei diesem Verfahren wird das aktuelle Bild und das Referenzbild fouriertransformiert. Gemäß dem Verschiebungssatz der Fouriertransformation entspricht eine Bewegung dann einem Phasenterm im Spektrum. Durch eine entsprechend normierte Kreuzkorrelation beider Fouriertransformierten kann dann die Verschiebung ermittelt werden. Für beide Verfahren gibt es Anwendungsbeispiele in der Super-Resolution [ZS02], [HJKH09].

Da die meisten Super-Resolution Algorithmen globale Bewegungsschätzer verwenden, sollen diese hier noch kurz erläutert werden. Im Gegensatz zur Bestimmung lokaler Objektbewegungen, konzentrieren sich globale Verfahren darauf, einige wenige Parameter zu schätzen, die einem globalen Bewegungsmodell entstammen [Smo01]. Im Falle einer rein translatorischen Bewegung besitzt das Modell zwei Parameter, darüber hinaus existieren noch Modelle zur Erfassung von Zoom und Rotation mit vier Parametern oder auch parabolische Modelle. Die Schätzung dieser Parameter basiert dabei entweder auf Optimierungskriterien bzw. dem Gradientenverfahren oder aber der Schätzung globaler Bewegungen mittels Matching von Merkmalspunkten (siehe [Smo01]). Da diese Verfahren aber in der vorliegenden Dissertation nicht mehr weiter untersucht werden, sei hier auf die entsprechende Literatur verwiesen.

# Kapitel 3

## Super-Resolution Verfahren

Dieses Kapitel beinhaltet wichtige Konzepte und konkrete Verfahren für die Super-Resolution, welche später als Basis für die neu entwickelten Verfahren dienen oder für Vergleiche herangezogen werden. Das Kapitel ist in zwei Hauptteile unterteilt. Im ersten Teil werden wichtige Verarbeitungskonzepte vorgestellt, auf denen die meisten Super-Resolution Algorithmen aufbauen, und die für eine allgemeine Analyse der Vor- und Nachteile genutzt werden können. Dazu zählt die direkte, nicht iterative Verarbeitung, welche das Ergebnis in einem Schritt und nur auf Vorwissen basierend berechnet. Im Gegensatz dazu nutzen iterative (Rekonstruktions-)Algorithmen eine Rückführungsschleife, um die Qualität des Ergebnisses schrittweise in mehreren Iterationen zu verbessern. Dadurch ist eine Kontrolle der Bildqualität am Ausgang möglich. Als Beispiele für diese Konzepte werden daran anschließend das Multiple Algorithm Blending (MABLE) und das Grundprinzip iterativer Super-Resolution Algorithmen angeführt. Aufbauend auf diesen Konzepten werden dann im zweiten Teil des Kapitels konkrete Verfahren vorgestellt. Den Schwerpunkt bilden hier die iterativen Rekonstruktionsalgorithmen zur Super-Resolution, insbesondere die Back-Projection und die Maximum-A-Posteriori Super-Resolution, da diese später als Grundlage für die Verbesserungen und Neuentwicklungen dienen.

### 3.1 Konzeptuelle Ansätze für die Super-Resolution

Bevor die konkreten Verfahren beschrieben werden, sollen hier kurz mögliche Konzepte diskutiert werden. Dadurch ist zum einen eine Klassifizierung der einzelnen Verfahren möglich, zum anderen können dadurch generelle Stärken und Schwächen leichter verdeutlicht werden. In der Super-Resolution sind vor allem zwei Grundkonzepte von Bedeutung. Die eine Klasse von Algorithmen berechnet die hoch aufgelösten Bilder ohne die Verwendung von Rekursionsschleifen. Diese Verfahren sollen im weiteren Verlauf als direkte Verfahren bezeichnet werden. Die andere Klasse von Algorithmen basiert oft auf der Optimierung einer mathematischen Gleichung oder eines Funktionals, wodurch sich ein rekursiv oder iterativ arbeitender Algorithmus ergibt. Diese Verfahren werden im weiteren Verlauf als

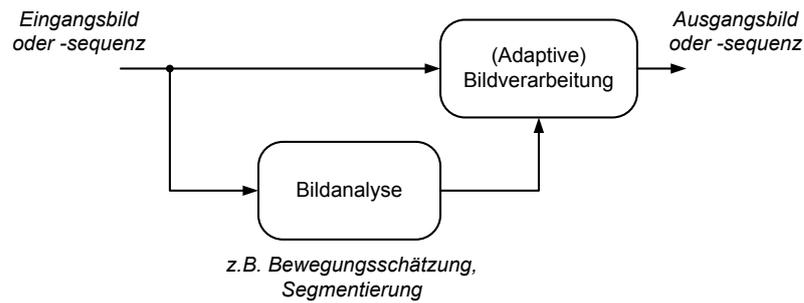


Abbildung 3.1: Schematische Darstellung der direkten (nicht-iterativen) Bildverarbeitung

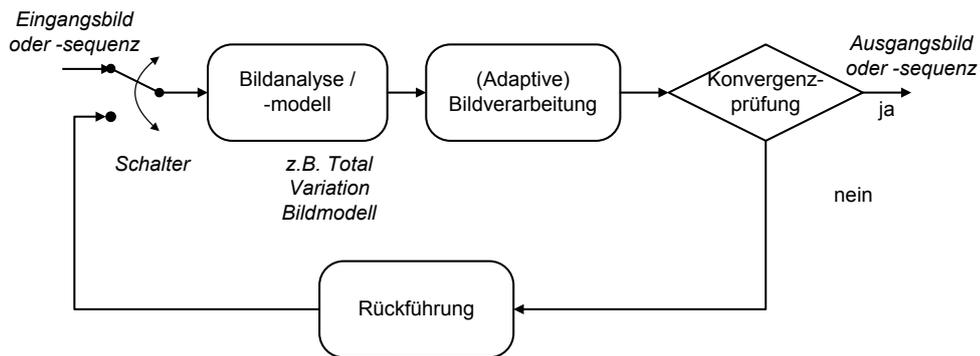


Abbildung 3.2: Schematische Darstellung der iterativen Bildverarbeitung

iterative Verfahren bezeichnet. Aufbauend darauf können dann weitere Verbesserungen und die Grundstruktur der Super-Resolution erklärt werden.

### 3.1.1 Direktes und iteratives Verarbeitungskonzept

Wie bereits erwähnt, können für SR-Verfahren zwei grobe Verarbeitungskonzepte unterschieden werden. Da bereits an diesen Konzepten wichtige Vor- und Nachteile verdeutlicht werden können, sollen diese Konzepte zuerst vorgestellt werden. Das erste Verarbeitungskonzept benutzt zur Berechnung der Ergebnissequenz nur Informationen, die aus der Eingangssequenz ermittelbar sind (siehe Abbildung 3.1). Da für dieses Konzept keine Rückführungs- oder Iterationsschleife benötigt wird, soll dieses Konzept als direktes Verarbeitungskonzept bezeichnet werden. Das iterative Verarbeitungskonzept verfügt im Gegensatz dazu über eine Rückführungsschleife, so dass das Ergebnis in einem oder mehreren Verarbeitungsschritten sukzessive verbessert werden kann. Ein Prinzipschaltbild dieser Verarbeitungsweise befindet sich in Abbildung 3.2. Werden Begriffe der Regelungstechnik zur Charakterisierung dieser Konzepte verwendet, so entspricht das direkte Konzept einer Steuerung und das iterative Konzept einer Regelung.

Tabelle 3.1: Eigenschaften der direkten und rekursiven Verarbeitung

	direkte Verfahren	iterative Verfahren
Anpassungsmöglichkeiten an Bildmaterial unterschiedlicher Qualität	verschiedene Parametersätze oder Algorithmen für unterschiedliches Material nötig	globale Optimierung bzw. Einstellung der Verarbeitung zur Laufzeit
Laufzeit	vorhersagbare Laufzeit durch keine oder feste Iterationsanzahl	schwankende Anzahl an Iterationen bis Optimum erreicht
Seiteneffekte	Fehldetektion in der Analyse kann zu Störungen führen; Ergebnisqualität wird nicht überprüft	Entfernung von Bildmerkmalen, die nicht durch die Modelle beschrieben sind
Beschreibungsform	Fallunterscheidungen und andere nichtlineare Operation möglich; teilweise sehr große Parameterzahl	Operationen müssen mathematische Rahmenbedingungen erfüllen, kleine Parameterzahl
Optimierung	unter Laborbedingungen	zur Laufzeit
Bildmodelle	Einbeziehung komplexer, auch nichtlinearer Effekte	beschränkt auf einfache, mathematische Terme

Die Vor- und Nachteile beider Konzepte sind in Tabelle 3.1 zusammengefasst. Ähnlich wie in anderen Bereichen kann auch bei der Wahl des Verarbeitungskonzeptes zwischen Rechenaufwand und erzielbarer Bildqualität ein passender Kompromiss gewählt werden. Das direkte Konzept besitzt dabei meist einen geringeren Rechen- und Speicheraufwand, da es über keine Rückführungsschleife verfügt, welche in Form eines Bildspeichers realisiert werden muss. Durch die fehlende Rückführungsschleife hat das direkte Konzept eine feste, vorhersagbare Laufzeit, so dass unter Verwendung entsprechend leistungsfähiger Hardware Echtzeitfähigkeit gewährleistet werden kann. Im Gegensatz dazu hängt die Laufzeit beim iterativen Konzept von der Eingangsbildqualität ab. Dabei wird das Ergebnis so lange korrigiert, bis ein Optimum erreicht ist. Die Korrekturschritte, die zu einem höherem Speicher- und Rechenaufwand führen, versprechen aber generell auch eine deutlich höhere Bildqualität. Diese lässt sich dadurch erklären, dass sichtbare Seiteneffekte am Ausgang, bei einer Super-Resolution z.B. ein zu scharfes Ergebnis oder Artefakte durch fehlerhafte Bewegungsvektoren, durch Bildmodelle erkannt und in einem weiteren Iterationsschritt korrigiert werden können. Da bei dem direkten Konzept diese Anpassungen durch das Konzept selbst nicht abgedeckt sind, müssen die Fehler entweder toleriert

oder durch eine Nachverarbeitung beseitigt werden. Zu diesem Zweck werden oft Fehlerverschleierungstechniken eingesetzt [FTS01]. Neben der besseren Anpassung bzw. Korrektur von Fehlern, kann durch die Rückführung auch besser auf Material unterschiedlicher Qualität reagiert werden. Dies liegt daran, dass der Optimierungsprozess selbst zur Laufzeit erfolgt. Beim direkten Verarbeitungskonzept muss diese Optimierung zuvor mittels geeigneter Testsequenzen durchgeführt werden. Zur Laufzeit kann dann lediglich eine Parameteranpassung oder Algorithmenauswahl erfolgen. Die Auswahl selbst hängt dabei nur von den aus der Eingangssequenz generierten Analyseinformationen ab. Bei der Super-Resolution können diese Analyseinformation beispielsweise Segmentierungsinformationen oder Bewegungsvektoren sein. Der Analyse kommt deshalb eine hohe Bedeutung zu, weil von ihr die nachfolgende Verarbeitung abhängt. Da diese Analyse nicht von einem mathematischen Modellansatz (z.B. einem Optimierungsproblem) abhängen muss, kann diese separat von der Verarbeitung entworfen werden und Operationen enthalten, die keine mathematischen Rahmenbedingungen erfüllen müssen. Um trotz der reinen Steuerung eine möglichst gute Bildqualität für unterschiedliches Eingangsmaterial zu erzielen, müssen die Parameter dieser Verfahren optimiert werden. Dies geschieht oft unter Zuhilfenahme von Bildqualitätsmaßen wie dem SSIM (Structural Similarity Index Measure, [WS04]) oder dem PSNR (Peak Signal to Noise Ratio) im Labor mittels geeigneter Testsequenzen. Die Auswahl der Testsequenzen ist dabei von entscheidender Bedeutung. Um möglichst alle denkbaren Eingangsqualitäten und -auflösungen geeignet verarbeiten zu können, werden ggfs. mehrere Parametersätze benötigt. Trotzdem stellt diese Optimierung immer nur einen Kompromiss dar, weil die Parametereinstellung nicht auf das Eingangsbildmaterial selbst abgestimmt werden kann. Weiterhin kann dies zu einer großen Parameterzahl führen. Im Vergleich dazu sind die Modelle bzw. Analyseschritte beim iterativen Konzept oft sehr einfach, da sich diese direkt aus dem mathematischen Ansatz ergeben. Das erschwert die Beschreibung komplexer Bildmerkmale (z.B. Texturen), so dass diese Merkmale durch die Verarbeitung entfernt werden. Dafür lassen sich diese Ansätze oft unter der Verwendung sehr weniger Parameter beschreiben.

### 3.1.2 Multiple Algorithm Blending (MABLE)

Beim Multiple Algorithm Blending handelt es sich um ein direktes Verarbeitungskonzept, bei dem durch eine geschickte Kombination mehrerer Algorithmen für jede Situation das jeweils beste Verfahren ausgewählt wird. Dieses Konzept wird in [INXO10] auch bereits auf die Super-Resolution angewendet. Die Verwendung dieses Konzeptes lässt sich dadurch motivieren, dass in der Bildverarbeitung nahezu kein Verfahren existiert, welches in jeder Situation optimale Ergebnisse erzielt. Die meisten Verfahren haben vielmehr in unterschiedlichen Situationen ihre Vor- und Nachteile. So zeichnen sich iterative Rekonstruktionsalgorithmen durch ihre im Vergleich zu direkten Verfahren recht einfache mathematische Beschreibungsform mit wenigen Parametern aus, was eine Einstellung die-

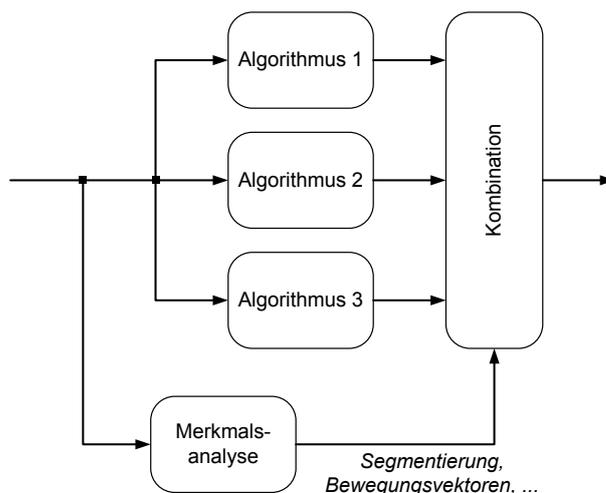


Abbildung 3.3: Prinzipschaltbild für das 'Multiple Algorithm Blending'

ser Algorithmen erleichtert. Allerdings ist es durch die einfachen, für das ganze Bild global formulierten Bildmodelle nur möglich, wesentliche Merkmale von Bildern und Sequenzen abzubilden. Komplexere Strukturen wie Texturen oder eine Unterscheidung von Artefakten und Bilddetails sind aber nur bedingt realisierbar, was in diesen Bereichen zu einer suboptimalen Verarbeitung führt. Da in diesen Bereichen ebenfalls Verfahren mit Stärken existieren, wurde in der Literatur ein Ansatz vorgeschlagen, die jeweiligen Vorteile zu nutzen, ohne die Nachteile in Kauf nehmen zu müssen [SMD09], [ZW09], [DRS09a], [RDS09a]. Bei dieser Strategie werden die Merkmale ermittelt, unter denen die jeweiligen Algorithmen ihre Vorteile besitzen, und basierend auf diesen eine Merkmalsanalyse / Klassifikation entwickelt, um die Sequenz nach diesen Kriterien segmentieren zu können. Alle Algorithmen werden dann parallel angewendet und mittels der Segmentierungsmaske das optimale Ergebnis ausgewählt. Abbildung 3.3 verdeutlicht dieses Prinzip. Das Umschalten zwischen den Algorithmen findet dabei nicht hart sondern mit einer weichen Überblendung bzw. Gewichtung zwischen den Verfahren statt, um eine höhere zeitliche Konsistenz zu erzielen und Umschaltartefakte zu vermeiden. Aus diesem Grund werden diese Verfahren als Multiple Algorithm Blending bezeichnet. Erkauft wird die deutlich bessere Bildqualität durch einen um ein Vielfaches erhöhten Hardware- und Rechenaufwand. Anwendung dieses Verfahren sind in der Interpolation [SMD09], der Rauschreduktion in Kombination mit adaptiver Bildschärfekorrektur [ZW09], der örtlichen Artefaktreduktion [DRS09a] und der örtlich-zeitlichen Artefaktreduktion [RDS09a], [RSS10b] zu finden. Beim letzten Verfahren wird zusätzlich auch eine örtlich-zeitliche Analyse zur Erkennung fehlerhafter Bewegungsvektoren verwendet. Die Kapitel 4 und 5 werden noch weitere Anwendungen für diese Verarbeitungsstrategie zeigen.

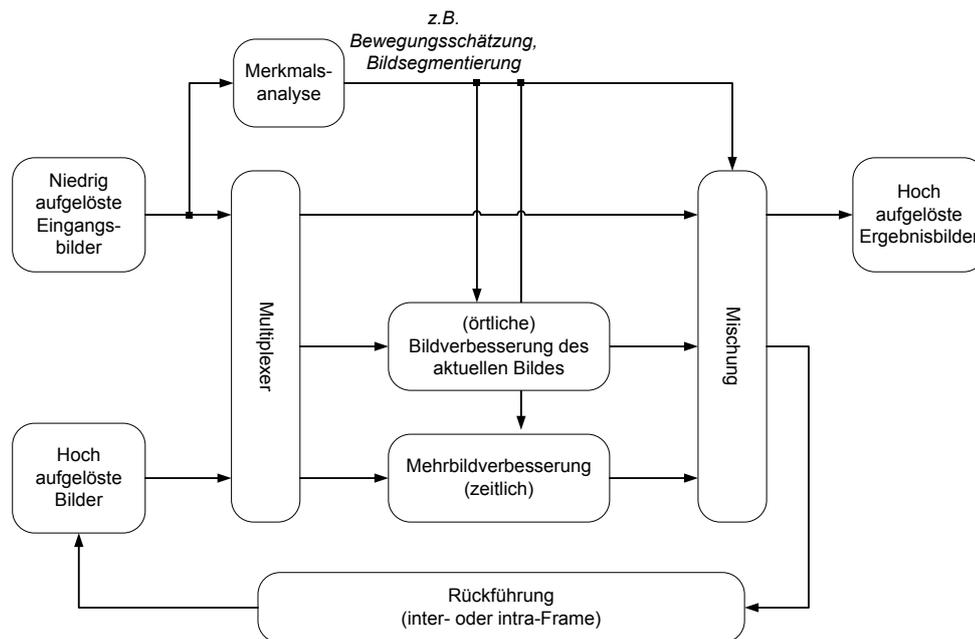


Abbildung 3.4: Generalisierte Beschreibung iterativer Super-Resolution Algorithmen

### 3.1.3 Grundaufbau iterativer Super-Resolution Algorithmen

Als Beispiel für das iterative Verarbeitungskonzept soll hier der Grundaufbau von Super-Resolution Algorithmen vorgestellt werden, die auf iterativen Optimierungsstrategien basieren. Abbildung 3.4 zeigt diese generalisierte Beschreibung. Charakteristisch ist eine Rückführungsschleife, die eine sukzessive Verbesserung der Ergebnisqualität unter Einbeziehung vorheriger Ergebnisse ermöglicht. Diese Rückführung kann dabei – wie im späteren Verlauf des Kapitels noch gezeigt werden wird – sowohl unabhängig pro Bild iterativ durchgeführt werden als auch zeitlich rekursiv. Die mathematische Beschreibung dieser Methoden lässt sich dabei durch auf Pixelbasis operierenden Operatoren beschreiben, die aus der Videosignalverarbeitung bereits bekannt sind. Im Allgemeinen führt dies zu einer rein örtlichen auf das hoch aufgelöste (Ergebnis-) Bild angewandten Bildverbesserung, die durch das Bild- und/oder Abbildungsmodell beschrieben ist. Bei Mehrbildverfahren findet zusätzlich noch eine örtlich-zeitliche Bildverbesserung durch die Berechnung eines Korrektursignals statt, welches auch die höherfrequenten Informationen enthält.

## 3.2 Super-Resolution Algorithmen

Aufbauend auf den generellen Konzepten werden nun konkrete Verfahren zur Auflösungserhöhung vorgestellt. Eine Übersicht über die wichtigsten Verfahren zeigt Abbildung 3.5.

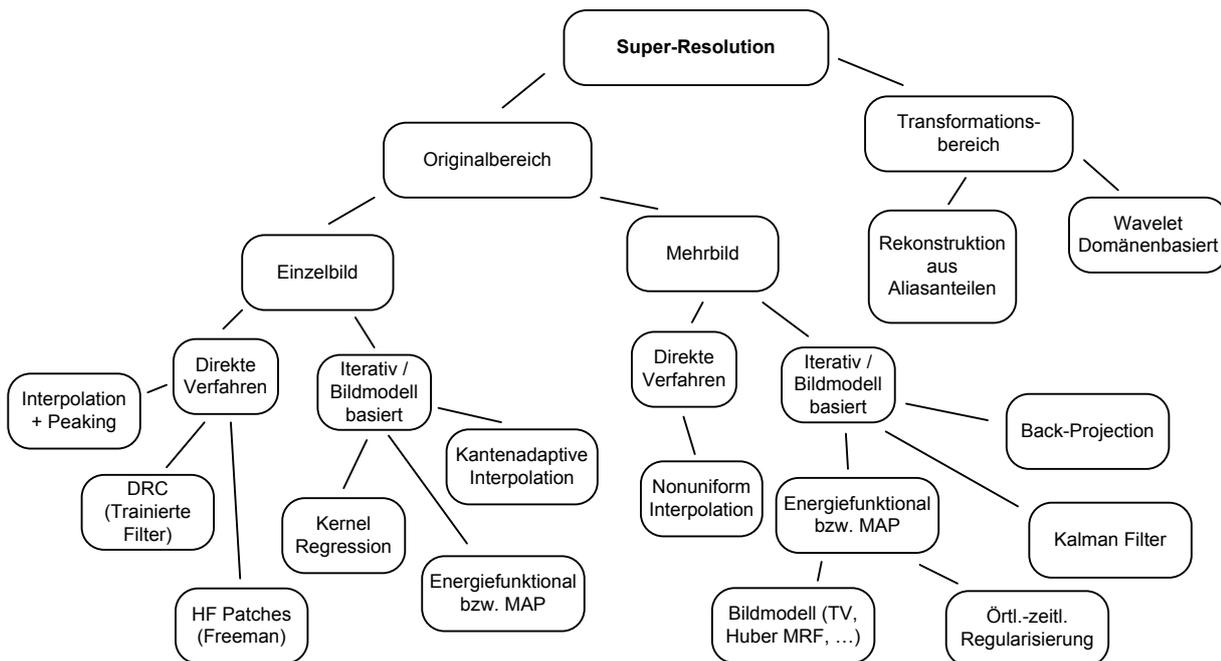


Abbildung 3.5: Übersicht über Verfahren zur Steigerung der Auflösung in jedem Einzelbild

Zwar ist neben der Unterscheidung von direkten und iterativen Verfahren auch eine Unterteilung in die Verarbeitung im Zeit- und Frequenzbereich möglich, allerdings wird im Frequenzbereich die Formulierung der Super-Resolution als Optimierungsproblem nur sehr selten angewandt [PPK03]. Deshalb können die meisten dieser Verfahren ebenfalls zu den direkten Verfahren gezählt werden. Abschnitt 3.2.4 behandelt diese Verfahren noch genauer. Weiterhin können Einzel- und Mehrbildverfahren unterschieden werden. Unabhängig von der Anzahl der Eingangsbilder kann eine Unterteilung in direkte und iterative Verfahren vorgenommen werden. Die folgenden Abschnitte werden sich nun zuerst den direkten Verfahren widmen. Bei den Verfahren, die nur ein einziges Eingangsbild zur Auflösungserhöhung benötigen, existiert die Interpolation mit anschließender Detailsignaladdition oder die Verwendung ganzer HF-Patches (Abschnitt 3.2.1) und eine bildinhaltsabhängige Interpolation z.B. durch DRC (Abschnitt 3.2.2). Der größte Vorteil dieser Verfahren ist ihre Speicher- und Recheneffizienz. So wird keine Bewegungsschätzung benötigt und es muss wenn überhaupt nur das aktuelle Eingangsbild zwischengespeichert werden. Dafür können diese Verfahren aber auch nur auf Informationen des aktuellen Bildes zurückgreifen, weshalb die erzielbare Auflösungserhöhung deutlich geringer ausfällt als bei Mehrbildverfahren. Verwendet man zur Detailsignaladdition aber synthetische, nicht im Bild enthaltene Informationen, so kann dieses Limit teilweise überwunden werden.

Bei den Mehrbildverfahren ist als direktes Verfahren die nonuniform Interpolation zu nen-

nen (Abschnitt 3.2.3). Diese beruht vor allem auf der Ausnutzung von Subpixelbewegungen, die ein einfaches Kopieren von Information aus benachbarten Bildern ermöglicht. Da dieser Kopierprozess aber nicht auf einem iterativen Verarbeitungskonzept basiert, werden Fehler nur verschleiert, die Korrektur, welche weiterhin zu einer erhöhten Auflösung auch im Fehlerfall führen würde, ist aber nicht vorgesehen. Dies kann nur durch die iterative Rekonstruktionsalgorithmen erreicht werden. Da der Aufbau dieser Verfahren für den Einzel- und Mehrbildfall sehr ähnlich ist, werden diese hier gemeinsam vorgestellt. Der Hauptunterschied liegt dabei in den Datentermen, der im Mehrbildfall aus einer Summe von Termen besteht, und für den Einzelbildfall meist nur einen Term umfasst. Die verwendeten Bildmodelle sind aber zumeist vollkommen identisch. Eine Unterscheidung ist dabei über den gewählten Ansatz möglich. So finden sich in der Literatur auf Energiefunktionalen basierende Algorithmen oder auf statistischen Annahmen beruhende Algorithmen (z.B. die Kernel Regression oder den MAP-Ansatz).

### 3.2.1 Auflösungserhöhung durch Detailsignaladdition

Eine sehr gängige Strategie, die Auflösung von Videosignalen und Bildern zu verbessern, ist die Addition (hochfrequenter) Detailsignale zum Eingangsmaterial. Dabei kann zwischen Verfahren, die die Detailsignale aus dem Bildinhalt selbst (z.B. durch Hochpassfilterung) generieren [KV04b] und Verfahren, die synthetische Detailsignale aus Datenbanken verwenden [Wu93], [Dol10], [RDS<sup>+</sup>08], unterschieden werden. Abbildung 3.6 verdeutlicht die Unterschiede zwischen den beiden Strategien im Spektrum. Während durch die Addition eines durch lineare Filterung erzeugten Hochpasssignals nur vorhandene Frequenzanteile angehoben werden können, kann eine nichtlineare oder synthetische Detailsignaladdition zu einer spektralen Erweiterung und damit zu einem echten Superresolutionseffekt führen.

#### Luminance Transient Improvement (LTI)

Die synthetische Detailsignaladdition wird in [SEF86] und [Wu93] entwickelt. Diese Detailsignaladdition beschränkt sich auf Kantenbereiche, da diese gut detektiert werden können und für Kanten zahlreiche Modelle existieren. Bei diesem Verfahren wird zuerst eine Kantendetektion durchgeführt, und dann in Abhängigkeit zum Kantenmittelpunkt phasenrichtig ein Korrektursignal hinzuaddiert. Dieses Korrektursignal wird unabhängig vom Bildsignal aus einem idealen Sprung erzeugt, wodurch eine Rauschverstärkung vermieden werden kann und die Detailgeneration nicht zur Laufzeit erfolgen muss, wodurch Rechenzeit gespart werden kann. Die Generierung dieses Detailsignals erfolgt über Degradationsmodelle, welche eine Dämpfung, hervorgerufen z.B. durch die Übertragung des Bildsignals über einen analogen Kanal, berücksichtigen. Die Vermeidung der Rauschverstärkung und die mögliche Kontrolle von Überschwüngen an Kanten sind Vorteile dieses Verfahrens. Dadurch entsteht ein sehr angenehmer Bildeindruck, da die Überschwünger an

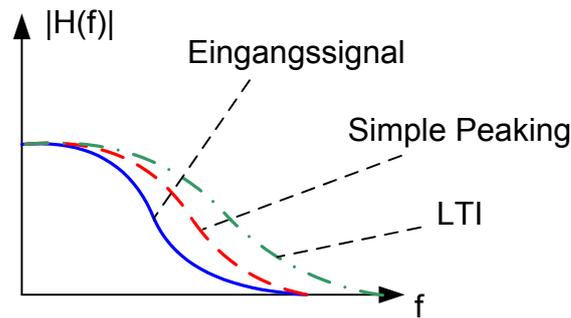


Abbildung 3.6: Beispielhafter Frequenzgang vor und nach der Bildschärfekorrektur

Kanten auf einen für das menschliche visuelle System optimalen Wert eingestellt werden [Wu93]. In [Dol10] und [RDS<sup>+</sup>08] wird gezeigt, wie die Detailsignaladdition dynamisch zur Bewegungsgeschwindigkeit der Pixel angepasst werden kann und das Verfahren somit zur Reduktion von Bewegungsunschärfe auf Hold-type Displays wie LCDs eingesetzt werden kann. Von Nachteil ist bei diesen Verfahren, dass es eine reine Kantenversteilerung durchführt, so dass Texturbereiche unscharf bzw. niedrig aufgelöst bleiben. Eine Lösung dieses Problem stellt das bildinhaltsabhängige Verarbeitungskonzept (MABLE, Abschnitt 3.1.2) dar. So wird zuerst das Bild in Kanten-, texturierte und homogene Bereiche unterteilt und dann das LTI nur auf Kantenbereiche angewendet. In den Texturbereichen kann dann ein einfacher Hochpass zur Detailerzeugung verwendet werden. In homogenen Flächen findet, um Rauschanhebung zu vermeiden, keine Bildschärfekorrektur statt. Diese Vorgehensweise ist in [RDS<sup>+</sup>08] detailliert beschrieben.

### Addition von Texturbausteinen

In [FJP02] wird ein Verfahren gezeigt, wie eine Detailsignaladdition auch für texturierte Bereiche verwendet werden kann. Das Grundkonzept ist dem des LTI sehr ähnlich. Zuerst müssen hochfrequente Detailsignale generiert werden, die dann in einer Datenbank abgespeichert werden können. Basierend auf einer Bildanalyse wird dann das für jeden Bereich am Besten passende Detailsignal ausgewählt und dieses schließlich zum Eingangsbild hinzuaddiert. In diesem Fall werden die Texturbausteine (die Detailsignale) aus hochfrequentem Testmaterial extrahiert. Die Bildanalyse / Klassifikation beruht auf der Annahme, dass zwischen den mittelfrequenten Anteilen, die im Eingangsmaterial vorhanden sind, und den hochfrequenten Anteilen der Texturpatches Ähnlichkeit besteht. Lenke [LE07] beschreibt eine alternative Möglichkeit zur Erzeugung hochfrequenter Texturinformation, welche auf der Resynthese basiert [Har02]. Bei diesem Verfahren werden Texturanteile räumlich an den Grenzen der Originalstruktur wiederholt, ohne dass sichtbare Übergänge an den Grenzen entstehen. Dadurch ist es möglich, texturierte Bereiche ohne eine echte

Interpolation zu vergrößern. In [LE07] werden dann so aus dem vergrößerten texturierten Bereich mittels Filterung die hochfrequenten Anteile extrahiert. Die Details können dann zum Interpolationsergebnis hinzugefügt werden. Dadurch entstehen neue hochfrequente Anteile bei einem gleichzeitigen Erhalt der Grundstruktur.

### 3.2.2 Bildinhaltsabhängige Interpolation

Ein guter Kompromiss zwischen einer hohen Interpolationsgüte und einem vergleichsweise geringen Rechenaufwand kann mit einer bildinhaltsabhängigen Interpolation erzielt werden. Trainierte Filter [Kon01], [Len09], [SHH08], Kernel Regression [TFM07], [Brü09] oder kantenadaptive Interpolationsverfahren [LO01] sind Beispiele für diese Verarbeitungsweise. Diese Verfahren werden nun im Einzelnen vorgestellt.

#### Trainierte Filter

Trainierte Filter wurden von Kondo [Kon01] eingeführt und später von Shao [SHH08] und Lenke [Len09] erweitert. Das Haupteinsatzgebiet dieser Filter ist die Interpolation, wobei sich dieses Filterverfahren von den Standardverfahren (siehe Abschnitt 2.1) dadurch unterscheidet, dass in diesem Fall alle Pixel im Ergebnisbild neu berechnet werden. Weitere Einsatzgebiete dieser Filter sind die Störungsreduktion oder Bildschärfekorrektur [SHH08]. Die Anwendung der trainierten Filter besteht aus zwei Schritten. Im ersten Schritt wird das Bildmaterial klassifiziert (z.B. durch Adaptive Dynamic Range Coding (ADRC), [Len09]), um für jede Situation die optimalen Filterkoeffizienten zu bestimmen. Danach erfolgt dann, basierend auf der Klassifikation, die Filterung mit den ausgewählten Koeffizienten. Die Koeffizienten werden dabei nicht zur Laufzeit selbst berechnet, sondern aus einer Look-up Tabelle geholt. Damit dies möglich ist, muss vor der erstmaligen Anwendung dieses Verfahrens ein Trainingsprozess zur Koeffizientenberechnung durchgeführt werden. Dieser Trainingsprozess und die dazu verwendeten Testsequenzen sind also völlig unabhängig von der späteren Anwendung des Algorithmus. Für den Trainingsprozess muss das verwendete Bildmaterial deshalb repräsentativ sein, um möglichst für jede Situation passende Filterkoeffizienten bereitzustellen. Für Fälle, die mit dem Testmaterial nicht abgedeckt sind, würden sonst später suboptimale Filterkoeffizienten ausgewählt.

Das Training entspricht einem Optimierungsprozess. Dazu wird Testmaterial benötigt, dass sowohl in niedrig als auch in hoch aufgelöster Form vorliegt. Dadurch ist dem Optimierungsprozess sowohl die Eingangssituation (Ist-Zustand) als auch die Ausgangssituation (Soll-Zustand) bekannt. Dann werden durch eine mathematische Approximation, z.B. der Methode des mittleren quadratischen Fehlers, die optimalen Filterkoeffizienten berechnet, indem das Ist-Material durch Variation der Filterkoeffizienten möglichst gut an den Soll-Zustand angenähert wird. Die besten Filterkoeffizienten werden dann abgespeichert. Die Speicherung der so ermittelten Koeffizienten in der Tafel erfolgt ebenfalls nach der

Klassifikation, um diese später richtig zuzuordnen zu können. Weil das Training und die spätere Anwendung voneinander getrennt sind, handelt es sich bei den Filterkoeffizienten aber immer nur um einen Kompromiss. Trotz dieses Kompromisses ist es aber möglich, Frequenzanteile oberhalb der ursprünglichen Nyquistfrequenz zu erzeugen, weshalb dieses Verfahren definitiv zu den Super-Resolution Algorithmen zu zählen ist. Die größten Qualitätsgewinne sind bei diesem Verfahren in Kantenbereichen zu beobachten, da hier die Klassifikation sehr gut funktioniert. Lenke hat dieses Verfahren durch die Einführung von Nebenbedingungen und dem Anlernen von Interpolationspolynomen anstelle von Filterkoeffizienten noch deutlich erweitert, wodurch eine bessere Interpolationsqualität erzielt werden kann und die Polynome nicht nur für einen festen Faktor, sondern für beliebige Interpolationsfaktoren eingesetzt werden können [Len09].

### Kernel Regression

Wie bereits erwähnt, stellt die einmalige Berechnung der Filterkoeffizienten für dediziertes Testmaterial einen Nachteil dar. Um diesen zu überwinden, wurden Verfahren entwickelt, die die Berechnung optimaler Koeffizienten zur Laufzeit durchführen [TFM07], [Brü09], [LO01]. Dies führt allerdings zu einem höheren Rechenbedarf. In [TFM07], [Brü09] finden sich Verfahren, die auf der Regressionsanalyse basieren. Dabei werden die Filterkoeffizienten mittels eines Optimierungsansatzes bestimmt. Die Optimierung wird dabei zwischen den Eingangsdaten und den zu bestimmenden Koeffizienten der Regressionskurve durchgeführt. Das Ergebnis der Optimierung führt zu einem polynomialen Verlauf, der einer Regressionskurve  $n$ -ter Ordnung entspricht. Die Ordnung ist im Optimierungsprozess frei wählbar, eine höhere Ordnung führt zwar zu einem besseren Ergebnis, allerdings auch zu einem höheren Aufwand. Die so ermittelte Regressionskurve dient dann als Ergebnis der Interpolation. Diese weist meist einen glatteren Verlauf mit höherer Auflösung als die Originalsequenz auf.

### Kantenadaptive Interpolation

Das NEDI Verfahren hat das Ziel, Kantenverläufe optimal zu interpolieren und ist vom Grundprinzip dem trainierten Filteransatz sehr ähnlich. Der Unterschied besteht darin, dass die Koeffizienten zur Laufzeit berechnet werden. Da zur Laufzeit aber das Material nicht in niedrig und hoch aufgelöster Form vorliegt, wird das Eingangsmaterial um den Interpolationsfaktor herunter getastet. Dieses dient dann als künstlich generiertes niedrig aufgelöstes Material im Optimierungsprozess und das ursprüngliche Material wird als hoch aufgelöstes Material herangezogen. Die auf diese Weise bestimmten Filterkoeffizienten werden dann für die eigentliche Interpolation verwendet. Dabei wird die Annahme getroffen, dass sich die Kantenverläufe auf allen Auflösungsstufen nicht ändern [LO01].

### 3.2.3 Nonuniform Interpolation

In [SVW02] wird ein Verfahren beschrieben, welches durch Kombination von Pixelwerten benachbarter Bilder unter Ausnutzung von (räumlichen) Alias und Subpixelbewegungen einen Auflösungsgewinn erzielt. Das Grundprinzip dieses Verfahren soll mittels Abbildung 3.7 verdeutlicht werden. Die Super-Resolution kann dabei durch die Schritte Bewegungsschätzung bzw. Bildregistrierung, Interpolation fehlender Pixel und Nachverarbeitung ausgedrückt werden. Für das Ergebnis der Super-Resolution ist eine Subpixelbewegung von entscheidender Bedeutung. Bei diesem Verfahren werden, wie es Abbildung 3.8 zeigt, wie bei der Standardinterpolation zuerst Nullen zwischen den vorhandenen Stützstellen des niedrig aufgelösten Rasters eingefügt. Die fehlenden Pixelstellen werden dann durch bewegungskompensierte Pixel aus benachbarten Eingangsbildern gefüllt. Stimmt diese Subpixelbewegung nicht genau mit einer (neuen) Pixelposition überein, so müssen die Pixel für das äquidistante Abtastraster berechnet werden. Dabei können sowohl Standardverfahren wie HRS verwendet werden als auch komplexere Techniken wie Kernel Regression [TFM07]. Eine erhöhte Auflösung kann nur dann erzielt werden, wenn diese Pixel neue Information tragen, was nur im Falle eines unterschiedlichen Störungsgrades oder Alias in der Eingangssequenz möglich ist. Von entscheidender Bedeutung für das Ergebnis ist eine korrekte Bewegungsschätzung. Fehlerhafte Bewegungsvektoren können zu deutlich sichtbaren Artefakten in Form von Kammstrukturen führen [HDRS09]. Aus diesem Grund werden für diese Algorithmen meist robuste globale Bewegungsschätzer eingesetzt [SVW02]. Weist die Eingangssequenz keine Subpixelbewegungen auf, so bleiben die zu interpolierenden Pixelpositionen unbesetzt und müssen durch andere Verfahren berechnet werden. Eine Nachverarbeitung, z.B. in Form einer Regularisierung [ROF92], kann diese Fehlstellen ebenfalls durch den Inpainting-Effekt des Bildmodells füllen und gleichzeitig Artefakte wie Rauschen beseitigen. Weiterhin ermöglicht dies eine Korrektur der Seiteneffekte durch falsche Bewegungsvektoren. Einen Sonderfall für diese Verfahren stellen Algorithmen dar, die ohne eine vorherige Interpolation eine reine zeitlich bewegungskompensierte Filterung durchführen. In [RDS09a], [Sch07],[Yan08] finden sich unterschiedliche Verfahren, welche für stark gestörte Sequenzen ohne eine Interpolation Detailgewinne bei gleichzeitiger Reduktion des Störungsgrades ermöglichen.

### 3.2.4 Verfahren im Transformationsbereich

Eine Berechnung des Super-Resolution Ergebnisses im Transformationsbereich hat den Vorteil, dass gezielt der Frequenzgang beeinflusst werden kann. In [PPK03] wird ein Verfahren beschrieben, welches im Frequenzbereich die Aliasanteile zu rekonstruieren versucht. Besonders attraktiv erscheint die Wavelettransformation für Super-Resolution Algorithmen, da diese als Bandaufspaltung aufgefasst werden kann [Mal99]. Somit können Frequenzbänder gezielt beeinflusst und die Verarbeitung auf die Bänder oberhalb der ur-

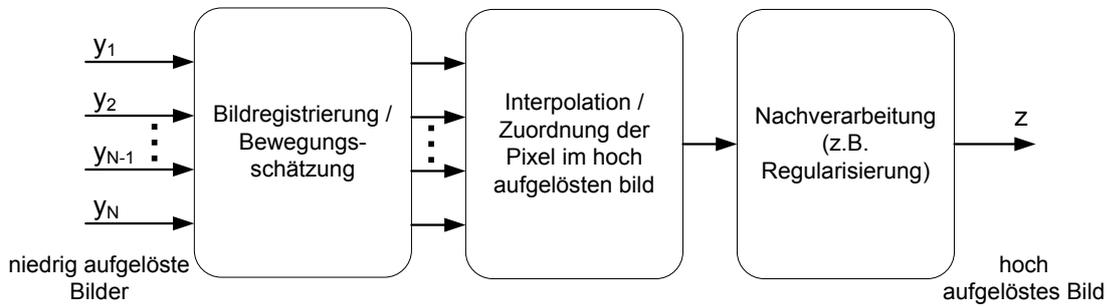


Abbildung 3.7: Nonuniform Interpolation Super-Resolution

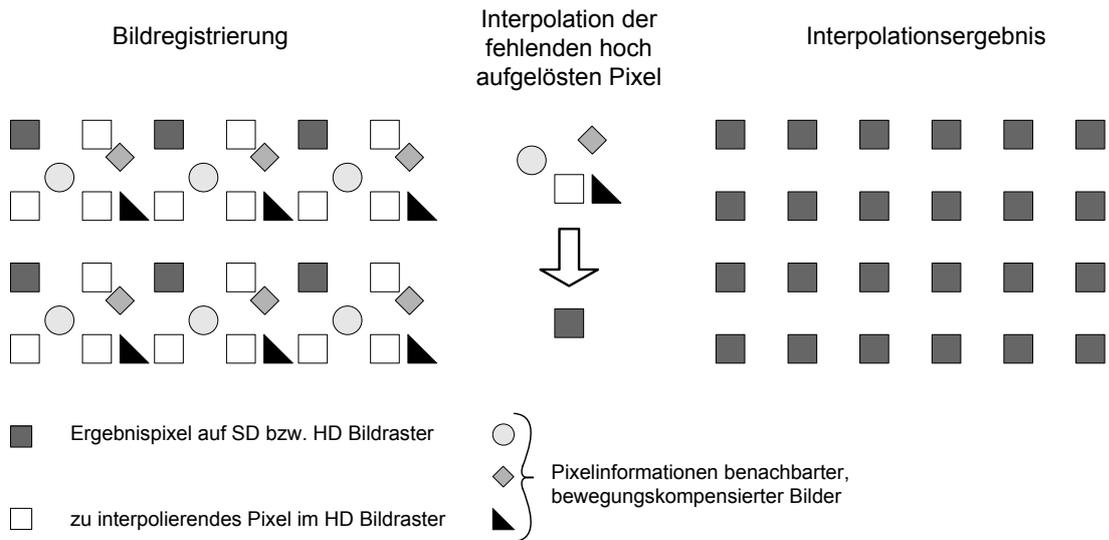


Abbildung 3.8: Registrierung und Interpolation für die nonuniform Interpolation Super-Resolution

sprünglichen Nyquistfrequenz konzentriert werden. Im Waveletbereich existieren sowohl einfache direkte Verfahren [LA08], die stark an die Nonuniform Interpolation im Ortsbereich angelehnt sind aber die Kombination der Pixelwerte in der Waveletdomäne erfolgt, als auch komplexe mathematische Konzepte basierend auf Energiefunktionalen [NM00].

### 3.2.5 Back-Projection Super-Resolution

Back-projection Super-Resolution erzeugt hochfrequente Komponenten durch einen iterativen Rekonstruktionsprozess. Wie später deutlich wird, entspricht dieses Verfahren der Lösung des Datenterms aus Abschnitt 2.3.2. Die Back-Projection Super-Resolution wurde von Irani und Peleg eingeführt [IP91]. Ein Blockschaltbild dieses Verfahrens ist in Abbildung 3.9 gezeigt. Für den iterativen Lösungsprozess wird eine initiale, hoch getastete Lösung  $\mathbf{y}_{int}$  benötigt, welche durch Standardinterpolationstechniken berechnet werden kann. Anschließend wird das jeweils aktuelle, hoch aufgelöste Bild der aktuellen Iteration einem Abbildungsprozess unterzogen (vergl. Abbildung 3.10), welcher die Verschiebung  $\mathbf{F}_n$ , Unschärfe  $\mathbf{B}$  und Unterabtastung  $\mathbf{D}$  beschreibt. Das dadurch entstehende 'modellerte', niedrig aufgelöste Bild wird dann mit dem aktuellen, tatsächlichen Eingangsbild verglichen. Die Differenz zwischen beiden Bildern liefert ein Korrektursignal, welches mit Hilfe eines 'Back-Projection' Operators  $\mathbf{H}_{BP}$  auf die finale Auflösung hochgerechnet wird. Dieser Back-Projection Operator kann das Korrektursignal zusätzlich noch einer Gewichtung unterziehen. Aus diesem Grund hat der Back-Projection Operator einen starken Einfluss auf das Konvergenzverhalten und das Aussehen der Lösung [PPK03]. In ihn kann sogar ein Bildmodell zur Beeinflussung des Aussehens der Lösung integriert werden. Nach der Addition des Korrektursignals zum Ergebnis der vorherigen Iteration steht das aktuelle Iterationsergebnis fest. Die Berechnung der Korrektursignale kann für eine beliebige Anzahl von Eingangsbildern erfolgen, wobei hierzu unterschiedliche Bewegungskompensationsmatrizen  $\mathbf{F}_n$  zur Verfügung stehen müssen, was den Einsatz einer Mehrreferenzbildbewegungsschätzung erforderlich macht. Der Iterationsprozess wird so lange durchgeführt, bis Konvergenz erreicht ist, was mathematisch durch  $|\mathbf{z}_q - \mathbf{z}_{q+1}| < \epsilon$  für ein hinreichend kleines  $\epsilon$  ausgedrückt werden kann. Für diesen Fall muss das Korrektursignal gegen Null gehen, so dass das modellierte niedrig aufgelöste Bild mit dem Eingangsbild identisch ist. Betrachtet man den zugehörigen mathematischen Ansatz für dieses Verfahren, so wird deutlich, dass mit diesem Verfahren keine eindeutige Lösung berechnet werden kann und das Ergebnis vom Anfangswert bzw. dem initialen Interpolationsergebnis abhängig ist. Der Ansatz (in Matrix-Vektor Schreibweise) ist eine Minimierung des mittleren quadratischen Fehlers zwischen den niedrig aufgelösten Bildern  $\mathbf{y}_n$  und dem den Beobachtungsmodell unterzogenen Ergebnisbild  $\mathbf{z}$ :

$$\min_{\mathbf{z}} \left\{ E = \sum_n (\mathbf{y}_n - \mathbf{DBF}_n \mathbf{z})^2 \right\} \quad (3.1)$$

Dass dieser Minimierungsprozess keine eindeutige Lösung liefert, kann anschaulich durch die Tiefpassfilterung des hoch aufgelösten Ergebnisses erklärt werden. Durch diese Tiefpassfilterung werden nur die niedrig- und mittelfrequenten Anteile miteinander verglichen bzw. im Minimierungsprozess berücksichtigt. Die hochfrequenten Anteile, z.B. verursacht durch ein hochfrequentes Rauschen, beeinflussen dabei die Lösung nicht und können ein beliebiges Aussehen haben. Wendet man für die Minimierung die Gradientenmethode an, so kann die Gleichung folgendermaßen umgeschrieben werden:

$$\begin{aligned} \mathbf{z}_{q+1} &= \mathbf{z}_q + \nabla E \\ \Rightarrow \mathbf{z}_{q+1} &= \mathbf{z}_q + \sum_n \mathbf{F}_n^T \mathbf{B}^T \mathbf{D}^T (\mathbf{y}_n - \mathbf{D} \mathbf{B} \mathbf{F}_n \mathbf{z}_q) \end{aligned} \quad (3.2)$$

$q$  entspricht dabei einer (örtlichen) Iterationsvariable. Für den Konvergenzfall sind die Lösungen  $\mathbf{z}$  für die Iterationsschritte  $q$  und  $q+1$  identisch, so dass die Ableitung des Energiefunktionals  $\nabla E$  aus Gleichung 3.1 gegen Null geht. Dies entspricht exakt der notwendigen Bedingung in einem Minimierungsprozess. Der Algorithmus wird für jedes hoch aufgelöste Ergebnisbild  $\mathbf{z}$  unabhängig gelöst. Die Berechnungsvorschrift aus Gleichung 3.2 kann so interpretiert werden, dass zur aktuellen Lösung ein Korrektursignal  $(\mathbf{y}_n - \mathbf{D} \mathbf{B} \mathbf{F}_n \mathbf{z}_q)$  hinzuaddiert wird, bis Konvergenz erreicht ist, wobei die Korrektur mit jedem Schritt kleiner wird. Das beruht auf der Konvexität von  $E$ . Die Transponierten der Matrizen können mit bildverarbeitenden Operatoren auch anschaulich beschrieben werden.  $\mathbf{D}^T$  entspricht einer Hochtastung durch Einfügen von Nullen und  $\mathbf{F}^T$  der inversen Bewegungskompensation, welche die Verschiebung rückgängig macht. Da sowohl eine Kompensation als auch inverse Kompensation nötig sind, kann dies im Falle komplexer Bewegungssituationen zu Mehrdeutigkeiten führen, wodurch in einem der beiden Vektorfehler Doppeldeutigkeiten oder Löcher, wie sie von der Vektorprojektion in der Bildratenkonversion bekannt sind. Ist der Unschärfe-Filter symmetrisch, was für die Beschreibung der Tiefpassfilterung vor einer Abtastung gültig ist, so ist die Transponierte der Matrix mit  $\mathbf{B}$  selbst identisch. In diesem Fall kann die Multiplikation mit der Matrix  $\mathbf{B}$  durch eine pixelweise arbeitende Faltung realisiert werden. Ist  $\mathbf{B}$  nicht symmetrisch, z.B. wenn lokal variierende Bewegungsunschärfe modelliert wird, so kann diese Vereinfachung nicht durchgeführt werden. Für den Fall einer symmetrischen Matrix  $\mathbf{B}$  beschreiben die Matrizen  $\mathbf{D}^T$  und  $\mathbf{B}^T$  einen Standardinterpolationsprozess (vergleiche Abschnitt 2.1). Abbildung 3.10 verdeutlicht dies auch noch einmal. Um das Ergebnis der Super-Resolution zu verbessern, kann für diesen Standardprozess ein höher qualitatives Interpolationsverfahren verwendet werden und die drei Matrizen  $\mathbf{D}^T$ ,  $\mathbf{B}^T$  und  $\mathbf{F}_n^T$  durch einen generalisierten Back-Projection Operator  $\mathbf{H}_{BP}$  ersetzt werden [PPK03]:

$$\mathbf{z}_{q+1} = \mathbf{z}_q + \sum_n \mathbf{H}_{BP} (\mathbf{y}_n - \mathbf{D} \mathbf{B} \mathbf{F}_n \mathbf{z}_q) \quad (3.3)$$

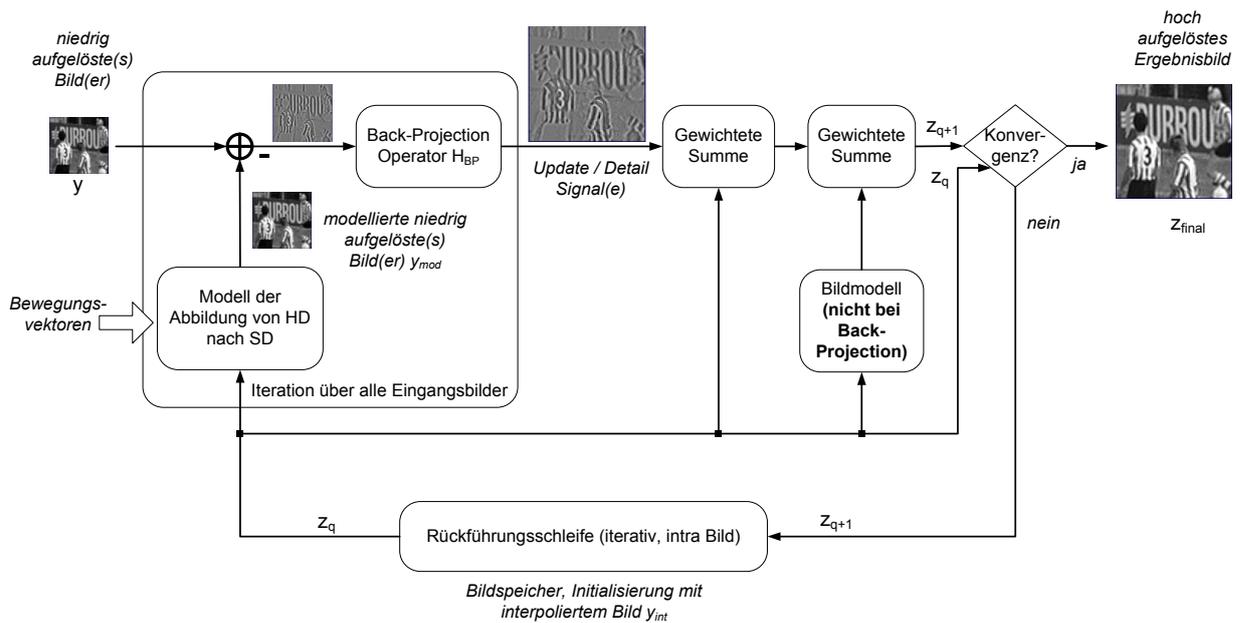


Abbildung 3.9: Blockdiagramm zur Realisierung der Back-Projektion Super-Resolution

Da die Back-Projektion die Lösung für den Datenterm bereitstellt, findet sich ihr Grundaufbau aus Abbildung 3.9 auch in anderen Algorithmen wieder. Dieser Algorithmus wird dann aber zumeist um ein Bildmodell erweitert, um eine mathematisch eindeutige Lösung mit sinnvoll geformten hohen Frequenzanteilen zu gewährleisten. Eine dieser Erweiterungen, die sich aus der Transformation eines statistischen Ansatzes in ein Energiefunktional ergibt, wird im nächsten Abschnitt vorgestellt.

### 3.2.6 Statistische Verfahren

Da bei der Super-Resolution für die Ergebnisbilder hoch frequente Informationen bestimmt werden müssen und für diese Bestimmung sowohl Störeinflüsse wie Rauschen als auch potenziell fehlerhafte Bewegungsvektoren zu berücksichtigen sind, eignen sich statistische Verfahren besonders gut zur Beschreibung und auch zur Lösung dieses Problems. Bei statistischen Super-Resolution Verfahren wird das Auftreten bzw. Aussehen der Eingangsbilder mit statistischen Zufallsprozessen beschrieben und die (zugehörigen) Ergebnisbilder ermittelt, welche mit größter Wahrscheinlichkeit der empfangenen Eingangsbildfolge zugrunde liegen. Im Folgenden werden zwei wichtige Verfahren vorgestellt, die MAP Super-Resolution und ein auf der Kalman-Filterung basierendes Verfahren. Bei der MAP Super-Resolution werden nur der Ansatz und ein erster möglicher Lösungsweg aufgezeigt. Eine exakte Beschreibung inklusive der Erweiterungen und eigene Lösungsansätze zu diesem Verfahren werden in Kapitel 5 vorgestellt.

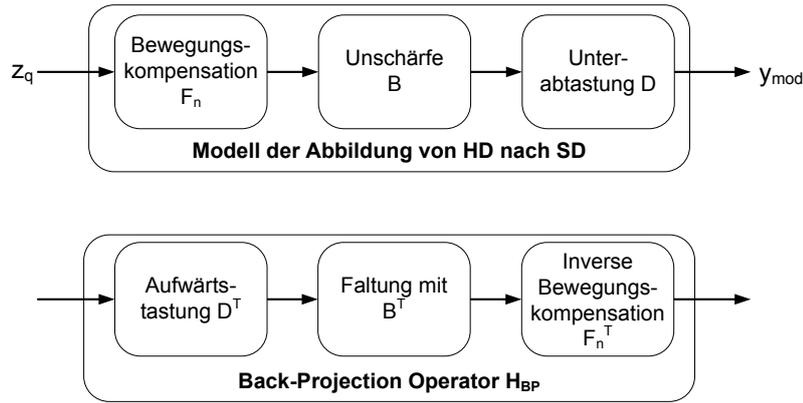


Abbildung 3.10: Realisierung der Modellierungen des Abbildungsprozesses (oben) und des Back-Projection Operators (unten)

### Maximum-A-Posteriori Probability Super-Resolution (MAP-SR)

Die Maximum-A-Posteriori Super-Resolution hat das Ziel, die Ergebnissequenz zu berechnen, welche mit der höchsten Wahrscheinlichkeit zur empfangenen Eingangssequenz gehört [BS99]. Zu diesem Zweck wird das Bayes' Theorem verwendet [BSMM01]:

$$P(\mathbf{z}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{z}) \cdot P(\mathbf{z})}{P(\mathbf{y})} \quad (3.4)$$

Zur Lösung dieses Problems muss die bedingte Wahrscheinlichkeit  $P(\mathbf{z}|\mathbf{y})$  maximiert werden. Um die Rechenvorschriften zu vereinfachen, kann auf beiden Seiten der Gleichung 3.4 logarithmiert werden, ohne das Maximum zu beeinflussen [BS99]. Da die empfangene Sequenz  $\mathbf{y}$  nicht beeinflusst werden kann, hat die Wahrscheinlichkeitsverteilung  $P(\mathbf{y})$  ebenfalls keinen Einfluss auf das Maximum. Nach Anwendung dieser Schritte auf obige Gleichung erhält man folgendes Optimierungsproblem:

$$\max \{ \ln [P(\mathbf{z}|\mathbf{y})] = \ln [P(\mathbf{y}|\mathbf{z})] + \ln [P(\mathbf{z})] \} \quad (3.5)$$

Für die Anwendung müssen nun geeignete Wahrscheinlichkeitsverteilungen  $P(\cdot)$  definiert werden. Unter der Annahme einer gauß'schen Wahrscheinlichkeitsverteilung lässt sich die Gleichung 3.5 in folgende Vorschrift transformieren.

$$\arg \min_{\mathbf{z}} \left\{ \sum_n (\mathbf{y}_n - \mathbf{DBF}_n \mathbf{z})^2 + \frac{1}{2\beta} \sum_{i=0}^3 \rho(d_i(\mathbf{z})) \right\} \quad (3.6)$$

Im Vergleich zur Back-Projection Super-Resolution besteht dieser Ansatz aus zwei Termen. Der erste Term ist vollkommen identisch mit dem Ansatz der Back-Projection Super-Resolution. Er garantiert die Ähnlichkeit zwischen Eingangs- und Ausgangssequenz und

stellt eine Bedingung an die nieder- und mittelfrequenten Anteile der Lösung. Im Gegensatz dazu wird hier noch ein Bildmodell formuliert, welches die Hochpassanteile beeinflusst. Mathematisch wird dadurch eine eindeutige Lösung unabhängig von der gewählten Initiallösung erzwungen. Als Bildmodell findet in diesem Fall das Huber-Markov Zufallsfeld (siehe Abschnitt 2.3.2) Anwendung. Dadurch bleiben sowohl Kanten als auch Texturen in der Ergebnissequenz erhalten, was beispielsweise durch das Total Variation Bildmodell nicht möglich ist. In Kapitel 5 wird gezeigt, wie dieser Ansatz effektiv gelöst werden kann und durch die Verwendung einer Merkmalsanalyse noch deutliche Verbesserungen erzielt werden können. Andere Lösungsstrategien sowie ein allgemeinerer Ansatz werden in [BS99] und [Cai06] diskutiert.

### Kalman Filter basierte Super-Resolution

Ein sehr leistungsfähiges Verfahren, das robust gegenüber Störungen oder Abweichungen der mathematischen Annahmen ist, stellt die Kalman Filter basierte Super-Resolution nach Farsiu, Elad und Milanfar dar [FREM04], [FEM06]. Da dieses Verfahren gute Möglichkeiten zur effizienten Realisierung von Super-Resolution Algorithmen aufzeigt, soll es hier kurz erläutert werden. Sein Haupteinsatzgebiet sind allerdings sehr niedrig aufgelöste Videosequenzen mit starken Kompressionsartefakten und einer globalen Bewegung. Solche Videos sind durch Handykameras oder auf Web-Portalen sehr zahlreich vorhanden. Durch die Spezialisierung der Bewegung auf reine Kameraschwenks ergibt sich die Möglichkeit, sehr robuste Bewegungsvektoren zu schätzen und den Kalman Filterprozess erheblich zu vereinfachen [FREM04], [FEM06]. Die zugrunde liegende Kalman Filter Theorie [Kal60] setzt sich aus zwei Schritten zusammen und ermöglicht die Bestimmung robuster Ergebnisse basierend auf unzureichenden oder ungenauen Schätzungen. Der Kalman Filter selbst bzw. seine Gewichtungen entstehen ebenfalls durch einen zur Laufzeit durchzuführenden Optimierungsprozess. Als Ausgangspunkt für den Kalman Filter dient der Zustandsraum. Die Gleichungen im Zustandsraum können aus dem Beobachtungsmodell direkt abgeleitet werden:

$$\mathbf{y}_n = \mathbf{D}_n \mathbf{z}_n + \mathbf{e}_{mess,n} \quad (3.7)$$

und

$$\mathbf{z}_n = \mathbf{F}_n \mathbf{z}_{n-1} + \mathbf{e}_{sys,n} \quad (3.8)$$

Die erste Gleichung wird als Messgleichung bezeichnet. Sie stellt den Zusammenhang zwischen der Messung (dem niedrig aufgelösten Eingangsbild  $\mathbf{y}_n$ ) und dem aktuellen, hoch aufgelösten Bild  $\mathbf{z}_n$  her. Der Verzicht auf eine Matrix zur Beschreibung der Unschärfe lässt sich durch eine Komplexitätsminderung erklären. Diese wird später aber noch genauer behandelt. Da bei diesem Vergleich Störungen mit berücksichtigt werden müssen, wird ein stochastischer Rauschprozess verwendet, der als Messrauschen  $\mathbf{e}_{mess}$  bezeichnet

wird. Die zweite Gleichung ist die Systemgleichung, welche die Dynamik des internen Systems beschreibt. In diesem Fall wird lediglich das Übergangsverhalten zwischen aktuellen und vorherigem hoch aufgelösten Ergebnisbild über eine Bewegungskompensation  $\mathbf{F}$  hergestellt. Neue Bildinformationen, hervorgerufen beispielsweise durch Auf- und Verdeckungen oder aber auch fehlerhafte Bewegungsvektoren, werden durch das Systemrauschen  $\mathbf{e}_{sys}$  beschrieben.

Nach Aufstellung der Zustandsgleichungen erfolgt die Lösung in zwei Schritten. Zuerst wird der Zustand  $\mathbf{z}_n$  aus dem vorherigen Ergebnis prädiziert. Danach wird durch einen Korrekturschritt der Zusammenhang zwischen der Prädiktion und der aktuellen Messung hergestellt. Nach Anwendung mehrerer Zwischenschritte, die z.B. in [FREM04] und [FEM06] ausführlich beschrieben sind, ergibt sich die finale Systemgleichung:

$$\hat{\mathbf{z}}_n = \mathbf{F}_n \hat{\mathbf{z}}_{n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{D}_n \mathbf{F}_n \hat{\mathbf{z}}_{n-1}) \quad (3.9)$$

Die Kalmanmatrix  $\mathbf{K}_n$  ist von den beiden Kovarianzmatrizen der Rauschprozesse  $\mathbf{e}_{sys,n}$  und  $\mathbf{e}_{mess,n}$  abhängig. Details zur Berechnung werden in [FREM04], [FEM06] und [RNS11] beschrieben. Da diese Matrix ebenfalls von der Bewegungsvektormatrix  $\mathbf{F}$  und dem Abtastungsprozess  $\mathbf{D}$  abhängig ist, kann eine pixelweise und damit recheneffiziente Lösung dieser Matrix nur dann gewährleistet werden, wenn es sich um eine Diagonalmatrix handelt, was nur bei translatorischen Verschiebungen erfüllt ist [FREM04]. Die Systemgleichung kann auch in Form eines Blockschaltbildes interpretiert werden, welches in Abbildung 3.11 zu sehen ist. Vergleicht man diese Abbildung mit dem Blockschaltbild der Back-Projection (Abbildung 3.9), so ist ersichtlich, dass der Grundaufbau sehr ähnlich ist. Unterschiede finden sich in der Ausführungsform der Rückführungsschleife sowie der Gewichtung der einzelnen Signalanteile. Ein weiterer Unterschied ist der Verzicht auf eine Unschärfematrix  $\mathbf{B}$ , die bereits in der Messgleichung fehlt. Diese Vereinfachung führt zu einer deutlich simpleren Berechnungsvorschrift [FEM06]. Deshalb wird aber nur eine unscharfe hoch aufgelöste Sequenz durch die Kalman Filter Super-Resolution berechnet. Für scharfe SR-Bilder ist eine Nachverarbeitung mit einem (iterativen) Deblurring-Algorithmus nötig [FREM04]. Verbesserungsmöglichkeiten, um konzeptionsbedingte Schwächen, z.B. die Beschränkung auf Sequenzen mit reinen Kameraschwenks zu überwinden, werden in [RNS11] und [Naf09] beschrieben. Die dort beschriebenen Verbesserungen basieren auf dem Konzept des Multiple Algorithm Blending und einer adaptiven Schätzung der Rauschvarianzen.

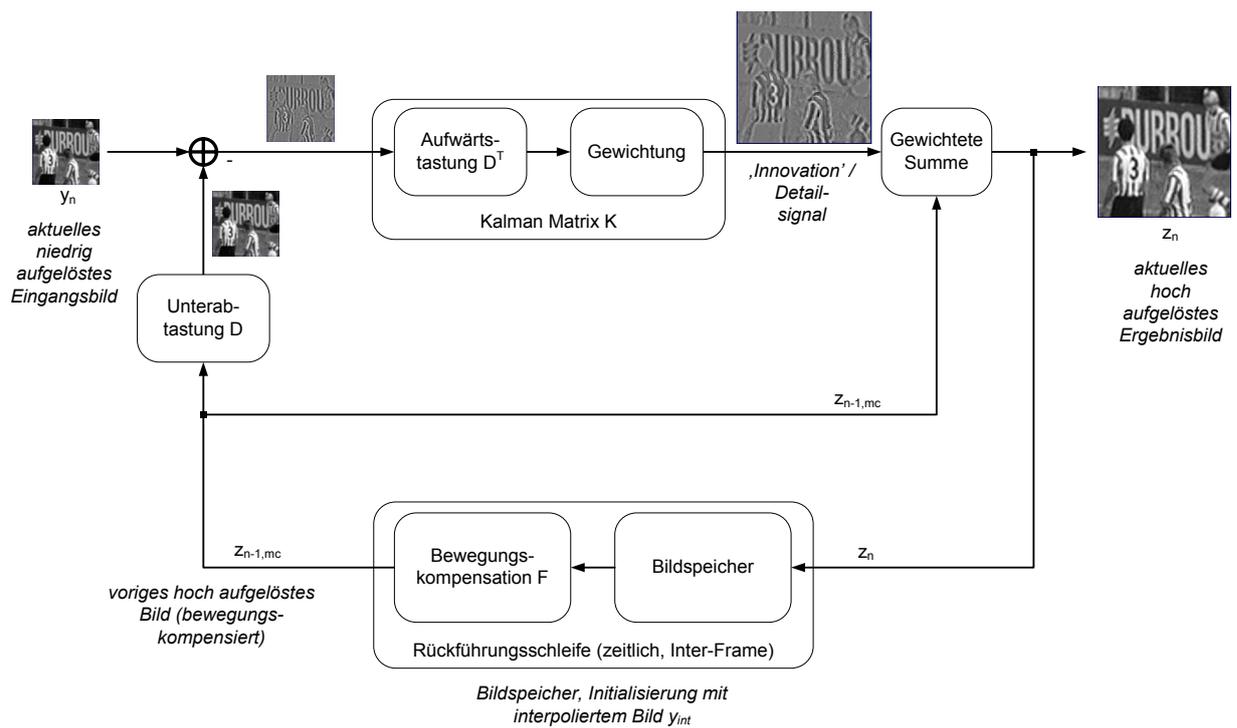


Abbildung 3.11: Kalman Filter basierte Super-Resolution

# Kapitel 4

## Entwurfskonzept für iterative Rekonstruktionsalgorithmen für die Videosignalverarbeitung

In diesem Kapitel wird verdeutlicht, wie iterative Rekonstruktionsalgorithmen so entworfen werden können, dass sie sich für die Videosignalverarbeitung eignen. Dazu wird der Entwurfsprozess so gestaltet, dass er einerseits zu einem speicher- und recheneffizienten Algorithmus führt und andererseits die Bildqualität der Ergebnisse auch zufriedenstellend ist, wenn die internen Modellannahmen der iterativen Rekonstruktionsalgorithmen nicht erfüllt sind. Zu Anfang des Kapitels wird deshalb zuerst das allgemeine Entwurfskonzept präsentiert. Danach wird dann als Entwurfsbeispiel dieses Konzept auf Algorithmen zur örtlichen und örtlich-zeitlichen Regularisierung angewendet. Die Anwendung auf mehrere Super-Resolution Algorithmen wird dann in Kapitel 5 gezeigt.

### 4.1 Formulierung des Entwurfskonzepts für iterative Rekonstruktionsalgorithmen

Iterative Rekonstruktionsalgorithmen basieren auf verschiedenen mathematischen Ansätzen, die bereits in Kapitel 2 beschrieben wurden. Diese Algorithmen und ihre Formulierung bietet entscheidende Vorteile, die hier noch einmal kurz wiederholt werden sollen. Die Algorithmen selbst treffen Modellannahmen und passen die Eingangssequenz gemäß dieser Modellannahmen an. Sind diese Modellannahmen für die Eingangssequenz erfüllt, so kann eine ausgezeichnete Bildqualität mit diesen Algorithmen erzielt werden. Da diese Modellanpassung bzw. Optimierung zur Laufzeit geschieht, wird die Ergebnissequenz mehrfach verarbeitet, bis alle Annahmen erfüllt sind und damit das Ergebnis gemäß dieser Kriterien optimal ist. Diese mehrfache Verarbeitung ermöglicht eine Kontrolle der Ergebnisqualität, welche eine Einstellung der Verarbeitung auf ein gewünschtes Zielergebnis bei gleichzeitiger Unterdrückung der Artefakte ermöglicht.

Da die meisten iterativen Rekonstruktionsalgorithmen aber nicht für die Videosignalverarbeitung entworfen wurden, ist eine Modifikation bzw. ein Neuentwurf dieser Algorithmen nötig, um sie an die dort vorherrschenden Rahmenbedingungen anzupassen. So muss in den meisten Fällen ein Kompromiss zwischen Bildqualität und Rechenaufwand eingegangen werden und der Algorithmus sehr unterschiedliches Material ohne störende Seiteneffekte verarbeiten können. Um diese Ziele zu erreichen, müssen konzeptbedingte Nachteile der iterativen Rekonstruktionsalgorithmen überwunden werden.

Um diesen Anforderungen zu genügen wird in dieser Dissertation ein vierstufiges Entwurfskonzept vorgeschlagen, welches in Abbildung 4.1 gezeigt wird. Dieses Entwurfskonzept ist weiterhin in zwei generelle Ebenen unterteilt, die (mathematische) Algorithmenableitung und die Verbesserung und Analyse des hergeleiteten Verfahrens. Während bei der Algorithmenableitung das Hauptaugenmerk auf der Reduktion des Speicher- und Rechenaufwandes liegt und diese dem allgemeinen Vorgehen bei der Entwicklung von iterativen Rekonstruktionsalgorithmen entspricht (siehe z.B. [Kau00], [AK06], [FREM04]), konzentrieren sich die Verbesserungen in der zweiten Ebene vor allem auf die Erhöhung der erzielbaren Bildqualität. Im Folgenden werden nun die vier Teilschritte genauer erläutert:

Der erste Schritt bei der Entwicklung iterativer Rekonstruktionsalgorithmen ist die Auswahl eines geeigneten mathematischen Ansatzes. Dazu kann entweder auf bereits bekannte Ansätze zurückgegriffen werden oder ein neuer Ansatz formuliert werden. Bei der ersten Strategie werden die Modifikationen also erst während des Entwurfsprozess vorgenommen, bei der zweiten kann über einen verbesserten Ansatz direkt auf eine möglichst recheneffiziente oder hoch qualitative Lösung geachtet werden. Im nächsten Abschnitt wird ersteres Vorgehen für die örtliche Regularisierung gezeigt und letztere Strategie bei der örtlich-zeitlichen Regularisierung angewendet werden.

Der zweite Schritt verfolgt das Ziel, eine optimale Beschreibung des Algorithmus in Form eines Blockdiagramms oder einer Filtervorschrift zu finden. Um das zu erreichen, können die in Abschnitt 2.3 beschriebenen mathematischen Werkzeuge verwendet werden. Durch die Wahl der Werkzeuge bzw. Lösungsstrategien kann so z.B. direkt Einfluss auf die Konvergenzgeschwindigkeit sowie den Speicheraufwand genommen werden. Die so ermittelte Beschreibungsform stellt die direkte Umsetzung des mathematischen Ansatzes in einen Algorithmus dar. Unterschiede bzgl. der Bildqualität werden deshalb durch diesen Schritt meist nicht erreicht.

Diese Verbesserungen können nun im dritten Schritt vorgenommen werden. Das ist nun möglich, da die Beschreibung mit einer Filtervorschrift bzw. einem Blockdiagramm die Analyse der Eigenschaften des Algorithmus in Hinblick auf seine Stärken und Schwächen erleichtert. Gerade ein Blockdiagramm ist hierzu besonders geeignet, weil die einzelnen Blöcke separat voneinander untersucht und optimiert werden können. Diese Analyse kann dann zu einem Austausch einzelner Komponenten durch leistungsfähigere oder das Einfügen zusätzlicher Komponenten in das Blockdiagramm zur gezielteren Beeinflussung des Gesamtsystems führen. Dieser Schritt fällt in jedem Algorithmus unterschiedlich aus und

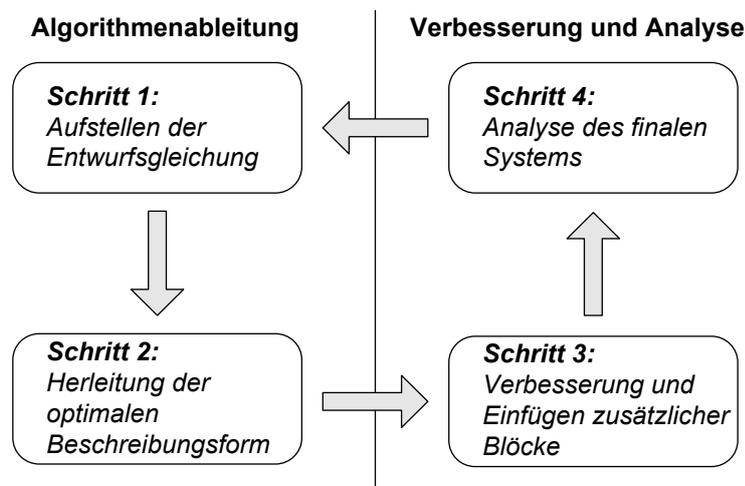


Abbildung 4.1: Entwurfskonzept für iterative Rekonstruktionsalgorithmen

ist durch Heuristiken bzw. Erfahrungswerte gesteuert. Dadurch können gezielt konzeptionsbedingte Schwächen beseitigt werden. Eine Verbesserung ist dann möglich, wenn die neuen Komponenten das gleiche Ziel haben wie die ursprünglichen, dieses aber besser umsetzen oder weniger Rechenaufwand dafür benötigen. In diesem Schritt können nun auch Komponenten eingesetzt werden, die sich mathematisch gar nicht oder nur sehr schwierig beschreiben lassen. Da sich durch die Modifikationen ein anderes Gesamtsystem ergibt, muss nun das Systemverhalten geprüft werden.

Das Systemverhalten wie z.B. die Stabilität wird im vierten Schritt überprüft. Dies kann entweder durch umfangreiche Experimente praktisch oder theoretisch, z.B. durch Verwendung der  $z$ -Transformation, geschehen. In einigen Fällen ist es sogar möglich, das finale System mit einer Gleichung zu beschreiben, aus der sich nachträglich ein neuer mathematischer Ansatz formulieren lässt. Dieser kann für Vergleiche mit anderen Algorithmen eingesetzt werden, oder der Entwurfsprozess noch einmal wiederholt werden.

## 4.2 Anwendung des Entwurfskonzepts auf Regularisierungsalgorithmen

In diesem Abschnitt wird exemplarisch die Anwendung des Entwurfskonzepts auf zwei Algorithmen zur Regularisierung erläutert. Während der Beschreibung wird gezeigt, welche Vorteile dieses Konzept bietet bzw. welche Möglichkeiten zur Verbesserung dieser Algorithmen bestehen. Die Regularisierung eignet sich zu dieser Darstellung besonders, weil der Ansatz als auch das finale Verfahren über eine vergleichsweise geringe Komplexität verfügen. Bei der örtlichen Regularisierung können durch diese Entwurfsmethodik Schwä-

chen beseitigt werden, die aus den Modellannahmen resultieren. Bei der örtlich-zeitlichen Regularisierung verteilen sich die Verbesserungen auf die Entwurfsschritte 1 und 3. Diese werden durch einen neuen Ansatz mit einer zusätzlichen zeitlichen Nebenbedingung zur Unterdrückung von Flackern sowie die Verwendung weiterer Komponenten erreicht.

## 4.2.1 Örtliche Regularisierung

### Aufstellen der Entwurfsgleichung

Die örtliche Regularisierung wurde ursprünglich von Rudin et al. [ROF92] für die Rauschreduktion entwickelt. Gemäß Schritt 1 des Entwicklungskonzepts wird als Ansatz ein (kontinuierliches) Energiefunktional formuliert. Dieses beruht auf der Überlegung, dass Rauschen kleinen (lokalen) Variationen im Bild entspricht. Eine Minimierung dieser Variationen führt also zu einer Rauschreduktion. Das Energiefunktional, welches aus diesen Überlegungen in [ROF92] hergeleitet wurde, lautet folgendermaßen:

$$\min_{\mathbf{z}} \left( \int_{\Omega} (\mathbf{z} - \mathbf{y})^2 d\Omega + \int_{\Omega} |\nabla \mathbf{z}| d\Omega \right) \quad (4.1)$$

Es besteht aus zwei Termen, wobei der Datenterm die Ähnlichkeit zwischen Eingangssequenz  $\mathbf{y}$  und Ausgangssequenz  $\mathbf{z}$  garantiert. Diese beiden Sequenzen werden bei diesem Ansatz als kontinuierliche Funktionen aufgefasst. Im Gegensatz zur Super-Resolution muss hier keine weitere Transformation der Ausgangssequenz durchgeführt werden, da sowohl die Auflösung als auch die Bewegungsphasen identisch sind. Das Bildmodell entspricht der Total Variation Norm, welche bereits in Kapitel 2 beschrieben wurde. Die weiter oben beschriebenen Variationen werden dabei mittels einer Filterung des (Iterations-) Ergebnisses mit dem (örtlichen) Gradienten und anschließender Betragsbildung extrahiert.

### Herleitung der optimalen Beschreibungsform

Da dieses Energiefunktional zwar anschaulich interpretiert werden kann, allerdings eine Lösungsstrategie sehr aufwändig ist, wird mittels der Variationsrechnung (siehe Abschnitt 2.3) das Energiefunktional in eine partielle Differentialgleichung überführt:

$$\mathbf{z} - \mathbf{y} - \lambda \operatorname{div} \left( \frac{1}{\sqrt{\mathbf{z}_x^2 + \mathbf{z}_y^2 + \epsilon}} \nabla \mathbf{z} \right) \quad (4.2)$$

Um diese kontinuierliche Vorschrift numerisch lösen zu können, muss eine Diskretisierung vorgenommen werden. Das gewählte Differenzschema hat dabei einen erheblichen Einfluss auf das Aussehen der Lösung, da dadurch die Stärke der Glättung beeinflusst werden

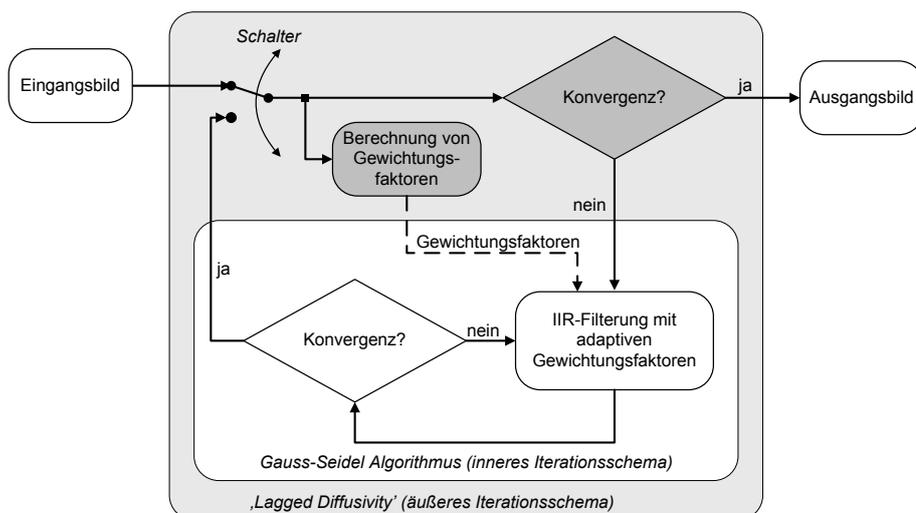


Abbildung 4.2: Schematischer Ablauf des Regularisierungsalgorithmus

kann, was sich durch die unterschiedlichen Übertragungsfunktionen der zugrunde liegenden Filter erklären lässt. Die Diskretisierung des Divergenztermes kann ebenfalls als eine Filturvorschrift beschrieben werden.

$$\begin{aligned} \operatorname{div}(\mathbf{b} \cdot \nabla \mathbf{z})_{k,l} \approx & 0,25 \cdot (\mathbf{b}_{k+1,l} \mathbf{z}_{k+1,l} + \mathbf{b}_{k-1,l} \mathbf{z}_{k-1,l}) + \\ & 0,25 \cdot (\mathbf{b}_{k,l+1} \mathbf{z}_{k,l+1} + 0,25 \cdot \mathbf{b}_{k,l-1} \mathbf{z}_{k,l-1}) - \\ & 0,25 \cdot (\mathbf{b}_{k+1,l} + \mathbf{b}_{k-1,l} + \mathbf{b}_{k,l+1} + \mathbf{b}_{k,l-1}) \mathbf{z}_{k,l} \end{aligned} \quad (4.3)$$

Die Gewichte  $\mathbf{b}$  entsprechen der Inversen des Betrages des örtlichen Gradienten an den jeweils mit den Subindizes angegebenen Stellen:

$$\mathbf{b}_{k,l} = \frac{1}{\sqrt{(\mathbf{z}_x)_{k,l}^2 + (\mathbf{z}_y)_{k,l}^2 + \epsilon}} \quad (4.4)$$

$\epsilon$  ist ein kleiner Wert, um eine Division durch Null zu vermeiden. Setzt man nun alle obigen Diskretisierungsvorschriften in Gleichung 4.2 ein, so entsteht eine nichtlineare Gleichung, da die Ableitungen von  $\mathbf{z}$  sowohl quadriert als auch einer Wurzeloperation unterzogen werden. Diese nichtlineare Gleichung wird nun in zwei Unterschritte unterteilt, die jeweils iterativ gelöst werden können. In einer äußeren Iteration werden die Gewichte  $\mathbf{b}$  gemäß Gleichung 4.4 bestimmt, wobei jeweils die Gradienten auf dem aktuellen Ergebnisbild berechnet werden. Anschließend werden die so berechneten Gewichte dazu verwendet, das Gleichungssystem zu linearisieren. Dieses kann dann durch den Gauß-Seidel Algorithmus iterativ gelöst werden. Dieses Vorgehen wird als Lagged-Diffusivity Fixpunktiteration bezeichnet [VO96]. Durch Anwendung dieser beiden Iterationsschemata ergibt sich dann

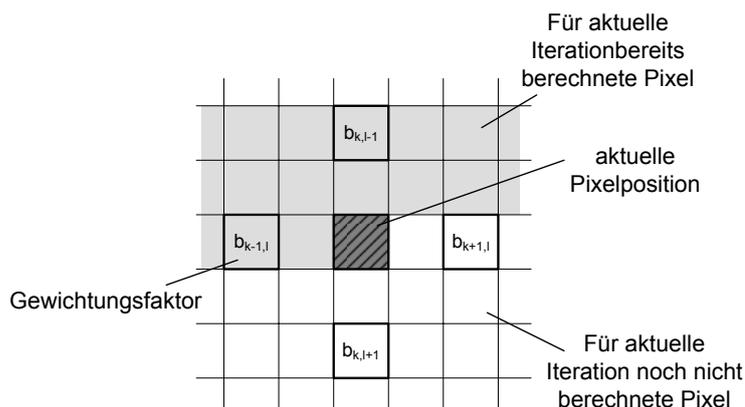


Abbildung 4.3: Position der verwendeten Filtergewichte für die Berechnung des aktuellen Ausgangspixels. Die Lage ergibt sich direkt aus der mathematischen Berechnungsvorschrift

die folgende Filtervorschrift mit der äußeren Iterationsvariable  $p$  zur Neuberechnung der Filtergewichte und der inneren Iterationsvariable  $q$  für den Gauß-Seidel Algorithmus:

$$(\mathbf{z}_{k,l})_{p,q+1} = c \cdot \mathbf{y}_{k,l} + c \cdot \lambda \sum_{i,j \in D_1} h_{k,l} \cdot (\mathbf{b}_{k+i,l+j})_p \cdot (\mathbf{z}_{k+i,l+j})_{p,q+1} + \quad (4.5)$$

$$c \cdot \lambda \sum_{i,j \in D_2} h_{i,j} \cdot (\mathbf{b}_{k+i,l+j})_p \cdot (\mathbf{z}_{k+i,l+j})_{p,q} \quad (4.6)$$

Die Regularisierung führt damit zu einer adaptiven Filterung, die abhängig zum inversen lokalen (örtlichen) Gradienten ist und die bis zur Konvergenz iterativ ausgeführt wird.  $c$  ist ein Normierungsfaktor der gewährleistet, dass die Summe aller Filtergewichte 1 beträgt:

$$c = \left( 1 + \lambda \sum_{i,j \in D_1 \cup D_2} h_{i,j} \cdot \mathbf{b}_{k+i,l+j} \right)^{-1} \quad (4.7)$$

Die Region  $D_1$  bezeichnet dabei Pixelpositionen von  $\mathbf{z}$ , die für die aktuelle innere Iteration  $q$  bereits berechnet wurden. Bei einer typischen Bearbeitungsreihenfolge von links oben nach rechts unten gehören also alle Pixel links und / oder oberhalb der aktuell zu berechnenden Pixelposition zu dieser Region. Die Region  $D_2$  beinhaltet alle noch nicht für die aktuelle Iteration verarbeiteten Pixel, die zur Lösung des aktuellen Pixels beitragen. Beide Regionen zusammen werden durch die Filtermaske  $h_{k,l}$  begrenzt. Die Pixelposition und Gewichte dieser (festen) Maske ergeben sich direkt aus der mathematischen Vorschrift und sie werden in Abbildung 4.3 gezeigt.

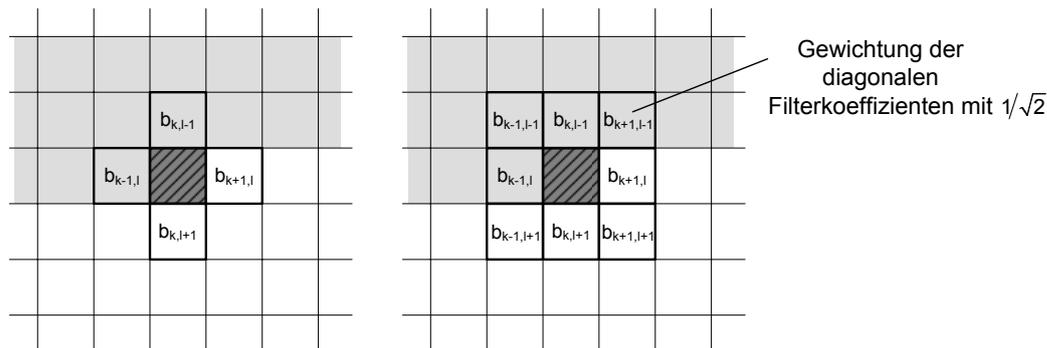


Abbildung 4.4: Position der Filtergewichte für eine verbesserte Filterung innerhalb der Regularisierung

### Verbesserung des Algorithmus und Einfügen zusätzlicher Blöcke

Die Regularisierung besitzt systembedingt zwei wichtige Eigenschaften, die sich aus den Filtergewichten, die proportional zur Inversen des Gradientenbetrags sind, ableiten lassen. Hohe Gradienten führen zu kleinen Filtergewichten, so dass in Kantenbereichen nur eine sehr schwache Filterung durchgeführt wird. Deshalb weist die TV-Regularisierung eine kantenerhaltende Eigenschaft auf, die auch als Median-Eigenschaft bezeichnet wird. Diese Eigenschaft ist in Kantenbereichen vorteilhaft. Allerdings werden dadurch auch Blocking-Artefakte gar nicht oder nur unzureichend reduziert. Des Weiteren führen kleinere Variationen im Bild, die auf Rauschen zurückzuführen sind, zu großen Filtergewichten, so dass dieses durch den Regularisierungsalgorithmus verringert wird. Da aber auch schwache Bild-details, z.B. Texturen mit niedrigem Kontrast zu kleinen Gradienten führen, werden diese ebenfalls entfernt. Dadurch entsteht der bereits in Abbildung 2.5 gezeigte Comic ähnliche Eindruck, der bei natürlichen Bildinhalten zu sehr künstlichen Ergebnissen führt.

Strategien, um diese Schwächen zu beseitigen ohne die Stärken aufgeben zu müssen, wurden bereits in [DRS09a] und [DRS09b] vorgestellt. Um die Regularisierung auch zur Reduktion von Codierartefakten einsetzen zu können, wird ein initiales bildinhaltsadaptives Deblocking benötigt, welches Texturen und Kanten schützt, Blockingartefakte aber entfernt. Hierfür kann beispielsweise das in [Pia08] entwickelte Verfahren verwendet werden. Durch die Glättung der Blockartefakte kann die Regularisierung die übrig gebliebenen Codierstörungen wie Ringing, aber auch verbleibendes Blocking mit kleinen und mittleren Gradienten effektiv entfernen. Da der Regularisierungsfiler iterativ arbeitet, wird mit jedem Iterationsschritt der Filterprozess verstärkt und es können auch weiter entfernte Pixel zum aktuellen Pixel beitragen, wodurch sich in homogenen Flächen ein sehr weicher und natürlicher Bildübergang frei von Blocking ergibt.

Texturen können durch die Anwendung von Multiple Algorithm Blending geschützt werden, indem die Regularisierung nur in den Bereichen eingesetzt wird, wo ihr internes

Bildmodell, welches homogene Flächen und steile Kanten beschreiben kann, erfüllt ist. In texturierten Bereichen wird hingegen nur das initiale Deblocking verwendet. Dies hat im Vergleich zu einer lokalen Beeinflussung der Regularisierungsrate [VO96], [ZS06] oder der Multiplikation der Gewichtungsfaktoren mit einer aus globalen Kriterien abgeleiteten Funktion den Vorteil [ZC10], dass der iterative Prozess nicht gestört wird und mathematisch schwierig beschreibbares, abruptes Umschalten der Verarbeitung von einem Pixel zum anderen möglich wird.

Neben der Verwendung dieser Zusatzkomponenten kann auch die Position der Filtergewichte verändert werden. Experimente haben gezeigt, dass durch die Verwendung einer modifizierten Maske für einige Bildsituationen bessere Ergebnisse erzielt werden können. Abbildung 4.4 zeigt zwei mögliche Masken, die ebenfalls untersucht wurden. Gerade für texturierte Bereiche oder bei diagonalen Kanten sind auch diagonal besetzte Pixel in einer Maske oder Pixel aus der direkten Nachbarschaft besser geeignet, da dadurch stärker korrelierter Bildinhalt zur aktuellen Lösung beiträgt. Durch den Abbruch des Algorithmus vor der Konvergenz, welcher mit der Wahl einer festen Anzahl an äußeren und inneren Iterationen möglich wird, kann weiterhin der Rechenaufwand auf Kosten einer schlechteren Bildqualität reduziert werden. Dies ist ein Beispiel, wie nachträglich auch noch der Aufwand vermindert werden kann.

### Analyse des finalen Systems

Da bei der örtlichen Regularisierung mit Ausnahme der Position der Filtergewichte und der Begrenzung der Iterationsanzahl keine weiteren Modifikation innerhalb des iterativen Rekonstruktionsalgorithmus vorgenommen wurden, soll für Möglichkeiten der Stabilitätsanalyse auf die Literatur verwiesen werden [AK06], [WS04]. Weitere Details, wie eine Stabilitätsprüfung vorgenommen werden kann, werden in Kapitel 5 für die Super-Resolution Algorithmen beschrieben. Bei der Bildqualität zeigt sich, dass durch die Kombination der Regularisierung mit einer Vorverarbeitung und der Anwendung dieser in einem MABLE-Konzept die Schwächen der Regularisierung beseitigt werden können, ohne große Modifikation im Algorithmus selbst vornehmen zu müssen.

### 4.2.2 Örtlich-zeitliche Regularisierung

Die im vorigen Abschnitt vorgestellte verbesserte Regularisierung ermöglicht eine sehr gute örtliche Verarbeitung. Zeitliche Störungen innerhalb einer Videosequenz wie beispielsweise Flimmern können durch diese aber nur sehr unzureichend und auf Kosten eines hohen Detailverlusts reduziert werden. Aus diesem Grund wurden im Rahmen dieser Arbeit ebenfalls die Möglichkeiten einer örtlich-zeitlichen Regularisierung untersucht. Dabei wurde ein neues Konzept erarbeitet, da bestehende Verfahren nicht für den Einsatz in der Videosignalverarbeitung geeignet sind. In [AK06] wurde ein erster Ansatz vorgeschlagen,

der allerdings zur Berechnung der optimalen Lösung die gesamte Videosequenz benötigt, was in einem sehr großen Speicher- und Rechenaufwand resultiert und für die Echtzeitverarbeitung völlig ungeeignet ist, weil sehr lange Latenzzeiten entstehen würden. Andere Ansätze sind aus der Medizintechnik [MTM<sup>+</sup>04] bekannt, die allerdings nur für das dort übliche Bildmaterial geeignet sind und mit komplexen Bewegungen in Videosequenzen nicht zurecht kommen. Lediglich das in [YFL<sup>+</sup>03] vorgestellte Verfahren wurde für die Videosignalverarbeitung entwickelt, allerdings arbeitet es in zeitlicher Richtung nur rein rekursiv und verwendet keine extra berechneten zeitlichen Filterkoeffizienten, sondern nur solche, die auf örtlichen Differenzbildungen beruhen. Ziel dieser Methode ist eine zeitliche Glättung, die im Falle von fehlerhaften Bewegungsvektoren zu einer erheblichen Unschärfe führen kann. Die nun vorgestellte Methode hat zwei entscheidende Vorteile. Durch die Berechnung echter zeitlicher Filtergewichte können bei diesem Algorithmus Vektoren eines prädiktiven Blockmatching Verfahrens verwendet werden. Weiterhin ist die Bildanzahl beliebig einstellbar, so dass eine Mehrreferenzbildschätzung eingesetzt werden kann, die zu einer Detailwiderherstellung und starken Artefaktreduktion führt.

### Aufstellen der Entwurfsgleichung

Ansatzpunkt für die neue örtlich-zeitliche Regularisierung ist eine Erweiterung der örtlichen Regularisierung, wobei die Formulierung direkt durch ein diskretes Energiefunktional  $E_n$  erfolgt:

$$E_n = \sum_{k,l} (\mathbf{z}_{k,l,n} - \mathbf{y}_{k,l,n+m})^2 + \lambda_{spat} \sum_{k,l} S_{spat}(\mathbf{z}_{k,l,n}) + \lambda_{temp} \sum_{k,l} S_{temp}(\mathbf{z}_{k,l,n+m}) \quad (4.8)$$

Dieses Funktional besteht aus drei Termen. Der erste Term entspricht dem Datenterm und garantiert die Ähnlichkeit zwischen Eingangs- und Ausgangssequenz. Bei dem zweiten Term handelt es sich um die örtliche Nebenbedingung, die – wie später gezeigt wird – exakt dem Bildmodell der Total Variation in örtlicher Richtung entspricht. Der dritte Term ist eine zeitliche Nebenbedingung, die im Rahmen dieser Arbeit entwickelt wurde. Diskret formuliert lautet die örtliche Nebenbedingung folgendermaßen:

$$S_{spat} = \frac{1}{N_{spat}} \sum_{i,j} (h_{spat})_{i,j} \cdot (\mathbf{b}_s)_{k+i,l+j} \cdot (\mathbf{z}_{k+i,l+j,n} - \mathbf{z}_{k,l,n})^2 \quad (4.9)$$

Mit  $h_{spat}$  wird wie bei der örtlichen Regularisierung die feste Filtermaske bezeichnet. Die Koeffizienten  $\mathbf{b}_s$  sind die adaptiven Koeffizienten, die gemäß Gleichung 4.4 berechnet werden müssen. Betrachtet man die quadratische Differenz zwischen aktuellem Pixel und seinen direkten Nachbarpixeln, so kann diese Nebenbedingung als örtliches Aktivitätsmaß interpretiert werden. Analog wird ein zeitliches Aktivitätsmaß formuliert, welches aus einer

Differenz zwischen aktuellem Pixel und zeitlich bewegungskompensierten Pixeln in verschiedenen Referenzbildern besteht.  $h_{temp}$  bezeichnet dabei eine feste Maske mit fixen Filterkoeffizienten, deren Gewichtung invers zum Abstand vom aktuellem Bild abnimmt. Die adaptiven zeitlichen Filterkoeffizienten  $\mathbf{z}_t$  müssen später noch geeignet berechnet werden.

$$S_{temp} = \frac{1}{N_{temp}} \sum_m (h_{temp})_m \cdot (\mathbf{b}_t)_{k,l,n+m} \cdot (\mathbf{z}_{k+v_{x,m},l+v_{y,m},n+m} - \mathbf{z}_{k,l,n})^2 \quad (4.10)$$

### Herleitung der optimalen Beschreibungsform

Nachdem nun das vollständige diskrete Energiefunktional formuliert ist, wird nun eine lokale Optimierung für jeden Bildpunkt vorgenommen. Damit ist es möglich, den Einfluss jedes Pixels auf das Ergebnis zu kontrollieren. Im Vergleich zur globalen Optimierung, wie sie bei der örtlichen Regularisierung durchgeführt wurde, entfällt die Summe über das gesamte Bild bei der Minimierung, wodurch zwar kein globales Optimum bestimmt werden kann und die Beziehung zwischen den Nachbarpixeln abnimmt, dafür das Funktional aber trotz der (unbekannten) Bewegungskompensationen in der zeitlichen Nebenbedingung für jedes Pixel einfach lösbar ist und sich so im Ergebnis ein adaptiver Filter ergibt.

Um das Minimum zu bestimmen, ist die pixelbasierte Vorschrift nach jedem Bildpunkt  $\mathbf{z}_{k,l,n}$  abzuleiten, was einer Reduktion an örtlicher und zeitlicher Aktivität innerhalb der Bildsequenz entspricht. Bei der Ableitung wird die Annahme getroffen, dass die örtlichen und zeitlichen Filterkoeffizienten nicht von  $\mathbf{z}_{k,l,n}$  abhängig sind, so dass sie in der Ableitung als Konstanten behandelt werden können. Diese Idee, Teile der mathematischen Formulierung während eines Lösungsprozesses konstant zu halten, wird ebenfalls bei der 'Lagged-Diffusivity' Strategie angewendet. Daraus ergibt sich dann folgende Ableitungsvorschrift:

$$\arg \min_{\mathbf{z}_{k,l,n}} (E_n) = \frac{\partial D}{\partial \mathbf{z}_{k,l,n}} + \lambda_{spat} \cdot \frac{\partial S_{spat}}{\partial \mathbf{z}_{k,l,n}} + \lambda_{temp} \cdot \frac{\partial S_{temp}}{\partial \mathbf{z}_{k,l,n}} \quad (4.11)$$

$D$  entspricht dem Datenterm aus Gleichung 4.8 ohne die äußere Summe. Da die Ableitung dieses Termes sowie ein Umsortieren der Gleichung sehr einfach möglich ist, soll hier darauf verzichtet werden und statt dessen nur die generelle Vorgehensweise geschildert werden. Nach Bildung der Ableitung in Gleichung 4.11 muss eine geeignete Lösungsstrategie angewendet werden, da alle Pixel der aktuellen Lösung  $\mathbf{z}$  unbekannt sind. Genau wie bei der rein örtlichen Regularisierung wird hier auf eine IIR-Filterstruktur gesetzt. Dazu wird die Gleichung nach dem aktuellen Pixel  $\mathbf{z}_{k,l,n}$  aufgelöst. Diese daraus resultierende Gleichung kann dann genauso wie die örtliche Regularisierung gemäß Abbildung 4.2 gelöst werden, wobei nun örtlich-zeitliche Gewichte zu berechnen sind und die Filterung eine örtlich-zeitliche ist. Im aktuellen Bild sind wie bei der örtlicher Regularisierung alle Pixel links und / oder über dem aktuellen Pixel bereits für die aktuelle Iteration berechnet, während alle anderen Pixel noch zu ermitteln sind. Die Pixel aus den vorherigen Bildern sind aber

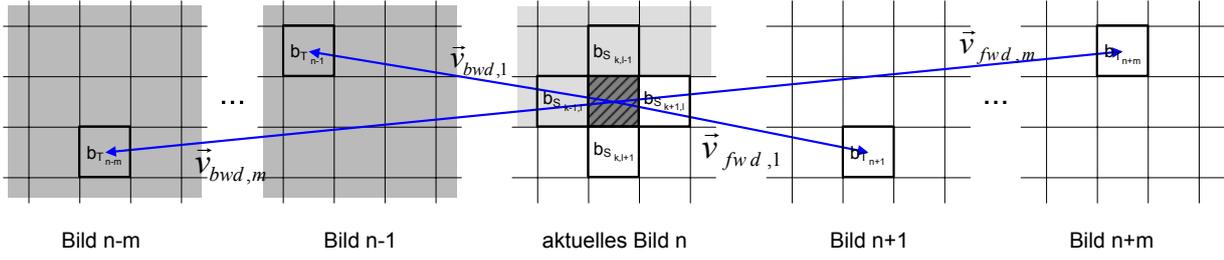


Abbildung 4.5: Lage der Filterkoeffizienten und -Gewichte bei der örtlich-zeitlichen Regularisierung

in jeder Iteration bereits vollständig berechnet und die zukünftigen Bilder bleiben während der Filterung des aktuellen Bildes komplett unverarbeitet. Abbildung 4.5 zeigt dieses Vorgehen inklusive aller Pixel, die zur Filterung beitragen. Der unterschiedliche Grad der Verarbeitung wird durch die Grautöne verdeutlicht, wobei die Pixel umso dunkler sind, je stärker die Verarbeitung bereits war (das aktuelle Pixel bildet hiervon eine Ausnahme). Durch die zeitliche Rekursion kann trotz einer begrenzten Bildanzahl eine sehr starke Filterwirkung erzielt werden. Im Ergebnis ergibt sich so folgende Gleichung zur Realisierung der örtlich-zeitlichen Regularisierung:

$$\begin{aligned} \mathbf{z}_{k,l,n} = & c \cdot \mathbf{y}_{k,l,n} + c \cdot \frac{\lambda_{spat}}{N_{spat}} \sum_{i,j} (h_{spat})_{i,j} \cdot (\mathbf{b}_s)_{k+i,l+j} \cdot \mathbf{z}_{k+i,l+j,n} + \\ & + c \cdot \frac{\lambda_{temp}}{N_{temp}} \sum_m (h_{temp})_m \cdot (\mathbf{b}_t)_{k,l,n+m} \cdot \mathbf{z}_{k+v_x,m,l+v_y,m,n+m} \end{aligned} \quad (4.12)$$

In der Filtergleichung findet eine Normierung mit  $c$  statt, damit die Summe aller Filterkoeffizienten gleich 1 ist:

$$c = \left( 1 + \frac{\lambda_{spat}}{N_{spat}} \sum_{i,j} (h_{spat})_{i,j} \cdot (\mathbf{b}_s)_{k+i,l+j} + \frac{\lambda_{temp}}{N_{temp}} \sum_m (h_{temp})_m \cdot (\mathbf{b}_t)_{k,l,n+m} \right)^{-1} \quad (4.13)$$

Um eine Abwärtskompatibilität zur rein örtlichen Regularisierung zu gewährleisten, werden die örtlichen Filtergewichte exakt wie für den örtlichen Fall berechnet. Bei den zeitlichen Filtergewichten wird eine ähnliche Strategie angewendet, wobei die Filtergewichte proportional zu einer inversen zeitlichen Differenz sind.

$$(\mathbf{b}_t)_{k,l,n+m} = \frac{1}{\sqrt{\epsilon^2 + diff\_t_{k,l,n+m}}} \quad (4.14)$$

Überträgt man die örtlichen Gewichte direkt auf die zeitlichen Gewichte, so entstünde eine bewegungskompensierte Differenz zwischen zwei benachbarten Referenzbildern. Dieses Verfahren hat zwar eine große Ähnlichkeit zur TV-Regularisierung, allerdings können

hierbei auch sehr kleine Differenzen und damit eine sehr starke Filterung entstehen, wenn beide Bewegungsvektoren vollkommen falsch sind, sie aber auf Bereiche mit ähnlichen Grauwerten zeigen.

$$diff\_t_{k,l,n+m} = \left| \mathbf{z}_{k+v_x,m,l+v_y,m,n+m} - \mathbf{z}_{k+v_x,m+1,l+v_y,m+1,n+m+1} \right| \quad (4.15)$$

### Verbesserung des Algorithmus und Einfügen zusätzlicher Blöcke

Das Einfügen zusätzlicher Blöcke betrifft wie bei der örtlichen Regularisierung nicht den Algorithmus selbst, sondern die Kombination des Verfahrens mit einer Vorverarbeitung und ggfs. dem MABLE-Konzept. Eine Veränderung innerhalb des Algorithmus bezieht sich auf die Berechnung der zeitlichen Filterkoeffizienten. Im Gegensatz zu den TV-Gewichten kann eine dem bilateralen Filter ähnliche Strategie angewendet werden [TM98], da diese eine höhere Robustheit gegen Bewegungsvektorfehler ermöglicht. Dabei wird die Differenz zwischen aktuellem Bild und bewegungskompensiertem Referenzbild bestimmt.

$$diff\_t_{k,l,n+m} = \sum_{i,j} w_{i,j} \cdot \left| \mathbf{z}_{k+i,l+j,n} - \mathbf{z}_{k+i+v_x,m,l+j+v_y,m,n+m} \right| \quad (4.16)$$

[TM98] Weiterhin kann diese Differenz über ein Fenster  $w_{i,j}$  gebildet werden, wodurch robuste und weniger stark variierende zeitliche Filtergewichte entstehen. Weitere Details und zu diesem örtlichen-zeitlichen Regularisierungsansatz sind in [RDS<sup>+</sup>11] beschrieben.

# Kapitel 5

## Entwurf von Super-Resolution Algorithmen für die Videosignalverarbeitung

In diesem Kapitel werden mittels des Entwurfskonzept mehrere Super-Resolution Algorithmen entwickelt, die für die Videosignalverarbeitung gut geeignet sind. Zuerst werden dazu Möglichkeiten aufgezeigt, wie der Back-Projektion Algorithmus aus Kapitel 3 weiter verbessert werden kann. Das Ergebnis dieser Verbesserungen ist ein vielseitig parametrisierbarer Algorithmus basierend auf der Back-Projection SR (VESBA - versatile super-resolution based on a back-projection algorithm), der die Zu- und Abschaltung der Einzelkomponenten erlaubt und somit für die Entwicklung weiterer verbesserter Algorithmen sowie für Untersuchungen besonders gut geeignet ist. Als eine weitere Verbesserung in diesem Algorithmus wird eine Mehrreferenzbildbewegungsschätzung beschrieben, die einen erheblichen Einfluss auf das Super-Resolution Ergebnis hat.

Daran anschließend wird basierend auf dem Maximum-A-Posteriori Ansatz nach Stevenson ein robustes und recheneffizientes Verfahren hergeleitet. Durch Verbesserungen und basierend auf den Erkenntnissen der vielseitig parametrisierbaren Back-Projection wird daraus ein doppelt rekursiver Mehrreferenzbild Super-Resolution Algorithmus (MUSIDORA - multi-reference frame super-resolution based on a double recursive algorithm) entwickelt. Als eine speicher- und recheneffiziente Alternative zu den auf mehreren Eingangsbildern basierenden Algorithmen wird schließlich ein rein rekursiver Algorithmus (ESTERA - efficient super-resolution based on a temporal recursive algorithm) hergeleitet. Die Komponenten aller Algorithmen können weiterhin durch eine Merkmalsanalyse gesteuert werden, so dass die Detailwiederherstellung und Bildschärfe gezielter beeinflusst werden können und die Algorithmen auch dann noch gute Ergebnisse liefern, wenn die (mathematischen) Modellannahmen nicht erfüllt sind. Weiterhin kann dadurch die Robustheit gegenüber Bewegungsvektorfehlern oder anderen Störungen erhöht werden. Das Kapitel endet mit Konvergenzbetrachtungen zu allen vorgestellten Algorithmen.

## 5.1 Vielseitig parametrisierbare Back-Projection Super-Resolution (VESBA)

Um die Einflüsse von Modifikationen beim mathematischen Ansatz zu untersuchen und die generelle Leistungsfähigkeit von Mehrreferenzbildverfahren im Vergleich zu einem Verfahren mit herkömmlicher Bewegungsschätzung und nur einem Bildspeicher bewerten zu können, wurde zusätzlich ein voll parametrisierbarer Back-Projection Super-Resolution Algorithmus implementiert. In Abschnitt 3.2.5 sind bereits die ersten beiden Schritte des Entwurfskonzeptes durchgeführt worden. Das Ergebnis dieser Ableitung ist Formel 3.2, welche in Abbildung 3.9 in ein Blockschaltbild umgesetzt wurde. Die Berechnung des Detailsignals der aktuellen Iteration erfordert die Anwendung eines Abbildungsmodell auf das Ergebnis der letzten Iteration, um es mit jedem Eingangsbild vergleichen zu können. Dazu ist gemäß Gleichung 3.2 eine Dezimation, eine Unschärfefilterung und eine Bewegungskompensation nötig. Nach dem Vergleich wird das Korrektursignal mittels dem Back-Projection Operator auf die Phasenlage und Größe der Ergebnissequenz umgerechnet. Dazu ist im Rückprojektionsschritt neben einer Interpolation und einer weiteren Unschärfefilterung eine inverse Bewegungskompensation nötig (vergleiche auch Abbildung 5.1). Nach diesem Vorgehen wird nun für jedes Eingangsbild ein Detailsignal errechnet, welche dann alle aufaddiert werden. Dieser Aufbau ermöglicht es nun, beim Abbildungsmodell, beim Back-Projection Operator und / oder bei der Addition aller Detailsignale Veränderungen vorzunehmen. Diese Veränderungen entsprechen Schritt drei im Entwurfskonzept und sollen nun vorgestellt werden. In diesem Kapitel werden dabei immer nur die Realisierungsaspekte aller Möglichkeiten diskutiert, die Auswertung erfolgt dann in Kapitel 6.

### 5.1.1 Modifikationen bei der Interpolation und Dezimation

Bei der Interpolation und Dezimation sind zwei unterschiedliche Strategien möglich. Auf der einen Seite kann auf die Interpolation und Dezimation innerhalb der Iterationsschleife verzichtet werden und statt dessen die Eingangsbilder interpoliert werden. Auf der anderen Seite kann die Interpolation durch eine höherwertige Strategie ausgetauscht werden. Die erste Variante hat den Vorteil, dass der gesamte Algorithmus nur noch mit einer Auflösung arbeiten muss. Weiterhin entfällt dadurch die nicht unerhebliche Zahl an Rechenoperationen für Hoch- und Runtertastung, die gerade bei komplexeren Interpolationsverfahren nicht unerheblich ist. Durch das Vorziehen der Interpolation muss diese nur einmal durchgeführt werden und deshalb kann dort eher ein höherwertiges Verfahren (z.B. eine bildinhaltsabhängige Interpolation, Abschnitt 3.2.2) eingesetzt werden. Weiterhin kann durch den Verzicht auf die interne Interpolation der Super-Resolution Algorithmus als reine Bildqualitätsverbesserung verstanden werden, die zu einer schärferen Sequenz bei gleichzeitig weniger Störungen führt.

Wird die Hoch- und Runtertastung eingesetzt, so kann gerade bei der Hochtastung eine höherwertige Strategie eingesetzt werden. Dazu ist es sinnvoll, den Unschärfefilter und das Einfügen von Nullen durch eine bessere Interpolation auszutauschen, z.B. eine High Resolution Spline Interpolation.

### 5.1.2 Veränderung der Unschärfefilter

Bei den beiden Unschärfefiltern kann die Verbesserung getrennt voneinander erfolgen, da der Filter im Abbildungsprozess und der im Back-Projection Operator unterschiedliche Wirkungen haben. Im Abbildungsprozess dient der Unschärfefilter dazu, hohe Frequenzanteile des aktuellen Ergebnisses zu entfernen bzw. deutlich zu dämpfen, damit diese dann beim Vergleich mit der Eingangssequenz weiter angeschärft werden können, was zu detailreicheren Ergebnissen führt. Der Unschärfefilter im Back-Projection Operator entfernt bzw. dämpft die hohen Frequenzanteile im Detailsignal. Dies kann zu einem natürlicheren Bildeindruck führen, aber auch zu einer nachträglichen Entfernung bereits wiederhergestellter hoher Frequenzen. Je nachdem, welche Interpolationsstrategie angewendet wird, kann dieser Tiefpass also auch weggelassen werden, weil er bereits in die Interpolation integriert ist oder aber ein schärferes Ergebnis gewünscht ist.

### 5.1.3 Alternativen bei der Bewegungskompensation

Wie bereits erwähnt, sind zwei Bewegungskompensationen nötig. Zum einen wird das aktuelle, hoch aufgelöste Bild (mit der Transformationsmatrix  $\mathbf{F}$ ) kompensiert, damit ein Vergleich mit jedem niedrig aufgelösten Eingangsbild möglich ist. Im Back-Projection Operator wird dann das Detailsignal der (inversen) Kompensation (realisiert durch die Transponierte von  $\mathbf{F}$ ) unterzogen, um es an die Phasenlage des aktuellen hoch aufgelösten Bildes anzupassen (siehe Abbildung 5.1). Technisch können diese zwei Kompensationsschritte auf unterschiedliche Weise realisiert werden. Entweder müssen für beide Kompensationsschritte je ein Bewegungsvektorfeld vorliegen, wobei die Vektoren des ersten Feldes im Referenzbild beginnen und die des Zweiten ihren Ursprung im aktuellen hoch aufgelösten Bild haben müssten. Dies erfordert evtl. zusätzliche Schätzungen, Umsortierungen oder Bildspeicher. Andernfalls muss für mindestens einen der beiden Kompensationsschritte eine Vektorprojektion angewendet werden. Problematisch ist hierbei die Tatsache, dass eine Bewegungsschätzung im mathematischen Sinne ein nicht eindeutiges Problem darstellt, da z.B. zwei Vektoren auf dieselbe Stelle im Referenzbild zeigen können (Doppeldeutigkeiten). Dies führt an anderen Stellen zu Löchern, die mit geeigneten Strategien sinnvoll gefüllt werden müssen. Aus diesem Grund wird ein alternatives Vorgehen vorgeschlagen: Statt der zweifachen Kompensation innerhalb der Iterationsschleife wird jedes niedrig aufgelöste Bild auf die Rasterlage des aktuellen hoch aufgelösten Bildes gebracht, wie in Abbildung 5.2 dargestellt. Durch die Kompensation der niedrig aufgelösten Bilder kann (allerdings

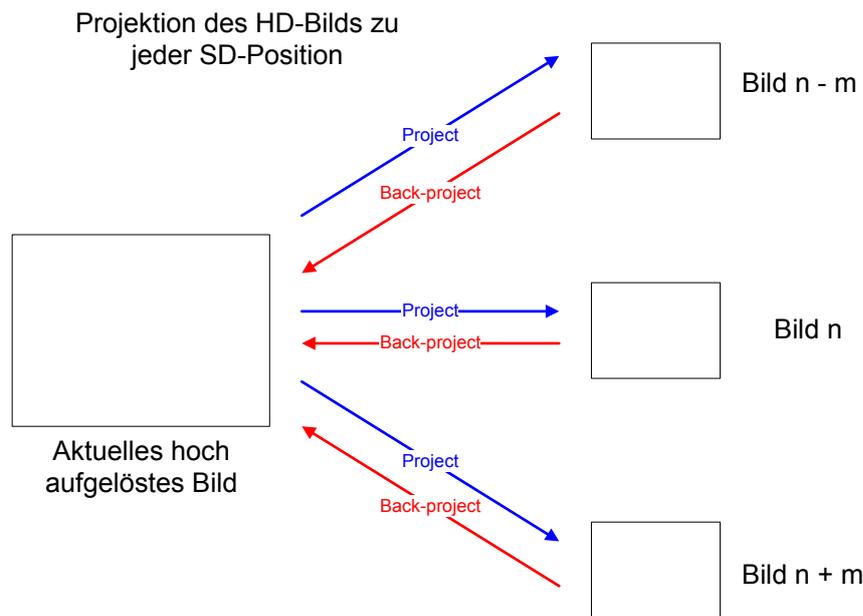


Abbildung 5.1: Ursprüngliches Schema der Back-Projektion mit Bewegungskompensation und der Inversen.

auf Kosten der Qualität, insbesondere bei der Detailwiederherstellung) weiterhin auf eine Anpassung der Vektorfelder an die Auflösung der Ergebnissequenz verzichtet werden.

### 5.1.4 Filterstrategien für die Detailsignale

Die Addition aller Detailsignale entspricht einer Mittelwertfilterung. Diese Filterart liefert zwar für korrekte Bewegungsvektoren gute Ergebnisse, Bewegungsvektorfehler, welche zu Details an der falschen Stelle und damit Seiteneffekten führen, lassen sich dadurch aber nicht vermeiden. Aus diesem Grund wird neben der Mittelwertfilterung eine adaptive FIR-Filterung und eine Medianfilterung vorgeschlagen. Eine Möglichkeit, adaptive Filtergewichte zu berechnen, wird in Abschnitt 5.5 erläutert werden.

### 5.1.5 Mehrreferenzbildbewegungsschätzung

Nach den internen Verbesserungen des Algorithmus wird nun als Letztes noch die verbesserte Mehrreferenzbildbewegungsschätzung präsentiert, welche die Qualität der Super-Resolution erheblich verbessern kann. Dieser Algorithmus wird auch für die anderen SR-Verfahren verwendet werden, wobei für die rein rekursive Super-Resolution ein rückwärts gerichtetes Vektorfeld ausreicht, so dass das Mehrreferenzbildverfahren selbst nicht nötig

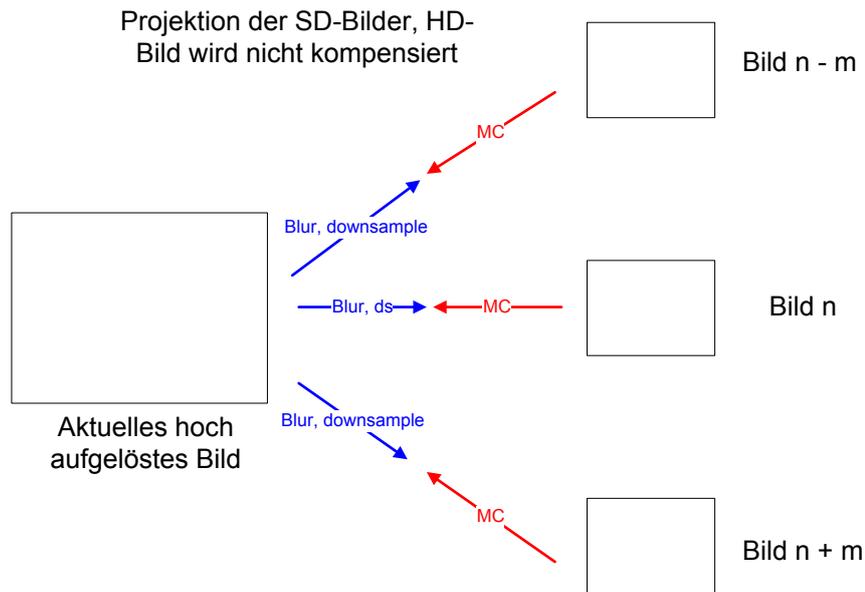


Abbildung 5.2: Back-Projektion Schema mit nur einer Bewegungskompensation.

ist. Für die zeitlich bewegungskompensierte Verarbeitung der Super-Resolution Algorithmen ist eine hochpräzise Bewegungsschätzung von entscheidender Bedeutung. Diese muss auch über mehrere Bilder zuverlässig sein. Aus diesem Grund wurde im Rahmen dieser Arbeit eine Mehrreferenzbildbewegungsschätzung implementiert, so dass Bewegungsinformationen auch für weiter entfernt liegende Referenzbilder zur Verfügung stehen, wie es bereits in Kapitel 2 für das Grundprinzip gezeigt wurde. Diese Technik hat sich bereits in ähnlicher Form in der Codierung bewährt [WSBL03]. Das hier eingesetzte Verfahren basiert aber auf dem Parallel Prädiktiven Blockmatching (PPBM, siehe [Blu97]), da prädiktive Verfahren zur Ermittlung der wahren Bewegung eines Objekts besser geeignet sind als Bewegungsschätzer, die bei der Codierung eingesetzt werden. Wenn die richtige Objektgeschwindigkeit geschätzt wird, was beim Einsatz des PPBM-Verfahrens sehr wahrscheinlich ist, können die Eingangspixel entlang der Bewegungstrajektorie zur Filterung herangezogen werden. Gerade bei der Super-Resolution ist die zuverlässige Schätzung über mehrere Bilder sehr wichtig, da sich in der Literatur kaum Maße finden, um zwischen einer neuen Information und einer fehlerhaften Bewegungsschätzung zu unterscheiden. Durch die Mehrreferenzbildschätzung stehen dabei deutlich mehr Kandidatenpixel für die Verarbeitung zur Verfügung, aus denen dann mit geeigneten Kriterien eine Teilmenge ausgewählt werden kann. Um diese hochpräzise Schätzung möglich zu machen, wurde die Mehrreferenzbildschätzung noch um zwei weitere Elemente erweitert: durch die Verwendung des optischen Flusses anstelle eines Korrektursterns können verbesserte Update-Vektoren berechnet werden. Weiterhin kann durch die Anwendung einer hierarchischen Schätzung mit

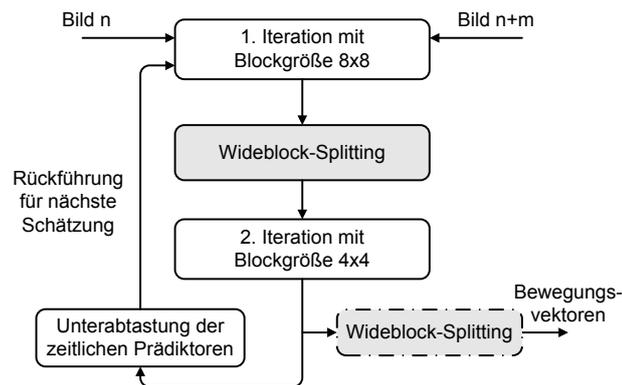


Abbildung 5.3: Iterationsschema der Mehrreferenzbildbewegungsschätzung

einem internen Block-Splitting zwischen den Iterationen die Güte der zeitlichen Prädiktoren verbessert werden und gleichzeitig die Auflösung des Vektorfelds gesteigert werden. Die folgenden drei Teilabschnitte fassen diese Strategien aus Gründen der Vollständigkeit zusammen. Sie wurden aber bereits in [RDS09b] vorgestellt.

### Hierarchischer Aufbau der Schätzstrategie

Für die Qualität des Blockmatchings ist die verwendete Blockgröße von entscheidender Bedeutung. Während bei einer zu kleinen Blockgröße lokale Störungen das Schätzergebnis beeinträchtigen können, führt eine zu große Blockgröße zu einer niedrigen Auflösung des Ergebnisvektorfelds und damit zu möglichen Artefakten an Grenzen bewegter Bildbereiche. Aus diesem Grund wird auf eine hierarchische Schätzstrategie zurückgegriffen, die in Abbildung 5.3 dargestellt ist. Pro Bild werden dazu zwei Iterationen durchgeführt, um die Qualität iterativ zu verbessern. Zwischen den beiden Schätzungen werden die zeitlichen Prädiktoren durch eine Splitting-Strategie verfeinert, so dass nun viermal so viele Prädiktoren zur Verfügung stehen und in der nächsten Iteration direkt ein feineres Prädiktorfeld Verwendung findet.

### Wideblock-Splitting

Generell handelt es sich beim Wideblock-Splitting um ein Verfahren zur Erhöhung der Vektorfeldauflösung nach einer Schätzung bzw. Iteration. Bei dieser Methode wird ausgenutzt, dass die Bestimmung des Fehlerkriteriums (meist SAD) beim PPBM-Verfahren auf so genannten 'Wideblocks' durchgeführt wird, um die Konvergenz der Schätzung zu verbessern [Blu97]. Dadurch ist ein guter Kompromiss zwischen der Auflösung des Vektorfelds und Robustheit der Schätzung gewährleistet, da ein größerer Block bei der Berechnung des Fehlerkriteriums stabilere Ergebnisse liefert. Die Dimensionen des Wideblocks sind grö-

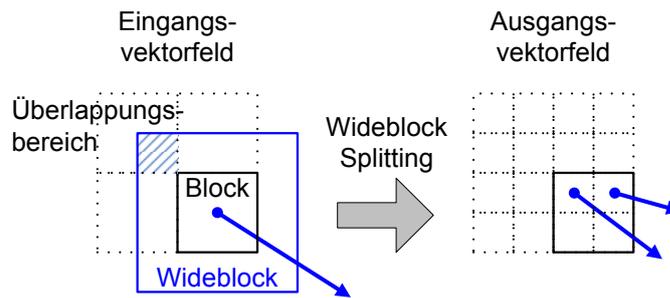


Abbildung 5.4: Wideblock-Splitting zur Verfeinerung des Vektorfelds

ßer als der Block, dem der Ergebnisvektor zugewiesen wird. Der Einfachheit halber sei hier angenommen, dass ein Wideblock in seinen Abmessungen die vierfache Fläche des normalen Blocks aufweist. Dadurch überlappen sich jeweils vier Teilbereiche der 'Wideblocks', wie es Abbildung 5.4 verdeutlicht. Als Folge davon sind auch pro Wideblock vier mögliche Kandidaten vorhanden, die für ein finales feiner aufgelöstes Vektorfeld zur Verfügung stehen. Nun wird diesem Blockviertel derjenige Vektor zugewiesen, der zu einem der vier Wideblocks gehört und der das kleinste Teilfehlerkriterium (Sub-SAD) aller vier Wideblocks für dieses Blockviertel aufweist. Somit ist die Erhöhung der Auflösung des Vektorfeldes recht effizient möglich, ohne dass viele zusätzliche Operationen nötig sind. Im Vergleich zu [PK04] hat dieses Verfahren den Vorteil, dass es bereits nach einer Iteration angewendet werden kann und nicht eine zweite Iteration wie beim Raster-Offset nötig ist.

### Einführung weiterer Prädiktoren

Ein Nachteil der prädiktiven Bewegungsschätzungsverfahren wie PPBM ist die Verwendung fester Update-Vektoren aus einem Korrekturstern, welche eine schnelle Konvergenz besonders bei diagonalen Bewegungen wegen der Belegung des Korrektursterns erschweren. Außerdem ergibt sich bei Prüfung nicht aller Update-Vektoren für den aktuellen Block die Möglichkeit, dass ein für die Bildsituation ungünstiger Update-Vektor aus dem Korrekturstern ausgewählt wird. Dadurch kann die Konvergenz erschwert bzw. bei schnellen Bewegungen evtl. gar nicht mehr möglich sein. Aus diesem Grund wird ein weiterer Kandidat mit Hilfe des optischen Flusses nach der Methode von Lucas und Kanade erzeugt [LK81]. Ein ähnlicher Ansatz ist bereits in [OTBH97] gewählt worden. Bei diesem Verfahren geschieht die Berechnung nach dem Gradientenverfahren pixelweise und separiert in x- und y-Richtung, wodurch mit einer höheren Stömpfindlichkeit zu rechnen ist. Das Grundprinzip dieser zusätzlichen Update-Vektorberechnung soll anhand von Bild 5.5 erklärt werden. Generell eignet sich der optische Fluss nur für die Bestimmung kleiner Geschwindigkeiten. Aus diesem Grund wird der optische Fluss zwischen dem aktuellen Block

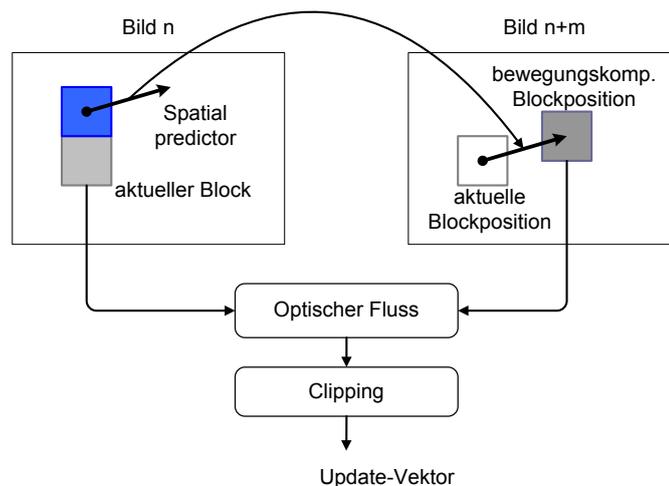


Abbildung 5.5: Prediktorgeneration mit dem optischen Fluss

und den um den örtlichen Prädiktor verschobenen Block im nächsten oder vorherigen Bild berechnet. Eine Begründung für dieses Vorgehen liefert die in [HBHO93] aufgestellte Annahme, dass Objekte meist größer als ein Block sind, wodurch von Block zu Block außer an Objektgrenzen nur sehr geringe Geschwindigkeitsänderungen zu erwarten sind. Deshalb kann mit dem optischen Fluss effektiv diese echte Bewegungsänderung bestimmt werden, welche normalerweise durch einen zufälligen Update-Vektor vorgegeben wird. Zum Schluss findet noch ein 'Clipping' auf eine maximale Vektorlänge statt, um Artefakte durch unwahrscheinliche, zu lange Bewegungsvektoren zu vermeiden. Neben diesem Kandidaten erfolgt eine Berechnung weiterer Kandidaten. Neben dem Nullvektor, dem örtlichen und dem zeitlichen Prädiktor können durch eine Verschränkung der verschiedenen Schätzungen für die einzelnen Referenzbilder (z.B. Vorwärts- und Rückwärtsschätzung, siehe [LBN07]) weitere Verbesserungen erzielt werden. Diese Verschränkung kann vor allem für die prädiktive Mehrreferenzbildschätzung eingesetzt werden. Dabei ist aber in Abhängigkeit des Abstands zwischen aktuellem Bild und Referenzbild eine Längen Anpassung dieses Prädiktors nötig. Durch diesen zusätzlichen Prädiktor sind die Schätzungen für die einzelnen Referenzbildern miteinander verbunden und bauen aufeinander auf.

## 5.2 Doppelt rekursive Mehrreferenzbild Super-Resolution (MUSIDORA)

In den vorigen Abschnitten wurde gezeigt, wie die Back-Projection Super-Resolution weiter verbessert werden kann. Zwar könnte der nun beschriebene doppelt rekursive Algorithmus ebenfalls aus der Back-Projection entwickelt werden, allerdings wird hier noch einmal der

gesamte Entwurfsprozess durchgeführt und gezeigt, wie die wichtigsten Modifikationen bereits beim Aufstellen der Ansatzes berücksichtigt werden können.

## 5.3 Aufstellen der Entwurfsgleichung

Basis für den Entwurf ist das in Gleichung 5.1 dargestellte, zwei Terme umfassende Energiefunktional. Der erste Term gewährleistet die Ähnlichkeit zwischen niedrig und hoch aufgelöster Bildfolge, die über das Beobachtungsmodell miteinander verknüpft sind. Da die niedrig aufgelöste Sequenz bis auf Aliasanteile nur Frequenzen unterhalb der Nyquistfrequenz enthält und die hoch aufgelöste Sequenz vor dem Vergleich tiefpassgefiltert wird, wird durch diesen Anteil nur das nieder- und mittelfrequente Spektrum der Ausgangssequenz kontrolliert. Für die höheren Frequenzen besteht durch diesen Term keine Beschränkung. Um auch für diese Frequenzanteile eine sinnvolle und eindeutige Lösung zu erhalten, wird eine Nebenbedingung (Term 2) eingeführt. Diese enthält einen Bestrafungsterm mit Hochpasscharakteristik:

$$\arg \min_{\mathbf{z}} \left( \int \left[ \sum_n (\mathbf{y}_n - \mathbf{DBF}_n \mathbf{z})^2 + \frac{1}{2\beta} \sum_{i=0}^3 \rho(d_i(\mathbf{z})) \right] \right) \quad (5.1)$$

Anstelle nun dieses (kontinuierliche) Funktional zu minimieren, wird ein modifizierter Ansatz verwendet, der die Zahl der Operationen stark reduzieren kann und gleichzeitig besser auf eine blockbasierte Bewegungsschätzung abgestimmt ist:

$$\arg \min_{\mathbf{z}} \left( \int \left[ \sum_n (\mathbf{y}_n^{int+mc} - \mathbf{Bz})^2 + \frac{1}{2\beta} \sum_{i=0}^3 \rho(d_i(\mathbf{z})) \right] \right) \quad (5.2)$$

In dieser Gleichung wird anstelle einer Bewegungskompensation und Unterabtastung der hoch aufgelösten Sequenz  $\mathbf{z}$  die Eingangssequenz  $\mathbf{y}$  bewegungskompensiert und hochgetastet, wie es bereits bei der vielseitig parametrierbaren Back-Projection erwähnt wurde. Dieses Vorgehen kann Seiteneffekte sowie den Aufwand deutlich reduzieren. Die initiale Interpolation des Eingangsbildes führt dazu, dass auf die Interpolation und Dezimation bei der Berechnung innerhalb des Iterationsschemas verzichtet werden kann. Gerade bei der Verwendung mehrerer Iterationen pro Bild kann bei einem höherwertigen Interpolationsverfahren (z.B. trainierte Filter) erheblich an Rechenaufwand gespart werden. Durch diese beiden Modifikationen ist der Algorithmus nun ein iteratives Deblurring-Verfahren, welchem eine zeitliche Filterung und Interpolation vorgeschaltet sind. Optional kann also auch die Interpolation weggelassen werden und das Verfahren als reines Enhancement-Verfahren verstanden werden, was den Einsatzzweck deutlich vergrößert.

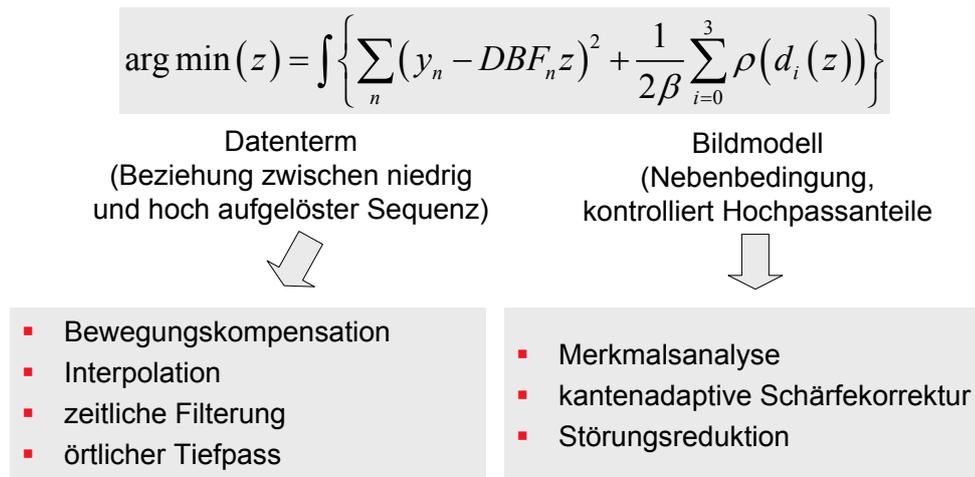


Abbildung 5.6: Beziehung zwischen mathematischem Energiefunktional und bildverarbeitenden Operatoren

### 5.3.1 Herleitung der Filtergleichung

Bevor eine exakte Ableitung der Filtergleichung durchgeführt wird, soll mit Abbildung 5.6 gezeigt werden, wie der Ansatz durch pixelweise arbeitende Operatoren realisiert werden kann. Die Matrizen zur Beschreibung der Unterabtastung, Unschärfefilterung und Bewegungskompensation können durch lokal arbeitende Abtastprozesse, Faltungen oder Verschiebungen ersetzt werden. Die Operationen des Bildmodells entsprechen einer kantenerhaltenden adaptiven Rauschreduktion, die durch die Größe der zweiten Ableitung gesteuert wird.

#### Diskretisierung des kontinuierlichen Energiefunktionals

Um nun die exakte Filtervorschrift zu bestimmen, wird zuerst das Energiefunktional umgeschrieben, indem das Integral durch eine diskrete Summe und die Matrix-Multiplikation mit  $\mathbf{B}$  durch eine diskrete Faltung mit einem lokalen Filter  $\mathbf{b}$  ersetzt wird. Diese beiden Schritte bedeuten keine Einschränkung für die Lösung, da das Eingangsbildmaterial  $\mathbf{y}$  sowieso in diskreter Form vorliegt und deshalb das Integral sowieso nur an diskreten Punkten ausgewertet werden könnte. Außerdem kann bekanntlich jede (diskrete) Faltung ebenfalls durch eine Matrixmultiplikation ausgedrückt werden [Sch00] und somit ist auch die inverse Operation möglich. Dadurch ergibt sich das folgende diskrete Funktional, in dem die Diskretisierung des Bildmodells noch nicht berücksichtigt ist:

$$\arg \min_{\mathbf{z}} \left( \sum_{k,l} \left[ \sum_m (\mathbf{y}_{k,l,n+m}^{int+mc} - \mathbf{b} * \mathbf{z}_{k,l,n})^2 + \frac{1}{2\beta} \sum_{i=0}^3 \rho(d_i(\mathbf{z}_{k,l,n+m})) \right] \right) \quad (5.3)$$

Die äußere Summe über  $k, l$  wird über das gesamte Bild gebildet. Das bedeutet, dass für jedes Eingangsbild ein einzelnes Optimum des Funktionals zu bestimmen ist. Für die Diskretisierung der zweiten Ableitung wurde mit dem diskreten Laplace Operator  $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$  eine sehr pragmatische Lösung gewählt bzw. diese aus der Literatur übernommen [BS99]. Höherwertige Diskretisierungen unter Einbeziehung weiterer Bildpunkte sind natürlich möglich. Wendet man nun die Diskretisierung auf das Bildmodell an und rechnet die innere Summe über  $m$  aus, so ergibt sich für jeden Bildpunkt folgende Vorschrift:

$$E(k, l, n) = \left[ \left( \sum_m y_{k,l,n+m}^{int+mc} \right) - m \cdot \sum_{i,j} \mathbf{b}_{i,j} \mathbf{z}_{k+i,l+j,n} \right]^2 + \frac{1}{2\beta} \sum_{d_i, d_j} \begin{cases} (\mathbf{z}_{k+d_i, l+d_j, n} - 2\mathbf{z}_{k, l, n} + \mathbf{z}_{k-d_i, l-d_j, n})^2 & \text{for } d_i(\mathbf{z}) \leq t_\alpha \\ 2t_\alpha \cdot |\mathbf{z}_{k+d_i, l+d_j, n} - 2\mathbf{z}_{k, l, n} + \mathbf{z}_{k-d_i, l-d_j, n}| & \text{sonst} \end{cases} \quad (5.4)$$

Im Vergleich zur Gleichung 5.3 ist die Summe über alle Eingangsbilder (Laufindex  $m$ ) bereits auf die einzelnen Summanden aufgeteilt worden, so dass die Struktur der zeitlichen Filterung über alle interpolierten und bewegungskompensierten Eingangsbilder und die Tiefpassfilterung des hoch aufgelösten Ergebnisbildes zu erkennen ist. Im zweiten Teil ist die diskrete Ableitung in alle vier Hauptrichtungen zu bilden. Dies wird durch die Summe über  $d_i, d_j$  zur Beschreibung der Drehung erreicht. Das diskrete Energiefunktional (pro Bild) ergibt sich dann durch Aufsummation:

$$E_n = \sum_{k,l} E(k, l, n) \quad (5.5)$$

### Minimierung des diskreten Energiefunktional

Die Minimierung erfolgt durch die Bildung der (Richtungs-) ableitung nach jedem Pixelwert  $\mathbf{z}_{k,l,n}$ . Das Ergebnis der Ableitung ist dann zu Null zu setzen, um die notwendige Bedingung für das Minimum zu erfüllen.

Anstelle nun aber das Energiefunktional des Gesamtbildes aus Gleichung 5.5 zu bilden, wie es beispielsweise bei der Back-Projection und der Total Variation gemacht wurde, soll hier nun (pixelweise) die lokale Energie aus Gleichung 5.4 minimiert werden. Durch diesen Fall ist eine einfache Ableitung mittels dem Gradientenverfahren auch für das Bildmodell möglich. Ansonsten wäre eine deutlich komplexere, auf Matrixoperationen aufbauende Lösung nötig gewesen [Cai06]. Durch diese Modifikation ergibt sich natürlich nicht das globale (mathematische) Optimum sondern nur ein Lokales, dafür aber eine erheblich weniger

aufwändige Rechenvorschrift, die sogar, wie später gezeigt wird, zu schärferen Ergebnissen führen kann. Die Ableitung nach jedem Pixel lautet für den lokalen Fall:

$$0 = \frac{\partial E(k, l, n)}{\partial \mathbf{z}_{k,l,n}} = \left[ \left( \sum_m \mathbf{y}_{k,l,n+m}^{int+mc} \right) - m \cdot \sum_{i,j} \mathbf{b}_{i,j} \cdot \mathbf{z}_{k+i,l+j,n} \right] (-\mathbf{b}_{k,l} \cdot 2) +$$

$$+ \frac{1}{2\beta} \sum_{d_i, d_j} \begin{cases} (\mathbf{z}_{k+d_i, l+d_j, n} - 2\mathbf{z}_{k,l,n} + \mathbf{z}_{k-d_i, l-d_j, n}) \cdot (-2 \cdot 2) & \text{for } d_i(\mathbf{z}) \leq t_\alpha \\ 2t_\alpha \cdot \text{sign}(\mathbf{z}_{k+d_i, l+d_j, n} - 2\mathbf{z}_{k,l,n} + \mathbf{z}_{k-d_i, l-d_j, n}) \cdot (-2) & \text{sonst} \end{cases} \quad (5.6)$$

Dieses Ergebnis besteht aus insgesamt 3 Termen, die Aufschluss über die Funktionsweise der Super-Resolution geben. Die ersten beiden Terme wurden bereits kurz erläutert. Dabei handelt es sich um die bewegungskompensierte, zeitliche Mittelwertfilterung und die örtliche Tiefpassfilterung des Ergebnisbildes. Der dritte Term, der sich durch das Bildmodell ergibt, ist ein Hochpassterm, der durch den Wert und das Vorzeichen der zweiten Ableitung eine unterschiedlich starke Reduktion dieser Anteile im Ergebnisbild bewirkt. Um Kanten und Bereiche hoher Aktivität vor der Filterung zu schützen, findet die Reduktion der Hochpassanteile nur statt, wenn die zweite Ableitung in die jeweilige Richtung unterhalb eines Schwellwerts  $t_\alpha$  liegt. Ansonsten wird der jeweilige Bereich nur um  $2 \cdot t_\alpha$  verändert, so dass diese Begrenzung der Verarbeitung als eine obere bzw. untere 'Clipping'-Schwelle wirkt. Diese Begrenzung durch das Bildmodell ermöglicht im Ergebnisbild erst hochfrequente Anteile und starke Kanten. Diese Interpretation des Bildmodells stellt auch einen Bezug zum statistischen MAP-Ansatz (Gleichung 3.4) her, der in [BS99] und [Bor96] bereits diskutiert wurde. Das Aussehen der Lösung ist dabei durch die Wahrscheinlichkeitsverteilungsfunktion  $p(\mathbf{z})$  gegeben. Durch die Verwendung einer gauß'schen Wahrscheinlichkeitsverteilungsfunktion in Kombination mit der Huber-Funktion  $\rho$ , die hohe zweite Ableitungen nur linear bestraft, sind hohe Frequenzanteile im Ergebnis statistisch möglich. Bei einer stärkeren Bestrafung auch der hohen Frequenzanteile wären diese im Ergebnis statistisch viel unwahrscheinlicher.

### Anwendung einer iterativen Lösungsstrategie zum Finden des Minimums

Für die Berechnung der Filterfunktion wird nun ein iteratives Gradientenverfahren angewendet, welches auf folgendem Schema beruht:

$$\mathbf{z}_{q+1} = \mathbf{z}_q + \tau \frac{\partial E(\mathbf{z}_q)}{\partial \mathbf{z}_{k,l,n}} \quad (5.7)$$

Bei dieser Strategie wird in jeder Iteration das in Gleichung 5.6 gezeigte Korrektursignal hinzuaddiert, wobei jeweils die jüngste vorhandene Lösung  $\mathbf{z}_q$  einzusetzen ist. Als Ergebnis erhält man die folgende, finale Gleichung:

$$\begin{aligned} \mathbf{z}_{q+1} = \mathbf{z}_q + \tau & \left[ \left( \sum_m \mathbf{y}_{k,l,n+m}^{int+mc} \right) - m \cdot \sum_{i,j} \mathbf{b}_{i,j} \cdot (\mathbf{z}_{k+i,l+j,n})_q \right] (m \cdot \mathbf{b}_{k,l}) + \\ & + \frac{\tau}{2\beta} \sum_{d_i,d_j} \begin{cases} \left( (\mathbf{z}_{k+d_i,l+d_j,n})_q - 2(\mathbf{z}_{k,l,n})_q + (\mathbf{z}_{k-d_i,l-d_j,n})_q \right) & \text{for } d_i(\mathbf{z}) \leq t_\alpha \\ 2t_\alpha \cdot \text{sign} \left( (\mathbf{z}_{k+d_i,l+d_j,n})_q - 2(\mathbf{z}_{k,l,n})_q + (\mathbf{z}_{k-d_i,l-d_j,n})_q \right) & \text{sonst} \end{cases} \end{aligned} \quad (5.8)$$

Alternativ bestehen auch andere Möglichkeiten zur Berechnung einer finalen Filtervorschrift. So kann Gleichung 5.6 einfach nach dem aktuellen Pixelwert  $\mathbf{z}_{n,k,l}$  aufgelöst werden, wodurch sich eine leicht modifizierte Filtervorschrift ergeben würde. Der wichtigste Änderungspunkt wäre dabei im Term des Bildmodells zu finden, bei dem anstelle der Subtraktion der hochfrequenten Anteile die Addition des Ergebnisses eines Mittelwertfilter (Tiefpass) eingesetzt wird. Faktisch wird dadurch aber ein sehr ähnlicher Effekt erzielt.

### Blockschaltbild zur Realisierung der modifizierten MAP-SR

Abbildung 5.7 zeigt das mit der Filtervorschrift korrespondierende Blockdiagramm. In diesem sind die verschiedenen Signalfade, die sich aus den Termen in Gleichung 5.8 ableiten lassen, gut zu sehen. Im Vergleich zum in der Literatur beschriebenen Verfahren [BS99] unterscheidet sich das hier vorgestellte Verfahren durch die vorher durchgeführte Interpolation und Bewegungskompensation. Diese Modifikationen ermöglichen eine deutlich verringerte Anzahl an Operationen in der Iterationsschleife, die – wie später noch gezeigt wird – trotzdem zu einer besseren Detailwiederherstellung führen. Ähnlich wie beim Back-Projektion Ansatz zur Super-Resolution wird auch hier durch eine örtliche Tiefpassfilterung des hoch aufgelösten Ergebnisses eine modellierte, niedrig aufgelöste Sequenz berechnet, die in diesem Fall mit dem zeitlich bewegungskompensierten Filterergebnis verglichen wird. Die dabei entstehende Differenz fungiert als Korrektursignal, um Bildschärfe und Details hervorheben zu können. Durch die lokale Minimierung anstelle der globalen ergibt sich hier aber eine weitere Modifikation, da nun das Detailsignal nicht mehr tiefpassgefiltert wird. Dies erklärt die höhere Bildschärfe im Ergebnis.

Durch die zeitliche Tiefpassfilterung am Eingang wird eine Rauschreduktion erreicht, die bei korrekten Bewegungsvektoren einen vollständigen Detailerhalt gewährleistet und sogar eine Überführung von Details aus anderen Eingangsbildern bewirkt. Durch die Verwendung anderer zeitlicher Filter, z.B. Medianfilter, könnte ein anderes Ergebnis mit z.B. höherer Robustheit gegen Vektorfehler oder die Unterdrückung gegen Salt- & Pepper Noise erzielt werden. Diese Art der Filterung kann durch eine Modifikation des Energiefunktional

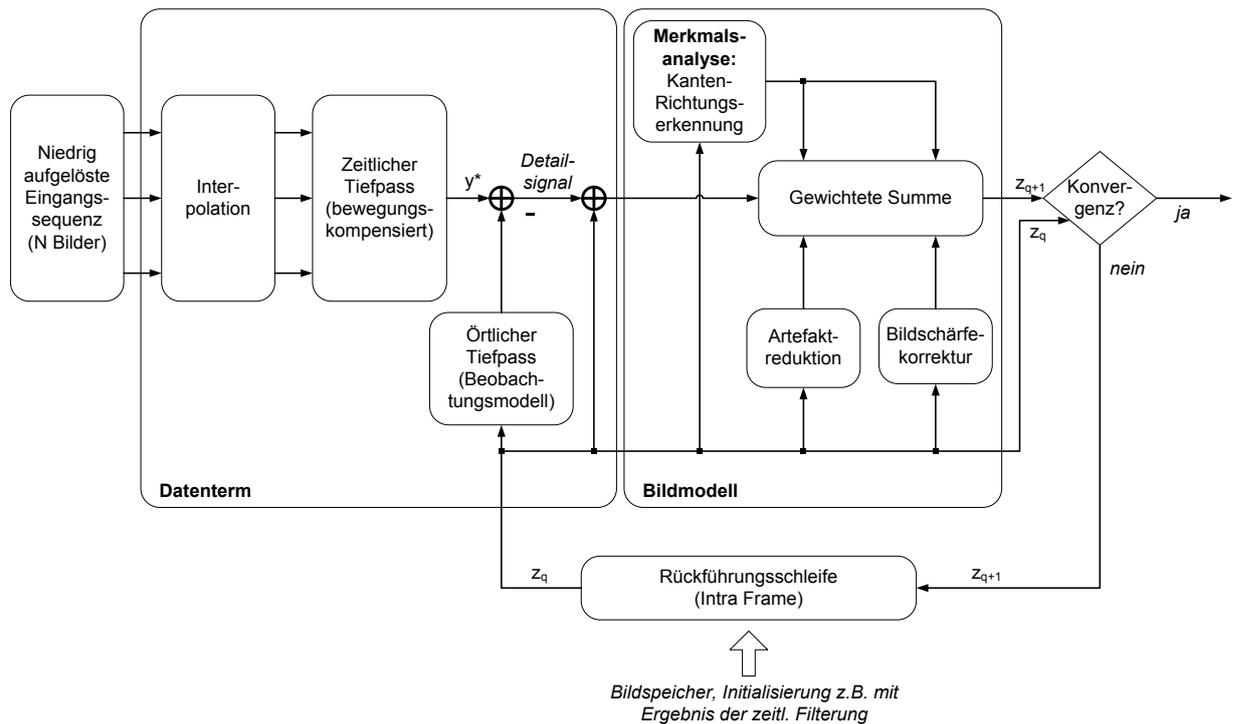


Abbildung 5.7: Blockschaltbild des Mehrreferenzbild Super-Resolution-Systems

auch mathematisch beschrieben werden, indem eine Maximumsnorm anstelle der euklidischen Norm verwendet wird [FREM04]. Die Clipping Operation im Bildmodell wird in diesem Fall über eine Merkmalsanalyse (Auswertung der zweiten Ableitung in jede der vier Richtungen) gesteuert.

Zusammenfassend zeichnet sich dieses Mehrreferenzbildverfahren dadurch aus, dass es im Falle korrekter Bewegungsvektoren und einer Anzahl von 5-11 Eingangsbildern sehr gute Ergebnisse liefert, sofern das Beobachtungs- und Bildmodell erfüllt sind. Das trifft u.a. dann zu, wenn am Eingang eine Interpolation durchgeführt wird, die zu einer Unschärfe im Interpolationsergebnis führt. Als Nachteile sind bei diesem Verfahren die hohe Zahl an Bildspeichern und die komplexe und aufwändige Mehrreferenzbildbewegungsschätzung anzusehen. Weiterhin können Bewegungsvektorfehler im Ergebnisbild bei Verwendung einer zeitlichen Mittelwertfilterung wahrgenommen werden, was zu einer Einschränkung in der Bildqualität führt. Fehlende Steuerungsmöglichkeiten erschweren zudem die Verwendung von unterschiedlichem Eingangsmaterial mit stark variierendem Schärfe- oder Störungsgrad, weil entweder das Enhancement zu stark ausfüllt oder eine Restunschärfe bleibt.

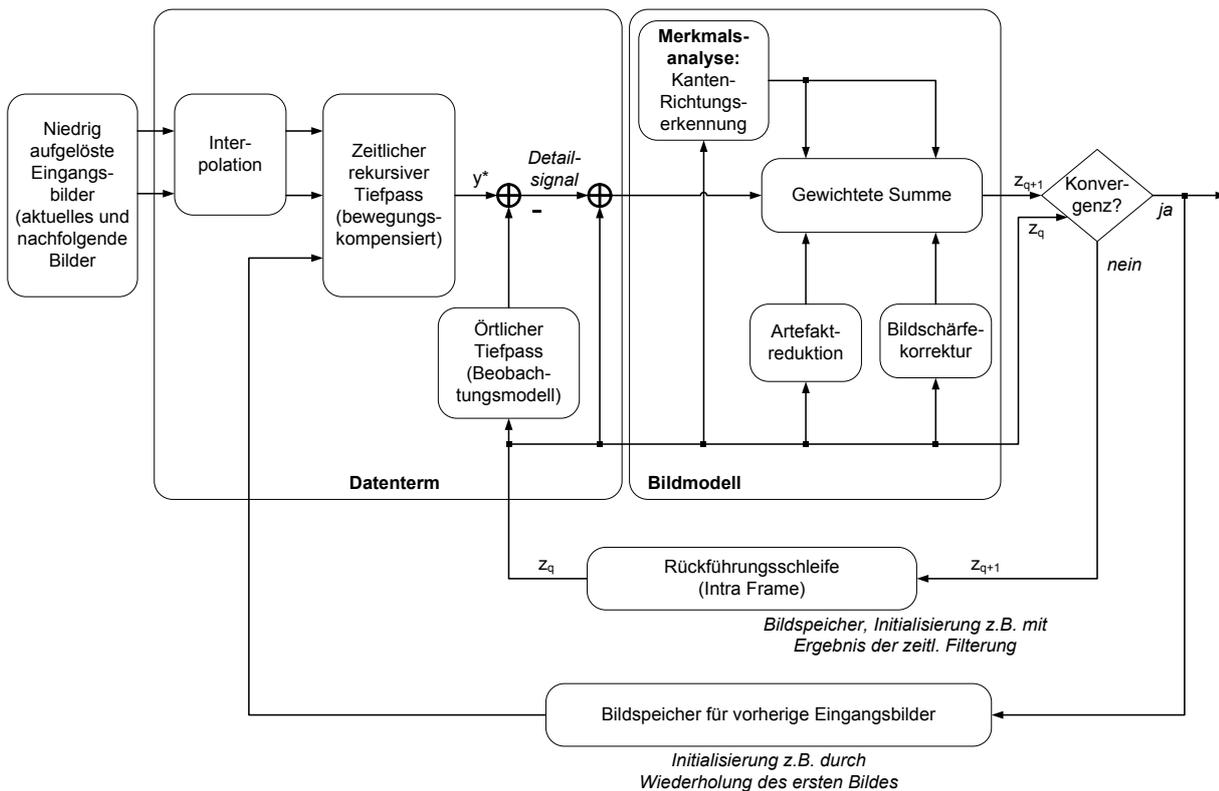


Abbildung 5.8: Verbessertes Mehrreferenzbildverfahren mit doppelt rekursiver Verarbeitung (MUSIDORA)

### 5.3.2 Verbesserung des entstandenen Verfahrens

#### Zeitlich rekursive Mehrreferenzbildfilterung

Nachdem nun die ersten beiden Schritte des Entwurfsprozess vollführt wurden und bereits erhebliche Verbesserungen in Bezug auf den Rechenaufwand erzielt wurden, sind noch weitere Verbesserungen möglich. Dadurch, dass die zeitliche Filterung außerhalb der Rekursionsschleife nur einmal durchgeführt wird, kann diese in eine zeitlich rekursive Mehrreferenzbildfilterung umgewandelt werden, wenn für die vorherigen Bilder bereits berechnete Ergebnisse verwendet werden (siehe Abbildung 5.8). Dadurch kann die aktuelle Lösung auch von vorherigen Ergebnissen profitieren, die Konvergenzgeschwindigkeit gesteigert werden und die Korrelation zwischen benachbarten Ergebnisbildern noch erhöht werden. Allerdings muss dann eine Stabilitätsprüfung vorgenommen werden, um die Detail- und Schärfenanhebung zu kontrollieren (vergleiche Abschnitt 5.6) oder das Detail-signal adaptiv durch eine Bildanalyse heruntergewichtet werden.

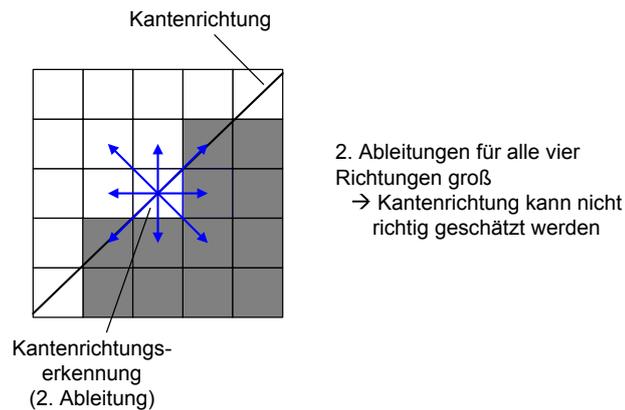


Abbildung 5.9: Strategie zur Erkennung der Kantenrichtung gemäß dem Huber Markov Zufallsfeld Bildmodell

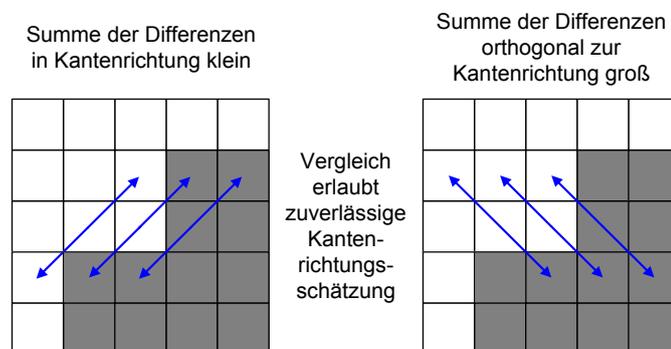


Abbildung 5.10: Verbesserte Kantenrichtungserkennung für das Bildmodell

### Verbesserte Kantenrichtungserkennung für das Huber-Markov Zufallsfeld Bildmodell

Das Bildmodell hat die Funktion, eine mathematisch eindeutige Lösung zu garantieren. Dies geschieht durch eine örtliche Bildverbesserung des Ergebnisbildes. Das modifizierte Huber-Bildmodell, das sich durch die lokale Minimierung ergibt, besteht aus zwei Termen, zwischen denen adaptiv umgeschaltet wird. Als Steuerung wird der Wert der zweiten Ableitung in die 4 Hauptrichtungen verwendet (Abbildung 5.9). Wird als initiale Interpolation ein sehr niederqualitatives Verfahren, z.B. eine Pixelwiederholung verwendet, so ist durch eine einfache Lokalisierung der zweiten Ableitung eine Richtungserkennung der Kante nicht mehr möglich, so dass durch die hohen zweiten Ableitungen in alle vier Richtungen nur eine sehr schwache Glättung erfolgt. Dies ist auch der Suche nach einem lokalen Optimum geschuldet. Ein ähnliches Problem wird aber bereits in [AK06] besprochen und als Lösung eine richtungsabhängige Berechnung des Gradienten vorgeschlagen. Dazu ist

eine Richtungsschätzung erforderlich, welche sich folgermaßen realisieren lässt: Anstelle eines einfachen Differenzschemas, wird die Berechnung eines dreifachen Differenzschemas vorgeschlagen, welches eine robustere Erkennung der Kantenrichtung ermöglicht, wie es in Abbildung 5.10 verdeutlicht wird. Zusätzlich zur Bestimmung der Differenzen in Kantenrichtung wird das Ergebnis der Differenzen noch mit der orthogonalen Richtung verglichen und mit Entscheidungsregeln die finale Kantenrichtung bestimmt. Dabei muss das Ergebnis des dreifachen Differenzschemas orthogonal zur Kantenrichtung über einem Schwellwert liegen und außerdem einen deutlich höheren Wert (z.B. um das 1,5 fache höher) als das Differenzschema in Kantenrichtung aufweisen. Dadurch ist eine effizientere Glättung von Treppenartefakten möglich als mit dem einfachen Schema. Eine weitere Verbesserung der Verarbeitung kann dadurch erzielt werden, dass nicht nur eine reine Glättung durch das Bildmodell, sondern auch eine Kantenanhebung orthogonal zur Kantenrichtung durchgeführt wird. Diese Anschärfung kann aber nur durchgeführt werden, wenn durch weitere Kontrollinstanzen, z.B. eine adaptive FIR-Filterung oder eine nachträgliche Korrektur des vollständigen Detailsignals die Stabilität des Verfahrens gewährleistet ist. Ansonsten besteht die Gefahr der Divergenz für dieses Verfahren.

Des Weiteren kann der Algorithmus noch durch Analysemaße kontrolliert werden. Da dieses Vorgehen aber mit dem für den rein rekursiven Algorithmus identisch ist, wird dies exemplarisch nur für den folgenden rein rekursiven Algorithmus gezeigt. Die Stabilitätsprüfungen aller Algorithmen wird abschließend nach dem Entwurf aller Algorithmen in Abschnitt 5.6 durchgeführt.

## 5.4 Zeitlich rein rekursive Super-Resolution (ESTERA)

Für die Anwendung der Super-Resolution in der Videosignalverarbeitung ist vor allem der Speicher- und Rechenaufwand von großer Bedeutung. Aus diesem Grund wurde neben Mehrreferenzbildverfahren auch das Potenzial eines rein rekursiven Algorithmus untersucht. Weiterhin wird bei der Beschreibung dieses rein rekursiven Algorithmus auch auf die Parametrisierung bzw. die verwendeten Kontrollinstanzen eingegangen, um den Algorithmus zu steuern und somit einen Einsatz für sehr unterschiedliche Videoquellen möglich zu machen. Da dieser Algorithmus auf der Grundstruktur der Back-Projection Super-Resolution aufbaut, sind in der Literatur ähnliche rein rekursive Ansätze bereits zu finden [INXO10]. Die Vorgehensweise bei der Entwicklung, das Weglassen der Hoch- und Runtertastung sowie viele der Kontrollinstanzen sind dort aber nicht zu finden.

### 5.4.1 Entwurf des rein rekursiven Algorithmus durch Modifikation der Filtergleichung

Um einen zeitlich rekursiven Algorithmus zu erhalten, muss das örtlich arbeitende Iterationsschema, welches zu einer längeren und nicht vorhersagbaren Laufzeit pro Bild führt, in ein zeitlich rekursives Schema überführt werden. Um die Bildspeicher für die Berechnung der zeitlichen Mehrreferenzbildfilterung einzusparen, wird außerdem immer nur ein Detailsignal pro Bild berechnet, so dass immer nur ein Abgleich mit dem aktuellen Eingangsbild durchgeführt wird. Die Aufsummierung der Detailsignale geschieht dann bildweise über die zeitliche Abfolge. Durch diese Modifikationen sinkt die Iterationszahl auf eine pro Bild. Die Folge davon ist, genau wie bei den direkten Verfahren, eine vorhersagbare Laufzeit. Um diese Änderungen umzusetzen, wird in Gleichung 5.8 das vorherige, zeitliche Iterationsergebnis  $\mathbf{z}_q$  durch das bewegungskompensierte, vorige Ergebnisbild  $\mathbf{z}_{n-1}^{mc}$  und das nächste Iterationsergebnis  $\mathbf{z}_{q+1}$  durch das aktuelle Ergebnisbild  $\mathbf{z}_n$  ersetzt:

$$\mathbf{z}_n = \mathbf{z}_{n-1}^{mc} + (\mathbf{y}_n - \mathbf{b} * \mathbf{z}_{n-1}^{mc}) + \Gamma(\mathbf{z}_{n-1}^{mc}) \quad (5.9)$$

Der Term für das Bildmodell wurde in diesem Fall durch  $\Gamma$  ersetzt. Diese direkte Umwandlung in ein zeitlich rekursives Schema ist nur für den Fall idealer Bewegungsvektoren geeignet. Bewegungsvektorfehler machen sich im Ergebnis der nächsten Iteration sehr stark bemerkbar. Da diese Vektorfehler mathematisch sehr schwer modellierbar sind, ist sogar nicht auszuschließen, dass im ungünstigsten Fall der Algorithmus divergiert bzw. Stabilität nicht gewährleistet werden kann. Aus diesem Grund wird die vorherige Iteration  $\mathbf{z}_{n-1}^{mc}$  durch eine gewichtete Auswahl (englisch: Weighted Selection (WS)) des aktuellen, niedrig aufgelösten und hoch interpoliertem Bild  $\mathbf{y}$  und der vorherigen, bewegungskompensierten Lösung  $\mathbf{z}_{n-1}^{mc}$  ersetzt:

$$\mathbf{z}_n = \mathbf{z}_n^{ws} + (\mathbf{y}_n - \mathbf{b} * \mathbf{z}_n^{ws}) + \Gamma(\mathbf{z}_n^{ws}) \quad (5.10)$$

Für die Berechnung von  $\mathbf{z}^{ws}$  gilt:

$$\mathbf{z}_n^{ws} = (1 - \alpha) \cdot \mathbf{y}_n + \alpha \cdot \mathbf{z}_{n-1}^{mc} \quad (5.11)$$

Die Auswahl wird durch den Gewichtungsfaktor  $\alpha$  gesteuert, der entweder fix oder adaptiv sein kann. Möglichkeiten zur adaptiven Berechnung von  $\alpha$  werden im Abschnitt 5.5 erläutert. Um den Einfluss der gewichteten Auswahl auf den Super-Resolution Algorithmus verstehen zu können, wird eine Grenzfallbetrachtung für  $\alpha$  durchgeführt. Für einen Wert von  $\alpha$  gegen 0 erhält man eine Neu-Initialisierung mit dem aktuellen, niedrig aufgelösten Bild  $\mathbf{y}$ , so dass die falschen Bewegungsvektoren gar keinen Einfluss auf die Lösung nehmen können. In diesem Fall errechnet sich die nächste Lösung durch eine Addition einer rein örtlichen Bildschärfekorrektur und der zusätzlichen Bildverbesserung (Artefaktreduktion) durch das Bildmodell:

$$\mathbf{z}_n = \mathbf{y}_n + (\mathbf{y}_n - \mathbf{b} * \mathbf{y}_n) + \Gamma(\mathbf{y}_n) \quad (5.12)$$

Für den anderen Grenzfall  $\alpha$  gegen 1 erhält man exakt die Vorschrift aus Gleichung 5.9 und somit eine volle Durchsteuerung der Lösung für das bewegungskompensierte vorige Ergebnis. Ein Wert von  $\alpha$  zwischen diesen beiden Extremwerten ermöglicht einen guten Kompromiss aus Detailwiederherstellung und Robustheit. Weiterhin kann durch eine geschickte Wahl von  $\alpha$  der Beitrag des aktuellen Bildes sehr hoch bleiben, so dass immer eine hohe Ähnlichkeit zur Eingangssequenz bestehen bleibt und der Schärfegrad in der Ergebnissequenz nicht zu stark ansteigt.

#### **5.4.2 Verwendung einer Merkmalsanalyse zur Verbesserung der Ablaufsteuerung**

Im Rahmen dieser Arbeit wurden mehrere Verfahren untersucht, um den Ablauf der Super-Resolution besser steuern zu können. Für eine recheneffiziente Realisierung wurden nur solche Maße betrachtet, die auf einfachen Rechenvorschriften basieren und nur eine kleine Umgebung um das aktuelle Pixel benötigen. Für die Berechnung globaler Maße bezogen auf je ein Bild oder die gesamte Sequenz können diese lokalen Maße einfach aufsummiert werden. Die Maße dienen dazu, Seiteneffekte durch falsche Bewegungsinformation zu vermeiden, eine adaptive Anpassung des Beobachtungsmodells an das jeweilige Eingangsmaterial zu realisieren und den Schärfegrad in der Ergebnissequenz zu kontrollieren. Die detaillierte Beschreibung der Maße findet sich in Abschnitt 5.5.

#### **5.4.3 Vereinfachte Darstellung des rekursiven Super-Resolution Algorithmus**

Abbildung 5.11 zeigt ein Blockschaltbild des vereinfachten rekursiven Super-Resolution Algorithmus, der sich nach Einfügen aller beschriebenen Verbesserungsschritte ergibt. Deutlich zu Erkennen ist die zeitliche Rekursionsschleife mit integrierter Bewegungskompensation, die Steuerungsmöglichkeiten für das adaptive Beobachtungsmodell und die Gewichtung der einzelnen Terme zum Gesamtergebnis.

### **5.5 Beschreibung der Super-Resolution als bildinhaltsabhängige Signalverarbeitung**

Der in dieser Arbeit entstandene, rein rekursive Super-Resolution Algorithmus (ESTERA) ist durch das Blockschaltbild 5.11 und die zugehörigen Gleichungen vollständig dargestellt. Für eine formalisierte Betrachtung und weitere Verbesserung kann der Algorithmus durch eine reine Umstellung der Systemgleichungen als bildinhaltsabhängige Signalverarbeitung (MABLE, siehe Abschnitt 3.1.2) beschrieben werden, wobei das Verhalten vollkommen

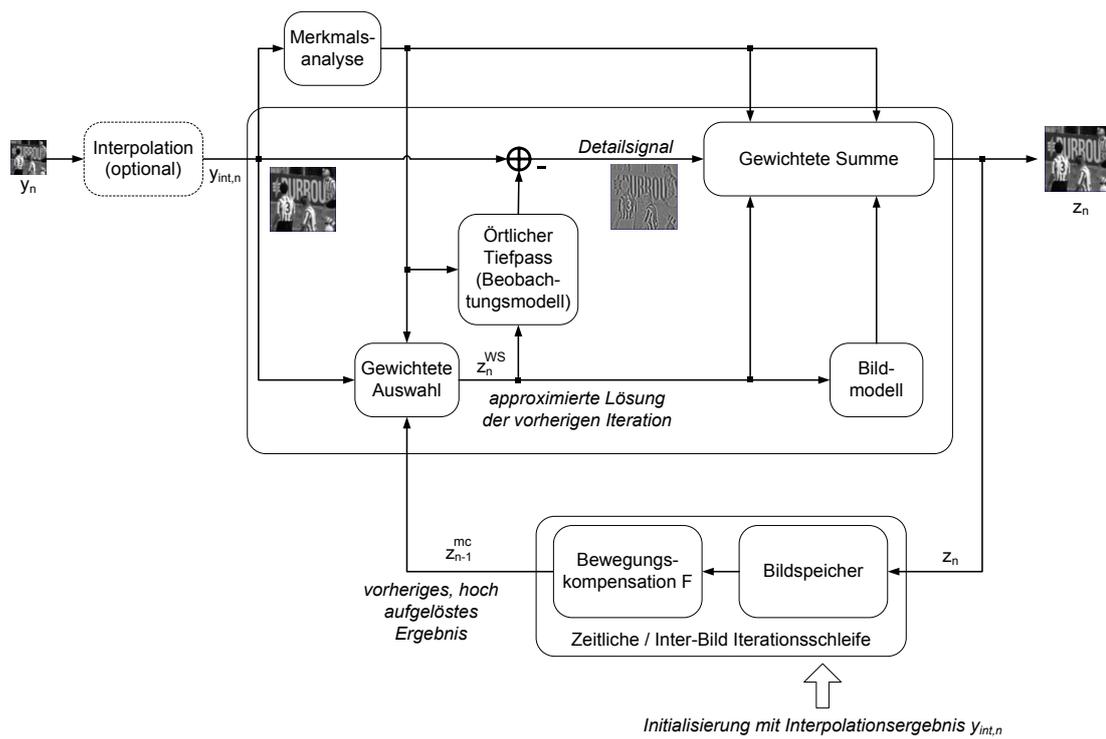


Abbildung 5.11: Vereinfachte Darstellung des rein rekursiven Super-Resolution Algorithmus mit integrierten Verbesserungen

identisch bleibt. Diese alternative Beschreibungsform lässt sich dadurch herleiten, dass man die Gleichung für die gewichtete Auswahl 5.11 in Gleichung 5.10 einsetzt:

$$\mathbf{z}_n = (1 - \alpha) \cdot \mathbf{y}_n + \alpha \cdot \mathbf{z}_{n-1}^{mc} + (\mathbf{y}_n - \mathbf{b} * [(1 - \alpha) \cdot \mathbf{y}_n + \alpha \cdot \mathbf{z}_{n-1}^{mc}]) + \Gamma(\mathbf{z}_n^{ws}) \quad (5.13)$$

Multipliziert man nun alle Terme aus, so erhält man die folgende Vorschrift:

$$\mathbf{z}_n = \mathbf{y}_n - \alpha \cdot \mathbf{y}_n + \alpha \cdot \mathbf{z}_{n-1}^{mc} + \mathbf{y}_n - \mathbf{b} * \mathbf{y}_n + \alpha \cdot \mathbf{b} * \mathbf{y}_n - \alpha \cdot \mathbf{b} * \mathbf{z}_{n-1}^{mc} + \Gamma(\mathbf{z}_n^{ws}) \quad (5.14)$$

Nachdem alle Terme umsortiert sind, entstehen vier funktionale Blöcke, die interpretierbare Bildverarbeitungsalgorithmen abbilden:

$$\mathbf{z}_n = \mathbf{y}_n + c_1 \cdot \left( \underbrace{\mathbf{y}_n - \mathbf{b} * \mathbf{y}_n}_{\text{Unsharp Masking}} + \underbrace{\alpha \cdot (\mathbf{z}_{n-1}^{mc} - \mathbf{y}_n)}_{\text{örtl.-zeitl. Detailsignal}} - \underbrace{\alpha \cdot \mathbf{b} * (\mathbf{z}_{n-1}^{mc} - \mathbf{y}_n)}_{\text{Gegenmassnahme Vektorfehler}} \right) + c_2 \cdot \Gamma(\mathbf{z}_n^{ws}) \quad (5.15)$$

Über Korrespondenzen ergibt sich als Beschreibung im Frequenzbereich:

$$\mathbf{Z}_n = \mathbf{Y}_n + c_1 \cdot (\mathbf{Y}_n - \mathbf{B} \cdot \mathbf{Y}_n + \alpha \cdot (1 - \mathbf{B}) (\mathbf{Z}_{n-1}^{mc} - \mathbf{Y}_n)) + c_2 \cdot F \{ \Gamma(\mathbf{Z}_n^{ws}) \} \quad (5.16)$$

Bei  $F \{ \}$  handelt es sich die Anwendung der Fouriertransformation. Folgende Terme werden zum aktuellen Eingangsbild hinzuaddiert: Bei dem ersten Term handelt es sich um eine rein örtliche Bildschärfekorrektur, die als *Unsharp Masking* bezeichnet wird. Der Term  $(\mathbf{z}_{n-1}^{mc} - \mathbf{y}_n)$  kann als eine örtlich-zeitliche Detailsignaladdition verstanden werden, unter der Annahme, dass das vorherige, bewegungskompensierte Ergebnisbild mehr und andere Details als das aktuelle Eingangsbild enthält. Dieser Term kann aber zusätzlich auch Artefakte enthalten, die durch eine fehlerhafte Bewegungskompensation entstanden sind. Besonders störend und stark sichtbar sind Bewegungsartefakte, die sehr niederfrequent sind und zu Helligkeitsschwankungen führen können. Um diese Störungen zu vermeiden, wird von diesem Detailsignal eine tiefpassgefilterte Version abgezogen. Transformiert man diese beiden das Detailsignal betreffenden Terme in den Frequenzbereich, so erhält man eine alternative Interpretation. Es entsteht eine frequenzselektive Gewichtung des Detailsignals mit  $1 - \mathbf{B}$ , wobei  $\mathbf{B}$  die Fouriertransformierte des Beobachtungsmodells ist. Diese frequenzselektive Gewichtung ist in Abbildung 5.12 für verschiedene Beobachtungsmodelle verdeutlicht. Die Gewichtung ergibt sich direkt aus der mathematischen Herleitung, durch die Verwendung anderer Filterkerne kann aber das Verhalten noch gezielter beeinflusst werden. Abbildung 5.13 zeigt das zur Gleichung 5.15 zugehörige Blockschaltbild in vereinfachter Form. Vergleicht man dieses Blockschaltbild mit dem MABLE-System aus Abbildung 3.3 so wird die Ähnlichkeit zu diesem deutlich. Als Erweiterung sind bei der Super-Resolution noch zusätzliche Analysemaßnahmen und eine Rückführungsschleife integriert. Aufbauend auf dieser Systembeschreibung können nun weitere systembedingte

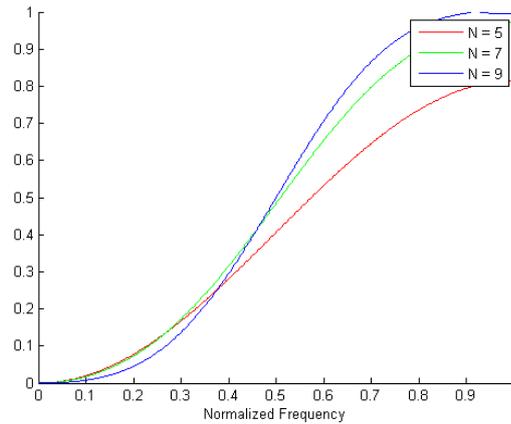


Abbildung 5.12: Frequenzselektive Gewichtung des Detailsignals.  $N$  bezeichnet hier ausnahmsweise einmal abweichend die Länge des Filters in x- bzw. y-Richtung.

Verbesserungen durchgeführt werden. So ist der Austausch einzelner Komponenten durch eine komplexere Verarbeitung sinnvoll, indem z.B. die örtliche Bildschärfekorrektur durch eine synthetische Detailsignaladdition ausgetauscht wird. Dadurch kann eine Bildschärfekorrektur ohne Rauschanhebung realisiert werden. Weiterhin kann die Gegenmaßnahme zur Behandlung von Bewegungsvektorfehlern durch die Verwendung einer hochqualitativen Bewegungsschätzung deutlich geringer gewichtet werden, um den Detailgrad in der Ergebnissequenz weiter zu steigern. Wird nun in Gleichung 5.15 jeder Term mit einem eigenen (adaptiven) Faktor gewichtet und für das Unsharp Masking und die 'Gegenmaßnahme Bewegungsvektorfehler' unterschiedliche Beobachtungsmodelle  $B1$  und  $B2$  verwendet werden, ergibt sich folgende modifizierte Systemgleichung:

$$\mathbf{z}_n = \mathbf{y}_n + c \cdot (\mathbf{y}_n - \mathbf{b}_1 * \mathbf{y}_n) + \alpha_1 \cdot (\mathbf{z}_{n-1}^{mc} - \mathbf{y}_n) + \alpha_2 \cdot \mathbf{b}_2 * (\mathbf{z}_{n-1}^{mc} - \mathbf{y}_n) + \Gamma(\mathbf{z}_n^{ws}) \quad (5.17)$$

Die Verwendung unterschiedlicher Beobachtungsmodelle hat dabei den Hintergrund, dass der Frequenzgang für das rein örtliche Enhancement und für das örtlich-zeitliche Enhancement nun getrennt voneinander eingestellt werden können. So kann das örtliche Enhancement mehr auf mittelfrequente Komponenten konzentriert werden, die in jedem Eingangsbild vorhanden sind, und das örtlich-zeitliche Enhancement auf deutlich höherfrequente Anteile. Das zugehörige Super-Resolution System mit allen Details ist in Abbildung 5.14 zu finden. In diesem Blockschaltbild ist ebenfalls zu sehen, welche Blöcke durch die Merkmalsanalyse gesteuert werden können.

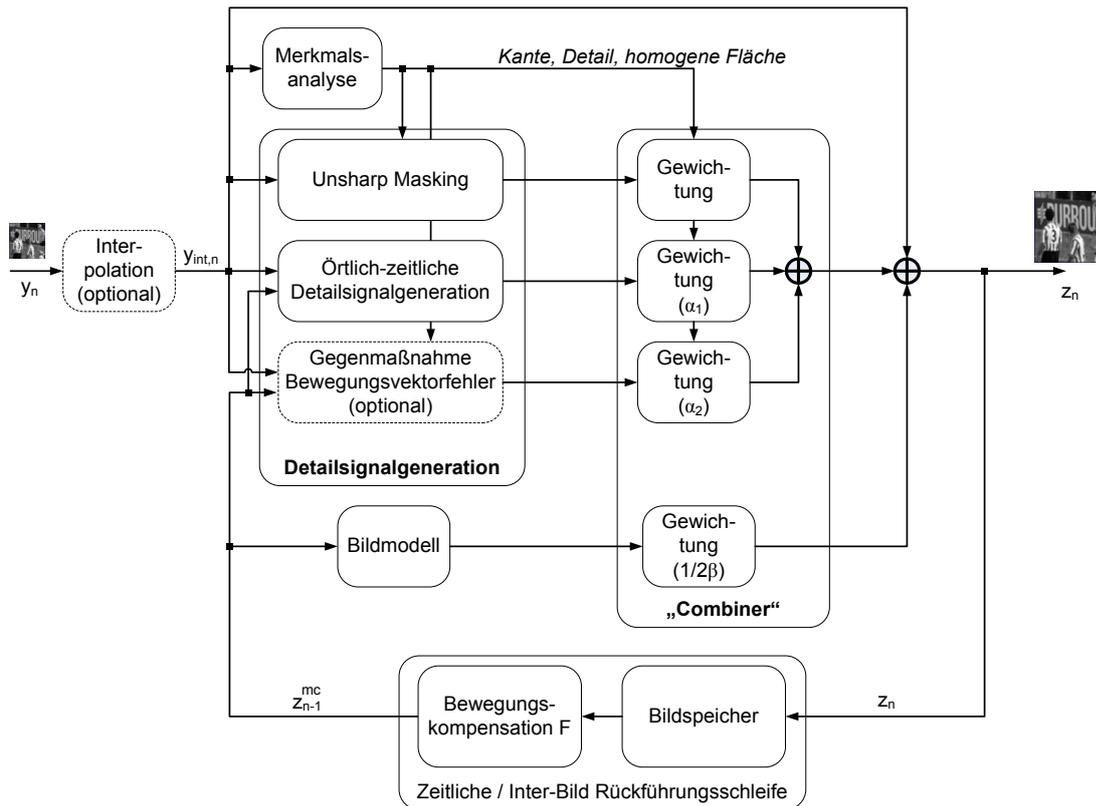


Abbildung 5.13: Vereinfachte schematische Darstellung des Super-Resolution Algorithmus als bildinhaltsabhängige Signalverarbeitung

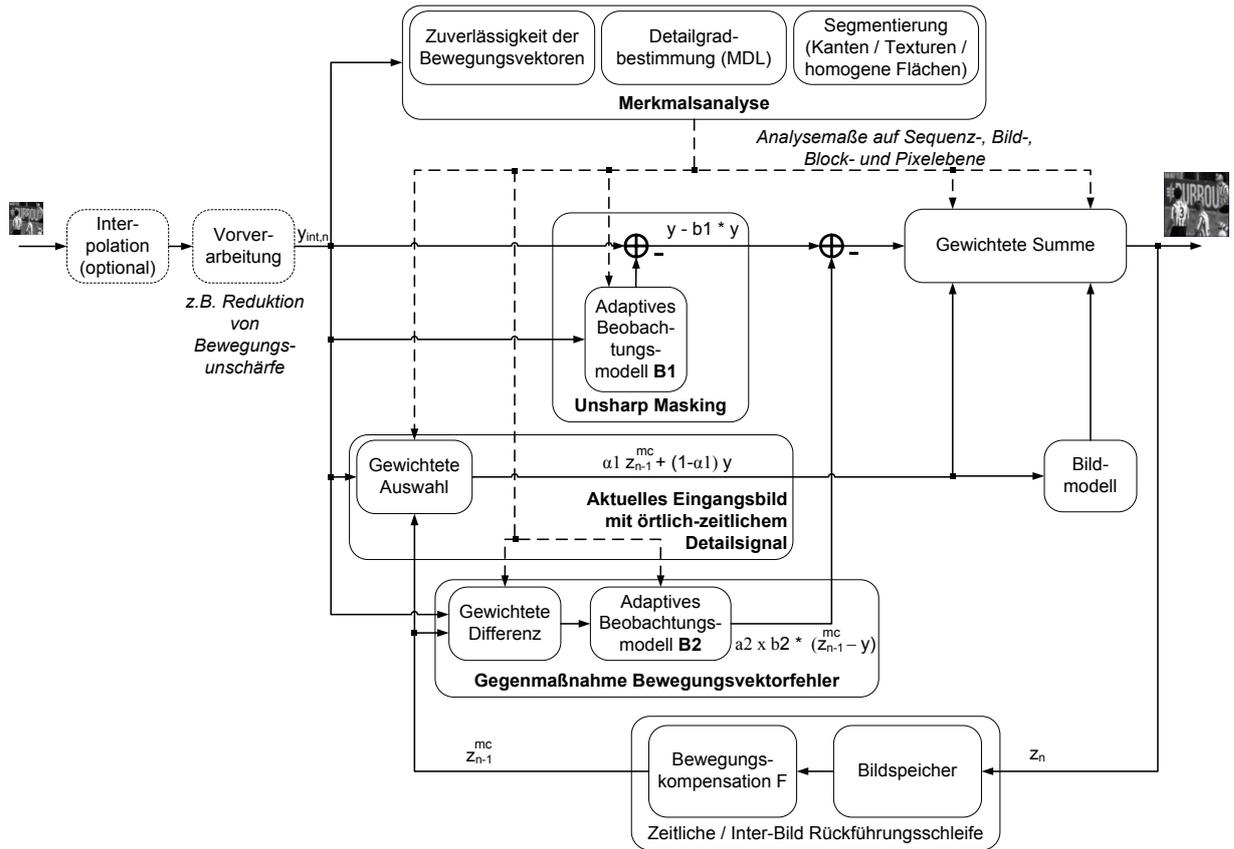


Abbildung 5.14: Blockschaltbild des finalen Systems (ESTERA) mit allen Blöcken

## Verwendete Maße für die Merkmalsanalyse

Zur Steuerung der Super-Resolution werden unterschiedliche Maße verwendet, die in folgende Klassen unterteilt werden können. Zwei Maße werden dazu eingeführt, um die Robustheit gegenüber Bewegungsvektorfehlern zu erhöhen. Dazu zählt eine gewichtete Summe absoluter Differenzen, die zwischen aktuellem Eingangsbild und bewegungskompensierten vorherigen Ausgangsbild berechnet wird und die Bewegungsvektorlänge, da die Geschwindigkeit einen Einfluss auf die Schärfewahrnehmung hat und kleine Bewegungen durch die meist geringere Bewegungsunschärfe und die Trägheit der prädiktiven Schätzverfahren zuverlässiger geschätzt werden können. Des Weiteren wird eine Segmentierung eingesetzt, um in jedem Bildbereich auf die unterschiedlich gute Sichtbarkeit von Störungen reagieren zu können. Weiterhin wird ein Maß zur Bestimmung des Schärfegrades eingeführt, um das Super-Resolution Verfahren optimal auf jede Eingangssituation einstellen zu können. Die Granularität der Maße ist dabei unterschiedlich zu wählen. Während lokal stark variierende Maße bzw. Artefakte wie Bewegungsvektorfehler eine sehr lokal steuerbare Analyse erfordern, wird die adaptive Steuerung des Beobachtungsmodell und die dadurch entstehende Schärfe in der Ergebnissequenz über einen deutlichen größeren Block oder sogar das gesamte Bild gleichmäßig ausgeführt, da die gewünschte Schärfe- und Detailzunahme sehr stark von der Eingangssequenz abhängt. Dadurch kann auf die unterschiedlichen Eigenschaften einer interpolierten Eingangssequenz und einer HD-Eingangssequenz gut reagiert werden. Um aber auch hybrides Material wie z.B. die Einbindung von scharfkantigen Details wie Texten (z.B. Untertitel) in unscharfem Material wie alten Filmen verarbeiten zu können, ist zusätzlich eine leicht variierende, lokale Anpassung sinnvoll.

## Bewegungsvektorlängen abhängige Gewichtung des Detailsignals

Gerade für kleine Geschwindigkeiten ist die Wiederherstellung eines hohen Detailgrades effektiver als für hohe Geschwindigkeiten. Das ist zum einen in den Objektverfolgungseigenschaften des menschlichen visuellen Systems begründet und zum anderen in den Displayeigenschaften. Wird mit der Super-Resolution verarbeitetes Material auf Hold-type Displays angezeigt, so entsteht mit zunehmender Geschwindigkeit ein immer höheres Maß an Bewegungsunschärfe hervorgerufen durch die Haltecharakteristik des Displays, was zu einer erheblich Reduzierung des Detailgrades führen kann, so dass der Mehrertrag dort durch Super-Resolution relativ begrenzt. Weiterhin können gerade für hohe Geschwindigkeiten durch fehlerhafte Kompensation sehr störende Artefakte durch die Super-Resolution entstehen. Aus diesem Grund ist es besser, für höhere Geschwindigkeiten auf die Super-Resolution in dieser Form zu verzichten und stattdessen Verfahren zur Reduktion von Bewegungsunschärfe einzusetzen [Dol10], [RDS<sup>+</sup>08]. Als Steuerparameter ist dazu die Ge-

wichtung des gesamten Detailsignals der Super-Resolution durch den lokalen Bewegungsvektor gut geeignet:

$$g_v = \sqrt{v_x^2 + v_y^2} \quad (5.18)$$

Dieser Gewichtungsfaktor wird dann einer Gewichtungsfunktion unterzogen, um ihn auf einen Bereich von 0..1 abzubilden. Dazu kann eine Gewichtungsfunktion mit linearem oder quadratischem Verlauf eingesetzt werden, wie es in Abbildung 5.17 für eine andere Anwendung gezeigt ist. Die Gewichtung eines Motion Deblurring Algorithmus kann dann mit  $1 - g_v$  erfolgen.

### Berechnung eines adaptiven Faktors für die gewichtete Auswahl

Beim zeitlich rein rekursiven Super-Resolution Algorithmus (ESTERA) kommt dem Gewichtungsfaktor zwischen aktuellem Eingangsbild und bewegungskompensiertem vorherigen Ergebnis eine entscheidende Bedeutung zu, da dieser Faktor einerseits Einfluss auf die Sichtbarkeit von Bewegungsvektorfehlern und andererseits auf die Übertragung von Details aus den vorherigen Bildern in das aktuelle Bild hat. Eine sehr einfache Strategie, um Bewegungsartefakte zu vermeiden und gleichzeitig Details in das aktuelle Bild zu übertragen, ist die Steuerung über die Summe der absoluten Differenzen zwischen aktuellem Eingangsbild und bewegungskompensiertem vorherigen Ergebnisbild:

$$\alpha = \frac{\kappa}{1 + \sum_{a,b} w_{a,b} |\mathbf{y}_{k+a,l+b,n} - \mathbf{z}_{k+a,l+b,n-1}^{mc}|} \quad (5.19)$$

Dieses Maß wird über ein gewichtetes Fenster oder ungewichtetes Fenster  $w_{a,b}$  berechnet, welches beispielsweise eine Ausdehnung von 3x3 Pixeln hat. Dadurch steht für jedes Pixel lokal ein eigener Gewichtungsfaktor zur Verfügung. Durch die Einbeziehung von Nachbarpixeln kann aber lokal eine Glättung dieses Maßes erzielt werden, ohne dass diese Glättung über Objektgrenzen hinaus zu stark wird. Da die Differenz im Gegensatz zum SAD-Maß aus der Bewegungsschätzung zwischen einem Eingangs- und Ausgangsbild berechnet wird, deutet eine hohe Differenz auf unterschiedliche Effekte hin. Zum einen kann ein hoher Wert auf einen Bewegungsvektorfehler hindeuten, zum anderen entsteht eine hohe Differenz, wenn das vorherige Ergebnisbild deutlich schärfer als das aktuelle Eingangsbild ist. Aber auch in diesem Fall ist eine Herunterwichtung des vorigen Bildes sinnvoll, um Überschärfe zu vermeiden. Mit  $\kappa$  ist eine Einstellung der Übertragung neuer Details in das aktuelle Bild und die Robustheit gegenüber Bewegungsvektorfehlern möglich, wobei beides nicht gleichzeitig erreicht werden kann und somit ein Kompromiss zu finden ist. Um diesen Kompromiss zu umgehen, kann anstelle eines festen Gewichtungsfaktors  $\kappa$  ein zum Bildinhalt adaptiver Faktor gewählt werden:

$$\alpha = \frac{\kappa_{fix} + \kappa_{adapt}}{1 + \sum_{a,b} w_{a,b} |\mathbf{y}_{k+a,l+b,n} - \mathbf{z}_{k+a,l+b,n-1}^{mc}|} \quad (5.20)$$

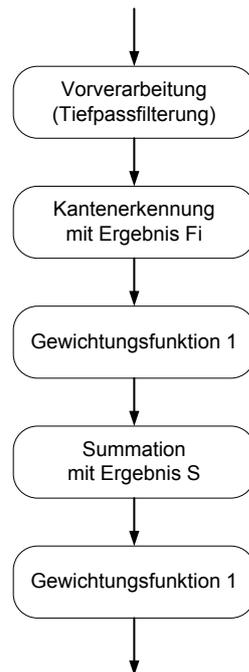


Abbildung 5.15: Ablauf der Segmentierung zur Erkennung homogener Flächen und Detailbereichen

In dieser Gleichung bezeichnet  $\kappa_{fix}$  einen festen Faktor, der unabhängig vom Bildinhalt ist und eine gewisse Abweichung zwischen aktuellem Eingangs- und vorigem Ergebnisbild erlaubt, welcher beispielsweise durch Rauschen hervorgehoben wird.  $\kappa_{adapt}$  ist ein adaptiver Faktor, der davon abhängig ist, ob die Berechnung des Faktors für einen homogenen Bereich, einen Detailbereich oder einen Kantenbereich erfolgt. So ist für einen homogenen Bereich generell eine geringere Abweichung zu erwarten, da in diesem Bereich kaum Aktivität herrscht, während in Kanten- oder Detailbereichen selbst für eine höhere Abweichung der Bewegungsvektor noch korrekt sein kann. Diese Berechnungsvorschrift ist ein Beispiel dafür, dass die Analysemaße nicht nur zur Steuerung der Bildsignale verwendet werden, sondern auch zur Korrektur der Analysemaße. Um auch für eine sehr hohe Differenz noch eine zeitliche Gewichtung gewährleisten zu können, kann außerdem noch der Faktor  $\alpha$  auf einen minimalen Wert  $\alpha_{min}$  begrenzt werden, um in jedem Fall eine Detailübertragung vom vorigen ins aktuelle Bild zu gewährleisten.

## Segmentierung

Die Verwendung eines bildinhaltsabhängigen Verarbeitungskonzeptes bei der Super-Resolution bietet genau wie in anderen Bereichen der Videosignalverarbeitung viele Vor-

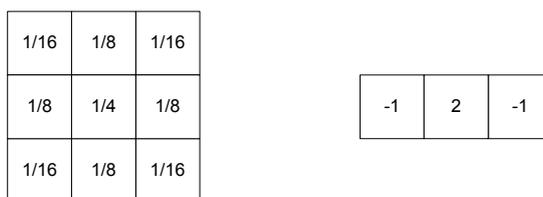


Abbildung 5.16: Filtermasken für die Segmentierung. Tiefpassfilter zur Vorverarbeitung (links) und Hochpassmaske zur Detailerkennung (rechts)

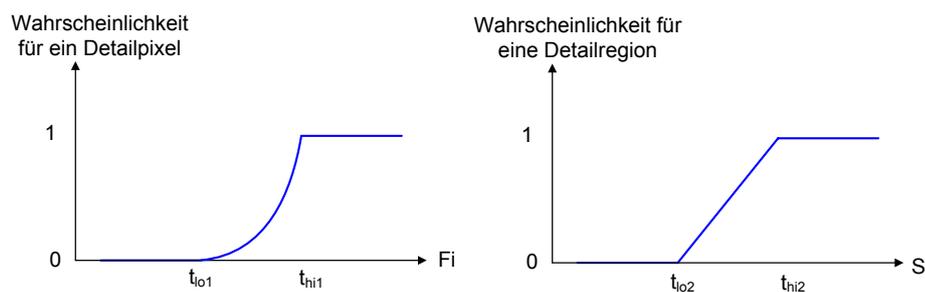


Abbildung 5.17: Lineare und quadratische Gewichtungsfunktion

teile. So können beispielsweise Artefakte deutlich besser reduziert oder zumindest maskiert werden. Diese sind nämlich in homogenen Flächen sehr gut sichtbar, in texturierten Bereiche fallen sie aber durch die vielen Details kaum auf. Für die Super-Resolution bedeutet dies, in homogenen Bereichen ohne Details nur eine sehr schwache Gewichtung des Korrektursignals vorzunehmen und den Algorithmus nur für die Reduktion von Störungen zu verwenden. Weiterhin können Kanten bereits sehr gut durch rein örtliche Verfahren verbessert werden. Zu starke Überschwinger müssen dort kontrolliert werden, da diese als sehr störend wahrgenommen werden können (siehe auch [SB00]). In texturierten Bereichen können die stärksten Gewinne durch die örtliche-zeitliche Super-Resolution erzielt werden, da diese Bereiche durch rein örtliche Verfahren nur bedingt verbessert werden können, so dass das Super-Resolution Ergebnis hier sehr stark zu wichten ist. Ein Verfahren, welches das Eingangsbild in homogene Flächen und Detailbereiche segmentiert, ist in Abbildung 5.15 gezeigt. Das grundlegende Vorgehen bei der Segmentierung hier ist dem Verfahren aus [Dol10] ähnlich. Zuerst wird eine Vorverarbeitung durchgeführt, um Rauschen zu reduzieren und somit die Zuverlässigkeit der Detektion erhöhen zu können. Daran anschließend findet eine Kantenerkennung statt. Abbildung 5.16 zeigt sowohl eine mögliche Maske für die Vorverarbeitung als auch für die Kantenerkennung. Sollen texturierte Bereiche erkannt werden, wird die zweite Ableitung für die Detail- bzw. Kantenerkennung verwendet, wobei die zweite Ableitung getrennt in x- und y-Richtung gebildet wird und anschließend eine pixelweise Betragsbildung erfolgt. Das Ergebnis dieser Betragsbildung,

bezeichnet mit  $F_i$ , wird dann für jedes Pixel einzeln einer Gewichtungsfunktion unterzogen, mit deren Hilfe die Wahrscheinlichkeit für ein mögliches Detailpixel ermittelt wird. Beispiele für Gewichtungsfunktionen sind in Abbildung 5.17 dargestellt, wobei sowohl ein linearer als auch quadratischer Verlauf möglich ist. Nachdem nun für jeden Pixelwert eine zugehörige Wahrscheinlichkeit ermittelt wurde, werden alle Wahrscheinlichkeiten über eine Fläche von  $5 \times 5$  Pixeln aufsummiert (Ergebnis ist die Summe  $S$ ), um für den jeweiligen Bereich eine Entscheidung über die Zugehörigkeit zu einer Region zu treffen. Auch hier wird keine harte Entscheidung getroffen, sondern ein Wahrscheinlichkeitswert bestimmt, ob es sich um einen Detailbereich oder eine homogenen Fläche handelt. Um zu starke Übergänge zwischen den segmentierten Bereichen zu vermeiden, wird nicht allen Pixeln im  $5 \times 5$  großen Analysebereich dieser Wert zugeordnet, sondern nur dem zentralen Pixel selbst oder alternativ, um Rechenzeit zu sparen, der inneren  $3 \times 3$  Umgebung. Für die anderen Pixel muss dieser Vorgang wiederholt werden, wobei diese nun im Zentrum der Analyseregion liegen. Nachdem nun eine Entscheidung zwischen einer homogenen Fläche und einem Detailbereich getroffen wurde, kann danach noch eine Entscheidung über eine Zugehörigkeit zu einem Kantenbereich getroffen werden. Diese Entscheidung läuft sehr ähnlich ab wie die Detailerkennung, nur dass ein anderer Kantendetektor (Gradientenoperator anstelle des Laplace-Operators) verwendet wird und höhere Schwellwerte für die Gewichtungsfunktionen verwendet werden.

Der Einsatzzweck für die so entstehende Segmentierungsmaske innerhalb der Super-Resolution ist die Steuerung der gewichteten Selektion in Abhängigkeit vom Bildinhalt (über  $\kappa$ ) sowie die unterschiedlich starke Gewichtung der Detailsignale der einzelnen Algorithmenblöcke zur Bestimmung des Gesamtsignals. So wird in homogenen Flächen zwecks Rauschreduktion des Ergebnis der gewichteten Selektion deutlich stärker gewichtet als das rein örtliche Enhancement. In Kantenbereichen hingegen wird das Detailsignal bestehend aus örtlichem und örtlich-zeitlichem Enhancement insgesamt sehr stark reduziert, um zu starke Überschwinger zu vermeiden.

### Schätzung der minimalen gauß'schen Varianz

Der bisher abgeleitete Super-Resolution Algorithmus basiert direkt auf dem Beobachtungsmodell und dem Bildmodell (siehe Kapitel 2). Sind diese beiden Modelle erfüllt, z.B. wenn eingangs eine Interpolation durchgeführt wird, so resultiert daraus eine Detailwiederherstellung in der Ergebnissequenz mit gut abgestimmten Schärfegrad. Liegt am Eingang aber bereits sehr detailreiches oder scharfes Material vor, so liefert der Super-Resolution Algorithmus unnatürliche und überscharfe Ergebnisse. Um diesen Seiteneffekt zu verhindern, wird nun ein Verfahren vorgestellt, welches den Schärfegrad der Eingangssequenz schätzen kann und damit eine adaptive Anpassung der internen Modelle möglich macht. Dieses Verfahren basiert auf der in [Est00] beschriebenen Strategie, einen Filterkern zu bestimmen, der den Kompromiss zwischen Detailerhalt und Störungsreduktion am besten

erfüllt. Dieses Maß, entwickelt für die Störungsreduktion, ist die so genannte Description Length (DL), welche den im Quellmaterial enthaltenen Informationsgrad in Bit angibt. Die Description Length errechnet sich folgendermaßen:

$$DL(\sigma) = \frac{\lambda_{DL}}{\sigma} + \sum_{a,b} (\mathbf{y} - \mathbf{b}_\sigma * \mathbf{y})^2 \quad (5.21)$$

Das Ergebnis besteht aus zwei Termen, die für jede Situation ein optimales Gleichgewicht bilden. Der erste Term ist eine varianzabhängige Bestrafung, die für größere Varianzen, also Filterkerne für eine stärkere Tiefpassfilterung, zu einer geringeren Bestrafung der Description Length führen. Der zweite Term ist eine Bestrafung der Energie der hochfrequenten Anteile, die durch die Filterung entfernt werden. Eine leichtere Filterung führt dabei zu einer geringen Bestrafung. Minimiert man diese Gleichung, so findet man den optimalen Filterkern für die jeweilige Bildsituation. In Abbildung 5.18 ist schematisch der Verlauf der Description Length für verschiedene Detailgrade gezeigt. Anstelle nun für jeden Bildpunkt eine Minimierung vorzunehmen und die MDL zu ermitteln, wird die Verwendung von drei festen Filterkernen vorgeschlagen, da so ein guter Kompromiss zwischen Aufwand und Nutzen erzielt werden kann. Dieses Vorgehen ist bereits in [Brü09] für eine bildinhaltsabhängige Regularisierung angewendet worden. Dazu ist jeder Bildausschnitt mit drei verschiedenen Filterkernen zu verarbeiten und die dabei wegfallende Hochpassenergie zu messen. Abbildung 5.19 zeigt für verschiedene Bereiche den jeweils optimalen Filterkern. Dieses Vorgehen kann direkt auf die Super-Resolution übertragen werden. Dabei wird für die Bestimmung der MDL das (evtl. hoch getastete) Eingangsmaterial verwendet und so für jedes Pixel die optimale Filtermaske bestimmt. Im Abbildungsmodell der Super-Resolution (Unschärfefilter) wird dann die so bestimmte Filtermaske verwendet. Dies führt zu einer deutlich besser angepassten Verarbeitung sowohl für sehr niedrig aufgelöstes als auch für hoch aufgelöstes Eingangsmaterial.

### Kontrastbestimmung und -anpassung

In vielen Fällen ist es wünschenswert, die Schärfe in der Ausgangssequenz gezielt durch nur einen Parameter steuern zu können. Da die SR-Algorithmen bzw. deren interne Analyseblöcke alle von ihren eigenen Parametern und Gewichtungskurven abhängig sind, wurde zusätzlich noch eine Kontrastmessung implementiert, die eine globale Gewichtung (bezogen auf jedes Bild) des finalen Detailsignals ermöglicht. Dadurch ist es möglich, alle anderen Gewichtungen global zu kontrollieren. Dadurch kann der Schärfegrad entweder herabgesenkt werden, wodurch selbst für ein divergierendes System noch Stabilität nachträglich gewährleistet werden kann oder der Schärfegrad nachträglich erhöht werden, wenn die Verstärkung zu gering ausgefallen ist. Der Aufbau dieser Kontrastmessung ist in Abbildung 5.20 zu sehen. Sowohl das aktuelle niedrig aufgelöste und interpolierte Eingangsbild

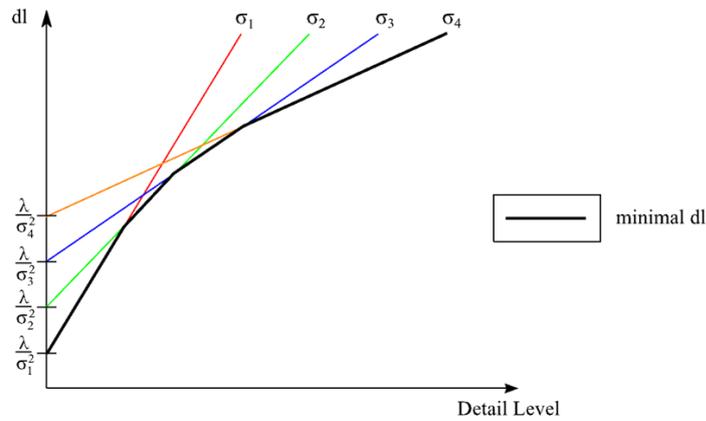


Abbildung 5.18: Beispielhafter Verlauf der Varianzschätzung mittels Minimal Description Length (vergl. [Brü09])

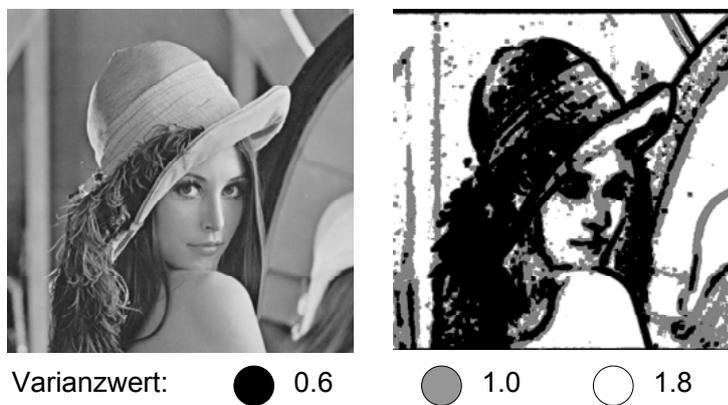


Abbildung 5.19: Ergebnis der Varianzschätzung mittels Minimal Description Length für unterschiedliche Bildbereiche

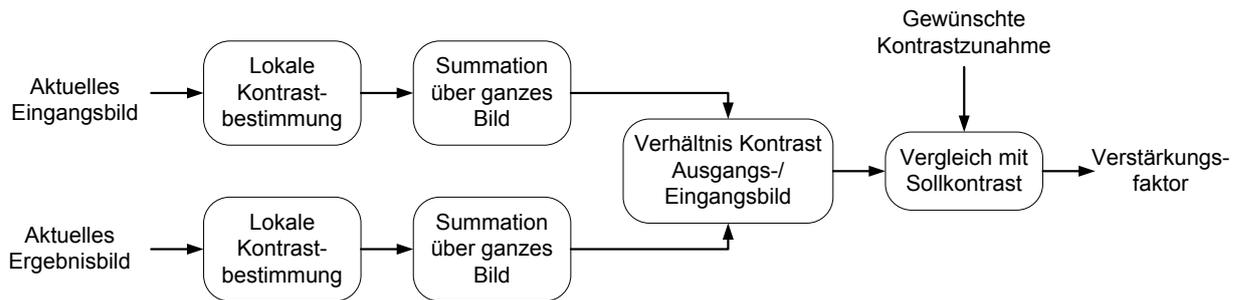


Abbildung 5.20: Blockschaltbild für die Kontrastmessung

als auch das hoch aufgelöste Ergebnisbild werden einer lokalen Kontrastmessung unterzogen. Dazu wird in einer  $3 \times 3$  Umgebung um das aktuelle Pixel das Minimum und das Maximum über alle beteiligten Pixelwerte bestimmt, und zwar für das Eingangsbild und das Ergebnisbild getrennt. Der lokale Kontrastwert ergibt sich dann in beiden Fällen aus der Differenz zwischen Maximum und Minimum. Die sich so ergebenden Werte werden für die gesamte Messumgebung getrennt aufsummiert. Um die Messung nicht zu stark zu verfälschen, wird diese nur auf Bereiche starker Texturen und Kanten beschränkt. Zu diesem Zweck wird auf die Ergebnisse der Detail- und Kantenmessung zurückgegriffen und nur Pixel verwendet, die in einer der beiden Karten einen Wert über 0.7 aufweisen. Nachdem nun beide Kontrastwerte für das Eingangs- und Ausgangsbild bestimmt sind, wird durch einfache Division das Kontrastverhältnis ermittelt. Dieser Wert wird nun mit einem vorgegebenen Zielkontrastwert verglichen und darauf basierend durch eine Division von aktuellem und Sollkontrastverhältnis ein globaler Verstärkungsfaktor ermittelt. Um eine zu starke Anhebung oder Dämpfung des Detailsignals zu vermeiden, wird der Verstärkungsfaktor auf einen sinnvollen Bereich (z.B. 0.2 bis 2) beschränkt.

### 5.5.1 Einsatz der Analysemaße innerhalb der Super-Resolution

Dieser Abschnitt soll verdeutlichen, welche Analysemaße auf welche Parameter Einfluss nehmen und welcher Effekt dadurch erzielt werden kann. Tabelle 5.1 fasst diese Sachverhalte zusammen. Die Parameter beziehen sich dabei auf die finale Super-Resolution Gleichung 5.17. Durch diese Maße kann eine erhebliche Steigerung der Bildqualität erzielt und die Robustheit gegenüber Seiteneffekten deutlich gesteigert werden. Daher soll hier auch beispielhaft auf die Grenzen der Steuerung eingegangen werden. Die Maße wurden vor allem in Hinblick auf ihre Recheneffizienz ausgewählt und sind deshalb als Konzept zu verstehen, wie ein iterativer Rekonstruktionsalgorithmus deutlich besser an die Bedürfnisse der Videosignalverarbeitung angepasst werden kann. Komplexere Maße, die z.B. im Frequenzbereich arbeiten oder für ihre Analyse den gesamten Bildinhalt zur Verfügung haben, können mit Sicherheit die Qualität noch weiter steigern. Ein weiterer wichtiger Analyseschritt wäre

Tabelle 5.1: Parameter der Super-Resolution und ihre Steuerung durch Analysemaße

Analysemaß	Einfluss auf folgende Parameter	Effekte
Homogenitätsmaske	starkes Heruntersetzen von $c$ für homogene Flächen	Algorithmus fungiert in homogenen Flächen als kombinierte örtlich-zeitliche Rauschreduktion
Kantenmaske	Herunterwichtung von $c, \alpha_1$ und $\alpha_2$ im Kantenfall	Deutlich geringere Überschwinger an Kanten führen zu angenehmen Bildeindruck
Geschwindigkeitabhängige Gewichtung $g_v$	Herunterwichtung von $\alpha_1$ und $\alpha_2$ zu höheren Geschwindigkeiten	Weniger Artefakte durch fehlerhafte Bewegungsvektoren für hohe Geschwindigkeiten
Kombinierte (Kanten / Textur / homogene Flächen)-Maske	Steuerung von $\kappa_{adapt}$ innerhalb der gewichteten Selektion	Höherer Detailgrad bei geringerer Sichtbarkeit von Artefakten im Ergebnisbild
Gewichtete Auswahl	$\alpha_1$ und $\alpha_2$	Automatische Gewährleistung der Stabilität des Algorithmus, Vermeidung von Überscharfe und Sichtbarkeit von Vektorfehlern
Minimal Description Length	Adaptive Steuerung von $B1$	Natürlichere Ergebnisse für unterschiedlichste Bildmaterialien
Kontrastabhängige Gewichtung	Heruntergewichtung aller Detailsignale	Anpassung des Ergebnissen an einen gewünschten Kontrastwert

eine intelligente Unterscheidung von Bewegungsvektorfehlern und neuer Details. Eine Literaturrecherche zur Erkennung von Bewegungsvektorfehlern hat aber gezeigt, dass es in diesem Bereich derzeit keine zuverlässigen Maße gibt, um diese Unterscheidung gut vorzunehmen. Aus diesem Grund wurde auf die Entwicklung eines entsprechenden Maßes verzichtet und statt dessen eine bestehende Bewegungsschätzung erheblich verbessert (siehe Abschnitt 5.1.5), so dass insgesamt sehr zuverlässige Vektoren vorliegen und deshalb einfachere Maße innerhalb des Algorithmus zur Erkennung von Vektorfehlern ausreichen.

## 5.6 Konvergenzbetrachtungen

Bei allen iterativen und rekursiven Algorithmen mit Rückführungsschleife ist eine Stabilitäts- bzw. Konvergenzbetrachtung nötig, um die korrekte Funktionsweise zu gewährleisten und zu scharfe Bilder und 'Clipping' auf die Minimal- bzw. Maximalwerte des Bildmaterials möglichst zu vermeiden. Diese Konvergenzbetrachtungen sind im vorgeschlagenen Entwicklungskonzept unter Schritt 4 berücksichtigt und sollen nun für alle Super-Resolution Algorithmen nachgeholt werden. Im Allgemeinen kann Stabilität auf mehrere Arten überprüft werden. So ist es möglich, die dem Algorithmus zugrunde liegende Gleichung als ein rekursives Filter zu schreiben und die Stabilität dann durch die Anwendung der z-Transformation und einer anschließende Pol-/Nullstellenbetrachtung zu zeigen (für Details siehe [Sch00]). In [WS04] wird ein alternativer Weg durch Überführung der Gleichung in eine geometrische Reihe beschrieben, der zu äquivalenten Ergebnissen führt. Der Vorteil dieser Vorgehensweise ist, dass damit das schrittweise Vorgehen der einzelnen Algorithmen besser gezeigt werden kann und somit neben dem Gesamtergebnis auch die Einflüsse jeder einzelnen Iteration gezeigt werden kann.

### 5.6.1 Einzelbild Back-Projection Super-Resolution

Um die Vorgehensweise bei der Konvergenzbetrachtung zu verdeutlichen, wird hier zuerst auf ein sehr einfaches iteratives Verfahren eingegangen, die Einzelbild Back-Projektion Super-Resolution mit einem bildweise arbeitenden Iterationsschema. Die Gleichung wird hier direkt im Transformationsbereich angegeben, so dass die Faltung im Originalbereich mit einer Multiplikation im Transformationsbereich korrespondiert:

$$\begin{aligned} Z_{p+1} &= Z_p + (Y - H \cdot Z_p) \\ \Leftrightarrow Z_{p+1} &= Y + (1 - H) \cdot Z_p \end{aligned} \tag{5.22}$$

$H$  bezeichnet dabei einen allgemeinen Degradationsoperator, der je nach Einsatzzweck ein Unschärfefilter, eine Bewegungskompensation und / oder Auf- und Abwärtstatstungen beschreiben kann. Um nun die Konvergenz zeigen zu können, wird die Strategie verfolgt,

die Lösung zu jedem Iterationsschritt  $p$  in Abhängigkeit von der initialen Lösung  $Z_0 = Y$  darzustellen, wobei  $Y$  die Fouriertransformierte des Eingangsbild darstellt:

$$\begin{aligned} Z_1 &= Y + (1 - H)Y \\ Z_2 &= Y + (1 - H) \cdot Y + (1 - H)^2 Y \\ &\vdots \\ Z_p &= \sum_{r=0}^p (1 - H)^r \cdot Y \end{aligned} \quad (5.23)$$

Da das Eingangsbild  $Y$  bzw. dessen Fouriertransformierte unabhängig von der Iterationsvariable  $p$  ist, kann diese vorgezogen werden. Die Summe  $\sum_{r=0}^p (1 - H)^r$  entspricht dabei exakt der geometrischen Reihe (siehe [BSMM01]). Diese Reihe konvergiert gegen den in der folgenden Gleichung angegebenen Wert:

$$\lim_{p \rightarrow \infty} \sum_{i=0}^p (1 - H)^i = \frac{1}{1 - (1 - H)} \quad \text{for } \|1 - H\| < 1 \quad (5.24)$$

Wobei  $(1 - H)$  im Betrag für die Konvergenz kleiner eins sein muss. Damit ist es möglich, für den Fall der Konvergenz den iterativen (rekursiven) Filter durch ein direktes Filter ohne Iterationsschema zu beschreiben. Abbildung 5.21 verdeutlicht die Entwicklung des Filters für jede Iteration und den in Summe resultierenden Gesamtfiler. Wie in dieser Abbildung zu sehen ist, konvergiert der Algorithmus für niedrige Frequenzen deutlich schneller als für hohe Frequenzen. Dieser Sachverhalt muss für den zeitlich rekursiven Algorithmus berücksichtigt werden, da für die Generation höherer Frequenzen über einen langen Zeitraum sehr präzise Bewegungsvektoren vorliegen müssen. Obwohl der iterative Algorithmus durch einen direkten, nicht iterativen Algorithmus ersetzt werden kann, hat die iterative Verarbeitung Vorteile. Durch die Rückführung mittels einer Schleife wird so die Bildqualität am Ausgang zurückgeführt und in einem 'Regelkreislauf' die Bildqualität schrittweise optimiert. Dies ermöglicht beispielsweise die Kontrolle des Rauschens im Ausgangsbild oder eine gezielte Einstellung der Bildschärfe, die durch eine nicht iterative Verarbeitung nicht oder nur sehr schwer möglich ist, da die Filterung nur einmal ausgeführt wird [WS04]. Durch die Einführung zusätzlicher Kontrollmechanismen, die durch eine einfache Gleichung wie oben skizziert nicht unmittelbar zu beschreiben sind, können die Vorteile der iterativen Verarbeitung noch gesteigert werden, da die Kontrolle und Analyse pro Iteration noch besser gesteuert werden kann.

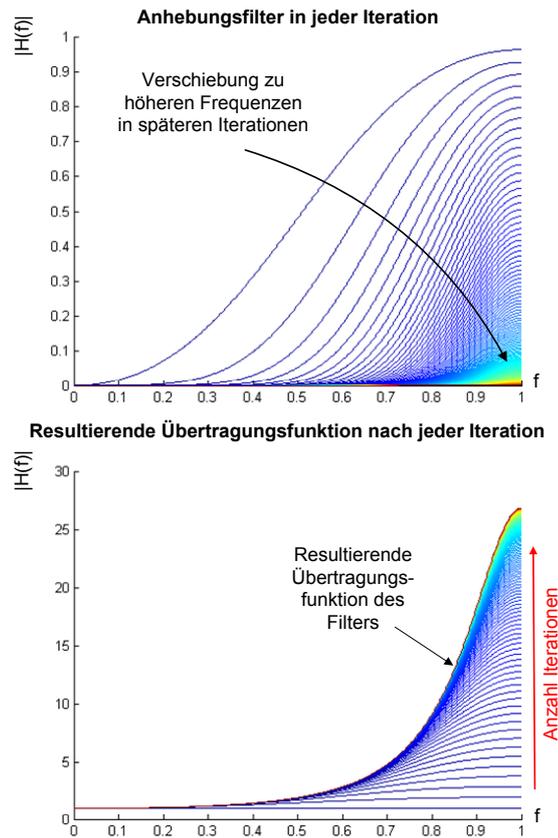


Abbildung 5.21: Filterübertragungsfunktion bei der Super-Resolution pro Iteration (oben) und resultierende, direkte Gesamtübertragungsfunktion (unten)

### 5.6.2 Mehrreferenzbild Back-Projection Super-Resolution

Im Vergleich zur Einzelbild Super-Resolution findet bei der Mehrreferenzbild Super-Resolution eine Wiederherstellung von Details statt, die nicht in jedem Eingangsbild vorhanden sind, was auch mathematisch gezeigt werden kann. In der Ausgangsgleichung wird der Beitrag von insgesamt  $n$  Eingangsbildern durch eine Summe beschrieben, so dass sich für jedes Eingangsbild ein entsprechendes Korrektursignal pro Iterationsschritt ergibt:

$$Z_{p+1} = Z_p + \sum_n (Y_n - H \cdot Z_p) \quad (5.25)$$

Wird die Summenoperation auf beide Summanden aufgeteilt und weiterhin berücksichtigt, dass der Degradationsoperator  $H$  und das Ergebnisbild  $Z_p$  nicht vom Zeitindex  $n$  der Eingangsbilder abhängig sind, kann die vorige Gleichung folgendermaßen umgeschrieben werden:

$$Z_{p+1} = Y_* + (1 - H_*) Z_p \quad (5.26)$$

Durch die Einführung von  $Y_*$  und  $H_*$  hat diese Gleichung ein identisches Schema wie Gleichung 5.22, wobei für die beiden Variablen  $Y_*$  und  $H_*$  gilt:

$$Y_* = \sum_n Y_n \text{ und } H_* = n \cdot H \quad (5.27)$$

$Y_*$  entspricht also einer zeitlichen Mittelwertfilterung, die normalerweise bewegungskompensiert erfolgt. Dadurch enthält  $Y_*$  weniger Rauschen bei gleicher Schärfe und auch mehr Details, sofern die Bilder einen unterschiedlichen Grad an Informationen an derselben Bildposition tragen. Führt man nun wie in Gleichung 5.23 gezeigt die aktuelle Iteration  $p$  auf die initiale Lösung (hier  $Y_*$ ) zurück, ergibt sich analog:

$$Z_p = \sum_{r=0}^{p-1} (1 - n \cdot H)^r \cdot Y_* + (1 - n \cdot H)^p \cdot Y \quad (5.28)$$

Wie nun ersichtlich ist, besteht das Ergebnis aus zwei Termen. Das Ergebnis der (bewegungskompensierten) zeitlichen Mittelwertfilterung wird dabei iterativ nur bis zu der Filtervorschrift verarbeitet, die der aktuellen Iteration einen Schritt zurück ist. Der Filter mit der höchsten Ordnung, der mit dem aktuellen Iterationsschritt korrespondiert, wird nur auf das aktuelle Eingangsbild  $Y$  angewendet. Konvergenz wird für den Fall  $|1 - n \cdot H| < 1$  erzielt, was in anderer Schreibweise

$$n \cdot H > 0 \text{ und } n \cdot H < 2 \quad (5.29)$$

bedeutet. Da  $H$  normalerweise 1 im Maximalfall erreicht, divergiert die Lösung ab  $n = 2$ . Das kann verhindert werden, indem eine geringere Iterationsschrittweite  $\tau$  gewählt wird, so

dass das Detail- bzw. Korrektursignal, welches in jeder Iteration addiert wird, schwächer gewichtet wird:

$$Z_{p+1} = Z_p + \tau \cdot \sum_n (Y_n - H \cdot Z_p) \quad (5.30)$$

Abhängig von  $\tau$  kann dann Konvergenz unter folgenden Bedingungen gewährleistet werden:

$$\tau \cdot n \cdot H > 0 \text{ and } \tau \cdot n \cdot H < 2 \quad (5.31)$$

### 5.6.3 Rekursive Super-Resolution

Auch für die (zeitlich) rekursive Super-Resolution (z.B. ESTERA) kann Konvergenz gezeigt werden, sofern einige Annahmen getroffen werden, die im Laufe der Herleitung beschrieben werden. Basis für die Konvergenzkontrolle ist Gleichung 5.10, wobei direkt die Darstellung im Frequenzbereich gewählt wurde und die beiden Terme für  $Z_{ws}$  zusammengefasst wurden:

$$\begin{aligned} Z_n &= Y_n + (1 - H) \cdot Z_{ws} \Leftrightarrow \\ Z_n &= Y_n + (1 - H) \cdot [(1 - \alpha_n) \cdot Y_n + \alpha_n \cdot F_n \cdot Z_{n-1}] \end{aligned} \quad (5.32)$$

In der unteren Gleichung sind für  $Z_{ws}$  die beiden Terme gemäß der gewichteten Auswahl eingesetzt worden, wobei die Bewegungskompensation durch die Matrix  $F_n$  beschrieben wird. Durch Ausmultiplizieren können die Beträge der vorherigen Iteration und dem aktuellen Eingangsbild getrennt werden:

$$Z_n = Y_n \cdot [1 + (1 - H)(1 - \alpha_n)] + \alpha_n \cdot F_n \cdot Z_{n-1} (1 - H) \quad (5.33)$$

Nun wird wieder die akute Iteration zum Zeitpunkt  $n$  in Abhängigkeit der Eingangsbilder ausgedrückt. Das Iterationsschema entspricht diesmal einem zeitlich rekursiven Einsetzen aller Eingangsbilder bis zum aktuellen Zeitpunkt. Da für das erste Eingangsbild kein Vorgänger existiert, ergibt sich so für die erste Lösung durch setzen von  $Z_{-1} = 0$ :

$$Z_0 = Y_0 \cdot [1 + (1 - H)(1 - \alpha_0)] \quad (5.34)$$

Die Lösung entspricht also einer rein örtlichen Bildschärfekorrektur durch Unsharp Masking. Alle weiteren Lösungen ergeben sich nun durch schrittweises Einsetzen dieser Gleichung in Gleichung 5.33:

$$\begin{aligned} Z_1 &= [1 + (1 - H)(1 - \alpha_1)] Y_1 && + \\ &\quad \alpha_1 \cdot F_1 \cdot [(1 - H) + (1 - H)^2 (1 - \alpha_0)] Y_0 \\ Z_2 &= [1 + (1 - H)(1 - \alpha_2)] Y_2 && + \\ &\quad \alpha_2 \cdot F_2 \cdot [(1 - H) + (1 - H)^2 (1 - \alpha_1)] Y_1 && + \\ &\quad \alpha_2 \cdot \alpha_1 \cdot F_2 \cdot F_1 \cdot [(1 - H)^2 + (1 - H)^3 (1 - \alpha_0)] Y_0 \\ &\quad \vdots && \end{aligned} \quad (5.35)$$

Wird nun berücksichtigt, dass  $(1 - H)^0 = 1$  gilt, lässt sich das Ergebnis zum Zeitpunkt  $n$  folgendermaßen ausdrücken:

$$Z_n = \sum_{m=0}^n F_{n-m}^s \cdot \alpha_{n-m}^s \cdot [(1 - H)^{n-m} + (1 - H)^{n-m+1} \cdot (1 - \alpha_m)] \cdot Y_m \quad (5.36)$$

Diese Gleichung zeigt auch mathematisch, dass das Ergebnis  $Z_n$  von jedem bis dahin empfangenen Eingangsbild  $Y_n$  abhängig ist und diese damit alle zur aktuellen Lösung beitragen. Weiterhin wird jedes Eingangsbild von genau zwei Filtern verarbeitet, wobei für ein großes  $\alpha_m$  in Näherung die Verarbeitung durch nur einen einzigen Filter pro Eingangsbild erfolgt. Für diesen Fall entspräche das Verfahren der einer rein zeitlich rekursiven Back-Projection gemäß Gleichung 5.9. Die Ordnung der jeweiligen Filter ist dabei von der zeitlichen Distanz des Eingangsbildes zum Schritt  $m$  zum aktuellen Eingangsbild zum Zeitpunkt  $n$  abhängig. Dieser Sachverhalt ist ebenfalls in Abbildung 5.22 verdeutlicht. Die Bewegungskompensationmatrix  $F_{n-m}^s$  und die Gewichtungsfaktoren  $\alpha_{n-m}^s$  fassen dabei die Gewichtungsfaktoren aller vorheriger Zeitschritte zusammen:

$$\alpha_{n-m}^s = \alpha_0 \cdot \alpha_1 \cdot \dots \cdot \alpha_{n-m} \text{ und } F_{n-m}^s = F_0 \cdot F_1 \cdot \dots \cdot F_{n-m} \quad (5.37)$$

Da der zeitliche Gewichtungsfaktor  $\alpha$  auf den Bereich 0..1 begrenzt ist, tragen zeitlich näher am aktuellen Bild liegende Bilder stärker zum aktuellen Ergebnisbild bei als weiter entfernte Eingangsbilder. Weiterhin werden nähere Bilder mit Filtern niedriger Ordnung verarbeitet, während bei weiter entfernten Bildern Filter höherer Ordnung zum Ergebnis beitragen. Dies ist ebenfalls direkt in Gleichung 5.36 und in Abbildung 5.22 zu sehen. Dabei wird jedes Eingangsbild durch genau zwei Filter verarbeitet, deren Ordnungen direkt aufeinanderfolgen und der zeitlichen Differenz zum aktuellen Bild bzw. einer Ordnung höher entsprechen. Um nun die Wirkungsweise noch genauer verstehen zu können, wird die Information der Eingangsbilder in einen festen, für jedes Bild unveränderten Anteil  $Y_{fix}$  und einen einzigartigen Anteil  $Y_{Var,n}$  unterteilt, der nur in einem Eingangsbild vorhanden ist, unterteilt. Durch diesen variablen Anteil fließen auch die Informationen der Bewegungskompensationsmatrizen mit ein, so dass diese ab nun weggelassen werden können.

$$Y_n = Y_{fix} + Y_{Var,n} \quad (5.38)$$

Dieser einzigartige Teil enthält sowohl neue Informationen, die für eine erhöhte Auflösung im Ergebnis unbedingt benötigt werden aber auch Artefakte, verursacht durch Rauschen oder eine fehlerhafte Bewegungsschätzung. Ersetzt man nun in Gleichung 5.36  $Y_n$  durch die beiden Summanden, so ergibt sich:

$$\begin{aligned} Z_n = & Y_{fix} \cdot \sum_{m=0}^n \alpha_{n-m}^s [(1 - H)^{n-m} + (1 - H)^{n-m+1} \cdot (1 - \alpha_m)] + \\ & + \sum_{m=0}^n \alpha_{n-m}^s [(1 - H)^{n-m} + (1 - H)^{n-m+1} \cdot (1 - \alpha_m)] \cdot Y_{Var,m} \end{aligned} \quad (5.39)$$

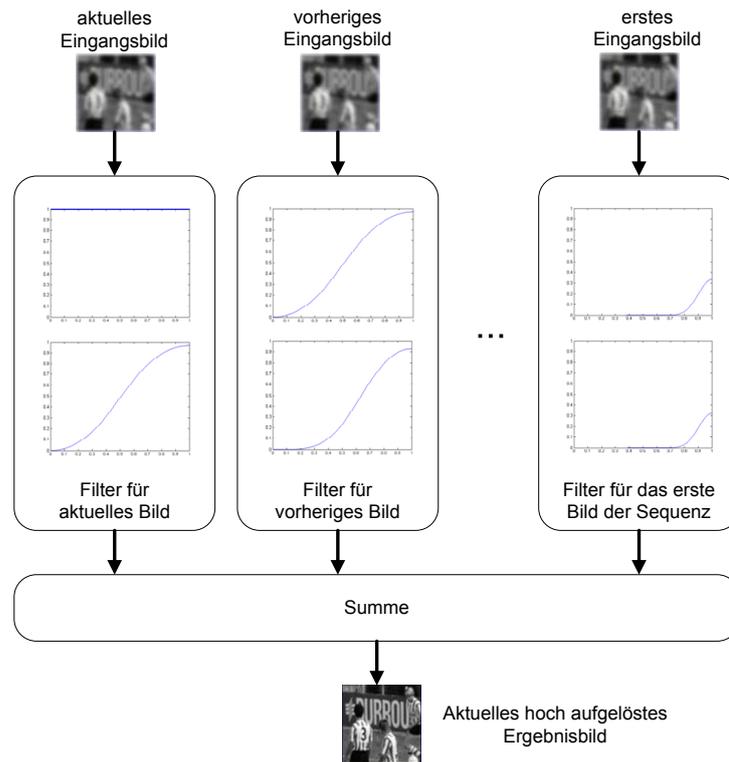


Abbildung 5.22: Darstellung der rekursiven Super-Resolution als direkte Verarbeitung unter Verwendung aller bisherigen Eingangsbilder. Auf die Gewichtung des zweiten Filters mit  $(1 - \alpha_m)$  wurde verzichtet.

Da die unveränderten Informationen unabhängig von den Zeitschritten  $m$  sind, kann dieser Term vor die Summe gezogen werden. Um nun die Konvergenz zu zeigen, können beide Summen jeweils in zwei Teilsommen unterteilt werden, so dass sich insgesamt vier Summen ergeben, für die einzeln Konvergenz gezeigt werden kann. Das SR-Verfahren ist dann stabil, wenn alle vier Teilsommen konvergieren. Die Konvergenz aller Teilsommen kann über das Majorantenkriterium gezeigt werden [BSMM01], indem für variable Terme eine Abschätzung nach oben vorgenommen wird. Kann Konvergenz für diese Abschätzung gezeigt werden, so konvergiert auch die Summe selbst, da sie definitiv kleiner ist. Beispielhaft wird dieses Vorgehen hier für den komplexesten Term gezeigt. Die Konvergenzprüfung für alle anderen Terme kann analog geschehen.

$$\sum_{m=0}^n \alpha_{n-m}^s [(1-H)^{n-m+1} \cdot (1-\alpha_m)] \cdot Y_{Var,m} \quad (5.40)$$

Setzt man nun alle Faktoren in  $\alpha_{n-m}^s$  auf einen Maximalwert  $\alpha_{max}$ ,  $Y_{var}$  auf 255 (8Bit Quantisierung angenommen) und  $\alpha_m$  auf den Minimalwert von 0, so ergibt sich für die Majorante als obere Schranke:

$$\begin{aligned} \sum_{m=0}^n \alpha_{n-m}^s [(1-H)^{n-m+1} \cdot (1-\alpha_m)] \cdot Y_{Var,m} \\ \leq 255 \cdot \sum_{m=0}^n \alpha_{max}^{n-m} [(1-H)^{n-m+1}] \end{aligned} \quad (5.41)$$

Die Majorante lässt sich dabei folgendermaßen umschreiben:

$$255 \cdot (1-H) \cdot \sum_{m=0}^n \alpha_{max}^{n-m} [(1-H)^{n-m}] \quad (5.42)$$

Für die hintere Summe lässt sich dann wieder über die geometrische Reihe die Konvergenzbedingung und damit die Stabilität herleiten:

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{m=0}^n \alpha_{max}^{n-m} [(1-H)^{n-m}] = \frac{1}{1 - \alpha_{max} \cdot (1-H)} \\ \text{for } \alpha_{max} \|1-H\| < 1 \end{aligned} \quad (5.43)$$

Generell können die einzigartigen Informationen beschrieben durch  $Y_{var}$  nicht manipuliert werden. Der Maximalwert von  $\alpha$  kann jedoch beeinflusst werden und auch auf einen Wert kleiner als 1 begrenzt werden. Das ermöglicht eine Kontrolle der Gewichtung, der Konvergenz und der Anhebung von Rauschen und Bilddetails. Je höher der Maximalwert von  $\alpha$  gewählt wird, desto höher ist der mögliche Detailgrad, aber auch die Rauschanhebung oder

die Sichtbarkeit von Artefakten durch eine fehlerhafte Bewegungskompensation. Im bereits beschriebenen Algorithmus werden noch weitere externe Gewichtungsfaktoren verwendet, welche aber nicht durch eine einfache Gleichung beschrieben werden können und die Konvergenzprüfung erschweren würden. Dadurch, dass diese Parameter aber die Detailsignale herunterwichten und nicht verstärken, wird deswegen die Konvergenz eher verbessert als verschlechtert. Im finalen Algorithmus bleibt aber der adaptive oder feste Gewichtungsfaktor  $\alpha$  der entscheidende Parameter. Da dieser aber von der absoluten Differenz zwischen aktuellem niedrig aufgelösten Eingangsbild und hoch aufgelösten, bewegungskompensierten vorherigen Ergebnisbild abhängt, wird dieser Faktor für eine starke Differenz, die auf eine mögliche Divergenz hindeutet, sehr klein. Im Grenzfall tendiert dieser Faktor zu Null, welches durch die Multiplikation aller Faktoren für jedes Eingangsbild noch beschleunigt wird (vergl. Gleichung 5.37). Aus diesem Grund wird in der Realität nur eine endliche Zahl an Eingangsbildern zum aktuellen Ergebnisbild beitragen. Weiterhin kann mit der Majorante, die den Filter für den Grenzfall und damit die maximale Verstärkung beschreibt, ein Parameter  $\alpha_{max}$  so bestimmt werden, dass die Anhebung auf einen gewünschten Sollwert begrenzt ist.

# Kapitel 6

## Auswertung

Dieses Kapitel beinhaltet Simulationsergebnisse der im Rahmen dieser Arbeit entwickelten Algorithmen. Zuerst wird in Abschnitt 6.1 unter kontrollierten Bedingungen die Fähigkeit der Detailwiederherstellung des rekursiven Super-Resolution Algorithmus (ESTERA) nachgewiesen. Dazu wird eine aus einem Einzelbild synthetisch erzeugte Sequenz mit kontrollierten Subpixelbewegungen und Alias verwendet. Zusätzlich wird gezeigt, welche quantitativen Verluste zu erwarten sind, wenn der Algorithmus mit robusteren Parametern zur Vermeidung von Seiteneffekten betrieben wird. Darauf folgend werden die experimentell ermittelten Ergebnisse mit Werkzeugen der Systemtheorie erklärt. Anschließend wird in Abschnitt 6.2 an Hand des mehrreferenzbildbasierten Back-Projection Algorithmus (VESBA) das Datenmodell inklusive aller vorgeschlagenen Modifikationen untersucht, wobei dieser Algorithmus auch die Möglichkeit bietet, dass doppelt rekursive Mehrreferenzbildverfahren MUSIDORA nachzubilden. In Abschnitt 6.3 folgt dann die Bewertung des rein rekursiven Algorithmus (ESTERA). Bei diesem werden schließlich auch die Bildanalysemaße zur Kontrolle der Parameter verwendet. Zum Abschluss des Kapitels werden die beiden neuen Algorithmen mit der höchsten Qualität – ESTERA und MUSIDORA – mit anderen bekannten Interpolationsverfahren nach objektiven und subjektiven Kriterien verglichen.

Bei allen Untersuchungen werden bis auf die jeweils erwähnten Modifikation die Standardparameter der Algorithmen aus Anhang A verwendet. Als Interpolationsfaktor wurde stets zwei verwendet. Die verwendeten Testsequenzen werden in Anhang B beschrieben. Sie bilden die wichtigsten Bewegungssituation und typische Merkmale aus der Praxis ab. Dadurch erlauben sie eine gute Einschätzung der verschiedenen Eigenschaften der Algorithmen. Den Schwerpunkt der Bewertungen bilden subjektive Eindrücke bzw. Bildbeispiele, da hiermit der erzielbare Gesamteindruck am Besten vermittelt werden kann. Ein besonderes Augenmerk liegt auf dem Detailerhalt bzw. der Wiederherstellung von Details. Auf einen subjektiven Test nach ITU-R BT 500.11 [ITU] wurde im Rahmen dieser Arbeit verzichtet, da die Menge der Algorithmen und die zu untersuchenden Fälle für einen Test zu umfangreich gewesen wäre. Zur objektiven Bewertung finden die Maße Peak Signal to Noise Ratio (PSNR) und Structural Similarity Index Measure (SSIM) Anwendung.

## 6.1 Potential der Super-Resolution

In der Literatur finden sich bereits zahlreiche Beispiele über die Rekonstruktionsfähigkeiten von Super-Resolution Algorithmen unter Idealbedingungen. So findet sich beispielsweise in [Mil05] eine Diskussion über mögliche Grenzen der Super-Resolution. Voraussetzungen wie Alias, Subpixelbewegungen oder ein variierender Störungsgrad werden dort ebenfalls bereits diskutiert. Deshalb soll in diesem Abschnitt nur der neue, rein rekursive Ansatz ESTERA betrachtet werden. Weiterhin wird der Versuch unternommen, die bekannten Bedingungen für eine erfolgreiche Super-Resolution durch ein fundiertes, auf der Systemtheorie basierendes Modell zu erklären.

### 6.1.1 Simulationsergebnisse für eine aliasbehaftete Sequenz

Um die Fähigkeit der Detailwiederherstellung zu zeigen, wurde aus dem Bild 'barbara' durch eine Verschiebung um  $(1,0)$ ,  $(1,1)$  und  $(0,1)$  eine synthetische Sequenz erzeugt. Anschließend wird eine Unterabtastung um den Faktor zwei durchgeführt. Bei dieser Unterabtastung kommt im ersten Experiment kein Vorfilter zur Anwendung, so dass die unterabgetastete Sequenz über sehr starke (örtliche) Aliasanteile verfügt. Da die Bewegung auf das Bild mit dem höherem Abtastraster bezogen ist, weist die niedriger aufgelöste Sequenz Subpixelbewegungen auf. Durch die gezielt gewählten Verschiebungen beinhaltet die Sequenz aber trotzdem noch alle Informationen des ursprünglichen Bildes, so dass eine vollständige Rekonstruktion von diesem möglich sein sollte. In Abbildung 6.1 sind die Ergebnisse des rein rekursiven Super-Resolution Algorithmus für diese synthetische Sequenz zu sehen. Die erste Reihe gibt dazu Referenzen an, wobei die Interpolation mit Pixelwiederholung unter B der niedrig aufgelösten Referenz entsprechen würde, wenn man davon ausgeht, dass durch die Umskalierung die Pixel eine höhere Ausdehnung besitzen. Die zweite Zeile veranschaulicht die Detailrekonstruktion durch Super-Resolution unter Verwendung unterschiedlicher initialer Interpolationsstrategien. Dabei muss beachtet werden, dass durch die Interpolation ein Rekonstruktionstiefpass zur Anwendung kommt, der einen Teil der nötigen Aliasinformation dämpfen oder entfernen kann. Dies wird deutlich, wenn man die beiden Ergebnisse D und E in der zweiten Zeile miteinander vergleicht. Die Dämpfung kann durch die Verwendung des Unschärfefilters im Degradationsmodell teilweise kompensiert werden (Teilbild F). Da in diesem Fall durch den starken Aliasanteil nur wenige Frequenzanteile durch die High-Resolution-Spline Interpolation gedämpft werden, findet in diesem Fall eine Überkompensation statt. Bei deutlich geringerem Aliasanteil oder für reale Sequenzen ist dieses Verhalten aber nicht zu erwarten (vergl. Abbildung 6.2). Die Überschärfe ist auch dadurch zu erklären, dass durch die Verwendung der (idealen) Bewegungsvektoren keine Unschärfe durch die zeitliche Filterung entstehen kann, die bei normalen Sequenzen ebenfalls durch den Unschärfefilter teilweise kompensiert werden kann / muss. Die untere Reihe in Abbildung 6.1 verdeutlicht, dass die Verwendung eines



Abbildung 6.1: Rekursive Super-Resolution mit unterschiedlichen Parametern für eine kontrollierte Sequenz mit Alias. Obere Zeile: Referenzbilder (A: hoch aufgelöste Referenz, B: Pixelwiederholung, C: HRS Interpolation). Mittlere Zeile: Einfluss der initialen Interpolation auf die SR (D: initiale Pixelwiederholung und ohne Beobachtungsmodell, E: initiale HRS Interpolation und ohne Beobachtungsmodell, F: initiale HRS Interpolation und Beobachtungsmodell). Untere Zeile: Einfluss des adaptiven Gewichtungsfaktors auf die SR (G: adaptiver Faktor und ideale Vektoren, H: adaptiver Faktor und reale Vektoren, I: fester Faktor und reale Vektoren). Wenn nicht anders angegeben, wurden ideale Bewegungsvektoren vorgegeben, ein fester Gewichtungsfaktor von 0.9 gewählt, eine initiale HRS Interpolation durchgeführt und das Beobachtungsmodell zugeschaltet.



Abbildung 6.2: Rekursive Super-Resolution mit unterschiedlichen Parametern für eine kontrollierte Sequenz ohne Alias. Obere Zeile: Referenzbilder (A: hoch aufgelöste Referenz, B: Pixelwiederholung, C: HRS Interpolation). Mittlere Zeile: Einfluss der initialen Interpolation auf die SR (D: initiale Pixelwiederholung und ohne Beobachtungsmodell, E: initiale HRS Interpolation und ohne Beobachtungsmodell, F: initiale HRS Interpolation und Beobachtungsmodell). Untere Zeile: Einfluss des adaptiven Gewichtungsfaktors auf die SR (G: adaptiver Faktor und ideale Vektoren, H: adaptiver Faktor und reale Vektoren, I: fester Faktor und reale Vektoren). Wenn nicht anders angegeben, wurden ideale Bewegungsvektoren vorgegeben, ein fester Gewichtungsfaktor von 0.9 gewählt, eine initiale HRS Interpolation durchgeführt und das Beobachtungsmodell zugeschaltet.)

adaptiven Gewichtungsfaktors zu zwei Effekten führt: Zum einen fällt die Detailwiederherstellung in Detailbereichen mit hohem Kontrast deutlich geringer aus. Das ist durch den stark unterschiedlichen Detailgrad in den Bildern und der damit verbundenen hohen zeitlichen Differenz zu erklären, die zu einem kleinen Gewichtungsfaktor führt. Dieses Verhalten wurde in dieser Abbildung bewusst übertrieben. Der Detailgrad kann durch eine Variation des Parameters  $\kappa$  auf Kosten einer geringeren Robustheit deutlich erhöht werden, wie der Grenzfall des festen Faktors in I zeigt. Für die gewählte adaptive Parametrisierung wird aber die eigentlich korrekte neue Information trotz perfekter Vektoren nicht zur Filterung verwendet. Zum anderen können durch die adaptive Gewichtung Seiteneffekte vermieden werden, was hier eindrucksvoll demonstriert wurde (vergleiche Abbildungen H und I). Dazu zählt die Vermeidung der Überschärfe und die Sichtbarkeit von Artefakten verursacht durch eine falsche Bewegungskompensation. Aber selbst mit den gewählten Einstellung für den adaptiven Gewichtungsfaktor ist eine Detailwiederherstellung möglich. Gerade in den weniger kontrastreichen Regionen (z.B. im Hintergrund) finden sich im Vergleich zur einfachen Pixelwiederholung oder der High Resolution Spline Interpolation mehr Details. Betrachtet man als letztes das Ergebnis unten rechts, so ist erkennbar, dass durch fehlerhafte Bewegungsvektoren Artefakte entstehen, die durch die scharfen Kanten zu einer subjektiv höher empfundenen Auflösung (in einem Einzelbild) führen können. Dieses Verhalten ist aber in den meisten Fällen nicht gewünscht, weil diese 'Pseudodetails' tatsächlich nicht vorhanden sind. Außerdem sind diese in der laufenden Sequenz selten zeitlich stabil und ohne große Korrelation zum Bildinhalt (mit Ausnahme pseudozufälliger Texturen), so dass die Entstehung in der Regel vermieden werden muss.

### 6.1.2 Simulationsergebnisse für eine Sequenz ohne Alias

Abbildung 6.2 zeigt die Ergebnisse derselben Algorithmen für den Fall, dass ein gauß'scher Tiefpass mit der normierten Grenzfrequenz 0.5 vor der Unterabtastung verwendet wird. Dadurch beinhaltet die niedrig aufgelöste Sequenz erheblich weniger örtlichen Alias. Gerade im Vergleich zum Fall mit Alias wird deutlich, dass nicht mehr alle Informationen aus dem hoch aufgelösten Bild wiederhergestellt werden können. Weiterhin sind die Unterschiede zwischen den einzelnen Verarbeitungsvarianten nun deutlich geringer. Durch das geringere Vorkommen hochfrequenter Anteile im Eingangsmaterial ist der Unschärfefilter im Beobachtungsmodell (Teilabbildungen F bis I) deshalb deutlich besser auf das Eingangsmaterial (bzw. das Degradationsmodell) abgestimmt, so dass die Überkompensation nicht mehr in Erscheinung tritt. Im Vergleich zu den Standardverfahren kann durch den rekursiven Super-Resolution Algorithmus aber selbst für diesen Fall eine bessere Qualität erzielt werden. Das lässt sich durch die Kombination aus rekursiver Verarbeitung, der Ausnutzung des geringen Aliasanteils und dem internen Unschärfefilter erklären. Für stark verrauschte Sequenzen, die hier nicht gezeigt wurden, führt die zeitliche Rückführung

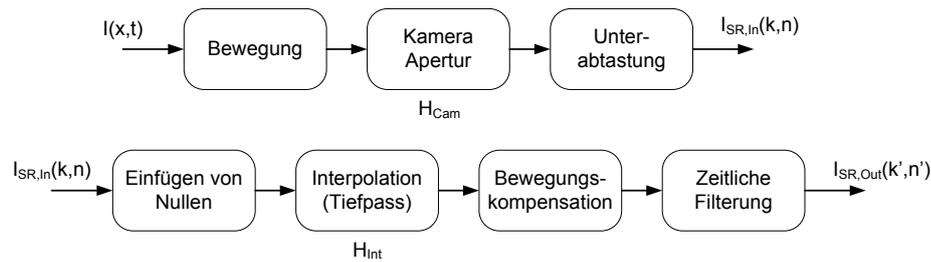


Abbildung 6.3: Bildverarbeitungskette bei Verwendung von Super-Resolution

im SR-Algorithmus weiterhin zu einer Rauschreduktion, so dass der Unterschied zu rein örtlichen Verfahren im Nutzfrequenzbereich wesentlich deutlicher ausfallen würde.

### 6.1.3 Systemtheoretisches Modell für den Super-Resolution Prozess

Die im vorigen Abschnitt dargestellten Simulationsergebnisse können auch systemtheoretisch erklärt werden. Zu diesem Zweck wird zuerst eine Bildverarbeitungskette vorgestellt, die den Prozess von der Bildaufnahme bis zur Rekonstruktion durch die Super-Resolution nachbildet. Zur Modellierung dieser Kette wurde auf Konzepte zurückgegriffen, die in [Sch00] für die allgemeine Bewegung und in [KV04a] und [Dol10] zur Modellierung von Bewegungsunschärfe bereits verwendet wurden. Für eine einfachere Beschreibung wird hier nur eine örtliche Ausdehnung in  $x$ -Richtung verwendet, die Übertragung auf den örtlich zweidimensionalen Fall ist aber sehr leicht möglich. Die so entstehende Kette zeigt Abbildung 6.3. Von der Kamera wird die reale und kontinuierliche Szene  $I(x, t)$ , die ggfs. eine Bewegung beinhaltet, örtlich und zeitlich abgetastet. Bei dieser Abtastung entsteht meistens Alias, wenn der (optische) Vorfilter keine genügende Bandbegrenzung zulässt. Als Eingangssequenz steht der Super-Resolution dann die diskrete Sequenzfolge  $I_{SR,in}(k, n)$  zur Verfügung. Zur Vereinfachung wird hier angenommen, dass alle beteiligten Elemente linear und zeit- bzw. ortsinvariant sind. Neue Informationen in Auf- und Verdeckungs-bereichen oder aber Nichtlinearitäten im Algorithmus wie z.B. Clipping oder die Gewichtung durch Bildanalysefunktionen werden deshalb hier nicht berücksichtigt. Für die Aufnahmeseite ergibt sich so:

$$I_{SR,in}(k, n) = [I(x, t) * \delta(x - vt) * h_{kam,s}(x)] \cdot \sum_{k,n} \delta(x - k\Delta x) \cdot \delta(t - n\Delta t) \quad (6.1)$$

Die Kamera wird zwecks Vereinfachung in zeitlicher Richtung als ideal angenommen, so dass keine zusätzliche Unschärfe durch Bewegung entsteht. Nur eine mögliche örtliche Unschärfe durch eine fehlerhafte Fokussierung mit  $h_{kam,s}$  wird berücksichtigt.  $\Delta x$  und  $\Delta t$

bezeichnen die Abstände zwischen den diskreten Abtastpunkten in örtlicher und zeitlicher Richtung. Die Bewegung in der Szene wird durch den Term  $\delta(x - vt)$  berücksichtigt [Sch00]. Der Super-Resolution Algorithmus selbst wird vereinfachend durch eine örtliche Interpolation gefolgt von einer zeitlich bewegungskompensierten Filterung modelliert. Diese Prozessschritte sind auch in beiden mehrreferenzbildbasierten Algorithmen (VESBA und MUSIDORA) und dem rein rekursiven Algorithmus (ESTERA) vorhanden. Mathematisch können diese Schritte folgendermaßen modelliert werden:

$$I_{SR,out}(k', n) = h_{int}(k') * h_{SR,t}(k' - (v_{est} - v)n, n) * I_{SR,in}(k, n) \quad (6.2)$$

Der örtliche Interpolationsfilter  $h_{int}$  arbeitet in dieser Gleichung bereits auf dem höher aufgelösten Raster  $k'$ . Die Filterpositionen des zeitlichen Filters der Super-Resolution,  $h_{SR,t}$  hängen von der diskretisierten, realen Bewegung der Szene  $v$  sowie der durch die Bewegungsschätzung ermittelten Bewegung  $v_{est}$  ab, wie die Abbildungen 6.4 und 6.5 verdeutlichen. Die Bewegung in der Szene macht sich durch eine Scherung bemerkbar [Sch00]. Durch die Abtastung in örtlicher und zeitlicher Richtung durch die Kamera entstehen Wiederhol-spektren bei Vielfachen der Abtastfrequenz. Bei der Interpolation durch Super-Resolution soll nun die Abtastfrequenz erhöht (verdoppelt) werden. Dazu müssen, wie in Abschnitt 2.1 bereits erläutert, zuerst Nullen eingefügt werden und anschließend die fehlenden Werte mit einem Interpolationsfilter errechnet werden. Im Spektrum korrespondieren diese Prozessschritte mit einer Umskalierung der Frequenzachse und einer (notwendigen) Entfernung des entstehenden Zwischenspektrums. Die Filterung erfolgt bei der Super-Resolution nun durch eine Kombination aus örtlichem und zeitlichem Filter. Der örtliche Filter  $H_{spat}$ , welcher aus einer Kombination aus unvermeidbarem örtlichen Tiefpass der Kamera und dem örtlichen Interpolationsfilter besteht, weist einen Verlauf entlang der örtlichen Frequenzachse auf. Der Durchlassbereich des zeitlichen Filters hängt dabei von der geschätzten und der tatsächlichen Bewegung der Szene ab.

### Filterverhalten der Super-Resolution für Subpixelbewegungen

Handelt es sich bei der geschätzten Bewegung um eine Subpixelbewegung, so ist der Durchlassbereich des zeitlichen Filters nur auf die Spektren und Wiederhol-spektren der hoch aufgelösten Szene beschränkt, wie es in der Abbildung 6.4 gezeigt wird. Dadurch ist allein mit dem zeitlichen Filter eine vollständige Unterdrückung der Störspektren möglich. Durch die Ausrichtung des Filters entlang der tatsächlichen Bewegung bleibt sogar der Aliasanteil (sofern vorhanden) erhalten. Dies ist durch eine rein örtliche Interpolation nicht möglich, da hier ein Kompromiss zwischen Unterdrückung der störenden (örtlichen) Wiederhol-spektren und dem Erhalt hoher Frequenzanteile gewählt werden muss. Um die Dämpfung durch den örtlichen Filter gering zu halten, könnte im Falle der korrekt geschätzten Subpixelbewegung sogar vollständig auf einen örtlichen Interpolationsfilter verzichtet oder ein

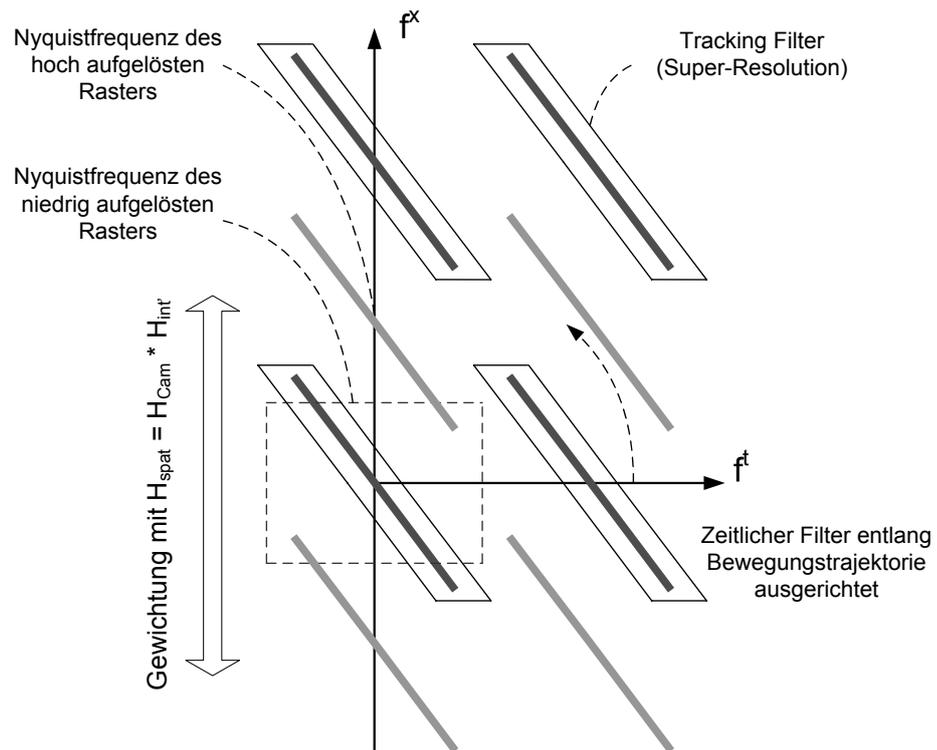


Abbildung 6.4: Spektrum für das Ergebnis der Super-Resolution für eine Sequenz mit Subpixelbewegung und korrekt geschätzter Bewegung

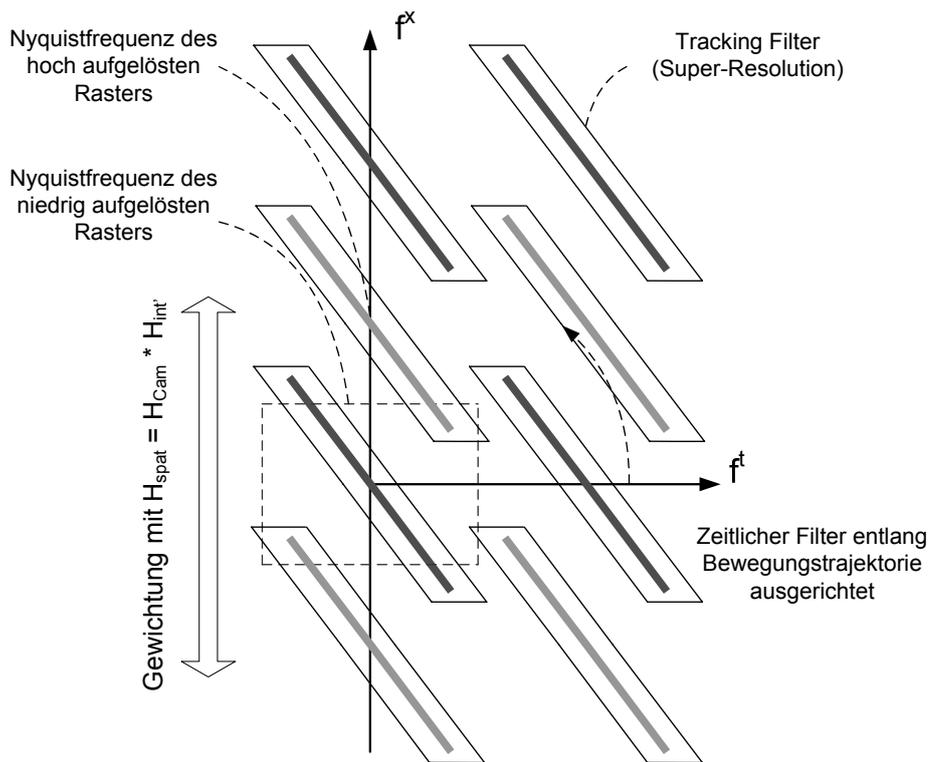


Abbildung 6.5: Spektrum für das Ergebnis der Super-Resolution für eine Sequenz mit Pixelbewegung und korrekt geschätzter Bewegung

örtlicher Filter mit geringerer Dämpfung eingesetzt werden. Die Pixelinformationen würden einfach aus vorherigen oder nachfolgenden Bildern an die Stellen des hoch aufgelösten Rasters kopiert. Die in diesem Fall unnötige Dämpfung ist auch in den mittleren drei Ausschnitten in den Abbildungen 6.1 und 6.2 zu sehen. Die mögliche Unschärfe durch den örtlichen Filter kann dabei durch das örtliche Filter im Super-Resolution Algorithmus teilweise kompensiert werden.

### Filterverhalten der Super-Resolution für Pixelbewegungen

Weist die Sequenz keine Bewegung oder nur Pixelbewegungen auf, so kann durch den zeitlichen Filter im Allgemeinen nur eine Rauschreduktion, aber keine Detailwiederherstellung durchgeführt werden. In diesem Fall arbeitet der Filter nur auf dem Abtastraster der niedrig aufgelösten Sequenz, was im Spektrum zu einem zusätzlichen Durchlassbereich bei den Störspektren führt (siehe Abbildung 6.5). Die neu eingefügten Pixelpositionen können also nicht gefüllt werden sondern müssen durch örtliche Interpolationsfilter berechnet werden. Für diesen Fall bietet sich ein Interpolationsfilter mit einer sehr

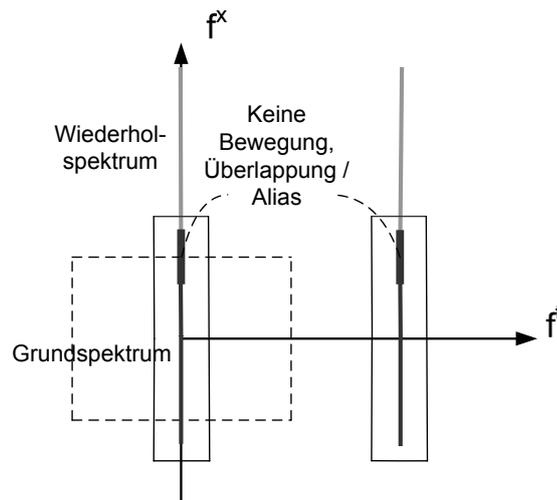


Abbildung 6.6: Alias im Falle von unbewegten Sequenzen

steilen Filterflanke im Bereich der Nyquistfrequenz an, um möglichst alle Störanteile zu unterdrücken. Hierbei zeigt sich, dass für beide Bewegungsarten ein Kompromiss für den örtlichen Filter gewählt werden muss, da zwei völlig gegensätzliche örtliche Filter optimal wären. Zwecks Robustheit wurde als Kompromiss für die hier entwickelten Algorithmen auf eine Interpolation mittels HRS gesetzt. Zukünftige Arbeiten könnten hier eine adaptive Auswahl des örtlichen Interpolationsfilters in Abhängigkeit der Bewegungsinformation untersuchen.

Eine Ausnahme für einen Gewinn auch bei Pixelbewegungen durch die zeitliche Filterung kann durch das lineare Modell nicht gezeigt werden. Dieser Gewinn ist bei zeitlich variierende Eingangsinformationen gegeben, z.B. Hervorgerufen durch eine unterschiedliche Kompressionsrate des Eingangsmaterials (vergleiche [RDS09a]).

### Filterverhalten der Super-Resolution für unbewegte Sequenzen

Liegt in der Sequenz keine Bewegung vor, so stellt dies einen Sonderfall der Pixelbewegungen dar. Auch hier erstreckt sich der zeitliche Filter über die gewünschten und ungewünschten Spektren, also über die gesamte örtliche Frequenzachse, so dass die Wiederhol-spektren durch diesen nicht entfernt werden können. Zusätzlich kann hier durch die Überlappung von Originalspektrum und dem ersten örtlichen Wiederhol-spektrum noch starker Alias entstehen (Abbildung 6.6). Die Entfernung der Störspektren sowie des Alias ist nur durch einen rein örtlichen Interpolationstiefpass auf Kosten eines eventuellen Detailverlust möglich.

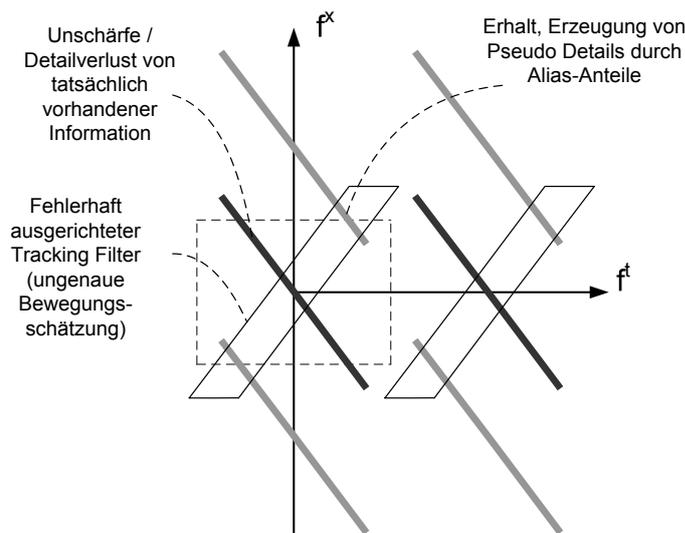


Abbildung 6.7: Erzeugung von falschen Details durch eine fehlerhafte Bewegungsschätzung

### Filterverhalten der Super-Resolution für eine fehlerhafte Bewegungsschätzung

Zuletzt soll noch die bereits angesprochene Erzeugung von zumeist ungewünschten 'Pseudodetails' mit Abbildung 6.7 erklärt werden. Eine fehlerhafte Bewegungsschätzung führt zu einem falsch ausgerichteten zeitlichen Filter. Dies führt zu einer Unschärfe hervorgerufen durch eine starke Dämpfung der im Signal enthaltenen hochfrequenten Anteile. Zum anderen werden nun Anteile aus dem Wiederholerspektren nicht mehr durch den zeitlichen Filter gedämpft. Im hoch aufgelösten Bild führen diese Frequenzanteile zu neuen, in der Sequenz nicht vorhandenen 'Pseudodetails' oberhalb der ursprünglichen Nyquistfrequenz (siehe auch Abbildung 6.1 unten rechts). Diese Pseudodetails tragen aber nur positiv zum Bildeindruck bei, wenn sie sich nicht von den tatsächlich vorhandenen Details unterscheiden lassen und sich sinnvoll und zeitlich stabil in die hochgetastete Sequenz einfügen. Hier stellt sich aber – wie bereits angesprochen – die Frage, ob die Erzeugung von nicht in der Sequenz vorhandenen Pseudoinformationen gewünscht ist.

## 6.2 Beurteilung der Mehrreferenzbild Back-Projection Super-Resolution

Um die Eigenschaften und die Vorteile der eigenen Modifikationen an den Super-Resolution Algorithmen untersuchen zu können, wurde der voll parametrisierbare Mehrreferenzbild-Algorithmus VESBA aus Abschnitt 5.1 verwendet, der eine Zu- und Abschaltung aller wichtigen Komponenten des Datenmodells ermöglicht. Hier sollen nun der Einfluss der

verschiedenen Interpolations- und Dezimationsstrategien und die Verwendung unterschiedlicher zeitlicher Filterverfahren auf das Gesamtergebnis verdeutlicht werden. Sofern nicht anders angegeben, werden hierzu die Standardeinstellungen verwendet, die in Anhang A aufgelistet sind. Nach der Untersuchung des Datenmodells wird noch kurz auf die Wirkungsweise des Bildmodells eingegangen. Für all diese Untersuchungen wurden fünf Eingangsbilder und vier auf jedes Bild bezogene Iterationen verwendet. Die Wahl dieser beiden Parameter stützt sich auf Erfahrungswerte und liefert einen guten Kompromiss zwischen Rechenaufwand und Bildqualität. Auf einen Nachweis für die Wahl dieser Parameter sowie der Konvergenzgeschwindigkeit wurde aus mehreren Gründen verzichtet: Zum einen ist eine Variation dieser Parameter, gerade in Bezug auf die Anzahl der Eingangsbilder, in einem Einzelbild nur unzureichend darzustellen, da sich die Gewinne vor allem in der laufenden Sequenz durch eine deutlich stärkere Rauschreduktion und Detailwahrnehmung bemerkbar machen. Zum anderen hängt die Zahl der benötigten Eingangsbilder sehr stark von der Eingangssituation ab, welche im vorherigen Abschnitt theoretisch betrachtet wurden. Liegt keine Bewegung oder kein Alias vor, so führt die Erhöhung der Anzahl an Eingangsbildern zu keinem nennenswerten Gewinn. Untersuchungen in Bezug zur optimalen Bildanzahl sind aber für andere zeitliche Algorithmen u.a. in [RDS09a] und [Stö09] durchgeführt worden. Ähnliche Bemerkungen können zur Konvergenzgeschwindigkeit der Verfahren gemacht werden, da diese gerade bei den rein rekursiven Algorithmen durch externe Einflüsse (wie fehlerhafte Vektoren) beeinflusst werden. Dieses Thema wird aber z.B. in [Naß09] und [RNS11] behandelt.

### 6.2.1 Untersuchung des Datenmodells

Das Datenmodell zeichnet sich für die Ähnlichkeit zwischen niedrig und hoch aufgelöster Sequenz verantwortlich. Zu diesem Zweck müssen Bilder unterschiedlicher Auflösung und von unterschiedlichen Zeitpunkten miteinander verglichen werden. Damit der Bildinhalt beider Sequenzen positionsgenau übereinstimmt, werden Interpolations-, Dezimations- und Bewegungskompensationstechniken benötigt. Beim ursprünglichen Back-Projection Algorithmus muss das aktuelle, hoch aufgelöste Bild in jedem Iterationsschritt zuerst unterabgetastet und dann bewegungskompensiert werden, um mit jedem Eingangsbild verglichen werden zu können. Aus dem Vergleich ergibt sich ein Korrektursignal, welches dann wiederum hochgetastet und der inversen Bewegungskompensation unterzogen werden muss. Bei diesen Prozessschritten bleiben die Eingangsbilder unangetastet. Alternative Lösungswege, die auf einer Interpolation und Bewegungskompensation der Eingangsbilder beruhen, können die Prozessschritte innerhalb einer Iteration deutlich vereinfachen (vergl. Abschnitt 5.1). Hier soll nun die Frage beantwortet werden, welche Strategie die qualitativ besten Ergebnisse liefert.



Abbildung 6.8: Interpolations- und Dezimationsstrategien für die Mehrreferenzbild-SR

### Interpolations- und Dezimationsstrategien

Abbildung 6.8 zeigt drei Möglichkeiten, mit den unterschiedlichen Auflösungen der Ein- und Ausgangsbilder umzugehen. Bei den ersten beiden Ergebnissen findet eine Runter- und Hochtastung innerhalb der Iterationsschleife statt. Als Heruntertastung (Beobachtungsmodell) wird beide Male ein gauß'scher Tiefpass gefolgt von einer Unterabtastung um den Faktor zwei verwendet. Wird die Hochtastung (Back-Projection Operator) mittels einer Pixelwiederholung mit einem anschließenden Unschärfefilter durchgeführt, so sieht man deutlich in manchen Bereichen die Pixelstrukturen der niedrig aufgelösten Sequenz, insbesondere an diagonalen Kanten wie der Schulter des gezeigten Spielers. Diese Strukturen können durch die Verwendung eines höherwertigen Interpolationsfilters für das Detailsignal (High Resolution Spline, mittlere Abbildung) deutlich verringert werden. Als dritte Möglichkeit ist eine initiale Interpolation des Eingangsbilds zu sehen. In diesem Fall fungiert der Super-Resolution Algorithmus selbst nur noch als örtlich-zeitlicher Deblurring-Algorithmus. Diese Strategie führt im Vergleich zu dem klassischen Konzept mit interner Veränderung der Pixelanzahl zu schärferen Ergebnissen, Artefakte durch eine fehlerhafte Bewegungsschätzung sind aber ebenfalls deutlich besser zu erkennen. Die Häufigkeit und Sichtbarkeit dieser Artefakte kann durch die Verwendung eines anderen zeitlichen Filters noch deutlich reduziert werden, was im nächsten Abschnitt erläutert wird.

### Bewegungskompensierte zeitliche Filterung

Wie bereits erwähnt, kann die bewegungskompensierte Mittelwertfilterung bei fehlerhafter Bewegungsinformation zu Seiteneffekten führen. Aus diesem Grund wurde neben dieser Strategie ein adaptiver FIR-Filter (siehe Abschnitt 5.5) und ein bewegungskompensierter



Abbildung 6.9: Einfluss unterschiedlicher zeitlicher Filterstrategien zur Bestimmung des Detailsignals



Abbildung 6.10: Einfluss unterschiedlicher Filterstrategien bei inkorrekturer Bewegungskompensation

zeitlicher Medianfilter implementiert. Die Ergebnisse für diese drei Filter finden sich in Abbildung 6.9. Im Vergleich zum Mittelwertfilter sind mit den beiden anderen Strategien deutliche Verbesserungen möglich. Die Unterschiede zwischen adaptivem FIR-Filter und Medianfilter sind nur marginal, was durch die sehr gute und präzise Bewegungsschätzung zu erklären ist. Außerdem ist es bei der Mehrreferenzbildbewegungsschätzung über 5 Bilder sehr unwahrscheinlich, dass alle Vektoren vollkommen fehlerhaft sind. Bei einem Median ergibt sich bereits dann ein sinnvolles Ergebnis für die Filterung, wenn mehr als die Hälfte der Schätzungen pro Bildpunkt (in diesem Fall 3) korrekt ist. Bei der FIR-Filterung werden fehlerhafte Schätzungen durch die adaptiven Koeffizienten einfach unterdrückt und nur die Pixel aus den Referenzbildern mit einem ähnlichen Luminanzwert wie das aktuelle Pixel zur Filterung verwendet. Aus diesem Grund sind deutliche Unterschiede erst dann wahrnehmbar, wenn alle Vektoren falsch sind. Ein Beispiel hierfür ist in Abbildung 6.10 dargestellt. In diesem Fall wurden bewusst nur Nullvektoren verwendet, die für nahezu alle Pixel im Bild nicht die korrekte Bewegungssituation abbilden. Wie erwartet, ergibt sich für die Mittelwertfilterung der Detailsignale durch die fehlerhafte Kompensation eine extreme Unschärfe. Der Medianfilter kann diese Unschärfe nahezu vollständig korrigieren, allerdings ergeben sich nur dann sinnvolle Ergebnisse, wenn die Mehrheit der Eingangspixel örtlich stark korreliert sind. Ansonsten werden Details entfernt oder aber falsche Bildinformationen von einer anderen Stelle eingefügt. Der adaptive Filter kann die Verwendung falscher Informationen am besten unterdrücken, neue Details können aber für diesen Fall selbstverständlich nicht generiert bzw. wiederhergestellt werden. Auf Grund dieser Ergebnisse ist für reale Sequenzen der adaptive FIR-Filter am robustesten. Die Detailwiederherstellung hingegen ist mit festen Koeffizienten höher, insbesondere wenn die Bewegung korrekt geschätzt wurde aber die neue Information im Referenzbild sehr stark vom Wert des aktuellen Pixels abweicht. In den weiteren Experimenten wird der Mittelwertfilter verwendet, da der Einfluss der anderen Komponenten und Parameter auf das Ergebnis mit diesem besser gezeigt werden kann. Für das neu entwickelte Mehrreferenzbildverfahren MUSIDORA wird aber der adaptive FIR-Filter genutzt.

### Ergebnisse der anderen Veränderungen im Datenmodell

Im Rahmen dieser Arbeit wurden drei Veränderungsmöglichkeiten erprobt, um die Bewegungskompensation entweder zu vereinfachen oder sogar bessere Ergebnisse möglich zu machen. Wird die Eingangssequenz an Stelle des hoch aufgelösten Ergebnisses kompensiert, so wird bei der ersten Iteration wenigstens eine Halbierung der Anzahl an Kompensationen erreicht (vergl. Abschnitt 5.1), da das Detailsignal dann direkt in der korrekten zeitlichen Phasenlage vorliegt. Bei allen weiteren Iterationen muss sogar keine erneute Kompensation durchgeführt werden, wenn man die kompensierte Eingangssequenz weiter vorhält (zwischenspeichert). Als positiver Nebeneffekt reicht ein Bewegungsvektorfeld mit der Pixelanzahl der niedrig aufgelösten Sequenz aus, wenn man auf Kosten einer geringe-

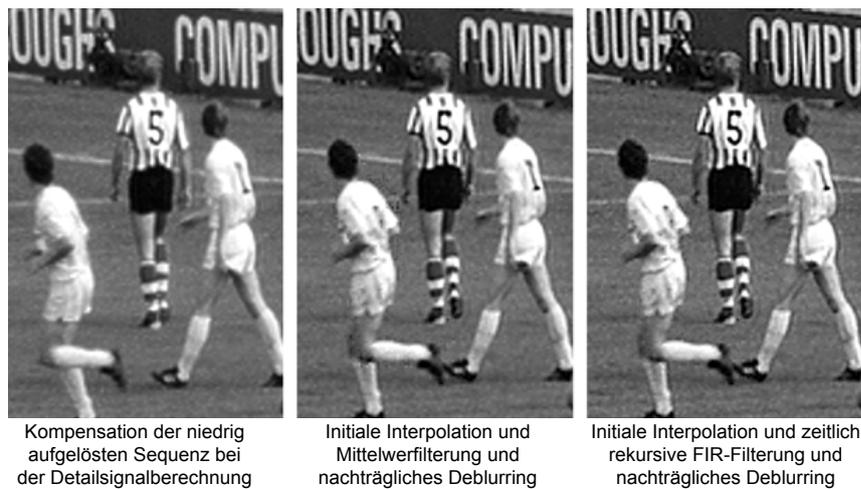


Abbildung 6.11: Alternative zeitliche Filtertechniken. Die rechte Teilabbildung entspricht den Einstellungen von MUSIDORA.

ren Auflösungserhöhung die Kompensation vor der Interpolation durchführt. Dadurch fällt entweder die Anzahl an Operationen innerhalb der Schätzung geringer aus oder es kann zumindest auf eine Hochtastung des Vektorfeldes verzichtet werden. An Hand von Abbildung 6.11 ist erkennbar, dass diese Vorgehensweise in manchen Bereichen zu einem minimal unschärferen Ergebnis führt. Unterschiede können beispielsweise im Bereich der Schuhe des linken Spielers beobachtet werden. Als zweite Möglichkeit kann die zeitliche Filterung vollkommen aus der Iterationsschleife herausgezogen werden. Auch dieses Vorgehen resultiert ähnlich wie das Herausziehen der Runter- und Hochtastung zu einer Verbesserung der Ergebnisqualität (Abb. 6.11 Mitte). In diesem Fall wird der Back-Projection Algorithmus in drei Stufen aufgeteilt, die sequentiell abgearbeitet werden: eine initiale Interpolation, eine zeitliche Mehrreferenzbildfilterung und ein iteratives Deblurring. Diese Vorgehensweise führt zu sehr robusten Ergebnissen ohne große Seiteneffekte. Negativ können sich aber der örtliche und zeitliche Tiefpass bemerkbar machen, die ggfs. hochfrequente Details entfernen und die Wiederherstellung dieser durch das nachträgliche Deblurring erschweren. In [FREM04] wurde dieses Vorgehen und die Auftrennung der einzelnen Stufen ebenfalls vorgeschlagen. Dieses beruhte aber eher auf der Komplexität des Kalman-Filters denn auf Überlegungen zur Bildqualität, da dieser ansonsten nicht effizient zu realisieren ist. Die letzte Möglichkeit, die in Abbildung 6.11 zu sehen ist, beruht auf einer zeitlich rekursiven Mehrreferenzbildfilterung außerhalb der Iterationsschleife. In diesem Fall werden für die Filterung bereits berechnete Ergebnisbilder anstelle der eigentlichen Eingangsbilder verwendet. Dadurch ergibt sich eine doppelt rekursive Struktur, da sowohl eine örtliche Rückführungsschleife zum Durchlaufen mehrerer auf jedes Bild bezogenen Iterationen als auch eine zeitlich rekursive Rückführung vorhanden ist, welche in Kapitel 5 als MUSIDORA



Abbildung 6.12: Eigenschaften des Huber MRF Bildmodells und des modifizierten HMRF auf das Ergebnis bei unterschiedlichen Schwellwerten  $\alpha$

bezeichnet wurde. Dieser Algorithmus liefert im Vergleich zum Standard Back-Projection Algorithmus deutlich bessere Ergebnisse mit höherer Bildschärfe und Detailrekonstruktion bei gleichzeitig weniger Seiteneffekten (Anwendung des adaptiven FIR-Filters vorausgesetzt). Der Aufwand bleibt für den rekursiven Filter bezogen auf die Bildspeicher gleich, durch das Herausziehen von Filterung und Interpolation aus der Iterationsschleife kann gleichzeitig sogar noch der Rechenaufwand reduziert werden. Die zeitlich rekursive Mehrreferenzbildfilterung kann bei der Verwendung eines Mittelwertfilters zu Instabilität führen, wenn die Detailverstärkung durch die zeitliche Rückführung zu groß wird. Aus diesem Grund sollte diese Filterstrategie nur mit dem adaptiven FIR-Filter verwendet werden, weil dann bei einem zu großem Unterschied zwischen aktuellem Bild und den Referenzbildern letztere von der Filterungen ausgeschlossen werden.

### Einfluss des Bildmodells auf das Ergebnis

Das Bildmodell gewährleistet eine eindeutige Lösung unabhängig vom der Initialisierung des hoch aufgelösten Bildes. Wie bereits in Kapitel 5 erwähnt, handelt es sich beim implementierten Huber Markov Zufallsfeld Bildmodell um eine adaptive, kantenerhaltende Rauschreduktion. Diese ist in den beiden linken Ausschnitten von Abbildung 6.12 erkennbar. Durch die Wahl des Steuerparameters  $t_\alpha$  kann die Stärke der Tiefpassfilterung und des Kantenerhalts gesteuert werden. In der Praxis hat sich eine Wahl dieses Faktors zwischen 5 und 15 als sinnvoll herausgestellt (vergleiche hierzu auch [Cai06]). In den rechten beiden Ausschnitten ist das modifizierte Bildmodell abgebildet. In diesem Fall wird anstelle des Kantenerhalts eine Anschärfung durchgeführt. Im Gesamtergebnis führt dies zu einer höheren Bildschärfe aber auch zu einer verminderten Wirkung der Rauschreduktion.



Abbildung 6.13: Feste oder adaptive Gewichtung des zeitlich vorigen Ergebnisbildes

Ähnlich wie die zeitlich rekursive Mehrreferenzbildfilterung muss auch dieses Bildmodell durch Analysemaße kontrolliert werden, da sonst eine Instabilität in Kantenbereichen möglich ist. Die einfachste Kontrolle wäre hier ein Heruntersetzen der Iterationszahl auf einen festen Wert.

### 6.3 Beurteilung der rein rekursiven Super-Resolution (ESTERA)

Da durch die Modifikationen des Datenmodells in Bezug auf die Bildgrößenänderung und Bewegungskompensation keine negativen, hingegen sogar positive Einflüsse beobachtet wurden, wird bei der Bewertung des rein rekursiven Algorithmus nur das modifizierte Datenmodell gemäß Abbildung 5.11 ohne interne Dezimation und Interpolation verwendet. In den nächsten Abschnitten soll nun zuerst auf die Kernkomponente des rekursiven Algorithmus, die gewichtete Selektion zwischen aktuellem niedrig aufgelöstem Eingangsbild und hoch aufgelöstem vorherigen Ergebnisbild eingegangen werden. Daran anschließend werden die systembedingten Verbesserungen evaluiert, die sich durch den Ersatz der Einzelkomponenten durch höherwertige Komponenten ergeben. Zum Schluss werden die Verbesserungen, die auf die Steuerung des Algorithmus mittels Bildanalysemaßen zurückzuführen sind, gezeigt.

### 6.3.1 Ergebnisse für unterschiedliche zeitliche Gewichtungsfaktoren

Der Einfluss unterschiedlicher fester oder adaptiver Gewichtungsfaktoren auf das Gesamtergebnis verdeutlicht Abbildung 6.13. In den ersten drei Ausschnitten sind Ergebnisse für eine feste Gewichtung mit unterschiedlichen Faktoren gezeigt. Wie zu erwarten, weisen die Bilder mit schwächerer Gewichtung des vorherigen Bildes eine geringe Schärfe und etwas weniger Details auf als die mit stärkerer Gewichtung. Dafür sind in diesen Bildern aber auch kaum Artefakte durch falsche Bewegungsvektoren wahrzunehmen. Bei höherem Beitrag des vorigen Ergebnisbildes verhält sich dies genau gegensätzlich. Unter Verwendung des adaptiven Gewichtungsfaktors kann ein ähnliches Resultat wie mit den beiden niedrigeren Werten für den festen Koeffizienten erreicht werden. Außerdem ist mit der adaptiven Gewichtung eine nahezu vollständige Unterdrückung der Seiteneffekte möglich.

### 6.3.2 Verbesserungen durch den Ersatz einzelner Komponenten im Algorithmus

In Abschnitt 5.5 wurde dargelegt, wie der rein rekursive SR-Algorithmus in eine erweiterte bildinhaltsabhängige Signalverarbeitung überführt werden kann. Hier sollen nun die Ergebnisse diskutiert werden, die sich durch die Verwendung höherqualitativer Komponenten oder die Abschaltung einzelner Module ergeben. Der linke Ausschnitt in Abbildung 6.14 verdeutlicht, dass durch das Weglassen der Gegenmaßnahme für Vektorfehler ein deutlich höherer Detailgrad ermöglicht wird. Im System findet dadurch keine frequenzselektive Gewichtung des Differenzsignals zwischen vorigem, bewegungskompensierten Ergebnisbildes und dem aktuellen Eingangsbildes statt. In der laufenden Sequenz ist dadurch vor allem auch eine deutlich stärkere Rauschreduktion wahrnehmbar, da jetzt nicht mehr die niederfrequenten Anteile des Differenzsignals unterdrückt werden. Diese Maßnahme ist aber nur im Falle sehr präziser Vektorfelder sinnvoll, da die Luminanzdifferenzen zwischen den Bildern, die sich zum Beispiel durch eine falsche Kompensation ergeben können, nicht mehr unterdrückt werden. Aus diesem Grund ist diese Maßnahme auch nur in Kombination mit dem adaptiven Gewichtungsfaktor sinnvoll. In Abbildung 6.14 ist im zweiten Ausschnitt von links zu sehen, welche Auswirkungen der Ersatz der örtlichen Bildschärfenanpassung (ursprünglich Unsharp Masking) durch ein bewegungsadaptives LTI [RDS<sup>+</sup>08] hat. Insbesondere Kanten können dadurch erheblich versteilert werden, ohne dass unnatürlich hohe Überschwinger entstehen. Allerdings führt dieses Verfahren nur in Kantenbereichen eine Manipulation der Pixel durch. Das verhindert einerseits eine Rauschanhebung in homogenen Flächen, andererseits werden nun Texturbereiche kaum verbessert.

Verzichtet man vollständig auf die örtliche Bildschärfekorrektur, so entsteht ein relativ unscharfes Ergebnis. In diesem Fall fungiert der Algorithmus nur noch als rekursive, örtlich-zeitliche Rauschreduktion, wobei die Gewichtung zwischen örtlicher und zeitlicher Filte-



Abbildung 6.14: Verbesserungen durch bildinhaltsabhängige Signalverarbeitung und adaptive Gewichtung der Komponenten

Die Gewichtung zwischen Daten- und Bildmodell hängt von der Gewichtung zwischen Daten- und Bildmodell ab. Der letzte Ausschnitt in Abbildung 6.14 ganz rechts zeigt die Verwendung der adaptiven örtlichen Bildschärfekorrektur gesteuert durch die Schätzung der minimalen gauß'schen Varianz, wobei in diesem Fall durch das sehr unscharfe Eingangsmaterial keine großen Gewinne zu verzeichnen sind. Auf Vorteile dieser Verarbeitungsweise wird in späteren Abschnitten aber noch genauer eingegangen.

### 6.3.3 Resultate unter Verwendung unterschiedlicher Vektorfeldauflösungen

Die Bewegungsschätzung hat eine erhebliche Wirkung auf die erzielbare Bildqualität. Nur bei korrekt geschätzter Bewegung ist eine Detailwiederherstellung überhaupt möglich. Gerade beim rekursiven Algorithmus kann aber auch die Auflösung des Vektorfeldes die Qualität sehr stark beeinflussen. Dies ist in Abbildung 6.15 illustriert. In dieser Arbeit kam bisher immer der verbesserte prädiktive Bewegungsschätzer zur Anwendung. Durch das interne Wideblock-Splitting entstehen bei diesem Verfahren sehr hoch aufgelöste Vektorfelder mit Ergebnisblockgrößen zwischen 1x1 und 2x2 Pixeln. In Bereichen um bewegte Objekte und insbesondere in Auf- und Verdeckungsrichtung entstehen aber örtlich stark variierende Vektoren durch diesen Algorithmus (Abbildung 6.15). Da in diesen Bereichen aber sowieso keine korrekte Schätzung möglich ist und Fehler unvermeidlich sind, ist dies nicht dramatisch. In der Super-Resolution führt das sogar zu einem positiven Nebeneffekt, weil sich dadurch bei der zeitlichen Filterung keine festen Fehlerraster ergeben, sondern



Abbildung 6.15: Einfluss der Vektorfeldauflösung auf das Ergebnis der Super-Resolution

ein eher zufällig verteiltes Rauschen. Verwendet man stattdessen ein sehr grob aufgelöstes Vektorfeld mit einer Blockgröße von 8x8, so sind die blockartigen Kompensationsfehler (bei Verwendung eines festen Gewichtungsfaktors) aus mehreren vorherigen Bildern im aktuellen Ergebnis wahrnehmbar. Weiterführende Untersuchungen zum Einfluss der Genauigkeit von Bewegungsvektorfeldern sind in [Naß09] durchgeführt worden.

### 6.3.4 Instabilität des Verfahrens und mögliche Gegenmaßnahmen

Eine Instabilität ist unter normalen Rahmenbedingungen bei Verwendung eines Gewichtungsfaktors im Bereich zwischen 0 und 1 nicht zu erwarten. Dies wurde auch bereits in Abschnitt 5.6 nachgewiesen. Möchte man aber eine besonders hohe Detailverstärkung durchführen, so kann es sinnvoll sein, den rekursiven Algorithmus außerhalb seiner mathematischen Stabilitätsgrenzen zu betreiben und die Stabilität nachträglich durch Kontrollinstanzen zu gewährleisten. In Abbildung 6.16 zeigt der linke Ausschnitt ein Ergebnisbild für einen instabilen Parametersatz. Dazu wurde eine extreme Verstärkung für die örtliche Bildschärfekorrektur und ein fester Gewichtungsfaktor mit einem hohen Wert von 0.9 gewählt. Durch eine Kontrastmessung (Abschnitt 5.5) und Heruntergewichtung des gesamten Detailsignals kann dieser Algorithmus aber trotzdem stabil gehalten werden, was die zweite Abbildung von links zeigt. Weiterhin kann alleine durch den adaptiven Gewichtungsfaktor die Tendenz zur Instabilität verringert werden, wobei ein sehr geringer Wert von  $\kappa$  in Gleichung 5.19 sogar Stabilität erzwingen kann, da dann die Gewichtung des vorigen Bildes sehr klein ist.



Abbildung 6.16: Gegenmaßnahmen gegen Divergenz und Kontrolle des Algorithmus durch Bildanalysemaße

## 6.4 Ergebnisse bei der Verwendung der Bildanalyse innerhalb der SR-Algorithmen

### 6.4.1 Beispiele für die Detail- und Kantenkarte

Wesentliche Bausteine zur adaptiven Gewichtung der einzelnen Signalanteile sind die Kanten- und Detailmaske, welche eine Einteilung des Bildes in unterschiedliche Bereiche ermöglichen. Ergebnisse für diese beiden Masken sind in Abbildung 6.17 dargestellt. Die Parameter zur Kantenerkennung sind so gewählt, dass nur sehr starke Kanten erkannt werden, da nur diese im SR-Algorithmus stark heruntergewichtet werden sollen. Hingegen ist die Detailerkennung sehr empfindlich eingestellt, um auch Texturen mit sehr niedrigem Kontrast detektieren zu können, so dass auch diese Bereiche durch die Super-Resolution verstärkt werden können.

### 6.4.2 Ergebnisse der Schätzung der minimalen gauß'schen Varianz für eine niedrig und eine hoch aufgelöste Sequenz

Um die Eignung der Schätzung der minimalen gauß'schen Varianz zur Erkennung unterschiedlich hoch aufgelösten Materials zu demonstrieren, wurde dieses Maß sowohl auf die hoch aufgelöste Sequenz 'seeking' als auch auf eine unterabgetastete und wieder interpolierte Version derselben Sequenz angewendet. Die Ergebnisse beider Messungen finden sich in Abbildung 6.18. Während für die interpolierte niedrig aufgelöste Sequenz in vielen Be-

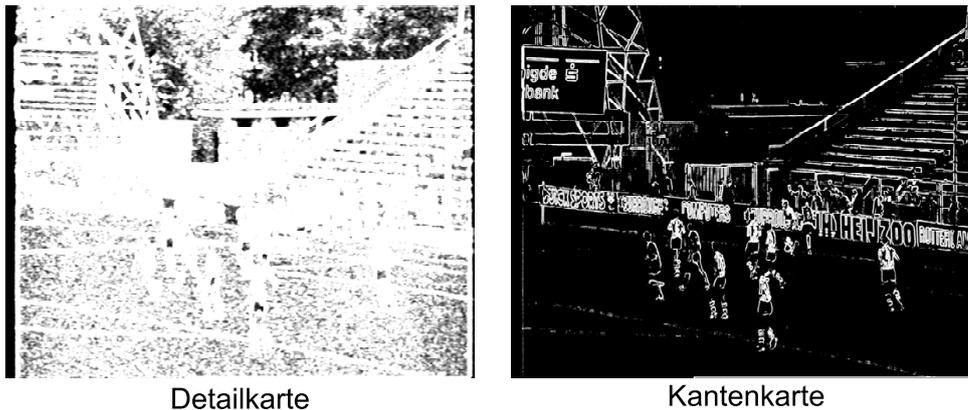


Abbildung 6.17: Detail- und Kantenkarte zur Steuerung des Algorithmus

reichen ein Filter mit einer niedrigeren Grenzfrequenz (weiße Pixelwerte) ausgewählt wird, liefert bei der bereits hoch aufgelösten Sequenz der Filter mit der höchsten Grenzfrequenz (schwarze Pixelwerte) meist den optimalen Wert, vor allem in Kanten- und Detailbereichen. Der Gewinn, der durch die Verwendung der minimalen Varianz zur Anpassung des Unschärfefilters entsteht, soll mit Abbildung 6.19 verdeutlicht werden. Während im Ausschnitt ohne Kontrollinstanzen eine deutliche Überschärfe gerade in Kantenbereichen und den Texturen wahrnehmbar ist, kann durch die Verwendung des adaptiven Degradaionsmodells ein deutlich angenehmerer Schärfegrad erreicht werden. Als Nebeneffekt bei der Verwendung eines Filterkerns mit einer besser angepassten Grenzfrequenz werden jetzt auch nur noch deutlich höherfrequente Anteile verstärkt, wodurch der hoch aufgelöste Eindruck erhalten bleibt. Werden hingegen auch die niedrigen und mittleren Frequenzanteile angehoben, entsteht der Eindruck einer niedriger aufgelösten Sequenz. Die Bildqualität kann zusätzlich noch weiter durch eine globale Kontrastanpassung gesteigert werden (Bild ganz rechts).

### 6.4.3 Bewegungsvektorlängenabhängige Gewichtung der Detailsignale

Besonders bei schnellen Bewegungen ist eine zuverlässige Bewegungsschätzung sehr schwierig. Aus diesem Grund wird zur Erhöhung der Robustheit das in der Super-Resolution berechnete Detailsignal mit zunehmender Geschwindigkeit heruntergewichtet. Abbildung 6.20 zeigt beispielhaft für eine Sequenz mit einer großen Bewegung, dass durch die Verwendung dieses Schutzmechanismus Seiteneffekte erheblich reduziert werden können. Trotzdem ist anzumerken, dass in diesem Fall eine rein örtliche Verarbeitung noch bessere Ergebnisse liefert, was die Grenzen der Möglichkeiten einer adaptiven Super-Resolution aufzeigt.

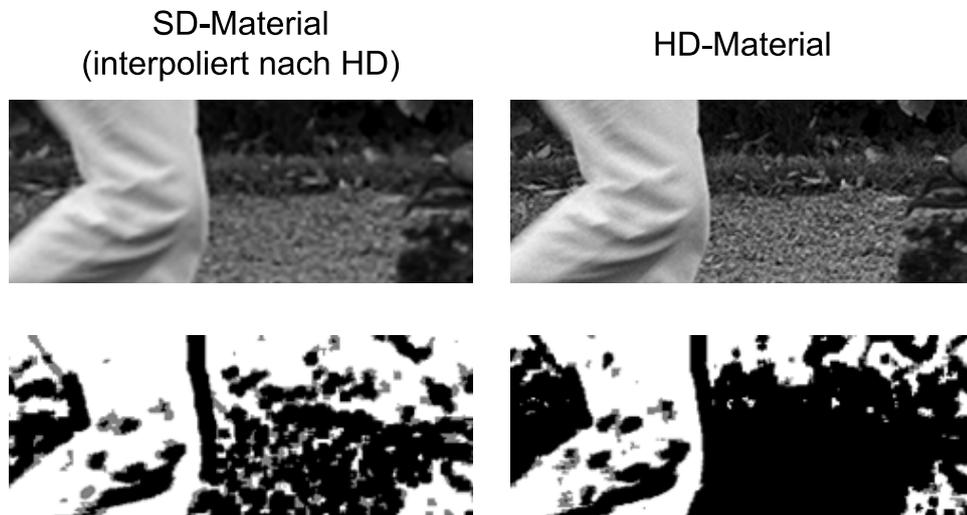


Abbildung 6.18: Bestimmung der minimalen lokalen Varianz für eine niedrig und eine hoch aufgelöste Sequenz. Bildbeispiel (oben) und Karte (unten) im Vergleich.



Abbildung 6.19: Anwendung der minimalen lokalen Varianzschätzung zur Kontrolle der Schärfe im Super-Resolution Algorithmus



Abbildung 6.20: Vermeidung von Bewegungsvektorfehlern bei hohen Geschwindigkeiten.

## 6.5 Vergleich mit anderen Interpolationsalgorithmen

Nachdem nun der Einfluss aller Parameter und der Analysemaße auf die Ergebnisse untersucht wurde, werden in diesem Abschnitt die entwickelten Super-Resolution Algorithmen mit anderen aus der Videosignalverarbeitung bekannten Interpolations- und Bildverbesserungsalgorithmen verglichen. Dazu wird neben Bildbeispielen auch auf objektive Messkriterien zurückgegriffen. Dabei kommen sowohl der PSNR als auch der SSIM [WS04] zum Einsatz. Da es sich bei beiden Maßen um Full-Reference Metriken handelt, muss neben dem eigenen Ergebnis auch die ideale Referenz vorliegen. Deshalb wurde für jede Sequenz mittels eines Dezimationstiefpasses und anschließendem Weglassen von Pixelwerten eine um den Faktor zwei verkleinerte Sequenz erzeugt. Diese konnte dann mittels der unterschiedlichen Algorithmen wieder auf die ursprüngliche Größe skaliert werden. Um zu starken Alias bei der Unterabtastung zu vermeiden, wurde ein Hamming-Filter als Dezimationstiefpass eingesetzt, dessen Grenzfrequenz auf die halbe Abtastfrequenz der hoch aufgelösten Sequenz eingestellt wurde.

### 6.5.1 Generelle Bemerkungen zur Verwendung der objektiven Maße

Bei beiden eingesetzten Maßen, dem PSNR und dem SSIM, handelt es sich um ein generelles Vergleichsmaß zwischen der verarbeiteten Sequenz und der Referenzsequenz. Während beim PSNR ein einfacher Zusammenhang zum mittleren quadratischen Fehler besteht, verwendet der SSIM komplexere Strukturmaße und bildet daher die subjektiv wahrgenommene Bildqualität besser ab [WS04]. Nachteilig ist aber bei beiden Maßen, dass zwischen

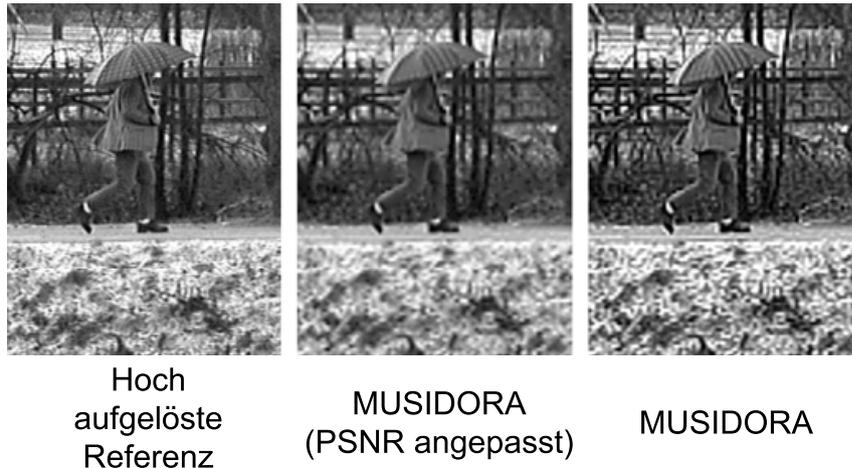


Abbildung 6.21: Vergleich von hoch aufgelöster Referenz und zwei SR-Ergebnissen mit unterschiedlichem Schärfe und Detailgrad

einer Abweichung nach unten (unschärfer als die Referenz) und nach oben (schärfer als die Referenz) bei beiden Maßen nicht unterschieden wird, subjektiv aber sehr unterschiedliche Bildeindrücke entstehen. Abbildung 6.21 verdeutlicht ein weiteres Problem der Maße exemplarisch für die Verwendung des PSNR. Wie dort zu sehen ist, können durch den relativ starken (und auch realistischen) Dezimationstiefpass die sehr hochfrequenten Anteile aus der Referenz nicht mehr hergestellt werden (Bildbeispiel ganz rechts). Um nun aber subjektiv trotzdem einen guten Bildeindruck zu erzielen, werden die mittleren Frequenzanteile und Details relativ stark angehoben. Da diese aber in der Referenzsequenz weniger stark ausgeprägt sind, führt diese Anhebung zu einem relativ großen Fehlerwert und damit zu niedrigen Bewertungen durch die objektiven Maße. Im Vergleich dazu ist in der Mitte eine Verarbeitung durch die Super-Resolution mit nach objektiven Gesichtspunkten besser geeigneten Parametern dargestellt. Es ist ersichtlich, dass diese Sequenz aber auch deutlich weniger Details enthält. Für die weiteren Betrachtungen wurden beide Parametersätze untersucht, um sowohl subjektive als auch objektive Kriterien besser berücksichtigen zu können.

### 6.5.2 Bildbeispiele und objektive Messwerte für unterschiedliche Algorithmen

In diesem Abschnitt werden Messergebnisse sowohl für direkte als auch für iterative Rekonstruktionsalgorithmen gezeigt. Als Basisvergleich dient in diesem Fall die High Resolution Spline (HRS) Interpolation, da diese auch innerhalb der Super-Resolution verwendet wird. Um einen faireren Vergleich zu ermöglichen, wird zusätzliche noch eine Variante mit einer

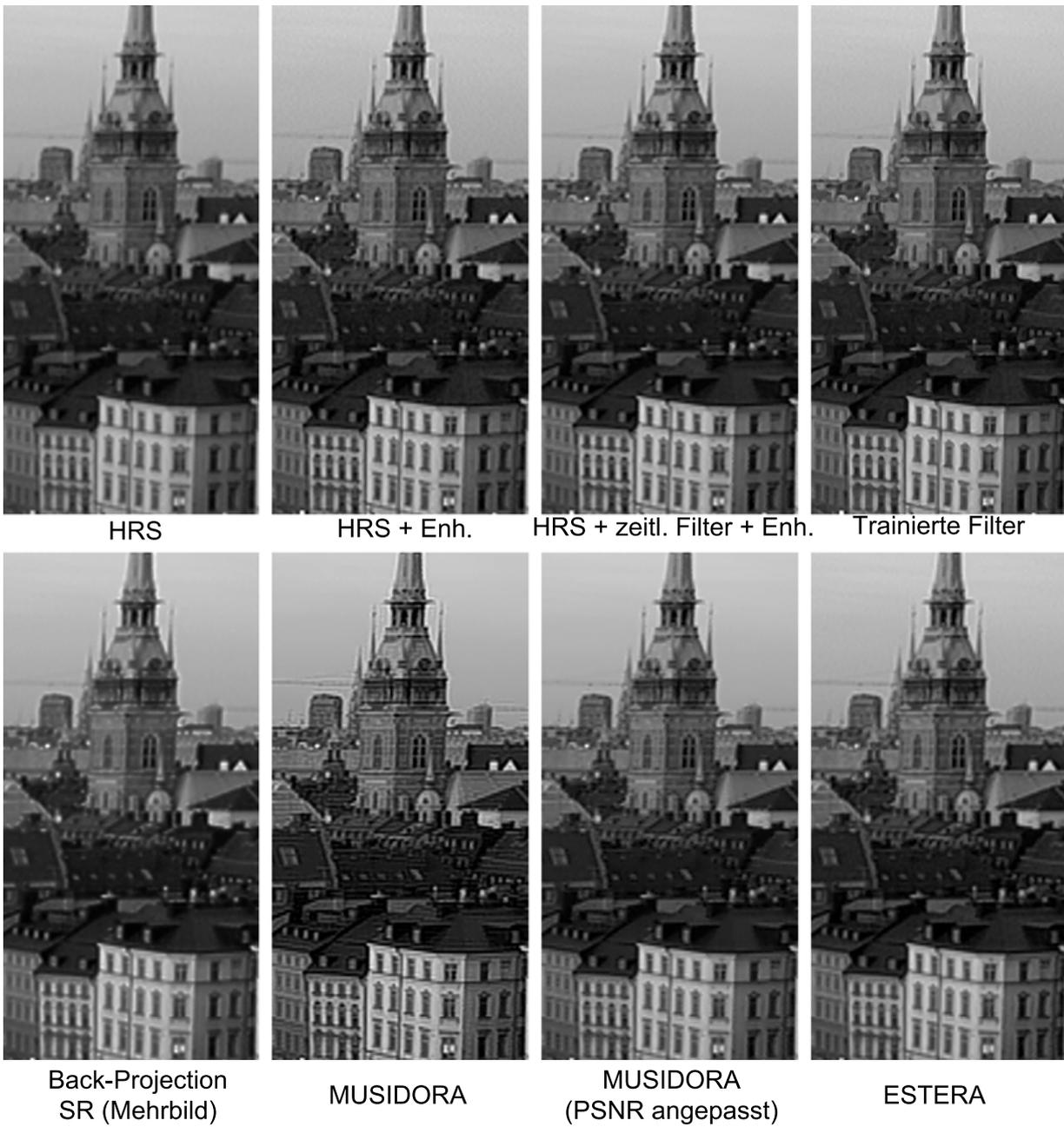


Abbildung 6.22: Ergebnisse der Sequenz 'stockholm' für unterschiedliche Interpolationsstrategien

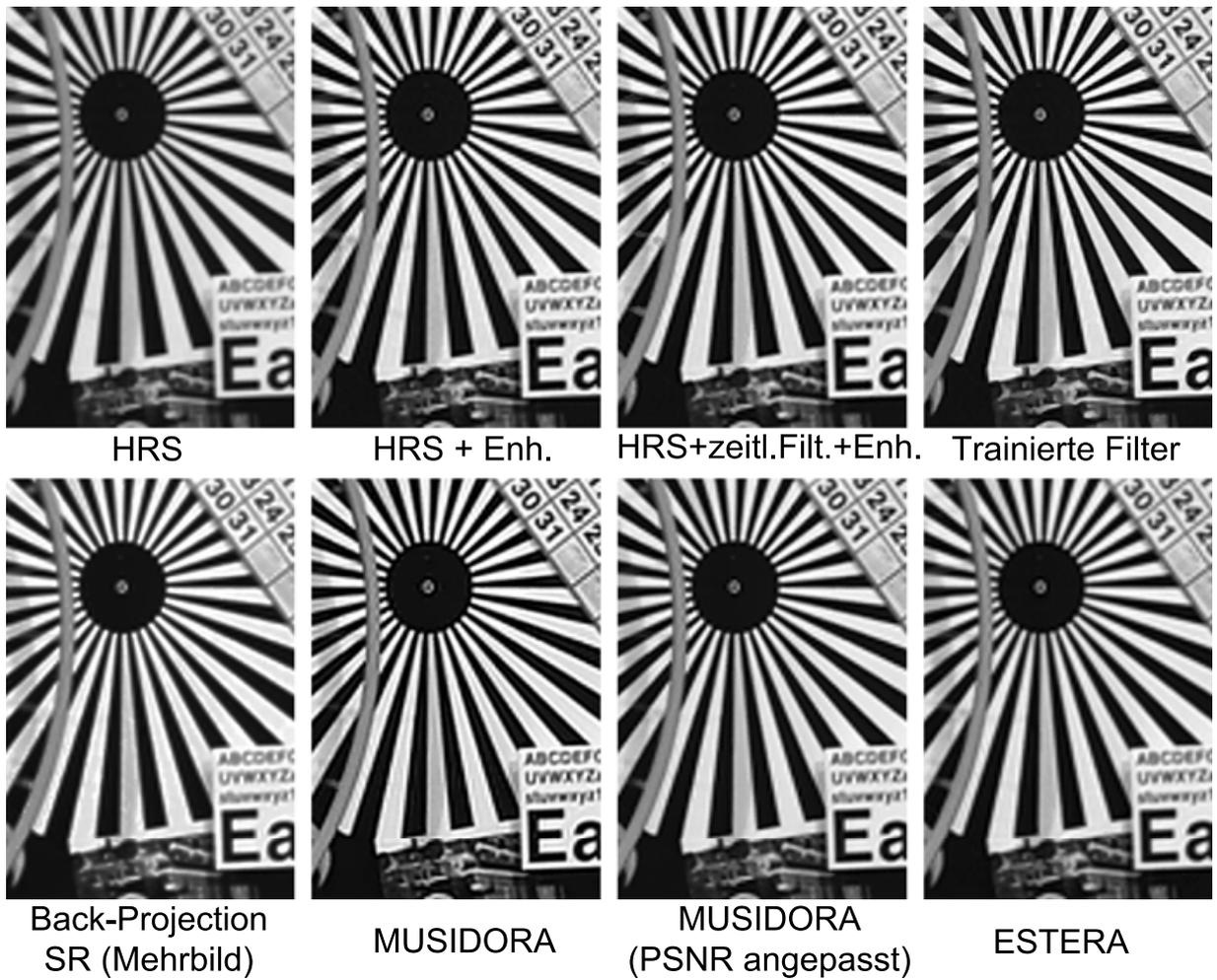


Abbildung 6.23: Ergebnisse der Sequenz 'wheel' für unterschiedliche Interpolationsverfahren

nachträglichen Bildschärfekorrektur (HRS + Enh.) und eine Kombination aus Interpolation, zeitlicher FIR-Filterung und nachträglicher Bildschärfekorrektur (HRS + Temp. Filt. + Enh.) herangezogen. Als beste rein örtliche Verfahren werden die trainierten Filter [Kon01], [Len09] verwendet. Dabei sei angemerkt, dass die letzten beiden Verfahren bereits zu den Super-Resolution Algorithmen gezählt werden können (siehe Kapitel 2 und 3).

Als Super-Resolution Algorithmen werden sowohl der voll parametrisierbare Back-Projection Algorithmus (VESBA) und der rein rekursive Algorithmus (ESTERA) verwendet, wobei jeweils eine Variante mit Standardeinstellungen aus Anhang A und eine Variante mit zugeschalteten Kontrollinstanzen und adaptiver Filterung untersucht wird. Bei der verbesserten Mehrreferenzbildschätzung handelt es sich um den doppelt rekursive MUSIDORA. Beide Mehrreferenzbildalgorithmen, VESBA und MUSIDORA, nutzen je fünf zeitlich aufeinanderfolgende Eingangsbilder, da das einen guten Kompromiss zwischen Rechenaufwand und Qualität darstellt. Neben dem nach subjektiven Gesichtspunkten optimierten MUSIDORA wird auch eine Variante, die nach objektiven Kriterien gute Ergebnisse liefert, mitsimuliert.

In Abbildung 6.22 sind für die Sequenz 'stockholm' die Ergebnisse aller Algorithmen gezeigt. Im Vergleich zur High Resolution Spline Interpolation liefern sowohl die Variante mit nachträglicher Bildschärfenanhebung als auch die trainierten Filter deutlich schärfere Ergebnisse. Jedoch kann bei diesen auch eine Rauschverstärkung wahrgenommen werden, die in der laufenden Sequenz noch stärker zu sehen ist. Dieses Rauschen kann durch die zeitliche Filterung (zweiter Ausschnitt von rechts in der oberen Zeile) merklich verringert werden. Die Ergebnisse der Super-Resolution weisen alle einen sehr geringen Rauschpegel auf, wobei die Verfahren mit mehr als einem Referenzbild erwartungsgemäß noch etwas stärker abschneiden. Den höchsten Detailgrad liefert mit Abstand der nach subjektiven Kriterien optimierte Mehrreferenzbild SR-Algorithmus MUSIDORA. Dabei ist anzumerken, dass der Detailgrad entlang der horizontalen Bewegungsrichtung deutlich höher ist als orthogonal dazu, was im Bereich der Dächer gut zu beobachten ist. Dieses Verhalten kann durch die rein horizontale Bewegung erklärt werden, weil nur bei Bewegung eine Detailverbesserung möglich ist. Vergleicht man die subjektiven Eindrücke mit den PSNR-Werten aus Tabelle 6.1 so wird deutlich, dass der subjektive Bildeindruck für HRS, HRS mit nachträglicher Bildschärfekorrektur und der Variante mit zusätzlicher zeitlicher Filterung gut wiedergegeben werden. Der sehr subjektiv eingestellte MUSIDORA-Algorithmus führt aber zu einer sehr hohen Bestrafung durch das objektive Maß. Hingegen ist der nach objektiven Kriterien beste SR Algorithmus in der Lage, für die meisten Sequenzen deutlich bessere Ergebnisse zu erzielen. Dieses Verhalten widerspricht aber den Eindrücken bezüglich der subjektiven Bildqualität.

Als zweites Beispiel sollen noch die Ergebnisse für die Sequenz 'wheel' gezeigt werden, da mit dieser die Interpolationsqualität diagonaler Kanten und das Verhalten der zeitlichen Verarbeitung für sehr komplexe Bewegungen gut gezeigt werden kann. Die diagonalen Kanten können erwartungsgemäß sehr gut mit den trainierten Filtern verarbeitet werden, da

Tabelle 6.1: PSNR-Werte für unterschiedliche Interpolationsalgorithmen

	parkrun	stockholm	wheel	matfight	football
HRS	23.88	31.47	28.82	39.08	29.33
HRS Enh.	24.36	31.68	28.78	38.81	29.62
HRS Enh. Temp	24.45	31.74	29.13	37.73	30.04
VESBA	24.11	31.57	28.19	38.30	29.51
MUSIDORA	22.91	27.71	26.96	33.74	26.72
MUSIDORA (PSNR opt.)	24.46	31.26	29.36	36.93	29.95
ESTERA	23.26	30.53	27.94	37.24	29.38
ESTERA (verb.)	24.06	30.49	28.53	34.82	29.18
Trainierte Filter	23.79	30.97	28.82	37.96	28.80

diese für solche Bildsituationen angelernt wurden. Es kann aber auch beobachtet werden, dass die integrierten Bildmodelle in den SR-Algorithmen und die gute Bewegungsschätzung im Bereich der Scheibe zu deutlich geringer ausgeprägten Treppenartefakten führen als bei der rein örtliche High Resolution Spline Interpolation alleine oder der Version mit nachträglicher Bildschärfekorrektur. Schwächen der Super-Resolution sind aber im Bereich der Speichen zu beobachten, da hier die Bewegung nicht exakt geschätzt werden konnte und dieser Bereich für die Schutzmaßnahme durch die adaptive zeitliche Gewichtung zu wenig Kontrast aufweist.

Zusätzlich zu den PSNR-Werten, die neben diesen beiden Sequenzen auch noch für andere Beispiele simuliert wurden, beinhaltet Tabelle 6.2 noch die SSIM-Werte. Beim Vergleich der PSNR- und SSIM-Werte fällt aber auf, dass diese zum Teil zu einer unterschiedlichen Bewertungsreihenfolge führen.

## 6.6 Zusammenfassung

In diesem Kapitel wurde verdeutlicht, welchen Einfluss die einzelnen Parameter innerhalb der Super-Resolution Algorithmen haben. Durch diese Untersuchungen wurde bestätigt, dass die in Kapitel 5 bei der Entwicklung der Algorithmen vorgenommen Modifikationen zur Reduktion des Rechenaufwandes entweder keinen Einfluss auf die Bildqualität haben oder sich sogar positiv auf diese auswirken. Dies gilt sowohl für den rein rekursiven Super-Resolution Algorithmus ESTERA als auch für den doppelt rekursiven Algorithmus mit mehreren Referenzbildern MUSIDORA. In zahlreichen Experimenten wurde gezeigt, dass durch die zusätzlich eingefügten Kontrollinstanzen die Robustheit gegenüber Seiteneffekten deutlich gesteigert werden konnte und somit für sehr unterschiedliches Eingangsmate-

Tabelle 6.2: SSIM-Werte für unterschiedliche Interpolationsalgorithmen

	parkrun	stockholm	wheel	matfight	football
HRS	0.790	0.858	0.937	0.976	0.863
HRS Enh.	0.8389	0.876	0.934	0.970	0.880
HRS Enh. Temp	0.835	0.867	0.936	0.957	0.879
VESBA	0.819	0.862	0.935	0.968	0.873
MUSIDORA	0.792	0.781	0.913	0.943	0.818
MUSIDORA (PSNR opt.)	0.828	0.858	0.941	0.966	0.875
ESTERA	0.823	30.53	0.865	0.966	0.886
ESTERA verb.	0.816	0.858	0.928	0.949	0.870
Trainierte Filter	0.819	0.866	0.939	0.966	0.869

rial hervorragende Ergebnisse erzielt werden können. Trotzdem muss angemerkt werden, dass Super-Resolution Algorithmen Grenzen gesetzt sind. Sind bestimmte Voraussetzungen nicht erfüllt, z.B. fehlen in den Sequenzen einmalige Informationen in jedem Bild die ihre Ursache in Subpixelbewegung, Alias oder einem unterschiedlichen Störungsgrad haben, so ist die Leistungsfähigkeit dieser Verfahren nicht wesentlich besser als die der Standardverfahren. Weiterhin wird die erhöhte Robustheit durch einen geringeren Grad an Detailwiederherstellung erkauft. Unter Idealbedingungen und für den Fall einer guten Bewegungsschätzung können aber, sowohl für synthetisch erzeugte als auch reale Sequenzen, Ergebnisse erzielt werden, die deutlich besser als die von Standardinterpolationsverfahren und sogar bekannter Super-Resolution Algorithmen sind. Verfahren, die mehr als ein Referenzbild verwenden, schneiden dabei erwartungsgemäß am besten ab. Dies gilt sowohl für die Detailwiederherstellung als auch für die Vermeidung von Seiteneffekten, die für Mehrreferenzbildalgorithmen deutlich geringer ist. Die Entscheidung für den Einsatz eines Algorithmus hängt also vor allem auch von den zur Verfügung stehenden Ressourcen ab.

Neben der Anwendung der Super-Resolution zur Interpolation von Bildsequenzen kann durch die zeitliche Filterung der Detailreichtum auch in anderen Situationen gesteigert werden. Gerade in der heutigen Zeit existiert durch die Verbreitung von Web-Portalen oder Handy-Videos einer sehr große Menge an Videomaterial sehr niedriger Qualität mit starken Kompressionsartefakten. Je nach Codierstandard entsteht so von Bild zu Bild eine stark variierende Informationsmenge, die durch die Super-Resolution in jedem Einzelbild erhöht werden kann. In Abbildung 6.24 ist zu sehen, dass bei alleiniger Anwendung der Super-Resolution Algorithmen in bewegten Bildbereichen die Sichtbarkeit der Artefakte verringert werden kann. Das feststehende Blockraster wird hingegen deutlich verstärkt, da es fälschlicherweise als Detail interpretiert wird. Schaltet man aber den Algorithmen



Abbildung 6.24: Deblocking und Interpolation niederqualitiver Videos

ein adaptives Deblocking-Verfahren (z.B. das aus [Pia08]) vor, so kann die Videoqualität erheblich gesteigert werden. Vergleicht man die Ergebnisse der Super-Resolution aber mit einem speziell für niederratige Videos entwickelten Algorithmus (örtlich-zeitliche Regularisierung, Abschnitt 4.2) so fällt auf, dass dieser noch deutlich leistungsfähiger in Bezug auf die Störungsreduktion ist. Zum Schluss soll noch einmal kurz die Vielseitigkeit der beiden Super-Resolution Algorithmen ESTERA und MUSIDORA erwähnt werden. Gerade der rein rekursive Algorithmus, der in Form einer bildinhaltsabhängigen Signalverarbeitung realisiert werden kann, kann durch Zu- und Abschaltung der einzelnen Blöcke für verschiedene Aufgabenbereiche eingesetzt werden. So ist neben der Interpolation durch die Heruntergewichtung der örtlichen Bildschärfekorrektur eine Rauschreduktion möglich oder das Verfahren kann für die generelle Bildverbesserung oder zur Reduktion von Bewegungsunschärfe eingesetzt werden, sofern dass bewegungsadaptive LTI in diesen Algorithmus integriert wird. Ergebnisse für all diese Anwendungsbereiche zeigt Abbildung 6.25



Abbildung 6.25: Allgemeine Bildverbesserung durch den Super-Resolution Algorithmus



# Kapitel 7

## Zusammenfassung und Schlußfolgerungen

In dieser Arbeit wurden die Einsatzmöglichkeiten von (iterativen) Super-Resolution Algorithmen in der Videosignalverarbeitung untersucht. Die Motivation für diese Untersuchungen waren sehr vielseitig. So versprechen iterative Rekonstruktionsalgorithmen im Allgemeinen und Super-Resolution Algorithmen im Speziellen eine sehr leistungsfähige Videosignalverarbeitung, da diese Algorithmen zumeist (interne) Bildmodelle verwenden und auf globalen Optimierungsstrategien basieren, weshalb eine Anpassung an und gute Verarbeitung von Material mit sehr unterschiedlicher Eingangsqualität möglich ist. Weiterhin liegt eine große Menge des Videomaterials in geringer Auflösung vor, so dass mit Super-Resolution Algorithmen die Auflösung dieses Materials deutlich erhöht werden kann, was sich positiv auf die Darstellung auf hoch auflösenden Displays auswirkt. Gerade die iterative Verarbeitungsweise der meisten SR-Algorithmen, die eine Korrektur der Ergebnisse und schrittweise Anpassung hin zu einer optimalen Qualität durch eine Rückführungsschleife ermöglicht, deutet auf großes Potential hin.

Diese Algorithmen haben ihren Ursprung aber zumeist in anderen Bereichen wie dem maschinellen Sehen, weshalb eine direkte Übernahme der Algorithmen nicht möglich ist, sondern eine Anpassung an die Rahmenbedingungen der Videosignalverarbeitung erfolgen muss. So sind möglichst geringe Hardwarekosten und die Verarbeitung von Videos mit sehr unterschiedlicher Qualität und Auflösung bei der Entwicklung zu berücksichtigen. Insbesondere die Robustheit gegenüber Störungen oder Verletzung von Annahmen der internen Bildmodelle stellt eine große Herausforderung dar. Weiterhin verursacht die iterative Verarbeitung bis hin zu einem Optimum eine nicht vorhersagbare Laufzeit, die einen Einsatz in der Videosignalverarbeitung verhindert. Möglichkeiten zur Überwindung dieser Probleme sowie die Verbesserung bekannter Super-Resolution Algorithmen bildeten deshalb den Schwerpunkt dieser Arbeit.

Die Kapitel 2 und 3 gaben einen Überblick über bereits bekannte iterative Rekonstruktionsalgorithmen sowie benötigte mathematische Werkzeuge zur effizienten Implementierung dieser Algorithmen. Mit diesen Werkzeugen ist dann eine Formulierung von Algorithmen

für unterschiedliche Anwendungsbereiche möglich, wobei in dieser Arbeit schwerpunktmäßig Algorithmen zur Auflösungserhöhung und Artefaktreduktion vorgestellt werden. Um die konzeptionellen Vor- und Nachteile iterativer Rekonstruktionsalgorithmen, zu denen die untersuchten Super-Resolution Algorithmen zählen, zu verstehen, wurden die bestehenden Algorithmen in zwei Klassen unterteilt: die direkten und iterativen Algorithmen. Während die direkten Algorithmen ihre Verarbeitung nur einmal auf die Eingangssequenz anwenden und eine nachträgliche Korrektur in einem angepassten Verarbeitungsschritt desselben Algorithmus nicht vorgesehen ist, können die iterativen Rekonstruktionsalgorithmen diese nachträgliche Korrektur und Anpassung über eine in den Algorithmus integrierte Rückführungsschleife vornehmen. Der Algorithmus wird deshalb so lange durchlaufen, bis das (objektiv) optimale Ergebnis erreicht wird. Dieser Qualitätsgewinn wird durch eine nicht vorhersagbare Rechenzeit erkauft. Die Formulierung dieser Algorithmen über einen mathematischen Ansatz führt zu weiteren Problemen: so müssen die verwendeten Modelle mathematische Rahmenbedingungen wie z.B. Differenzierbarkeit erfüllen und dies erschwert die Beschreibung komplexer Bildmerkmale wie Texturen. In weiteren Schritten müssen diese Ansätze dann in ausführbare Algorithmen überführt werden.

Um die Rahmenbedingungen der Videosignalverarbeitung zu erfüllen und die konzeptionsbedingten Nachteile zu überwinden, wurde in dieser Arbeit ein an die Videosignalverarbeitung angepasstes Entwurfskonzept für iterative Rekonstruktionsalgorithmen vorgeschlagen. Dieses Entwurfskonzept besteht aus insgesamt vier Schritten. Die ersten beiden Schritte sind dabei bereits beim Entwurf iterativer Rekonstruktionsalgorithmen zu finden. In diesen beiden Schritten wird der Algorithmus selbst abgeleitet, so dass durch sie das Einsatzgebiet sowie der Hardwareaufwand festgelegt wird. Die beiden darauf folgenden Schritte wurden in dieser Arbeit definiert. In diesen wird der Algorithmus weiter verbessert, wobei diese Verbesserungen nun vor allem die Bildqualität sowie die Vielseitigkeit des Algorithmus betreffen. Durch die Modifikationen ist dann noch eine abschließende Analyse sowie Stabilitätsprüfung nötig.

Im Detail wird bei diesem Konzept folgendermaßen vorgegangen: Im ersten Schritt ist ein geeigneter mathematischer Ansatz, beispielsweise ein Energiefunktional, auszuwählen, mit dem die gestellte Aufgabe beschrieben werden kann. Die Auswahl dieses Ansatzes kann je nach Problemstellung auch variieren. Dann wird in einem zweiten Schritt dieser mathematische Ansatz mittels entsprechender Werkzeuge in eine rekursive, adaptive Filterstruktur transformiert. Variationsrechnung, partielle Differentiation oder iterative Optimierungsstrategien sind nur einige Beispiele dafür, wie diese Transformation durchgeführt werden kann. Die so hergeleitete Filterstruktur kann dann durch ein Blockschaltbild beschrieben werden, welches eine einfache und übersichtliche Analyse hinsichtlich der Stärken und Schwächen des Algorithmus ermöglicht. Basierend auf Erfahrungswerten oder auch anderen Modellen aus der Videosignalverarbeitung, kann dann der Algorithmus im dritten Schritt verbessert werden. Das kann durch den Austausch einzelner Komponenten im Algorithmus durch höher qualitative Komponenten geschehen. Zusätzlich kann der gesamte

---

Algorithmus bzw. einzelne Komponenten durch Maße, die durch eine Bildanalyse bestimmt werden, kontrolliert werden. Der Vorteil ist nun, dass diese Komponenten nicht mehr unbedingt an die zuvor geltenden mathematischen Rahmenbedingungen gebunden sind und somit z.B. auch nichtlineare Komponenten leicht verwendet werden können. In einem vierten Schritt ist dann aber zu prüfen, ob der so entstehende Algorithmus weiterhin stabil ist. Dies kann durch Experimente und Abschätzungen bzw. theoretische Betrachtungen erfolgen.

Mit diesem Entwurfskonzept wurde dann in Kapitel 4 zuerst ein bereits bekannter Regularisierungsalgorithmus abgeleitet und durch nachträgliche Verbesserungen Schwächen überwunden. Darauf aufbauend wurde gezeigt, wie mit einem verbesserten Ansatz und einem weiteren zeitlichen Bildmodell ein neuer, örtlich-zeitlicher Regularisierungsalgorithmus entwickelt werden kann. Nachdem exemplarisch zwei Verfahren mit diesem Entwurfskonzept hergeleitet wurden, erfolgte die Anwendung auf die Algorithmen zur Super-Resolution in Kapitel 5. Für einen Back-Projection Algorithmus, der mehrere Eingangsbilder für die Berechnung eines Ausgangsbild verwendet, wurde dazu gemäß der ersten beiden Schritte der Ansatz in eine Filterstruktur umgeformt, die sich so bereits in der Literatur finden lässt. Im dritten Schritt erfolgte dann die Verbesserung des Algorithmus, indem das interne Modell (beschrieben durch Dezimation, Interpolation und Bewegungskompensation) modifiziert wurde, so dass sich ein Algorithmus mit geringerem Rechenaufwand bei gleichzeitig besserer Bildqualität ergibt. Diese Verbesserungen ergaben sich durch ein Herausziehen der Interpolation und zeitlichen Verarbeitung aus der Iterationsschleife.

Basierend auf diesen Erkenntnissen konnte ein noch leistungsfähigerer Algorithmus entwickelt werden, der auf einem doppelt rekursiven Algorithmus beruht (MUSIDORA). Dieser Algorithmus baut auf einem modifizierten Maximum-A-Posteriori Ansatz nach Stevenson auf, der die bereits untersuchten Modifikationen des vorherigen Back-Projection Algorithmus bereits im Ansatz berücksichtigt. Als Ergebnis ergibt sich so in der inneren Rekursionsschleife ein iterativer Deblurring-Algorithmus, der Details und Bildschärfe deutlich verstärkt. Durch ein in dieses Iterationsschema zusätzlich integriertes Bildmodell kann weiterhin die Rauschanhebung kontrolliert werden. Die äußere Rekursionsschleife besteht aus einem zeitlich rekursiven adaptiven Mehrreferenzbildfilter (bewegungskompensiert), der als Eingangsbilder die bis auf eine evtl. nötige initiale Interpolation unverarbeiteten aktuellen und nachfolgenden Eingangsbilder sowie die bereits verbesserten hoch aufgelösten vorherigen Ergebnisbilder verwendet. Somit profitiert das Ergebnis dieser Filterung auch von der Berechnung vorheriger Super-Resolution Ergebnisse, wodurch mehr Details wiederhergestellt werden können und sich ein zeitlich homogeneres Verhalten (z.B. weniger Unruhe durch Flimmern oder Detailschwankungen) in der Ergebnissequenz ergibt. Durch zusätzliche Kontrollinstanzen, welche später noch kurz zusammengefasst werden, kann eine weitere Qualitätsverbesserung erreicht werden.

Um für die Videosignalverarbeitung auch eine noch aufwandgünstigere Alternative bereitzustellen, wurde zusätzlich die Möglichkeiten einer zeitlich rein rekursiv arbeitenden

Super-Resolution untersucht, die nur das aktuelle niedrig aufgelöste Eingangsbild und das vorherige Ergebnisbild benötigt. Dieses als ESTERA bezeichnete Verfahren setzt erst in Schritt drei des Entwurfskonzeptes an, indem die Filtergleichung der Back-Projection modifiziert wird und das bildweise arbeitende Iterationsschema in eine zeitlich rekursive Verarbeitung überführt wird. Zu diesem Zweck muss die vorherige Lösung in bewegungskompensierter Form gebildet werden. Wendet man pro Bild nur diese eine Iteration an, so kann der Rechenaufwand erheblich gesenkt werden und es entsteht sogar eine feste, vorhersagbare Laufzeit des Algorithmus. Im Falle von perfekter Bewegungskompensation lassen sich im Vergleich zum rein örtlichen Iterationsschema exakt identische Ergebnisse erzielen. Da die perfekte Kompensation in der Praxis aber unmöglich ist, müssen zusätzliche Kontrollinstanzen zur Erhöhung der Robustheit integriert werden. Dazu muss die Rückführungsschleife verbessert werden. Durch eine gewichtete Selektion aus vorherigem, bewegungskompensiertem Ergebnisbild und aktuellem Eingangsbild kann die Sichtbarkeit von Artefakten verringert und der Detail- und Schärfegrad besser eingestellt werden.

Die Qualität der Ergebnisse kann noch weiter gesteigert werden, wenn die einzelnen Komponenten des Algorithmus durch eine Merkmalsanalyse gesteuert werden. So kann durch eine Heruntergewichtung des Signals an Kanten eine zu hohe Schärfe vermieden werden oder die Kontrastzunahme zwischen Eingangs- und Ausgangsbild gezielt eingestellt werden. Eine adaptive Steuerung der internen Modelle im Algorithmus ermöglicht weiterhin eine Anpassung des Algorithmus an Eingangsmaterial von unterschiedlichen Quellen. So kann der Algorithmus sowohl für interpolierte und damit unscharfe Sequenzen als auch für HD-Eingangsmaterial die nötige Schärfe- und Detailzunahme gezielt steuern. Diese Kontrollinstanzen machen den Einsatz der Super-Resolution in der Videosignalverarbeitung überhaupt erst möglich, da ohne diese die Seiteneffekte einfach zu stark sind und dann die Nachteile überwiegen würden. Dieses Problem konnte selbst durch die hier beschriebenen relativ einfachen Analysemaße behoben werden.

Durch den letzten Schritt des Entwurfskonzeptes, dem Analyseschritt, konnte durch eine Umstellung der Systemgleichung für den rekursiven Algorithmus gezeigt werden, dass die Super-Resolution in eine Form überführt werden kann, die eine bildinhaltsabhängige Signalverarbeitung repräsentiert bzw. als solche interpretiert werden kann. Bei der bildinhaltsabhängigen Signalverarbeitung wird das Bild durch mehrere Algorithmen parallel verarbeitet und dann – basierend auf einer Merkmalsanalyse – die beste Verarbeitung für die jeweilige Situation ausgewählt. Diese Beschreibungsform ermöglicht weitere Verbesserungen, beispielsweise das Weglassen bestimmter Komponenten im Algorithmus oder die Verwendung höherwertiger Komponenten. So kann der Algorithmus zur Rauschreduktion, zur generellen Bildschärfeverbesserung oder zur Reduktion von Bewegungsunschärfe eingesetzt werden und damit der Einsatzbereich erheblich erweitert werden.

Die Stabilität aller entwickelten SR-Algorithmen konnte mit Hilfe einer geometrischen Reihe gezeigt werden. Unter Berücksichtigung eines festen Wertebereiches der Parameter kann Stabilität für beide neu entwickelten Algorithmen (ESTERA und MUSIDORA) und auch

---

für den Back-Projection Algorithmus (VESBA) garantiert werden. Diese Ergebnisse konnten auch experimentell bestätigt werden. Darüber hinaus ist es mit den Kontrollinstanzen sogar möglich, den Algorithmus selbst bei einem zur Instabilität führenden Parametersatz zu kontrollieren. Dies kann wünschenswert sein, wenn die Detailzunahme sehr hoch ausfallen soll, was in manchen Fällen allerdings nur bei sehr großen, zur Instabilität führenden Parametersätzen möglich ist.

In der Auswertung (Kapitel 6) wurde mittels einer synthetisch erzeugten Sequenz die Fähigkeit zur Detailwiederherstellung unter bestimmten Rahmenbedingungen gezeigt. Der Einfluss von (örtlichem) Alias in der Sequenz und Bewegung konnte so gezielt untersucht werden. Diese Messungen konnten ebenfalls über ein einfaches lineares Modell systemtheoretisch erklärt werden. Darüber hinaus wurde dargelegt, dass durch die Einführung der neuen Komponenten die Robustheit deutlich gesteigert werden kann, allerdings auch der Grad der Detailwiederherstellung sinkt. Für diese beiden Kriterien muss ein Kompromiss gefunden werden, der sehr stark von der Zuverlässigkeit der Bewegungsschätzung abhängt. Je zuverlässiger diese ist, desto geringer kann die Robustheit des SR-Algorithmus ausfallen.

Weitere Untersuchungen zeigten den Einfluss aller wichtigen Parameter auf das Ergebnis sowie die Wirkungsweise der Einzelkomponenten. Beim Vergleich mit anderen örtlichen und örtlich-zeitlichen Algorithmen zur Interpolation und Bildverbesserung wurde deutlich, dass die Super-Resolution zu besseren subjektiven Ergebnissen führen kann. Dieser Eindruck lässt sich aber durch objektive Maße (SSIM, PSNR) nur teilweise bestätigen. Dies ist auf zwei Gründe zurückzuführen. Zum einen verfügen die verwendeten Referenzsequenzen über Rauschen, was durch die SR entfernt wird, jedoch zu einer Abweichung von der Referenz führt und diese im Maß bestraft wird, obwohl die Qualität eigentlich besser wird. Zum anderen können sehr hochfrequente Details bei Fehlen von Alias nicht mehr hergestellt werden, jedoch werden mittelfrequente Details deutlich angehoben. Das führt zu einem guten Bildeindruck, allerdings entstehen so ebenfalls starke Abweichungen von der Referenzsequenz, weil die mittelfrequenten Anteile dort nicht so stark ausgeprägt sind. Diese Abweichungen fallen mindestens genauso groß aus wie bei den anderen Algorithmen, die diese mittelfrequenten Details nicht anheben können, so dass die subjektiv höher aufgelöste Sequenz meistens schlechter bewertet wird.

Abschließend soll nun mit Hilfe von Abbildung 7.1 ein qualitatives Fazit gezogen werden. Unterschieden wird dabei zwischen zwei Szenarien, der Anwendung der Algorithmen auf Material mit sehr niedriger und Material mit sehr hoher Qualität. Ersteres Material findet sich sehr oft auf Web-Portalen oder bei Aufnahmen von Videos durch Handykameras und es weist oft eine niedrige Auflösung und starke Kompressionsartefakte auf. Erwartungsgemäß können Mehrreferenzbildverfahren (MUSIDORA, örtl.-zeitl. Regularisierung) deutlich mehr Details wiederherstellen sowie Artefakte beseitigen, da die zeitliche Filterwirkung deutlich stärker als bei rein rekursiven Algorithmen ist. Um die Verstärkung von Artefakten durch die SR-Algorithmen zu vermeiden, insbesondere wenn diese über mehre-

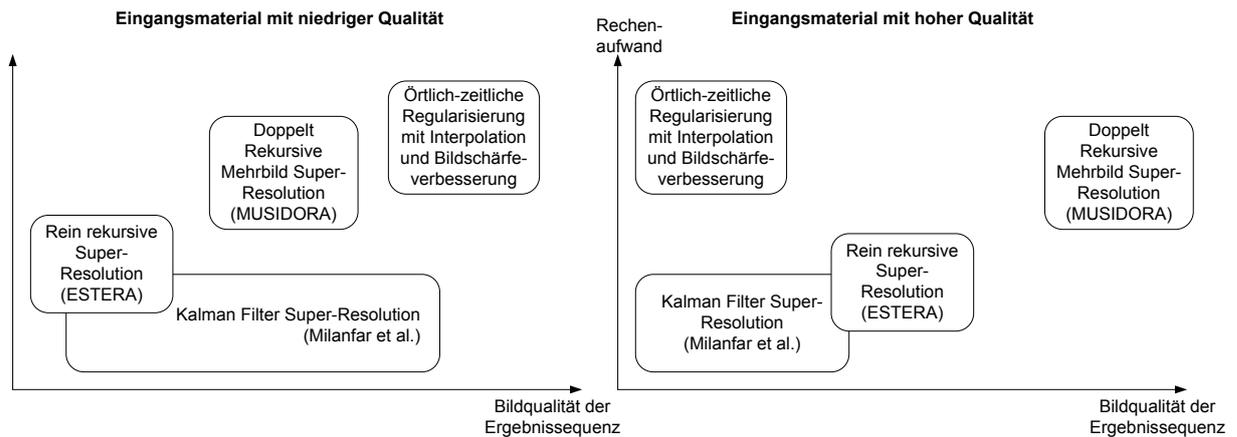


Abbildung 7.1: Qualitative Beurteilung der Leistungsfähigkeit der entwickelten Algorithmen im Vergleich zu einem bestehenden SR-Ansatz

re aufeinanderfolgende Bilder nicht variieren und somit durch zeitliche Verarbeitung nicht entfernt werden können, wird eine Vorverarbeitung durch ein Deblocking-Verfahren empfohlen. Trotzdem reichen die SR-Algorithmen nicht an die speziell für diesen Einsatzzweck entwickelten Regularisierungsalgorithmen heran. Die Qualität eines sehr vielversprechenden Verfahrens aus der Literatur (Kalman Filter Super-Resolution), hängt sehr stark von der Eingangssequenz ab, da dieses Verfahren nur für Videos mit rein translatorischer globaler Bewegung ausgelegt ist. Liegt diese Bewegungsart vor, so können aber hervorragende Ergebnisse erzielt werden, die die der anderen Algorithmen deutlich übertreffen. Ansonsten verhält sich dies aber genau umgekehrt. Die Entscheidung, ob nun ein Verfahren mit mehr als einem Referenzbild als Eingang zur Anwendung kommt (z.B. MUSIDORA) oder ein rein rekursives Verfahren bevorzugt wird, hängt vor allem von den zur Verfügung stehenden Ressourcen ab. Es ist aber definitiv anzumerken, dass selbst der Qualitätssprung von einem rein örtlichen auf ein örtlich-zeitliches Verfahren mit nur einem Bildspeicher bereits erheblich ist. Bei der Verwendung eines zweiten Bildspeichers ist noch einmal ein gewaltiger Qualitätssprung (z.B. von ESTERA zu MUSIDORA) zu beobachten, danach nimmt der Gewinn aber nur noch wenig zu. Ab mehr als fünf Eingangsbildern erscheint der Gewinn in den meisten Fällen nicht mehr gerechtfertigt.

Für Eingangsmaterial mit hoher Qualität, z.B. SD- oder HD-Material, sind die Super-Resolution Algorithmen den Verfahren zur Artefaktreduktion deutlich überlegen. Beim Regularisierungsalgorithmus macht sich hier ein Detailverlust bemerkbar, so dass schlussfolgernd unter der Begutachtung beider Einsatzgebiete die Super-Resolution leichte Vorteile besitzt. Diese werden noch größer, wenn zusätzlich noch eine Parameteranpassung an unterschiedliches Material sowie optionales Hinzuschalten eines Deblocking-Algorithmus möglich sind.

## 7.1 Ausblick

Die in dieser Arbeit durchgeführten Tätigkeiten lassen sich in Zukunft noch weiter fortsetzen. Da den Kontrollinstanzen bzw. der Merkmalsanalyse eine essentielle Bedeutung zukommt, können weitere oder höherwertige Maße die Ergebnisse der Algorithmen noch weiter verbessern. Insbesondere die gewichtete Selektion und Bewegungskompensation sind entscheidend. Ohne korrekte Bewegungsinformation können keine Details wiederhergestellt werden. Aus diesem Grund könnten Maße, die einerseits den Grad der gefundenen Information oder andererseits die Präzision der Bewegungsschätzung bewerten können, die Steuerung der zeitlichen Verarbeitung noch weiter verbessern. In der vorliegenden Arbeit wurden für diesen Zweck nur relativ einfache, aber trotzdem effektive Maße verwendet. Auch die Interaktion zwischen örtlicher und zeitlicher Filterung, die den Super-Resolution Algorithmus bestimmt, könnte so gezielt verbessert werden. Gerade bei Vorliegen von Subpixelverschiebung könnte so die örtliche Verarbeitung, die immer zu einem Detailverlust führt, reduziert werden. Als Ergebnis könnte sich so ein adaptiv arbeitender örtlicher Interpolationsfilter innerhalb der SR ergeben.

Zuletzt soll nun noch ein Ausblick auf ein weiteres mögliches Verarbeitungskonzept gegeben werden, dass in [Stö09] und [RSS10a] bereits für die Artefaktreduktion eingesetzt wurde und auch für die Super-Resolution großes Potential verspricht. Bei diesem bildqualitätsgesteuerten Verarbeitungskonzept werden grundlegende Gedanken der bildinhaltsadaptiven Signalverarbeitung aufgegriffen sowie Ideen aus den direkten und iterativen Verarbeitungskonzepten kombiniert. Somit können die Vorteile jedes Konzeptes genutzt werden, ohne Kompromisse bei der Bildqualität eingehen zu müssen. Erkauft werden diese Vorteile aber durch einen Mehraufwand bei der Hardware. Bei diesem Konzept werden mehrere Algorithmen, die in unterschiedlichen Situationen vorteilhaft sind, parallel auf das Bildmaterial angewendet. Anstelle nun die Ergebnisse wie bei MABLE nur basierend auf Vorwissen (z.B. durch eine Merkmalsanalyse) zusammenzumischen, wird die Bildqualität am Ausgang für jeden Algorithmus gemessen und darauf basierend die beste Verarbeitung ausgewählt. Erweitert man diese Konzept noch mit einer Rückführungsschleife, so kann die Qualität in nachfolgenden Schritten weiter verbessert werden. Entscheidend für die Qualität der Ergebnisse sind hierbei die Bildqualitätsmessungen. Diese können ähnlich wie bei der Formulierung von Ansätzen für iterative Rekonstruktionsalgorithmen auf zwei Termen basieren, einem Ähnlichkeitsterm und einer oder mehreren Nebenbedingungen. Anstelle diesen Ansatz aber zu minimieren, was bei komplizierten oder nichtlinearen Termen nur sehr schwer möglich ist, muss das Ergebnis dieser Funktion lediglich für alle Algorithmen bestimmt werden. Um dieses Verfahren aber in der Videosignalverarbeitung einsetzen zu können, muss auf leistungsfähige Metriken zurückgegriffen werden, die ohne das ideale Referenzmaterial auskommen. Möglichkeiten, um nicht auf das Referenzmaterial angewiesen zu sein, sind aber für die Artefaktreduktion bereits in [Stö09] und [RSS10a] untersucht worden, wo noch weitere Details zu diesem Konzept zu finden sind.



# Literaturverzeichnis

- [AK06] AUBERT, G. ; KORNPORBST, P.: *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer Verlag, 2006
- [BH04] BRASPENNING, R.A. ; HAAN, G. de: True-motion estimation using feature correspondence. In: *SPIE, Proceedings of VCIP*, 2004, S. 396–407
- [Blu97] BLUME, H.: *Nichtlineare fehlertolerante Interpolation von Zwischenbildern*, Universität Dortmund, Diss., 1997
- [Bor96] BORMAN, S.: *On Parameter Estimation for a Class of Random Field Image Models*. 1996
- [Brü09] BRÜGGEMANN, M.: *Total Variation Regularization with Adaptive Filter Kernels*. 2009. – Diplomarbeit
- [Bru01] BRUHN, A.: *Regularization in Motion Estimation*. 2001. – Department of Mathematics and Computer Science, Universität Mannheim, Diplomarbeit
- [Bru06] BRUHN, A.: *Variational Optic Flow Computation: Accurate Modelling and Efficient Numerics.*, Department of Mathematics and Computer Science, Saarland University, Saarbrücken, Diss., 2006
- [BS98] BORMAN, S. ; STEVENSON, R.L.: Super-resolution from image and sequences - a review. In: *Midwest Symposium on Circuits and Systems*, 1998, S. 374–378
- [BS99] BORMAN, S. ; STEVENSON, R.: Simultaneous Multi-frame MAP Super-Resolution Video Enhancement using Spatio-temporal Priors. In: *IEEE Int. Conference on Image Processing*, 1999
- [BSMM01] BRONSTEIN, I.N. ; SEMENDJAJEW, K.A. ; MUSIOL, G. ; MÜHLIG, H.: *Taschenbuch der Mathematik, 5. überarbeitete und erweiterte Auflage*. Verlag Harri Deutsch, 2001
- [Cai06] CAI, Y.: *Super-Resolution zur Interpolation von Bildsequenzen*. 2006. – Diplomarbeit D10-2006, TU Dortmund

- [CDLL98] CHAMBOLLE, A. ; DEVORE, R.A. ; LEE, N.-Y. ; LUCIER, B.J.: Nonlinear wavelet image processing: variational problem, compression and noise removal through wavelet shrinkage. In: *IEEE Trans. Image Process.* 7 (1998), S. 319–335
- [CS05] CHAN, T. ; SHEN, J.: *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Models*. SIAM, 2005
- [Dol10] DOLAR, C.: *LCD-Modelle und ihre Anwendung in der Videosignalverarbeitung*, Technische Universität Dortmund, Diss., 2010
- [DRS09a] DOLAR, C. ; RICHTER, M. ; SCHRÖDER, H.: Total Variation Regularization for Video Signal Processing. In: *IEEE International Symposium on Consumer Electronics*, 2009
- [DRS09b] DOLAR, C. ; RICHTER, M. ; SCHRÖDER, H.: Total Variation Regularization for Video Signal Processing Applications. In: *13. Dortmunder Fernsehseminar*, 2009
- [Est00] ESTRADA, GG.: Local smoothness in terms of variance: the adaptive Gaussian filter. In: *Proc. of the 11th British Machine Vision Conference, Bristol Bd. 2*, 2000, S. 815–824
- [FEM06] FARSIU, S. ; ELAD, M. ; MILANFAR, P.: Video-to-Video Dynamic Superresolution for Grayscale and Color Sequences. In: *EURASIP Journal of Applied Signal Processing, Special Issue on Superresolution Imaging* (2006)
- [FJP02] FREEMAN, T ; JONES, T.R. ; PASZTOR, E.C.: Example Based Super-Resolution. In: *IEEE Computer Graphics and Application* 22 (2002), Nr. 2
- [FREM04] FARSIU, S. ; ROBINSON, D. ; ELAD, M. ; MILANFAR, P.: Advances and Challenges in Super-Resolution (Invited Paper). In: *Special Issue on High Resolution Image Reconstruction* 14 (2004), Nr. 2, S. 47–57
- [FTS01] FRANZEN, O. ; TUSCHEN, C. ; SCHRÖDER, H.: Intermediate Image Interpolation using Polyphase Weighted Median Filters. In: *IS&T/Spie Electronic Imaging*, 2001, S. 306–317
- [HA78] HSIEH, H. ; ANDREW, H.: Cubic splines for image interpolation and digital filtering. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 25 (1978), Nr. 6, S. 508–517
- [Haa00] HAAN, G. de: *Video Processing for Multimedia Systems*. University Press Eindhoven, 2000

- [Har02] HARRISON, P.: Patchwork texture synthesis, Technical Report 2002/119 / Monash University School of Computer Science and Software Engineering. 2002. – Forschungsbericht
- [HBHO93] HAAN, G. de ; BIEZEN, P.W.A.C. ; HUIJGEN, H. ; OJO, O.A.: True-Motion Estimation with 3-D Reducursive Search Block Matching. In: *IEEE Trans. On Circ. and Systems for Video Techn.* 3 (1993), Okt., Nr. 5
- [HDRS09] HERBORT, S. ; DOLAR, C. ; RICHTER, M. ; SCHRÖDER, H.: Anwendung der Superresolution im TV-Endgerät. In: *ITG Fachtagung für Elektronische Medien 'Systeme, Technologien, Anwendungen', 13. Dortmunder Fernsehseminar, TU Dortmund, 2009*
- [HJKH09] HA, H.-G. ; JANG, I.-S. ; KO, K.-W. ; HA, Y.-H.: Robust subpixel shift estimation using iterative phase correlation of a local region. In: *SPIE IS&T Electronic Imaging* Bd. 7241, 2009
- [HS81] HORN, B.K.P. ; SCHUNCK, B.G.: Determining optical flow. In: *Artificial Intelligence* 17 (1981), S. 185–203
- [Int07] INTEL: *White Paper: Video Image Reconstruction and Enhancement - A Tera-scale Computing Application.* 2007
- [INXO10] ICHIHASHI, H. ; NAGUMO, T. ; XAVIER, M. ; ORYOJI, H.: *Image Processing Apparatus, Image Processing Method, and Program.* Sony Corporation, US Patent Application US 2010/0119176A1, 2010
- [IP91] IRANI, M. ; PELEG, S.: Improving resolution by image registration. In: *CVGIP: Graphical Models and Image Processing* 53 (1991), Mai, S. 231–239
- [ITU] Recommendation ITU-R BT.500-11: Methodology for the subjective assessment of the quality of television pictures
- [Kal60] KALMAN, R.E.: A New Approach to Linear Filtering and Prediction Problems. In: *Transactions of the ASME–Journal of Basic Engineering* 82 (1960), S. 35–45
- [Kau00] KAUP, A.: Image restoration for frame- and object-based video coding using an adaptive constrained least-squares approach. In: *Signal Processing* 80 (2000), S. 2337–2345
- [KH75] KUGLIN, C.D. ; HINES, D.C.: The phase correlation image alignment method. In: *IEEE Conference on Cybernetics and Society*, 1975, S. 163–165

- [KLF08] KELLER, S. ; LAUZE, F. ; FROMENT, J.: Deinterlacing using variational methods. In: *IEEE Trans. on Image Proc.* 17 (2008), Nr. 11, S. 2015–2028
- [Kon01] KONDO, T. et a.: *Picture conversion apparatus and learning method.* Sony Corporation, US Patent 6323905, 2001
- [KV04a] KLOMPENHOUWER, M. ; VELTHOVEN, L.: LCD Motion Blur Reduction with Motion Compensated Inverse Filtering. In: *SID Symposium Digest* Bd. 35, 2004, S. 1340–1343
- [KV04b] KLOMPENHOUWER, M. ; VELTHOVEN, L.: Motion blur reduction for liquid crystal displays: Motion compensated inverse filtering. In: *proc. SPIE, VCIP*, 2004, S. 690–699
- [KYKR98] KIM, S.D. ; YI, J. ; KIM, H.M. ; RA, J.B.: A Deblocking Filter with Two Separate Modes in Block-Based Video Coding. In: *IEEE Trans. on Circuits and Systems Video Techn.* 9 (1998), März, Nr. 1, S. 2–14
- [LA08] LIYAKATHUNISA ; ANANTHESHYANA, V.K.: Super Resolution Blind. Reconstruction Of Low Resolution Images using wavelets based Fusion. In: *World Academy of Science, Engineering and Technology* (2008), Nr. 40
- [LBN07] LIVONIUS, J. ; BLUME, H. ; NOLL, T.G.: *Analyse und Pareto-Optimierung eines parametrisierten Algorithmus zur Bewegungsschätzung.* 2007
- [LE07] LENKE, S. ; ERDLER, O.: Detail- und Texturwiedergabe bei der SDTV -> HDTV Interpolation. In: *ITG-Fachbericht 199, ITG/FKTG-Fachtagung Elektronische Medien, 12. Dortmunder Fernsehseminar, Dortmund, Germany*, 2007, S. 145–150
- [Len09] LENKE, S.: *Klassifikationsbasierte Poloyphasen Bildinterpolation*, Technische Univesität Dortmund, Diss., 2009
- [LK81] LUKAS, B. ; KANADE, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. In: *Internat. Joined Conference on Artificial Intellegence, Vancouver, Kanada*, 1981
- [LO01] LI, X. ; ORCHARD, T.: New Edge-Direction Interpolation. In: *IEEE Transactions on Image Processing* 10 (2001), Okt., Nr. 10
- [Mal99] MALLAT, S.: *A Wavelet Tour of Signal Processing.* Academic Press, 1999

- [Mil05] MILANFAR, P.: (Super)Resolution: Statistical Definition, Computation, and Fundamental Limits. In: *International Conference on Super-Resolution Imaging, Hong Kong, 2005*
- [MTM<sup>+</sup>04] MESSNARTZ, B. ; TILG, B. ; MOORE, R. ; FISCHER, G. ; HANSER, F.: A New Spatiotemporal Regularization Approach for Reconstruction of Cardiac Transmembrane Potential Patterns. In: *IEEE Trans. on Biomedical Engineering* 51 (2004), Nr. 2
- [Naß09] NASSE, F.: *Untersuchung eines rekursiven Super-Resolution Ansatzes*. 2009. – Technische Universität Dortmund, Diplomarbeit
- [NM00] NGUYEN, N. ; MILANFAR, P.: A wavelet-based interpolation-restoration method for superresolution. In: *Circuits, Systems, and Signal Processing* 19 (2000), Nr. 4
- [NSLZ07] NG, M.K. ; SHEN, H. ; LAM, E.Y. ; ZHANG, L.: A Total Variation Regularization Based Super-Resolution Reconstruction Algorithm for Digital Video. In: *EURASIP Journal on Advances in Signal Processing* (2007)
- [OTBH97] OHM, J.R. ; TALMI, M. ; BRAUN, M. ; HAHN, M.: Motion Compensating Real-Time Format Converter for Video on Multimedia Displays. In: *IEEE* (1997)
- [Pia08] PIASTOWSKI, P.R.S.: *Objektive Bewertung von Deblocking-Verfahren*, Technische Universität Dortmund, Diss., 2008
- [PK04] PIASTOWSKI, P. ; KOHLMAYER, G.: Neues Verfahren zur kombinierten örtlich/zeitlichen Rauschreduktion und verbesserte Bewegungsschätzung für neue Display-Technologien. In: *FKTG Jahrestagung, 2004*
- [PPK03] PARK, S.C. ; PARK, M.K. ; KANG, M.G.: Super-Resolution Image Reconstruction: A Technical Overview. In: *IEEE Signal Processing Magazine* 20 (2003), Mai, Nr. 3, S. 21–36
- [RBSK11] RICHTER, M. ; BRÜGGEMANN, M. ; SCHRÖDER, H. ; KAYS, R.: Feature Analysis Based Super-Resolution for Video Signal Processing. In: *1th ITG Conference on Electronic Media Technology, 2011*
- [RDS<sup>+</sup>08] RICHTER, M. ; DOLAR, C. ; SCHRÖDER, H. ; ERDLER, O. ; SARTOR, P.: Reduktion von Bewegungsunschärfe durch bewegungsabhängige synthetische Detailsignaladdition. In: *FKT* (2008), November

- [RDS09a] RICHTER, M. ; DOLAR, C. ; SCHRÖDER, H.: Coding Artifact Reduction by Temporal Filtering. In: *IEEE International Symposium on Consumer Electronics*, 2009
- [RDS09b] RICHTER, M. ; DOLAR, C. ; SCHRÖDER, H.: Zeitliche Verarbeitungsstrategien zur Reduktion von Codierartefakten. In: *ITG/FKTG-Fachtagung Elektronische Medien, 13. Dortmunder Fernsehseminar, Dortmund, Germany*, 2009
- [RDS<sup>+</sup>11] RICHTER, M. ; DOLAR, C. ; SCHRÖDER, H. ; ERDLER, O. ; SPRINGER, S.: Spatio-Temporal Regularization featuring Multiple Input Frames and Novel Spatio-Temporal Priors. In: *IEEE International Symposium on Broadcast Multimedia Systems and Broadcasting, Nürnberg*, 2011
- [RNS11] RICHTER, M. ; NASSE, F. ; SCHRÖDER, H.: Motion adaptive Kalman filter for super-resolution. In: *IS&T/SPIE Electronic Imaging, San Francisco, USA*, 2011
- [ROF92] RUDIN, L.I. ; OSHER, S. ; FATEME, E.: Nonlinear total variation based noise removal algorithms. In: *Physica D: Nonlinear Phenomena* 60 (1992), Nr. 1-4, S. 259–268
- [RSS10a] RICHTER, M. ; STÖBER, M. ; SCHRÖDER, H.: Quality Controlled Artifact Reduction. In: *European Signal Processing Conference (EUSIPCO), Aalborg, Dänemark*, 2010
- [RSS10b] RICHTER, M. ; STÖBER, M. ; SCHRÖDER, H.: Rekursive und direkte Verfahren in der Artefaktreduktion. In: *FKTG Jahrestagung*, 2010
- [SB00] SCHRÖDER, H. ; BLUME, H.: *Mehrdimensionale Signalverarbeitung - Band 2: Architekturen und Anwendungen für Bilder und Bildsequenzen*. B.G. Teubner, 2000
- [Sch00] SCHRÖDER, H.: *Mehrdimensionale Signalverarbeitung - Band 1*. B.G. Teubner, 2000
- [Sch07] SCHEWIOR, G.: *Temporal Noise Reduction based on multi-frame motion estimation in a MPEG to AVC transcoding environment*. 2007. – Diplomarbeit
- [SEF86] SCHRÖDER, H. ; ELSLER, H. ; FRITSCH, M.: Nonlinear Picture Enhancement Techniques for Vertically Interpolated TV-Signals. In: *EUSIPCO Conf. Proceedings*, 1986, S. 841–844

- [SHH08] SHAO, L. ; H., Zhang ; HAAN, G. de: An Overview and Performance Evaluation of Classification-Based Least Squares Trained Filters. In: *IEEE Transactions on Image Processing* 10 (2008), Okt., S. 1772–1782
- [SMD09] SPEIGLE, J.M. ; MESSING, D.S. ; DALY, S.: Application of a visual model to the design of ultra-high definition upscaler. In: *Proc. SPIE* 7240 (2009), Nr. 1
- [Smo01] SMOLIC, A.: *Globale Bewegungsschätzung und Video Mosaiking unter Verwendung Parametrischer 2-D Modelle, Schätzverfahren und Anwendungen*, RWTH Aachen, Diss., Mai 2001
- [Stö09] STÖBER, M.: *Bildqualitätsmaße zur Steuerung von Verfahren für die Codierartefaktreduktion*. 2009. – Technische Universität Dortmund, Diplomarbeit
- [SVW02] SMOLIC, A. ; VATIS, Y. ; WIEGAND, T.: Long-Term Global Motion Compensation Applying Super-Resolution Mosaics. In: *IEEE Int. Symposium on Consumer Electronics, Ilmenau, Germany, 2002*
- [TFM07] TAKEDA, H. ; FARSIU, S. ; MILANFAR, P.: Kernel Regression for Image Processing and Reconstruction. In: *IEEE Transactions on Image Processing* 16 (2007), Nr. 2
- [TH84] TSAI, R.Y. ; HUANG, T.S.: Multipleframe image restoration and registration. In: *Advances in Computer Vision and Image Processing*, AI Press Inc., 1984, S. 317–339
- [Tik43] TIKHONOV, A.N.: On the stability of inverse problems. In: *Dokl. Akad. Nauk SSSR* 39 (1943), Nr. 5, S. 195–198
- [TM98] TOMASI, C. ; MANDUCHI, R.: Bilateral filtering for gray and color images. In: *IEEE ICCV*, 1998, S. 839–846
- [UG92] UR, H. ; GROSS, D.: Improved resolution from sub-pixel shifted pictures. In: *CVGIP: Graphical Models and Image Processing* 54 (1992), März, S. 181–186
- [VO96] VOGEL ; OMAN: Iterative methods for total variation de-noising. In: *SIAM J. Sci. Comput* 17 (1996), S. 227–238
- [VSVV07] VANDEWALLE, P. ; SBAIZ, L. ; VANDEWALLE, J. ; VETTERLI, M.: Super-Resolution From Unregistered and Totally Aliased Signals Using Subspace Methods. In: *IEEE Transactions on Signal Processing* 55 (2007), Nr. 7

- [WS04] WANG, B. ; SHEIKH, S: Image Quality Assessment: From Error Visibility to Structural Similarity. In: *IEEE Transactions on Image Processing* 13 (2004), Nr. 4
- [WSBL03] WIEGAND, T. ; SULLIVAN, G.J. ; BJNTEGAARD, G. ; LUTHRA, A.: Overview of the H.264/AVC Video Coding Standard. In: *IEEE Trans. Circuits Syst. Video Technol.* 13 (2003), Juli, Nr. 7, S. 560–576
- [Wu93] WU, X.: *Synthetische Kantenversteigerung zur Verbesserung der Bildschärfe*, Universität Dortmund, Diss., 1993
- [Yan08] YAN, R.: *Untersuchung der Leistungsfähigkeit zeitlicher De-Blocking Verfahren*, Technische Universität Dortmund, Diplomarbeit, 2008
- [YFL<sup>+</sup>03] YAO, S. ; FENG, G. ; LIN, X. ; LIM, K.P. ; LIN, W.: A Coding Artifact Removal Algorithm Based on Spatial and Temporal Regularization. In: *IEEE* (2003)
- [YST<sup>+</sup>09] YAMASHITA, M. ; SAKAI, R. ; TANAKA, A. ; IMADA, K. ; TAKAHASHI, Y. ; TAKASHI, I. ; MATSUMOTO, N. ; KATO, N.: Image Super-Resolution and Multiple Video Playback on TV sets using Cell Processor. In: *IEEE International Symposium on Consumer Electronics, Kyoto, Japan, 2009*, S. 305–306
- [ZC10] ZHANG, X. ; CHAN, T.F.: Wavelet Inpainting by Nonlocal Total Variation. In: *Inverse Problems and Imaging* 4 (2010), Nr. 1
- [ZP01] ZOMET, A. ; PELEG, S.: Super-Resolution from Multiple Images Having Arbitrary Mutual Motion. In: *Super-Resolution Imaging*. Kluwer Academic, 2001, Kapitel 2
- [ZS02] ZHAO, W.Y. ; SAWHNEY, H.S.: Is Super-Resolution with Optical Flow Feasible? In: *Computer Vision ECCV* Bd. 2350/2002, Springer, 2002, S. 599–613
- [ZS06] ZEEVI, Y.Y. ; SOCHEN, N.: Texture preserving variational de-noising using an adaptive fidelity term. In: *Intern. J. Computer Vision* 67 (2006), Nr. 1, S. 111–136
- [ZW09] ZAFARIFAR, B. ; WITH, P. de: Content-adaptive image enhancement based on sky and grass segmentation. In: *IEEE International Conference on Consumer Electronics, 2009*, S. 1–2

# Anhang A

## Standardeinstellungen der Super-Resolution Algorithmen

### Mehrreferenzbild Algorithmus

Die im Rahmen dieser Arbeit verwendeten Algorithmen wurden unter der Programmierumgebung VisiQuest implementiert. Für die Untersuchungen wurde dazu der voll parametrisierbare Mehrreferenzbild Back-Projection Algorithmus VESBA geschaffen, der in Abschnitt 5.1 bereits ausführlich beschrieben wurde. Um den Einfluss aller Parameter und Modifikationen testen zu können, können die Komponenten, die zum Datenmodell gehören, zu- und abgeschaltet werden. Dazu zählen der Unschärfe-Filter, die Runter- und Hochtastung und die Art der Bewegungskompensation. Weiterhin ist es möglich, die Interpolation und die zeitliche Filterung vor den eigentlichen Algorithmus zu ziehen und den SR-Algorithmus nur noch als iteratives Blurring zu nutzen. Zusätzlich kann die Filterung dann zeitlich rekursiv durchgeführt werden, so dass sich die Einstellungen für den doppelt rekursiven Algorithmus MUSIDORA ergeben.

In Abbildung A.1 ist die Benutzeroberfläche für den voll parametrisierbaren Algorithmus zu sehen. In der Standardeinstellung sind die Bewegungskompensation und das Up- and Downsampling in der Iterationsschleife aktiviert, wie es die mathematische Ableitung auch vorsieht. Als einzige Modifikation wird für die Interpolation des Detailsignals ein High Resolution Spline anstelle dem Einfügen von Nullen und einer anschließenden Tiefpassfilterung durch einen Gauß'schen Tiefpass verwendet. Für die zeitliche Filterung wird ein Mittelwertfilter über fünf Eingangsbilder angewendet. Jegliche Kontrolle durch Bildanalysemaße ist ausgeschaltet.

Im Vergleich dazu müssen bei MUSIDORA lediglich die Kästchen 'merge in Advance', 'temporally Recursive' und 'Initial upsampling' sowie die 'Temporal Merging' von 'Mean' (Mittelwertfilter) auf 'FIR-Filter' umgestellt werden.

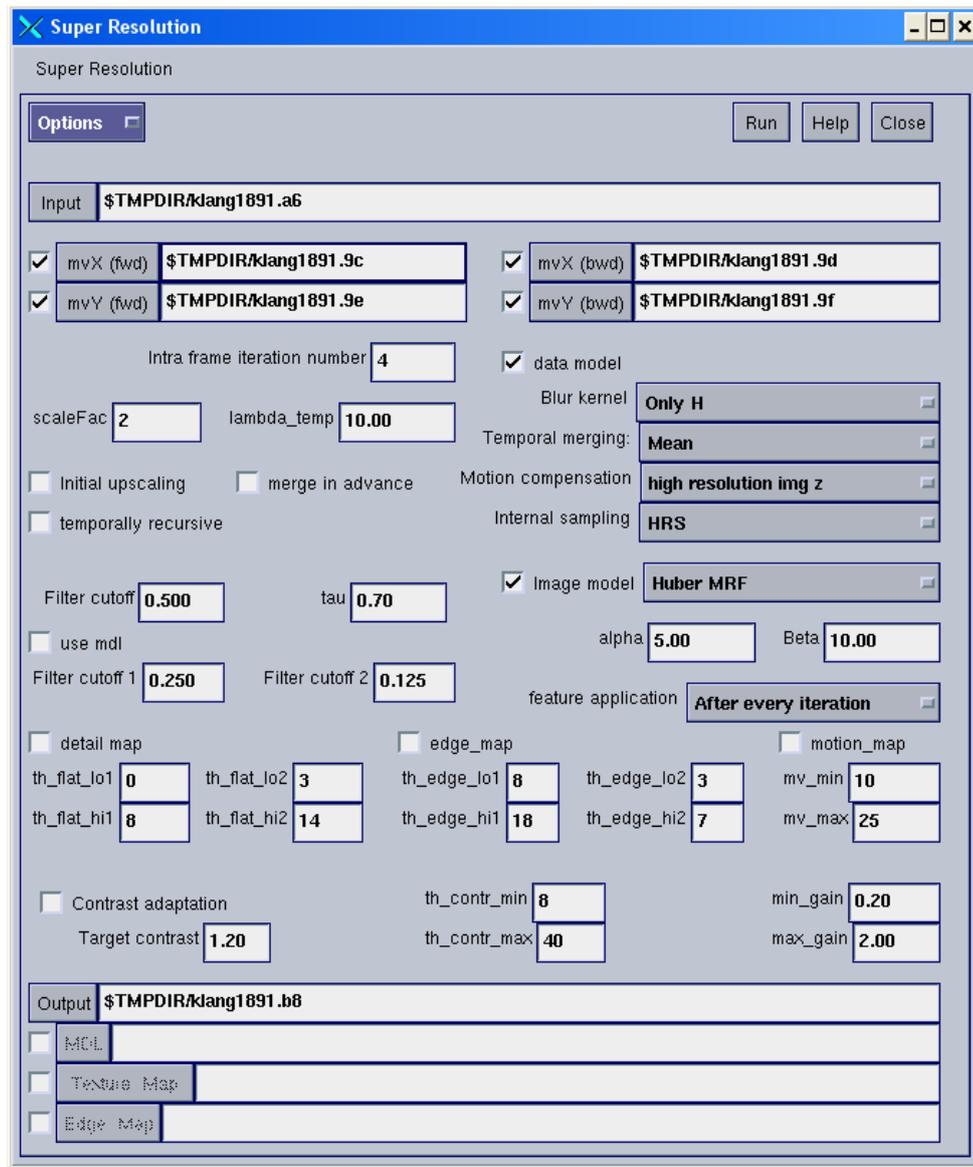


Abbildung A.1: Parametereinstellmöglichkeiten der Mehrreferenzbild Super-Resolution

---

## Zeitlich rein rekursiver Algorithmus

Die Benutzeroberfläche des rein rekursiven Algorithmus ESTERA findet sich in Abbildung A.2. In diesem Algorithmus wird die niedrig aufgelöste Sequenz immer zu Anfang mit einem High Resolution Spline auf die Ergebnisgröße interpoliert, so dass auf eine interne Hoch- und Runtertastung verzichtet werden kann. Für die zeitliche Filterung kommt ein fester Gewichtungparameter von 0.7 zum Einsatz, so dass das aktuelle Bild mit 30% und das vorige Ergebnisbild zu 70% zur initialen Schätzung des nächsten Ergebnisses beiträgt. Die rein örtliche Bildschärfekorrektur und die Gegenmaßnahme zur Verringerung der Sichtbarkeit von Vektorfehlern ist eingeschaltet und wird ebenfalls mit 0.7 gewichtet, wie es sich aus der mathematischen Berechnungsvorschrift ergibt. Modifikation kommen erst in den verbesserten Varianten des Algorithmus zu tragen.

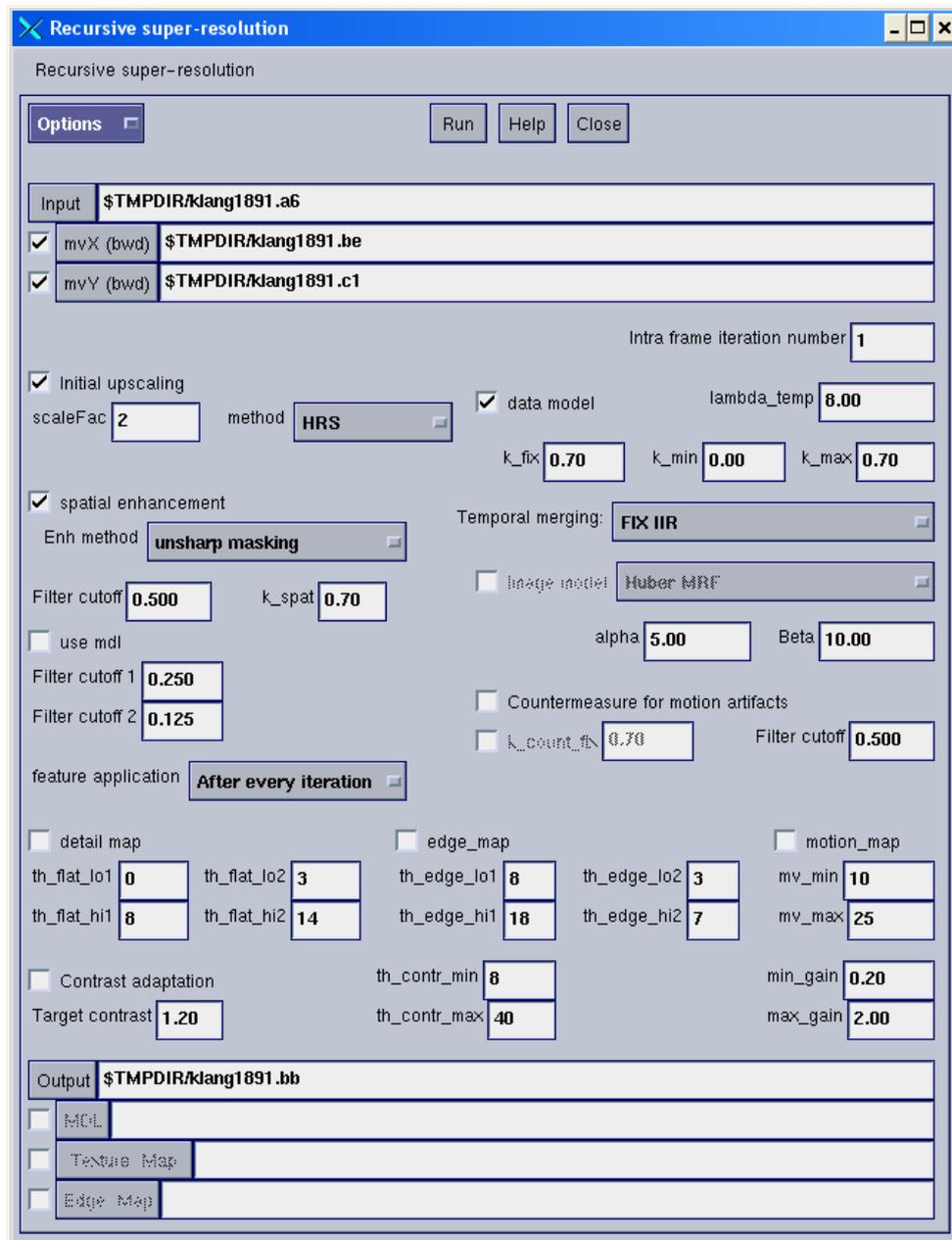


Abbildung A.2: Parametereinstellungsmöglichkeiten der rein rekursiven Super-Resolution

# Anhang B

## Verwendete Testsequenzen

### football

Die Sequenz 'football' hat eine Auflösung von 720x576 Pixeln. Sie zeigt die Szenerie eines Fussballspiels. In der Sequenz gibt es sehr viele texturierte Flächen (Rasen, Bäume im Hintergrund) aber auch Kanten mit sehr steilen Übergängen und viele Bilddetails. Die Sequenz selbst ist relativ stark verrauscht, wobei die Trennung zwischen Rauschen und Bilddetails gerade im Vordergrund im Einzelbild schwierig zu unterscheiden ist. Die Kamera zoomt in dieser Szene herein, so dass eine globale Bewegung vorherrscht. Daneben vollführen die Spieler teilweise komplexe Bewegungen. Diese Sequenz wird im Kapitel 5 sehr intensiv zur Diskussion der Parameter der Super-Resolution Algorithmen verwendet.

### parkrun

Die Sequenz 'parkrun' besitzt eine Auflösung von 1280x720 Bildpunkten. In ihr verfolgt die Kamera eine laufende Person, so dass eine globale Schwenkbewegung vorherrscht. Die Person selbst gehorcht dieser globalen Bewegung nicht. Insgesamt ist die Szene sehr detailreich und verfügt nur über sehr leichtes Rauschen.

### stockholm

Diese Sequenz zeigt einen Kameraschwenk über die Stadt Stockholm. Sie hat eine Auflösung von 1280x720 Pixeln. Sie besitzt ein leichtes Hintergrundrauschen, welches teilweise sehr niederkontrastige Details im Einzelbild verdeckt. Insgesamt ist die Szene sehr detailreich, besitzt im Bildhintergrund große homogene Flächen und im Bildvordergrund auch sehr scharfe Kanten.



Abbildung B.1: Sequenz 'football'



Abbildung B.2: Sequenz 'parkrun'



Abbildung B.3: Sequenz 'stockholm'

## wheel

Die Sequenz 'wheel' hat eine Standardauflösung von 720x576 Bildpunkten und zeigt verschiedene Bildgegenstände. Die Scheibe als auch das Speicherrad drehen sich entgegengesetzt zueinander, was für die Bewegungsschätzung eine große Herausforderung darstellt. Das Klalenderblatt als auch die Scheibe weisen sehr viele unterschiedliche Kantenrichtungen auf. Der Hintergrund der Szene steht still und er weist ein Hintergrundrauschen auf. Sie wird in der Auswertung dazu verwendet, die Leistungsfähigkeit der Interpolationsalgorithmen in Bezug auf die Interpolation vor allem diagonalen Kanten zu testen.

## matfight

Die Sequenz 'matfight' zeigt eine Szene aus dem Film Matrix. Sie hat eine Auflösung von 720x576 Bildpunkten. In dieser Szene herrschen sehr schnelle Objektbewegungen vor, die zu einer starken Bewegungsunschärfe einzelner Objekte führt. Die Szene besitzt überwiegend homogene Flächen und starke Kanten und nur in einzelnen Bereichen (z.B. Gesichter und Kleidung) Bilddetails mit schwachem Kontrast. Mit dieser Szene wird die Qualität der Super-Resolution bei sehr komplexen Bewegungssituationen getestet.

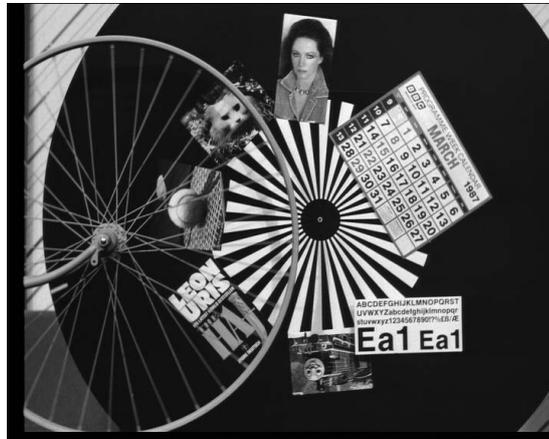


Abbildung B.4: Sequenz 'wheel'



Abbildung B.5: Sequenz 'matfight'



Abbildung B.6: Sequenz 'seeking'

## seeking

In der Sequenz 'seeking' ist eine sehr detailreiche Szene zu sehen. Die Auflösung dieser Sequenz beträgt 1280x720 Bildpunkte. Die im Bildausschnitt sichtbaren Personen bewegen sich aufeinander zu und die Kamera vollführt einen leichten Schwenk nach rechts.

## foreman

Die Sequenz 'foreman' hat eine Auflösung von 320x288 Bildpunkten. Sie wurde mittels ffmpeg mit dem MPEG-4 Standard bei fester Qualitätsstufe 18 codiert, so dass in ihr starke Codierartefakte sichtbar sind. Die Kameraführung ist unruhig, so dass der Szene zufällige, kleine globale Bewegungen überlagert sind. Zusätzlich gibt es lokale Bewegungen der Person im Vordergrund. Diese Szene wurde insbesondere dazu verwendet, die Qualität der Super-Resolution bei stark komprimierten Sequenzen zu testen, da diese heutzutage durch Handy-Videos oder Web-Portale einen typischen Anwendungsfall darstellen.



Abbildung B.7: Sequenz 'foreman'

# Lebenslauf

## Persönliche Daten

Name: Martin Richter  
Geburtsdatum: 22.01.1981  
Geburtsort: Essen

## Werdegang

2000 Abitur am Gymnasium an der Wolfskuhle in Essen  
2001-2006 Studium der Elektrotechnik an der Technischen Universität Dortmund  
Dez. 2006 Abschluss des Studiums mit dem Diplom  
Jan. 2007- wissenschaftlicher Mitarbeiter am Arbeitsgebiet  
Jun. 2011 Schaltungen der Informationsverarbeitung, TU Dortmund

