

**Bewegungskompensierende Teilbandcodierung
von Bewegtbildsequenzen
mit sehr niedrigen Übertragungsbitraten**

Vom Fachbereich Elektrotechnik und Informationstechnik
der Universität Hannover

zur Erlangung des akademischen Grades

Doktor–Ingenieur

genehmigte

Dissertation

von

M.S. Deyu Qian

geboren am 31. März 1957 in Zhejiang

2003

Referent: Prof. Dr.–Ing. H. G. Musmann
Korreferent: Prof. Dr.–Ing. H.–P. Kuchenbecker
Vorsitzender: Prof. Dr.–Ing. J. Gaul
Tag der Prüfung: 05. 05. 2003

Kurzfassung

Es wird eine bewegungskompensierende Teilbandcodierung von Bewegtbildsequenzen mit sehr niedrigen Übertragungsbitraten vorgestellt. Dieses Codierungsverfahren stellt eine Weiterentwicklung bekannter Teilbandcodierungsverfahren dar. Der Prädiktionsfehler der bewegungskompensierenden Prädiktion wird blockweise verarbeitet. Zur Auswahl der zu codierenden Blöcke wird eine Vorwärtsanalyse der Prädiktionsfehler eingefügt. Im Gegensatz zu bekannten Teilbandcodierungsverfahren werden die Teilbandanalyse und –synthese hier nur auf die zu codierenden Blöcke angewendet, wobei in jeder Filterstufe eine symmetrische periodische Erweiterung der zu filternden Signale durchgeführt wird, um eine Erhöhung der Anzahl zu übertragender Abtastwerte zu vermeiden. Eine neue Methode zur Steuerung der orts- und zeitadaptiven Quantisierung der Teilbandsignale wird unter Anwendung eines psychovisuellen Modells entwickelt und eingesetzt. Experimentelle Untersuchungen zeigen, daß im Vergleich mit dem Hybrid-DCT-Coder TMN5, der während des Entwurfs des internationalen Standards H.263 als Referenz dient, der Signal-Rausch-Abstand in den Rekonstruktionsbildern bei 8 kbit/s durch Anwendung des vorgestellten Codierungsverfahrens um ca. 1 dB im Mittel erhöht werden kann.

Stichworte: Vorwärtsanalyse, Blockbasierte Teilbandcodierung,
Adaptive Quantisierung

Abstract

A motion compensated subband coding of video sequences at very low bit-rates is presented. This coding technique is a further development of the well-known subband coding. The prediction error of the motion compensated prediction is processed on the block basis. A forward analysis is introduced to select the blocks to be coded. In contrast to the conventional subband coding, the subband analysis and synthesis here are only applied to the blocks to be coded. A symmetric-periodic extension of the signals to be filtered is applied in each filtering stage, in order to avoid an increase in the number of samples to be transmitted. A new method for controlling the spatial and temporal adaptive quantization of the subband signals is developed and applied by using a psychovisual model. Experimental evaluations show, that in comparison to the hybrid DCT coder TMN5, which is used as reference during the development of the international standard H.263, a mean PSNR gain of about 1 dB in the reconstructed images can be achieved at 8 kbit/s by using the presented coding technique.

Keywords: Forward analysis, Block-based subband coding
Adaptive quantization

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als Stipendiat und wissenschaftlicher Mitarbeiter am Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung der Universität Hannover.

Dem Leiter des Instituts, Herrn Professor Dr.-Ing. H. G. Musmann, danke ich für die Anregung zum Thema der Arbeit, seine wissenschaftliche Betreuung bei der Durchführung und die Übernahme des Hauptreferats. Er sorgte darüber hinaus für hervorragende Arbeitsmöglichkeit.

Herrn Prof. Dr.-Ing. H.-P. Kuchenbecker danke ich für die freundliche Übernahme des Korreferats.

Die Durchführung eines Teils der Arbeit wurde ermöglicht durch ein Stipendium der Konrad-Adenauer-Stiftung. Ich danke Herrn H. Fuchs und Herrn Dr. R. Zülch für die Betreuung während meiner Stipendiatenzeit.

Alle Meinen ehemaligen Kolleginnen und Kollegen, die durch Diskussionen und Anregungen zum Gelingen der Arbeit beigetragen haben, danke ich herzlich. Insbesondere gilt mein herzlicher Dank Herrn Dr.-Ing. G. Seelmann, Herrn Dr.-Ing. B. Edler für die stets entgegengebrachte Hilfsbereitschaft und für viele konstruktive Diskussion.

Zum Schluß möchte ich meiner Familie für die von ihr aufgebrachte Unterstützung danken.

Inhalt

1. Einleitung	1
2. Prinzipien der bewegungskompensierenden Teilbandcodierung ...	8
2.1 Bewegungskompensierende Interframe-Prädiktion	9
2.2 Teilbandcodierung	11
2.3 Hybrid-DCT-Coder TMN5	16
2.4 Struktur des entwickelten Teilbandcoders	18
3. Vorwärtsanalyse	22
3.1 Ansätze und Aufgaben der Vorwärtsanalyse	22
3.2 Klassifizierung der Prädiktionsfehler	26
3.3 Algorithmen zur Vorwärtsanalyse	28
3.3.1 Stufe 1: Irrelevanzreduktion	29
3.3.2 Stufe 2: Relevanzreduktion	31
3.4 Codierung der Ausgangssignale der Vorwärtsanalyse	37
4. Teilbandcodierung der ausgewählten Blöcke	39
4.1 Blockbasierte Teilbandcodierung	39
4.1.1 Probleme bei der konventionellen Teilbandcodierung mit sehr niedrigen Datenraten	39
4.1.2 Teilbandzerlegung von Blockbündeln	41
4.2 Auswahl der Teilbandfilterbank	49
4.2.1 Auswahl der Struktur der Filterkaskaden	50
4.2.2 Auswahl der Teilbandfilter	53
4.3 Umschaltung zwischen einer Inter- und einer Intraframe- Codierung in jedem Teilband	57
4.4 Codierung der quantisierten Teilbandsignale	59

5. Adaptive Quantisierung unter Anwendung eines psychovisuellen Modells	60
5.1 Ein Psychovisuelles Modell der menschlichen visuellen Wahrnehmung	60
5.2 Potentielle Sichtbarkeit der Rekonstruktionsfehler bei der Teilbandcodierung	65
5.3 Optimale Aufteilung der Datenrate in den Teilbändern	73
5.4 Realisierung der adaptiven Quantisierung der Teilbandsignale ...	80
5.5 Codierung der Quantisiererschrittweiten	87
6. Experimentelle Untersuchungen	89
6.1 Beschreibung der verwendeten Testsequenzen	89
6.2 Gewinne der einzelnen Komponenten des Coders	91
6.2.1 Vorwärtsanalyse	92
6.2.2 Blockbasierte Teilbandcodierung	94
6.2.3 Adaptive Quantisierung	96
6.3 Vergleich des Teilbandcoders mit dem Hybrid-DCT-Coder TMN5	97
7. Zusammenfassung	104
8. Literaturverzeichnis	109
A. Impulsantwortfunktionen der in dieser Arbeit verwendeten QM-Filter	115
B. Filterimpulsantwortfunktionen und Parameter des w-Modells	117
B.1 Filterimpulsantwortfunktionen	117
B.2 Parameter	118
C. Herleitung der optimalen Bitaufteilung auf die Teilbandsignale	119

Abkürzungen und Formelzeichen

AF	Analysefilter
(A)SPE	(Anti-)Symmetrische Periodische Erweiterung
BAS	Basis-Analysefilterstufe
BSS	Basis-Synthesefilterstufe
BBRF	Bandbreitereduktionsfaktor
BFRF	Bildfrequenzreduktionsfaktor
c^* , c_{inh}^*	kompanierte Signale in verschiedenen Stufen des Modells der nichtlinearen Signalverarbeitung in der Fovea
Δc	Ganglienzellstörsignal
Δc_m	Ganglienzellstörsignal aus dem Quantisierungsfehlersignal des Teilbandes m
CCITT	Comité Consultatif International de Téléphonie et Télégraphie
CQF	Conjugate Quadrature Mirror Filter
DCT	Diskrete Cosinus Transformation
DPCM	Differenz-Pulsmodulation
e_m	Quantisierungsfehlersignal des Teilbandes m
f_{Bild}	Eingangsbildfrequenz
$g_0[n]$	Impulsantwort des Tiefpaßfilters in einer Teilbandsynthese mit zwei Teilbändern
$G_0(e^{j\omega})$	Übertragungsfunktion des Tiefpaßfilters in einer Teilbandsynthese mit zwei Teilbändern
$g_1[n]$	Impulsantwort des Hochpaßfilters in einer Teilbandsynthese mit zwei Teilbändern
$G_1(e^{j\omega})$	Übertragungsfunktion des Hochpaßfilters in einer Teilbandsynthese mit zwei Teilbändern
$g_{gl}(\zeta)$	Impulsantwort des Ganglienzellfilters
$G_{gl}(\Omega)$	Übertragungsfunktion des Ganglienzellfilters
$g_{inh}(\zeta)$	Impulsantwort des Inhibitionsfilters
$G_{inh}(\Omega)$	Übertragungsfunktion des Inhibitionsfilters
$g_{optic}(\zeta)$	Point-Spread-Function der Augenoptik
$G_{optic}(\Omega)$	Übertragungsfunktion der Augenoptik

$g_{rez}(\zeta)$	Impulsantwort der Rezeptoren
$G_{rez}(\Omega)$	Übertragungsfunktion der Rezeptoren
G_{SBC}	Teilbandcodierungsgewinn
$g_{syn,m}(\zeta)$	Synthesefilter für das Teilband m
$G_{syn,m}(\Omega)$	Übertragungsfunktion des Synthesefilters für das Teilband m
$g_{\theta}(\zeta)$	Impulsantwort der unterschwelligen Summation
$G_{\theta}(\Omega)$	Übertragungsfunktion der unterschwelligen Summation
$h_0[n]$	Impulsantwort des Tiefpaßfilters in einer Teilbandanalyse mit zwei Teilbändern
$H_0(e^{j\omega})$	Übertragungsfunktion des Tiefpaßfilters in einer Teilbandanalyse mit zwei Teilbändern
$h_1[n]$	Impulsantwort des Hochpaßfilters in einer Teilbandanalyse mit zwei Teilbändern
$H_1(e^{j\omega})$	Übertragungsfunktion des Hochpaßfilters in einer Teilbandanalyse mit zwei Teilbändern
HP	Hilfsparameter zur Bestimmung der optimalen Quantisierungsschrittweiten $\underline{\Delta}_{q,m}$
ISDN	I ntegrated S ervices D igital N etwork
ITU-T	T elecommunication Standardization Sector of I nternational T elecommunication U ion
k_{SAT}	Sättigungsparameter des Modells der Signalverarbeitung in der Fovea
k_{θ}	Summe der Koeffizienten von $g_{\theta}(\zeta)$
l^*	Bildschirmleuchtdichte
Δl	Änderung der Bildschirmleuchtdichte
L_{AD}	Adaptionsleuchtdichte
L_{MON}	Parameter der Bildröhrenggradation
l_{retina}^*	Retinaleuchtdichte
Δl_{retina}	Änderung der Retinaleuchtdichte
l_{rez}^*	Leuchtdichte nach Berücksichtigung der zeitlichen Trägheit der Rezeptoren in der Retina
Δl_{rez}	Änderung von l_{rez}^*
M	Anzahl der Teilbänder

$MON[\cdot]$	nichtlineare Bildröhrenggradation
MPEG	M oving P icture E xpert G roup
MTF	Modulationsübertragungsfunktion (M odulation T ransfer F unction)
N_m/N	Decimation-Faktor des Teilbandes m
OBMC	O verlapped B lock M otion C ompensation
PSB	P otentielle S ichtbarkeit
PSB_{\min}	Minimum von PSB bei einer optimalen Bitaufteilung
\hat{PSB}_{\min}	Schätzung von PSB_{\min}
PCM	P uls C ode M odulation
PSNR	Signal-Rausch-Abstand (P eak S ignal to N oise R atio)
QCIF	Q uarte C ommon I ntermediate F ormat
QMF	Q uadrature M irror F ilter
Q, Q_C	Quantisierungsparameter für die Luminanz- bzw. Chrominanzkomponenten
R	mittlere Datenrate aller Teilbänder in bit/Abtastwert
R_{ist}	resultierende mittlere Datenrate aller Teilbänder in bit/Abtastwert
R_k	mittlere Datenrate aller Teilbänder in bit/Abtastwert aus dem Gebiet k mit konstanten w -Parametern
R_m	Datenrate in bit/Abtastwert des Teilbandes m
\underline{R}_m	Datenrate R_m bei einer optimalen Bitaufteilung
R_{soll}	Soll-Wert der mittleren Datenrate R
R_{ii}	Übertragungsbitrate in bit/s
s	zu codierende Bewegtbildsequenz
s'	rekonstruierte Bewegtbildsequenz aus der Teilbandcodierung
$s^*(\xi)$	Nutzsignal
$\Delta s(\xi)$	Störsignal
$S_{\Delta C_m}(\Omega)$	Leistungsdichtespektrum von Δc_m
$S_{e_m}(\Omega)$	Leistungsdichtespektrum des Quantisierungsfehlersignals e_m
$S_{h_m}(\Omega)$	Leistungsübertragungsfunktion des Teilsystems h_m zwischen e_m und Δc_m
s_{MON}	Parameter der Bildröhrenggradation
$SAT[\cdot]$	Sättigungscharakteristik des Modells der Signalverarbeitung in der Fovea

SF	Synthesefilter
$SF_{\text{äqu}}$	äquivalentes Synthesefilter einer Synthesefilterkaskade
TBC	Teilbandcodierung
w_1, w_2, w_3	orts–zeitveränderliche Faktoren des w -Modells
$\bar{w}_1(j_b), \bar{w}_2(j_b)$	Mittelwert von w_1 und w_2 innerhalb eines zu codierenden Blocks J_b
$W_{SB,m}$	Gewichtungsfaktoren zur Berechnung von PSB aus $\sigma_{q,m}^2$, ($m = 1, \dots, M$)
$w_{SB,m}$	orts– und zeitunabhängiger Anteil des Gewichtungsfaktors $W_{SB,m}$
$x[n]$	Eingangssignal der Teilbandanalyse mit zwei Teilbändern
$X(e^{j\omega})$	Spektrum von $x[n]$
$x'[n]$	Ausgangssignal der Teilbandsynthese mit zwei Teilbändern
$X'(e^{j\omega})$	Spektrum von $x'[n]$
y_m	Teilbandsignal des Teilbandes m
y'_m	quantisiertes Teilbandsignal des Teilbandes m
$z_0[n], z_1[n]$	Teilbandsignale aus der Teilbandanalyse mit zwei Teilbändern
$z'_0[n], z'_1[n]$	quantisierte Teilbandsignale $z_0[n]$ und $z_1[n]$
γ	Exponent der Bildröhrencharakteristik
$\Delta_{q,m}$	orts– und zeitabhängige Quantisiererschrittweite des Teilbandsignals y_m
$\underline{\Delta}_{q,m}$	optimale Einstellung von $\Delta_{q,m}$ bei einer optimalen Bitaufteilung
ε_m^{*2}	Quantisierer–Performance–Faktor [26] für das Teilbandsignal y_m
$\zeta = (x, y, t)$	Orts–zeit–Vektor im Bildschirmkoordinatensystem
θ	Detektionsschwelle für die sichtbaren Störsignale
$\xi_{\odot}(t)$	Winkel der visuellen Achse
σ_{pe}^2	Prädiktionsfehlervarianz
$\sigma_{q,m}^2$	Varianz des Quantisierungsfehlersignals e_m
$\underline{\sigma}_{q,m}^2$	Quantisierungsfehlervarianz $\sigma_{q,m}^2$ bei einer optimalen Bitaufteilung
σ_{re}^2	Rekonstruktionsfehlervarianz
$\sigma_{y,m}^2$	Varianz des Teilbandsignals y_m
$\Omega = (\omega_x, \omega_y, \omega_t)$	örtlich–zeitlicher Kreisfrequenzvektor im Bildschirmkoordinatensystem

1. Einleitung

Begünstigt durch den Fortschritt der Mikroelektronik und der Bildcodierung ist die Bedeutung der digitalen Übertragung von Bewegtbildsequenzen während der letzten Jahre in immer stärkerem Maße gestiegen. Für Übertragung von Bewegtbildsequenzen über das digitale Kommunikationsnetz ISDN (Integrated Services Digital Network) mit niedrigen bis mittleren Übertragungsbitraten von $n \cdot 64$ kbit/s hat die internationale Standardisierungsorganisation CCITT bereits im Jahr 1990 ein Codierungsverfahren mit der Bezeichnung H.261 [9] standardisiert. Nach der Standardisierung der Übertragung von Bewegtbildsequenzen über ISDN gewinnt die digitale Übertragung von Bewegtbildsequenzen über öffentliche analoge Telefonnetze und Mobilkanäle mit sehr niedrigen Übertragungsbitraten von 8 bis 64 kbit/s immer größeres Interesse. Die internationale Standardisierungsorganisation ITU-T hat für diese Anwendung einen neuen Standard H.263 [25] entwickelt, der als Weiterentwicklung des Verfahrens H.261 angesehen werden kann. Darüber hinaus hat sich die internationale Standardisierungsorganisation ISO im Rahmen der Aktivität MPEG-4 mit der Standardisierung eines Verfahrens zur Codierung von Bewegtbildsequenzen mit sehr niedrigen Datenraten befaßt, das weitere Funktionalitäten [23], wie z.B. objekt-orientierte Bildverarbeitung in Multimedia-Anwendungen, ermöglichen soll.

Die Übertragung von Bewegtbildsequenzen mit sehr niedrigen Übertragungsbitraten geht von Bildsequenzen mit dem Bildformat QCIF (Quarter Common Intermediate Format) [9] aus. Eine PCM-Übertragung (Puls Code Modulation) solcher Bewegtbildsequenzen benötigt bei einer Bildfrequenz von 8,33 Hz eine Datenrate von 2,5 mbit/s. Um solche Bewegtbildsequenzen mit einer Übertragungsbirtrate von z.B. 16 kbit/s übertragen zu können, ist eine Bilddatenkompression von 158 zu 1 erforderlich. Zur Erreichung dieser sehr hohen Bilddatenkompression müssen leistungsfähige Codierungsverfahren auf Basis der *Redundanz- und Irrelevanzreduktion* [40][34] angewendet werden. Durch die Redundanzreduktion werden die Teile der Bildinformation von der Übertragung ausgeschlossen, die ein Decoder aus der bereits übertragenen Information rekonstruieren kann. Durch die Irrelevanzreduktion werden die Teile der Bildinformation entfernt, die von einem Beobachter nicht wahrgenommen werden können. Diese irrelevanten Teile der Bildinformation werden durch die Eigenschaften des menschlichen visuellen Systems [17] bestimmt. Wenn durch die Redundanz- und Irrelevanzreduktion allein die erforderliche Datenkompression nicht erreicht werden kann, müssen zusätzliche relevanzreduzierende Verfahren angewendet, oder die Art der zu codierenden Sequenzen eingeschränkt werden. Bei Bildtelefon-Anwendungen wird es beispielsweise

auf Sequenzen mit Kopf–Schulter–Szenen eingeschränkt, um eine ausreichende Bildqualität bei der vorgegebenen Übertragungsbitrate zu erzielen.

Die zur Zeit bekannten Codierungsverfahren zur Übertragung von Bewegtbildsequenzen mit sehr niedrigen Bitraten können in zwei Gruppen eingeteilt werden. Die erste Gruppe bilden die sogenannten *blockbasierten* Codierungsverfahren [9][10][33]. Dabei wird jedes Bild einer Bewegtbildsequenz in Blöcke mit $N \times N$ Bildpunkten zerlegt. Anschließend wird auf jeden Block eine bewegungskompensierende Interframe–Prädiktion angewendet und der Prädiktionsfehler blockweise einer Transformationscodierung [34] unterworfen. Die zweite Gruppe bilden die sogenannten *objektbasierten* Codierungsverfahren [2][13][35][38], in denen die Bilder einer Bewegtbildsequenz in beliebig berandete Bewegtoobjekte zerlegt werden. Jedes Objekt wird durch drei Parametersätze beschrieben, die die Bewegung, Form und Textur des Objektes definieren [22].

Als ein repräsentatives Beispiel der blockbasierten Codierungsverfahren dient das im Standard H.261 angewendete Hybrid–Codierungsverfahren. Dabei wird die zeitliche Redundanz in einer zu übertragenden Bewegtbildsequenz durch die bewegungskompensierende Interframe–Prädiktion reduziert. Durch eine anschließende Intraframe–Verarbeitung des Prädiktionsfehlers mittels einer Diskreten Cosinus Transformation (DCT) wird eine Reduktion der örtlichen Redundanz erzielt. Experimentelle Untersuchungen zeigen, daß bei mittleren bis höheren Übertragungsbitraten das blockbasierte Hybrid–Codierungsverfahren H.261 bezüglich der Bilddatenkompression und der subjektiven Qualität der Rekonstruktionsbilder gute Codierungsergebnisse liefert, weil in diesem Fall die bewegungskompensierten Prädiktionsfehler mit einer ausreichenden Qualität mit Hilfe der DCT codiert werden können. Bei niedrigen bis sehr niedrigen Übertragungsbitraten verursacht dieses Hybrid–Codierungsverfahren jedoch zunehmend subjektiv sehr störende Codierungsartefakte, wie z.B. Block– und Moskitoeffekte. Die Ursachen dafür liegen in der blockbasierten Verarbeitung. Durch die blockbasierte Bewegungskompensation können im Prädiktionsbild Diskontinuitäten des Bildsignals an Blockkanten entstehen. Wenn die Prädiktionsfehler aufgrund der geringen Datenrate nicht hinreichend genau codiert werden können, verursachen diese Diskontinuitäten sichtbare Blockeffekte im Rekonstruktionsbild. Weiterhin können Block– und Moskitoeffekte erzeugt werden, wenn DCT–Koeffizienten infolge grober Quantisierung unterdrückt werden und die zugehörigen 2–dimensionalen Basisfunktionen der DCT im Rekonstruktionsbild fehlen. Beispielsweise werden bei typischen Bildtelefon–Testsequenzen für etwa 30 % der Blöcke DCT–Koeffizienten übertragen. Bei einer Übertragungsbitrate von 16 kbit/s steht dann im Mittel nur 0,15 bit je Koeffizient zur

Verfügung. Bei einer so niedrigen Datenrate sind große Quantisierungsfehler der DCT-Koeffizienten und demzufolge Block- und Moskitoeffekte unvermeidlich.

Zur Vermeidung der durch die DCT-Codierung verursachten Blockeffekte wurde in zahlreichen Veröffentlichungen [3][8][29][62] vorgeschlagen, die blockbasierte DCT durch eine Teilbandcodierung [58][63] zu ersetzen. Bei einer Teilbandcodierung wird auf der Coder-Seite das Spektrum des zu codierenden Bildsignals durch Filterung und anschließende Unterabtastung in mehrere Teilbänder zerlegt. Jedes Teilband enthält nur einen Teil des gesamten Spektrums und kann daher an dessen jeweilige Eigenschaften und die Eigenschaften der menschlichen Wahrnehmung angepaßt codiert und übertragen werden. Auf der Decoder-Seite werden die Teilbandsignale nach der Decodierung, Interpolation und Filterung wieder zu einem Rekonstruktionsbild zusammengefaßt. Der Vorgang der Teilbandzerlegung wird auch als Teilbandanalyse bezeichnet. Dementsprechend wird der Vorgang, in dem ein Bild aus seinen Teilbändern wieder rekonstruiert wird, Teilbandsynthese genannt. Im Gegensatz zur DCT enthält die Teilbandcodierung keine blockweise Verarbeitung des Bildsignals, so daß die Blockeffekte vermieden werden können. Außerdem ermöglicht die Teilbandcodierung einen höheren Codierungsgewinn, weil sie die Korrelation zwischen den benachbarten Blöcken zusätzlich ausnutzen kann.

Zu Beginn dieser Arbeit lagen zwei Veröffentlichungen vor, die sich mit der Teilbandcodierung von Bewegtbildsequenzen bei niedrigen Übertragungsbitraten in der Größenordnung von 64 kbit/s befassen. Beide Verfahren basieren auf einer Teilbandcodierung des Prädiktionsfehlersignals. In einem Coder [62] wird jedes Prädiktionsfehlerbild zunächst mittels einer Filterkaskade bestehend aus Quadrature-Mirror-Filtern (QMF) [27] in 13 Teilbänder zerlegt. Anschließend werden Blöcke mit 4x4 Abtastwerten in einzelnen Teilbändern gebildet, wobei jedes Teilband als ein 2-dimensionales Datenfeld betrachtet wird. Zur Reduzierung der Datenrate werden nur diejenigen Blöcke übertragen, deren Varianz eine vorgegebene Schwelle überschreitet. In [62] wird angeführt, daß durch die dort verwendete Bewegungskompensation keine Vorteile bezüglich der Gesamtdatenrate erzielt werden. Deswegen wird dort eine einfache Interframe-Prädiktion ohne Bewegungskompensation bevorzugt. Durch die Teilbandcodierung werden in diesem Coder die Blockeffekte vermieden. Der Verzicht der Bewegungskompensation verliert aber gewissen Prädiktionsgewinn, der durch Anwendung einer effizienteren Bewegungskompensation erzielt werden kann. Zur Erhöhung der Codierungseffizienz wird im anderen Coder [3] eine effizientere Bewegungskompensation angewendet. Dazu wird eine Displacementschätzung eingesetzt, die eine Kombination von Phasenkorrelation und Blockmatching ist. Für die Teilbandzerlegung des Prädikti-

onsfehlerbildes werden die in [30] vorgestellten symmetrischen Filter mit kurzen Längen verwendet. Diese Filter ermöglichen eine perfekte Rekonstruktion des zu codierenden Signals [59], wenn keine Quantisierungsfehler der Teilbandsignale vorhanden sind. Zur Codierung der Teilbandsignale wird in diesem Fall eine Vektorquantisierung [19][31] angewendet. Dabei werden Interband-Vektoren aus den 16 Teilbändern der Luminanzkomponente und den 8 Teilbändern der beiden Chrominanzkomponenten gebildet. Übertragen werden nur solche Vektoren, deren zugehöriges Gebiet im Prädiktionsfehlerbild nur wenige Null-Werte enthält. Im Vergleich zum ersten Coder erzielt dieser Coder durch die effizientere Bewegungskompensation und die Vektorquantisierung eine Verbesserung der Codierungsergebnisse.

Bei sehr niedrigen Datenraten weisen die erläuterten Teilbandcoder jedoch folgende Mängel auf.

- In diesen Codern wird die bewegungskompensierende Interframe-Prädiktion weiterhin blockweise durchgeführt. Dadurch werden in den Prädiktionsfehlerbildern Diskontinuitäten an den Blockkanten verursacht. Aufgrund der sehr geringen Datenrate können diese Prädiktionsfehler nicht hinreichend genau codiert werden. Nach der Teilbandsynthese erscheinen sie als Überschwinger (englisch: Ringing) in den Rekonstruktionsbildern.
- Aufgrund der sehr niedrigen Datenrate kann aus allen Abtastwerten der Teilbänder nur ein relativ kleiner Teil übertragen werden. Bei der Auswahl der zu übertragenden Teilbandabtastwerte werden die unterschiedlichen Einflüsse der verschiedenen Prädiktionsfehler auf die subjektiven Bildqualität sowie die aktuelle verfügbare Datenrate nicht berücksichtigt. Oft werden sehr viele Teilbandabtastwerte ausgewählt, so daß sie nur mit großen Quantisierungsfehlern codiert werden können.
- Ein Prädiktionsfehlerbild kann in Gebiete, die nur Nullen enthalten, und die restlichen Gebiete mit beliebigen Abtastwerten unterteilt werden. Zu übertragen braucht man nur die Position und die Größe der Abtastwerte in diesen restlichen Gebieten. Führt man auf dem gesamten Prädiktionsfehlerbild eine Teilbandfilterung, wie in den Veröffentlichungen beschrieben, durch, so entstehen Teilbandabtastwerte ungleich Null, die auch außerhalb dieser Gebiete liegen. Zur Rekonstruktion dieser Gebiete durch Teilbandsynthese werden alle diese Teilbandabtastwerte benötigt. Demzufolge ist die Anzahl der zu übertragenden Teilbandabtastwerte größer als die Anzahl der Bildpunkte in diesen Gebieten. Hier könnte die Ursache liegen, warum die theoretischen Vorzüge der Teilbandcodierung im Vergleich zur blockbasierten DCT-Codierung nicht verifiziert werden konnten.

-
- Die Quantisierung der Teilbandsignale wird in den zitierten Teilbandcodern ohne Berücksichtigung der Eigenschaften der menschlichen Wahrnehmung durchgeführt. In den anderen Veröffentlichungen wird die mögliche Irrelevanzreduktion nur zum Teil ausgenutzt, indem sich auf die Ausnutzung der Modulationsübertragungsfunktion des menschlichen visuellen Systems eingeschränkt wird.

Infolge der oben beschriebenen erst drei Mängel der Teilbandcoder entstehen in deren Rekonstruktionsbildern bei sehr niedrigen Übertragungsbitraten Codierungsartefakte, die in Form von Ringing auftreten [1]. Diese Ringing-Effekte sind insbesondere in der Nähe von Luminanzsprüngen sichtbar und können subjektiv als ebenso störend empfunden werden wie die von der DCT hervorgerufenen Block- und Moskitoeffekte. Die Ursachen der Ringing-Effekte liegen in großen Quantisierungsfehlern, die durch die Teilbandsynthese in Ringing umgewandelt werden. Die Form des Ringings entspricht der Impulsantwortfunktion der Synthesefilter.

Ziel der vorliegenden Arbeit ist die Weiterentwicklung der Teilbandcodierungsverfahren zur Übertragung von Bewegtbildsequenzen mit sehr niedrigen Übertragungsbitraten. Dabei sollen in erster Linie die subjektiv sehr störenden Codierungsartefakte, wie Ringing-Effekte, eliminiert werden, so daß eine bessere subjektive Qualität der Rekonstruktionsbilder erreicht werden kann. Ferner soll das neue Codierungsverfahren in der Lage sein, seine Parameter automatisch der vorgegebenen Übertragungsbitrate und den Inhalten der zu codierenden Bewegtbildsequenzen anzupassen, so daß das Verfahren für unterschiedliche Anwendungen sowie für verschiedene Übertragungsbitraten ohne zusätzliche manuelle Anpassung direkt eingesetzt werden kann.

Zur Erreichung des Ziels wird als Ansatz eine blockbasierte bewegungskompensierende Teilbandcodierung ausgewählt. Dabei soll zur Reduzierung der Diskontinuitäten an den Blockkanten eine bekannte Interframe-Prädiktion mit überlappenden Blöcken eingesetzt werden. Um große Quantisierungsfehler zu vermeiden, soll vor der Teilbandcodierung eine Vorwärtsanalyse zur Einschränkung der Anzahl der zu codierenden Blöcke durchgeführt werden. Im Gegensatz zur konventionellen Teilbandcodierung soll die Teilbandanalyse und -synthese in diesem Coder nur auf die ausgewählten Blöcke angewendet werden und keine Vergrößerung der Anzahl der zu codierenden Abtastwerte verursachen. Zur Erzielung einer hohen Irrelevanzreduktion soll die Quantisierung der Teilbandsignale mit Hilfe eines psychovisuellen Modells gesteuert werden, so daß sich die resultierenden Rekonstruktionsfehler möglichst dicht der lokalen Sichtbarkeitschwelle nähern.

Aus diesem Ansatz ergeben sich folgende Teilaufgaben:

- Es ist ein Algorithmus für die Vorwärtsanalyse zur Auswahl der zu codierenden Blöcke zu entwickeln. Dabei ist zu erklären, wie die Prädiktionsfehler bezüglich ihrer subjektiven Bedeutung bewertet werden sollen, so daß die Prädiktionsfehler, die subjektiv besonders störend sind, zuerst codiert und übertragen werden. Es soll weiterhin untersucht werden, wie die Bildfrequenz und die Bandbreite des zu übertragenden Bildsignals abhängig von der verfügbaren Datenrate und der Anzahl der Blöcke mit hoher Priorität verändert werden können.
- Es ist ein Verfahren zur Teilbandanalyse und –synthese der ausgewählten Blöcke zu entwickeln. Dabei ist vor allem zu erklären, wie die vorhandene Technik modifiziert und erweitert werden soll, so daß die Anzahl der zu übertragenden Abtastwerte sich nicht erhöht und kein systematischer Rekonstruktionsfehler verursacht wird.
- Es ist eine lokal–adaptive Quantisierung der Teilbandsignale unter Anwendung eines psychovisuellen Modells zu entwickeln. Dazu soll zunächst beantwortet werden, wie die Quantisierungsfehler der einzelnen Teilbandsignale bei einer vorgegebenen Gesamtdatenrate am besten zu verteilen sind, so daß der Rekonstruktionsfehler dicht an der lokalen Sichtbarkeitsschwelle des Modells liegt.

Im Verlauf dieser Arbeit wurden Codierungsverfahren unter Anwendung von sogenannten "Zerotrees" der Wavelet–Koeffizienten [45][48] veröffentlicht. In [64] wird gezeigt, daß für die Codierung der Bewegtbildsequenzen diese Verfahren keinen eindeutigen Vorteil gegenüber der Hybrid–DCT–Codierung haben. Deshalb wird in dieser Arbeit nicht weit auf die "Zerotrees"–Codierung eingegangen.

Die vorliegende Arbeit gliedert sich wie folgt:

Als Vorbereitung werden in **Kapitel 2** die *Prinzipien der bewegungskompensierenden Teilbandcodierung* beschrieben. Anschließend wird die Struktur des in dieser Arbeit zu entwickelten Teilbandcoders sowie des Referenzcoders TMN5 dargestellt.

In **Kapitel 3** wird eine *Vorwärtsanalyse* zur Auswahl der zu codierenden Blöcke vorgestellt. Nach einer Analyse der unterschiedlichen subjektiven Bedeutungen der verschiedenen Prädiktionsfehler werden die Algorithmen zur Vorwärtsanalyse im Detail beschrieben.

Für die Codierung der ausgewählten Blöcke wird in **Kapitel 4** eine 2–dimensionale *blockbasierte Teilbandcodierung* entwickelt. Als Grundelement dieser Teilbandcodierung wird eine eindimensionale Teilbandanalyse– und –synthesestufe mit zwei Teilbän-

dern unter Anwendung einer symmetrischen periodischen Erweiterung des Signals vorgeschlagen. Anschließend wird die Auswahl der Bandaufteilung und der dafür verwendeten Filter diskutiert.

Kapitel 5 behandelt die *lokal-adaptive Quantisierung der Teilbandsignale*, die anhand eines psychovisuellen Modells gesteuert wird. Dazu wird zunächst ein bekanntes Modell der menschlichen visuellen Wahrnehmung zur Irrelevanzreduktion von Fernsehuminanzsignalen kurz vorgestellt. Anschließend wird erklärt, wie die lokal-adaptiven Quantisierschrittweiten der Teilbandsignale mit Hilfe dieses psychovisuellen Modells berechnet werden können.

In **Kapitel 6** werden die Ergebnisse der *experimentellen Untersuchungen* beschrieben. Zunächst wird der Gewinn der einzelnen in den letzten drei Kapiteln beschriebenen Komponenten dargelegt. Anschließend wird der vorgestellte blockbasierte Teilband-coder mit dem in ITU-T entwickelten Testmodell TMN5 anhand objektiver und subjektiver Gütekriterien verglichen.

2. Prinzipien der bewegungskompensierenden Teilbandcodierung

In Bild 2.1 wird die Grundstruktur der bewegungskompensierenden Teilbandcodierung dargestellt. Diese Grundstruktur ist vergleichbar mit der einer bewegungskompensierenden DCT-Codierung, in der anstelle der Teilbandzerlegung die Cosinus-Transformation angewendet wird. Solche Codierung wird in der Fachliteratur als Hybrid-Codierung bezeichnet. Dabei wird zunächst die zeitliche Redundanz in den zu übertragenden Bewegtbildsequenzen durch eine bewegungskompensierende Interframe-Prädiktion reduziert. Anschließend wird anhand einer Intraframe-Codierung, z.B. Teilbandcodierung oder DCT, die verbleibende örtliche Redundanz im Prädiktionsfehlersignal weiter eliminiert.

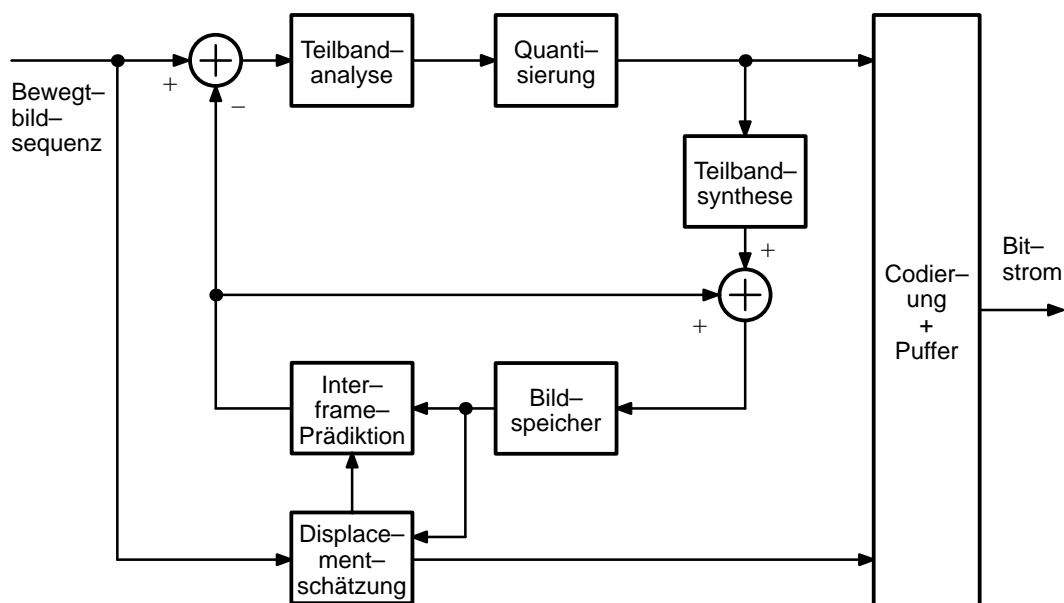


Bild 2.1 Grundstruktur der bewegungskompensierenden Teilbandcodierung

Im folgenden werden zunächst die Grundlagen der bewegungskompensierenden Interframe-Prädiktion sowie der Teilbandcodierung kurz beschrieben. Anschließend wird ein als TMN5 bezeichneter Hybrid-DCT-Coder, der in dieser Arbeit als Referenz dient, vorgestellt. Im Abschnitt 2.4 wird dann die Struktur des in dieser Arbeit zu beschreibenden Coders vorgestellt.

2.1 Bewegungskompensierende Interframe-Prädiktion

In einer Bewegtbildsequenz weisen zwei aufeinander folgende Bilder, abgesehen von der Verschiebung bewegter Objekte, eine hohe Ähnlichkeit auf. Diese Ähnlichkeit kann durch Anwendung einer Interframe-Codierung zur Datenratereduktion ausgenutzt werden. Dabei wird das aktuelle Bild zuerst aus dem vorangegangenen Bild prädiziert und dann statt des Originalbildes das Prädiktionsfehlerbild codiert übertragen. Das Prädiktionsfehlerbild ist die Differenz zwischen dem Originalbild und seinem Prädiktionsbild. Der Gewinn dieser Interframe-Codierung ergibt sich daraus, daß die Signalleistung im Prädiktionsfehlerbild in meisten Fällen wesentlich kleiner als die im Originalbild ist. Die Genauigkeit der Prädiktion und demzufolge der Codierungsgewinn lassen sich erhöhen, wenn bei der Prädiktion die Bewegungen im Bild berücksichtigt werden. Solche Prädiktion wird als bewegungskompensierende Interframe-Prädiktion bezeichnet.

In den internationalen Standards zur Bewegtbildcodierung werden verschiedene Varianten der in Bild 2.2 dargestellten *blockbasierten bewegungskompensierenden Interframe-Prädiktion* eingesetzt. Dazu wird das aktuelle Bild zunächst in Blöcke mit $N \times N$

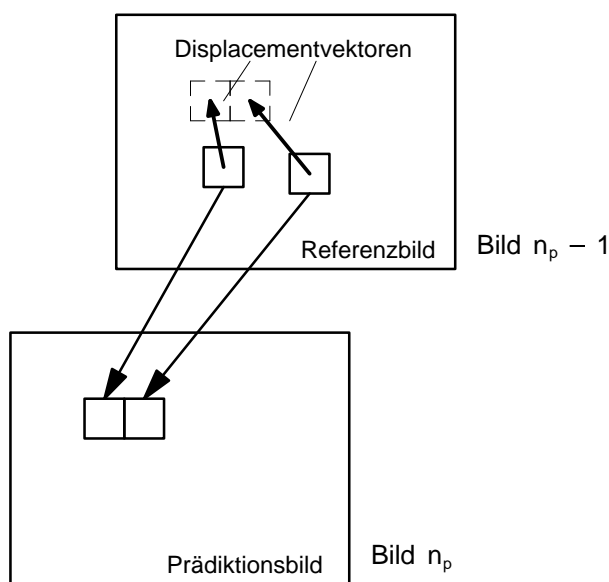


Bild 2.2 Blockbasierte bewegungskompensierende Interframe-Prädiktion

Bildpunkten aufgeteilt. Für jeden Block wird dann einen sogenannten Displacementvektor geschätzt, der die Verschiebung zwischen dem zu prädizierenden Block im aktuellen Bild und dem entsprechenden Referenzblock im letzten Rekonstruktionsbild (Referenzbild) beschreibt. Der Referenzblock stellt die Prädiktion des aktuellen Blocks dar. Das gesamte Prädiktionsbild wird so gewonnen, indem alle Referenzblöcke gemäß der entsprechenden Displacementvektoren zu jeweiligen neuen Positionen verschoben werden. Dieser Technik liegt das Quellenmodell *Block mit translatorischer Verschiebung* zugrunde. Demnach werden die Bewegungsparameter auf die 2D-Displacementvektoren reduziert.

Ein bekanntes Problem dieser herkömmlichen blockbasierten Interframe-Prädiktion ist die Diskontinuität der Signalamplitude, die häufig an den Blockgrenzen im Prädiktionsbild auftritt. Die Ursache liegt darin, daß die Prädiktionen zweier benachbarten Blöcke unterschiedliche Werte an der Blockgrenze aufweisen können. Diese Diskontinuität wirkt subjektiv als störende blockartige Artefakte aus. Daher wird sie in der Fachliteratur als Blockeffekte bezeichnet.

Eine effektive Lösung zur Eliminierung der Blockeffekte ist die Anwendung überlappender Blöcke bei der bewegungskompensierenden Interframe-Prädiktion (Overlap-

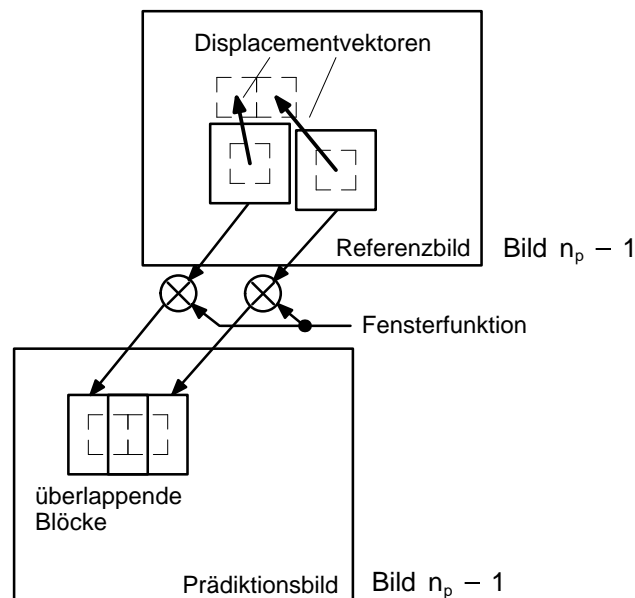


Bild 2.3 Blockbasierte bewegungskompensierende Interframe-Prädiktion mit überlappenden Blöcken

ped Block Motion Compensation, kurz OBMC) [4][37]. Die Funktionsweise dieser Technik läßt sich an einem in Bild 2.3 dargestellten Beispiel erklären. Wie bei der herkömmlichen blockbasierten Interframe-Prädiktion wird das aktuelle Bild ebenfalls in Blöcke mit $N \times N$ Bildpunkten aufgeteilt. Für jeden Block wird ein Displacementvektor geschätzt. Anhand dieses Displacementvektors wird dann ein Referenzblock im letzten Rekonstruktionsbild bestimmt. Der Unterschied zur herkömmlichen blockbasierten Prädiktion liegt darin, daß die Referenzblöcke hier größer als die zu prädizierenden Blöcke sind, so daß sie sich nach Verschiebung zu neuen Positionen örtlich überlappen. Die Prädiktion eines Bildpunktes ergibt sich demnach aus der Summe aller an dieser Stelle sich überlappenden Referenzblöcke, die zuvor mit einer Fensterfunktion multipliziert werden.

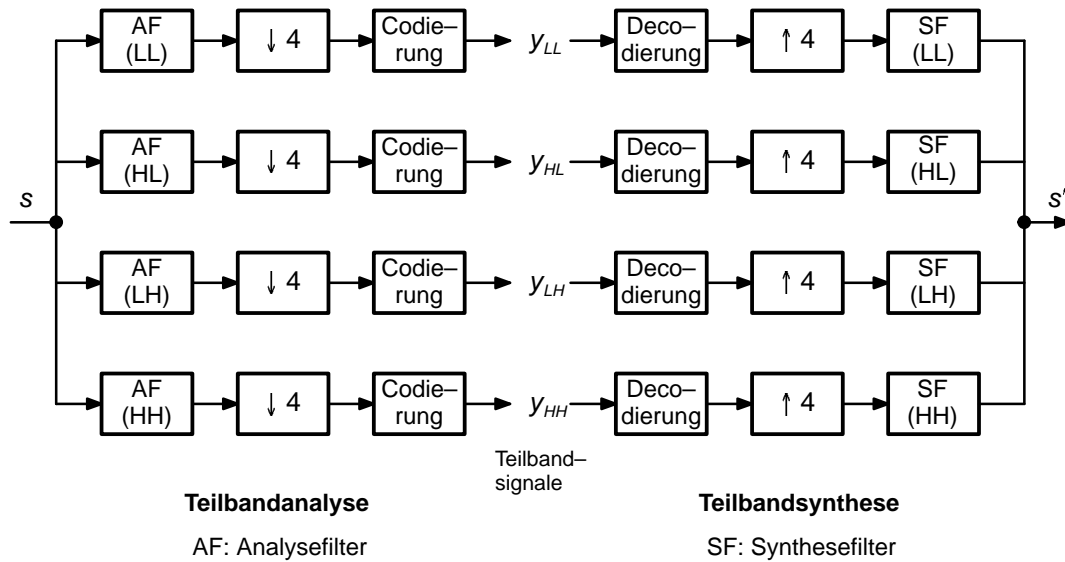
Der Vorteil der Interframe-Prädiktion mit überlappenden Blöcken besteht darin, daß bei richtiger Auswahl der Fensterfunktion die störenden Blockeffekte sich vollständig eliminieren lassen. Das Prädiktionsbild hinterläßt einen natürlicheren Eindruck. Dabei wird auch die Signalleistung der Prädiktionsfehler in der Regel reduziert.

In Referenzen [3][6][34] werden verschiedene Verfahren zur Bewegungsschätzung und bewegungskompensierenden Interframe-Prädiktion ausführlich beschrieben.

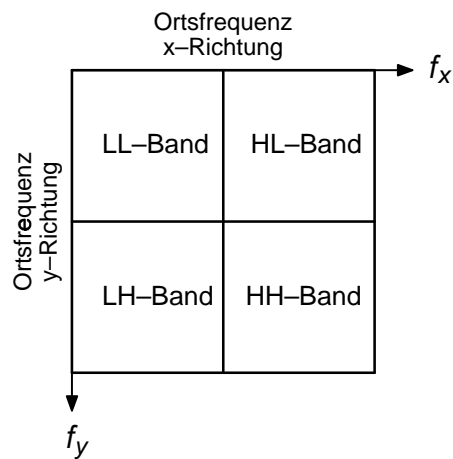
2.2 Teilbandcodierung

Die Teilbandcodierung basiert auf einer spektralen Zerlegung des zu codierenden Signals [12][63]. Das Prinzip der Teilbandcodierung läßt sich am Beispiel in Bild 2.4 erklären.

Bei der Teilbandanalyse wird auf der Coder-Seite das Spektrum der Eingangsbilder durch eine 2-dimensionale Analysefilterung und anschließende Reduktion der Abtastfrequenz in mehrere Teilbänder zerlegt. Jedes Teilband enthält nur einen Teil des gesamten Spektrums und kann daher an dessen jeweilige Eigenschaften angepaßt codiert und übertragen werden. Bei der Teilbandsynthese werden auf der Decoder-Seite die Bilder nach der Decodierung, der Erhöhung der Abtastfrequenz, der Synthesefilterung und Summierung der Teilbänder rekonstruiert. In Bild 2.4 wird beispielsweise das Spektrum der Eingangsbilder in der horizontalen sowie der vertikalen Richtung in je zwei Teilbänder zerlegt. Die dadurch entstandenen 4 Teilbänder werden als LL-, HL-, LH- und HH-Teilband gekennzeichnet. Die Buchstaben L und H stehen für "Lowpass" und "Highpass". Zum Beispiel enthält das HL-Band horizontal hochfrequente und vertikal niederfrequente Ortsfrequenzen. Diese Bandaufteilung ist in Bild 2.4 (b) dargestellt.



(a) Blockschaltbild



(b) Bandaufteilung

Bild 2.4 Teilbandcodierung mit 4 Teilbändern

Der Vorteil der Teilbandcodierung liegt darin, daß man für einzelne Teilbänder unterschiedliche Codierungsverfahren, die genau an die Eigenschaften der jeweiligen Teil-

bandsignale angepaßt sind, verwenden kann. Dadurch wird einerseits die ungleiche Verteilung der Signalleistung im Spektrum und andererseits die unterschiedliche Wahrnehmungsfähigkeit des menschlichen visuellen Systems für verschiedene Ortsfrequenzen ausgenutzt [56]. Man kann daher einen großen Codierungsgewinn aufgrund der Teilbandcodierung erwarten. Im Gegensatz zur DCT werden bei der Teilbandcodierung die Blöcke des Bildes nicht getrennt sondern abhängig voneinander behandelt, so daß die Blockeffekte vermieden werden können. Außerdem ermöglicht die Teilbandcodierung einen höheren Codierungsgewinn, weil sie die Korrelation zwischen den benachbarten Blöcken zusätzlich ausnutzen kann. Allerdings ist dieser Vorteil mit einem erhöhten Realisierungsaufwand verbunden.

Wie oben dargestellt, werden die Teilbandanalyse und –synthese durch Anwendung von Analyse– und Synthesefiltern realisiert [55]. Dabei können die Filter so wohl *parallel* als auch *kaskadiert* eingesetzt werden. Bei einer parallelen Filterung wird für jedes Teilband ein separates Filter verwendet. Dadurch läßt sich die Bandaufteilung flexible gestalten. Bei einer kaskadierten Filterung wird das 2D–Spektrum des Eingangssignals stufenweise in beiden Richtungen zerlegt. Jede Filterstufe unterteilt das Spektrum ihres Eingangssignals in einen tieffrequenten und einen hochfrequenten Anteil. Der Realisierungsaufwand einer solchen Filterkaskade ist im Vergleich zur Anwendung von parallelen Filtern geringer. Damit können jedoch nur Bandaufteilungen mit den Bandbreiten aus $1/(2^n)$ der gesamten Bandbreite je Richtung erzielt werden, wobei n die Anzahl der Filterstufen in der Filterkaskade dieser Richtung ist.

Ein Grundelement der Filterkaskaden ist die in Bild 2.5 dargestellte Basisfilterstufe, die eine Teilbandanalyse und –synthese mit zwei Teilbändern auf das 1–dimensionale Eingangssignal ausführt. Ein wichtiges Thema beim Entwurf solcher Filterstufe ist die

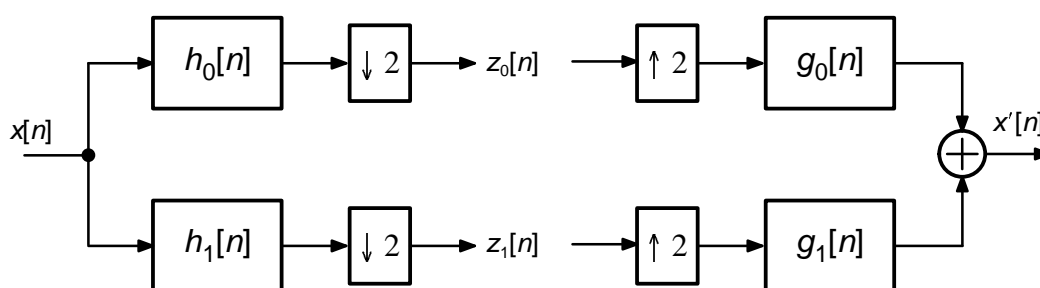


Bild 2.5 Basisfilterstufe für eine eindimensionale Teilbandanalyse und –synthese mit zwei Teilbändern

Behandlung der Aliasings in den Teilbandsignalen, die aufgrund der nicht idealen Bandtrennung durch realisierbare Filter immer vorhanden sind.

Der Zusammenhang zwischen dem Ausgangssignal $X'(e^{j\omega})$ und dem Eingangssignal $X(e^{j\omega})$ in Bild 2.5 läßt sich mit Hilfe des in Bild 2.6 dargestellten Ersatzschaltbildes ermitteln. In Bild 2.6 werden die Reduktion der Abtastfrequenz um den Faktor 2 und die anschließende Erhöhung der Abtastfrequenz um den Faktor 2 zu einer Multiplikation mit $[1 + (-1)^n]/2$ zusammengefaßt. Unter Anwendung der Beziehung

$$f[n] \cdot [1 + (-1)^n] \quad \bullet \text{---} \circ \quad F(z) + F(-z) \quad (2.1)$$

$$f[n] \cdot [1 + (-1)^n] \quad \bullet \text{---} \circ \quad F(e^{j\omega}) + F(e^{j(\omega+\pi)}) \quad (2.2)$$

ergibt sich aus Bild 2.6

$$\begin{aligned} X'(e^{j\omega}) &= \frac{1}{2} [X(e^{j\omega}) H_0(e^{j\omega}) + X(e^{j(\omega+\pi)}) H_0(e^{j(\omega+\pi)})] G_0(e^{j\omega}) \\ &\quad + \frac{1}{2} [X(e^{j\omega}) H_1(e^{j\omega}) + X(e^{j(\omega+\pi)}) H_1(e^{j(\omega+\pi)})] G_1(e^{j\omega}) \\ &= \frac{1}{2} [H_0(e^{j\omega}) G_0(e^{j\omega}) + H_1(e^{j\omega}) G_1(e^{j\omega})] X(e^{j\omega}) \\ &\quad + \frac{1}{2} [H_0(e^{j(\omega+\pi)}) G_0(e^{j\omega}) + H_1(e^{j(\omega+\pi)}) G_1(e^{j\omega})] X(e^{j(\omega+\pi)}) , \end{aligned} \quad (2.3)$$

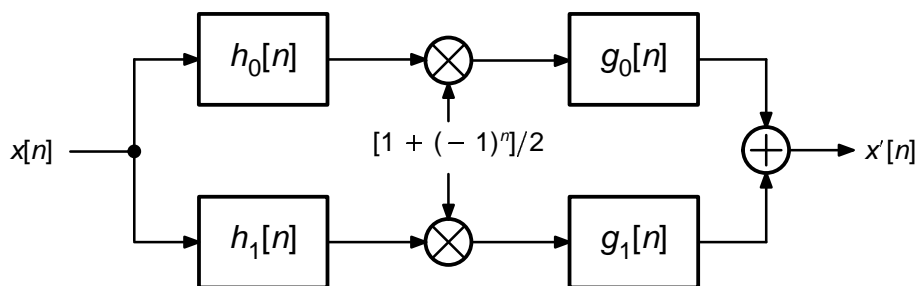


Bild 2.6 Zusammengefasstes Ersatzschaltbild der Analyse- und Synthesestufe in Bild 2.5

wobei $H_0(e^{j\omega})$ und $H_1(e^{j\omega})$ die Übertragungsfunktionen des Tiefpaß- bzw. Hochpaß-Analysefilters sind. $G_0(e^{j\omega})$ und $G_1(e^{j\omega})$ sind die der entsprechenden Synthesefilter. Wie in Gl. (2.3) gezeigt, setzt sich das Ausgangssignal $X'(e^{j\omega})$ aus zwei Teilen zusammen. Der erste Teil stammt vom ursprünglichen Eingangssignal $X(e^{j\omega})$, während der zweite Teil sich aus dem Aliasing $X(e^{j(\omega+\pi)})$ ergibt. Um eine Aliasings-freie und nahezu exakte Rekonstruktion zu erzielen, müssen die Analyse- und Synthesefilter die folgenden Bedingungen erfüllen:

$$H_0(e^{j(\omega+\pi)}) G_0(e^{j\omega}) + H_1(e^{j(\omega+\pi)}) G_1(e^{j\omega}) = 0 \quad (2.4)$$

$$\frac{1}{2} [H_0(e^{j\omega}) G_0(e^{j\omega}) + H_1(e^{j\omega}) G_1(e^{j\omega})] \approx e^{-j\omega k} \quad (2.5)$$

Dabei bezieht sich die erste Bedingung in Gl. (2.4) auf die Aliasing und die zweite in Gl. (2.5) auf die Amplituden- und Phasenverzerrungen des Ausgangssignals. Gleichung (2.4) bedeutet, daß die in den Teilbandsignalen enthaltenen Aliasings bei der Teilbandsynthese wieder gegenseitig kompensiert werden, so daß sich schließlich ein Aliasing-freies Ausgangssignal wieder ergibt.

Unter Berücksichtigung der Bedingungen in Gl. (2.4) und (2.5) wurden für die Teilbandanalyse und -synthese mit zwei Teilbändern bereits verschiedene Typen von speziellen Filtern entwickelt [27][30][49]. Dazu zählen insbesondere die Quadraturspiegelfilter (Quadrature Mirror Filter, kurz QMF) [11][12][27] und die konjugierten Quadraturspiegelfilter (Conjugate Quadrature Mirror Filter, kurz CQF) [49]. Für die Aliasing-Kompensation bestehen zwischen den Analyse- und Synthesefiltern dieser Typen die folgenden Beziehungen.

$$G_0(e^{j\omega}) = 2 H_0(e^{j\omega}) \quad (2.6)$$

$$H_1(e^{j\omega}) = H_0(e^{j(\omega+\pi)}) \quad (2.7)$$

$$G_1(e^{j\omega}) = -2 H_0(e^{j(\omega+\pi)}) \quad (2.8)$$

Wie in Gl. (2.7) und Bild 2.7 gezeigt, ist die Übertragungsfunktion des Hochpaßfilters $H_1(e^{j\omega})$ um die Kreisfrequenz $\pi/2$ spiegelsymmetrisch zu der des Tiefpaßfilters $H_0(e^{j\omega})$.

Die Quadraturspiegelfilter unterscheiden sich von den konjugierten Quadraturspiegelfiltern durch den Phasengang und die Möglichkeit der exakten Rekonstruktion. Die

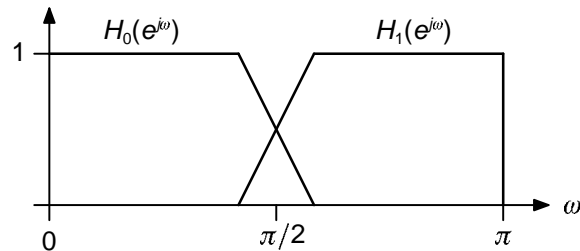


Bild 2.7 Schematische Darstellung der Quadraturspiegelfilter

QMF besitzen eine symmetrische Impulsantwortfunktion und demzufolge einen linearen Phasengang. Mit QMF läßt sich aber theoretisch keine exakte Rekonstruktion des zu codierenden Signals erzielen, da die QMF die Bedingung in Gl. (2.5) nur annähernd erfüllen. In der Praxis sind diese Rekonstruktionsfehler jedoch so klein, daß sie bei einer 8-bit-Darstellung der Rekonstruktionsbilder kaum noch meßbar sind. Im Gegensatz zu QMF haben die CQF keinen linearen Phasengang, erlauben aber eine exakte Rekonstruktion des zu codierenden Signals.

Während der Entwicklung des zu beschreibenden Teilbandcoders wurde sich mit verschiedenen Themen befaßt, die relevant zur Teilbandcodierung sind. Dazu zählen u.a.

- Symmetrische periodische Erweiterung der Signale,
- Auswahl der Bandaufteilungsschema,
- Auswahl der Analyse- und Synthesefilter,
- Quantisierung und Codierung der Teilbandsignale.

Diese Themen werden in Kapitel 4 und 5 ausführlich behandelt.

2.3 Hybrid-DCT-Coder TMN5

In dieser Arbeit wird zur Beurteilung des Teilbandcoders der Hybrid-DCT-Coder TMN5 als Referenz herangezogen. Der Coder TMN5 wurde während des Entwurfs des internationalen Standards H.263 für Videocodierung unter 64 kbit/s von einer ITU-T-Expertengruppe als Testmodell entwickelt. Dieser Coder stellt der Zeit den Stand der Technik dar.

In Bild 2.8 wird das Blockschaltbild des Referenzcoders TMN5 dargestellt. Dieser Coder besteht aus folgenden Basiskomponenten:

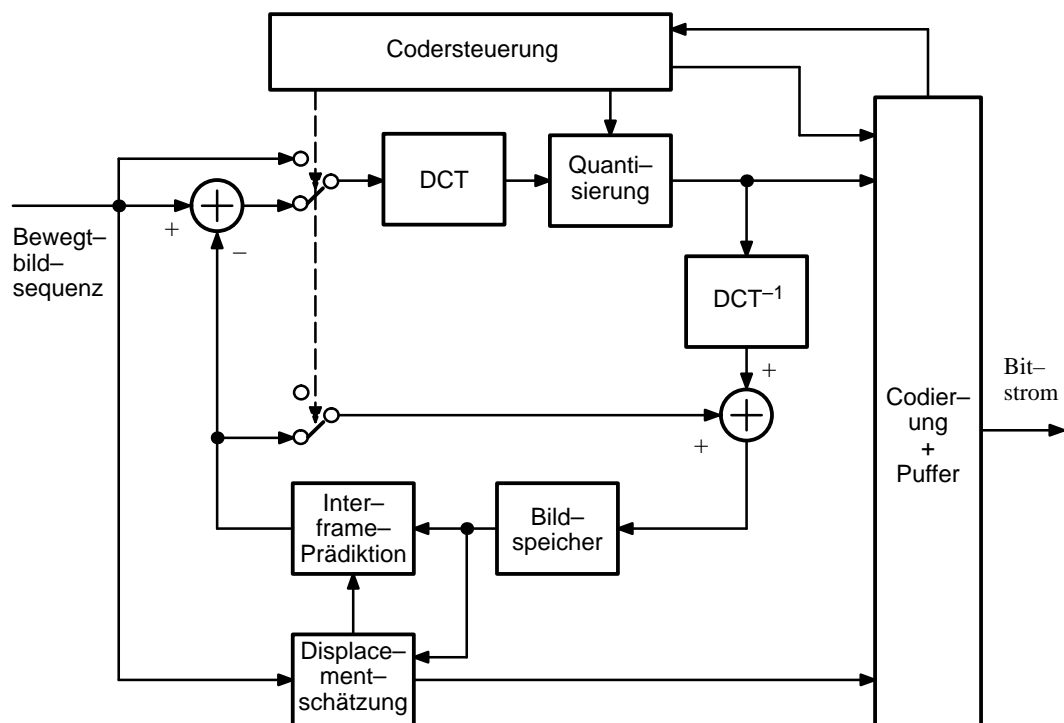


Bild 2.8 Blockschaltbild des Hybrid-DCT Coders TMN5

- Bewegungsschätzung und bewegungskompensierende Interframe-Prädiktion mit einer Halbpel-Genauigkeit und einem Wertbereich von -16 bis $+15.5$ Bildpunkten, wobei die Bewegungskompensation standardmäßig auf Makroblöcke angewendet wird. Ein Makroblock enthält vier benachbarte 8×8 Blöcke der Luminanzkomponente und zwei zugehörige 8×8 Blöcke der beiden Chrominanzkomponenten.
- Entscheidung für jeden Makroblock, ob das Prädiktionsfehlersignal oder das Originalbildsignal codiert werden soll.
- DCT des Prädiktionsfehlersignals bzw. Originalbildsignals mit einer Blockgröße von 8×8 Bildpunkten.
- gleichförmige Quantisierung der DCT-Koeffizienten. Die Quantisierungsschrittweite beträgt einen geraden Wert zwischen 2 und 62.
- Zickzack-Scan und anschließende 2D-Lauflängencodierung der quantisierten DCT-Koeffizienten.

Neben diesen Basiskomponenten können folgende vier optionale Modi zur Verbesserung der Codierungseffizienz unabhängig voneinander eingesetzt werden:

1. *Unbeschränkte Bewegungsvektoren*: Die Bewegungsvektoren dürfen auch auf einen Punkt außerhalb eines Bildes zeigen. Der Wertebereich der Bewegungsvektoren wird von $(-16, +15.5)$ auf $(-31.5, +31.5)$ erweitert, damit größere Bewegungen kompensiert werden können.
2. *Syntax-basierte arithmetische Codierung*: In diesem Modus ersetzt die syntax-basierte arithmetische Codierung die ursprüngliche Codierung mit variablen Codewortlängen.
3. *Fortschrittliche Prädiktion*: In diesem Modus wird die Interframe-Prädiktion mit *überlappenden* Blöcken (OBMC) auf die Luminanzkomponente angewendet. Statt für jeden Block mit 16×16 Bildpunkten wird hier für jeden Block mit 8×8 Bildpunkten ein Displacementvektor geschätzt und übertragen. Die erste Technik eliminiert die Blockeffekte. Die zweite ermöglicht eine präzisere Bewegungskompensation.
4. *PB-Frames*: Ein PB-Frame besteht aus einem P-Bild und einem B-Bild. Für ein P-Bild wird nur eine zeitliche Vorwärtsprädiktion durchgeführt. Für ein B-Bild werden hingegen sowohl eine Vorwärts- als auch eine Rückwärtsprädiktion verwendet. In diesem Modus werden die beiden Bilder in einem PB-Frame gemeinsam codiert. Gegenüber eine separate Codierung der beiden Bilder als P-Bilder wird eine weitere Datenratereduktion erzielt.

Von den optionalen Modi werden im Referenzcoder die Modi 1 und 3 verwendet, da der Teilband vergleichbare Komponenten enthält.

2.4 Struktur des entwickelten Teilbandcoders

In Bild 2.9 wird die Struktur des entwickelten bewegungskompensierenden Teilbandcoders dargestellt. Im Vergleich zur in Bild 2.1 vorgestellten Grundstruktur enthält diese Struktur folgende neue Komponenten:

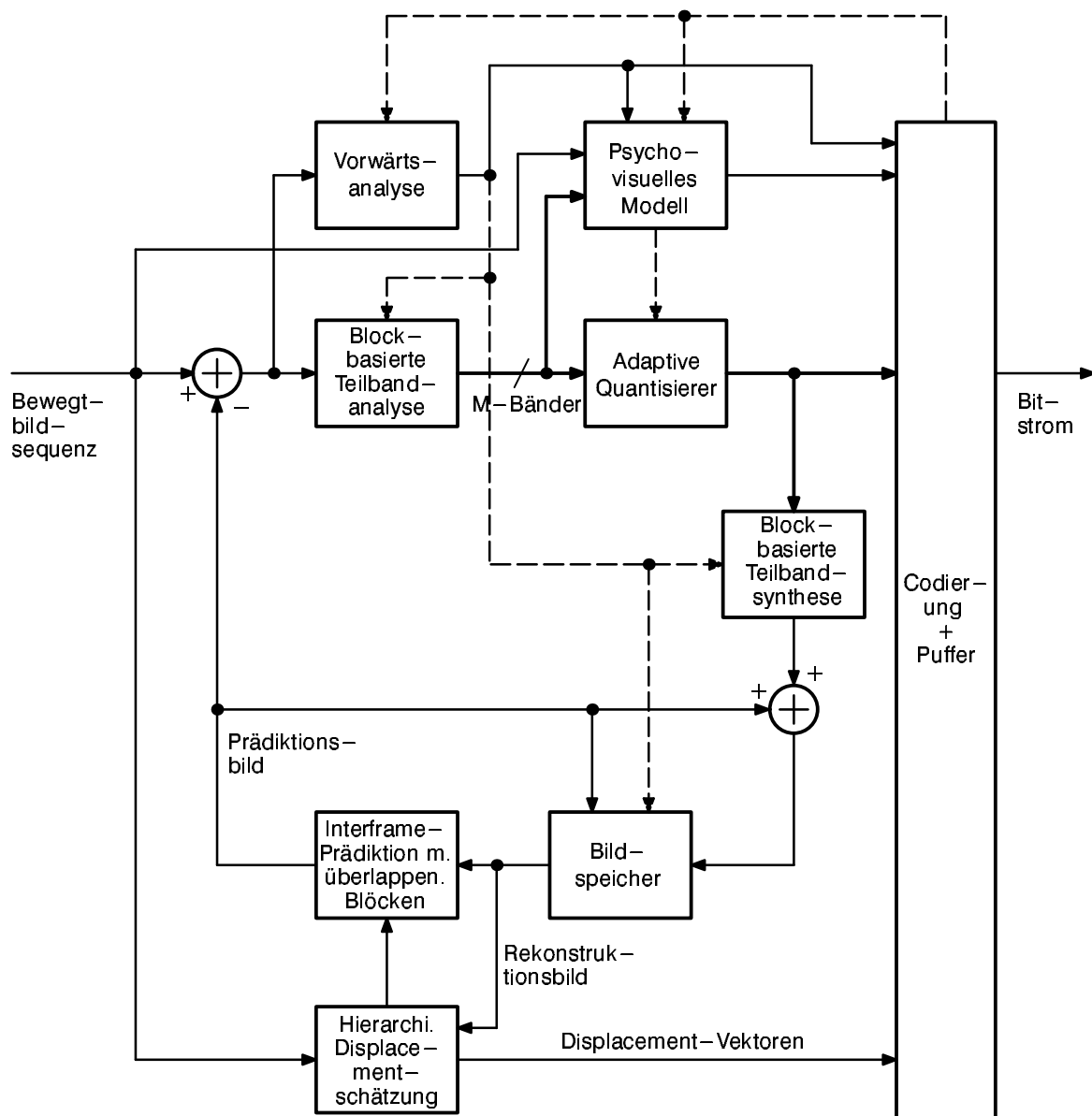


Bild 2.9 Blockschaltbild des bewegungskompensierenden Teilbandcoders

1. Vorwärtsanalyse,
2. Blockbasierte Teilbandcodierung,
3. Adaptive Quantisierung der Teilbandsignale unter Anwendung eines psychovisuellen Modells.

Die *Vorwärtsanalyse* dient u.a. zur Auswahl der zu codierenden Blöcke eines Bildes, die nur große, relevante Prädiktionsfehler enthalten. Dadurch soll die Anzahl der zu codierenden Blöcke eines Bildes deutlich reduziert werden.

Mit der *blockbasierten Teilbandcodierung* sollen die Prädiktionsfehler dieser ausgewählten Blöcke komprimiert werden. Im Gegensatz zur konventionellen Teilbandcodierung wird die Teilbandanalyse- und -synthese hier nur auf die ausgewählten Blöcke angewendet.

Die Anwendung *eines psychovisuellen Modells* zur Steuerung der adaptiven Quantisierung soll die Irrelevanzreduktion durch Ausnutzung der Eigenschaften des menschlichen visuellen Systems erhöhen.

Zusätzlich wird eine Bewegungskompensierende Interframe-Prädiktion mit überlappenden Blöcken in Verbindung mit einer hierarchischen Displacementschätzung [6] eingesetzt, um die Blockeffekte in den Prädiktionsbildern zu eliminieren und den Prädiktionsgewinn zu erhöhen. Anhand der in [6], Kapitel 3.2.1, vorgeschlagenen Parameter wird für jeden Luminanzblock mit 16x16 Bildpunkten ein halbpel-genauer Displacementvektor ermittelt. Zur Prädiktion dieser Blöcken werden Referenzblöcke mit 24x24 Bildpunkten verwendet.

Die Funktionsweise dieses Coders läßt sich wie folgt zusammenfassen:

Zunächst werden zwischen dem aktuellen Eingangsbild und dem letzten Rekonstruktionsbild im Bildspeicher die Displacementvektoren geschätzt. Anhand dieser Displacementvektoren wird dann das aktuelle Prädiktionsbild aus dem letzten Rekonstruktionsbild berechnet. Die Differenz zwischen dem Eingangsbild und dem Prädiktionsbild liefert das Prädiktionsfehlerbild. Vor der Teilbandcodierung wird eine Vorwärtsanalyse auf die Prädiktionsfehler durchgeführt. Dabei werden u.a. Blöcke mit großen relevanten Prädiktionsfehlern ausgewählt. Anschließend werden diese ausgewählten Blöcke der Prädiktionsfehler anhand einer blockbasierten Teilbandanalyse in Teilbänder zerlegt. Im nächsten Schritt werden die Teilbandsignale skalar-quantisiert, wobei die Quantisierungsschrittweiten mit Unterstützung eines psychovisuellen Modells lokal-adaptiv eingestellt werden. Für die Übertragung werden die quantisierten Teilbandsignale mittels einer 2D-Lauf längencodierung codiert. Zur Rekonstruktion des Bildes werden zuerst

für die ausgewählten Blöcke die Prädiktionsfehler aus den quantisierten Teilbandsignalen durch eine blockbasierte Teilbandsynthese berechnet. Für die restlichen Blöcke werden die Prädiktionsfehler auf Null gesetzt. Daraus ergibt sich ein rekonstruiertes Prädiktionsfehlerbild. Dieses Prädiktionsfehlerbild wird dann mit dem Prädiktionsbild addiert. Schließlich wird das resultierende Rekonstruktionsbild in dem Bildspeicher abgelegt und steht für die Prädiktion des nächsten Bildes zur Verfügung.

Zur Codierung des ersten Bildes in einer Szene arbeitet dieser Coder in einem *Intraframe-Mode*. Dabei werden die Funktionsblöcke für die Displacementschätzung, die Interframe-Prädiktion und die Vorwärtsanalyse deaktiviert. Das Prädiktionsbild enthält nun den Wert Null. Damit werden die Teilbandanalyse und -synthese auf das gesamte Eingangsbild angewendet. Zusätzlich wird im Teilband mit tiefsten Ortsfrequenzen eine weitere intra-bande Prädiktion durchgeführt, um weitere Redundanzreduktion zu erzielen.

Die Syntax des Bitstroms lehnt sich in diesem Teilbandcoder eng an den Standard H.263 an. Zur Codierung zusätzlicher Information aus diesem Coder wird die vorhandene Syntax entsprechend erweitert oder angepaßt. Diese Änderungen werden in Abschnitten 3.4, 4.4 und 5.5 angegeben.

In den nächsten Kapiteln werden die einzelnen Komponenten des Codes, die im Rahmen dieser Arbeit entwickelt werden, detailliert beschrieben.

3. Vorwärtsanalyse

Bei sehr niedrigen Übertragungsbitraten ist es meistens nicht möglich, die Prädiktionsfehler in allen Bildgebieten vollständig zu codieren und übertragen. In diesem Kapitel wird gezeigt, wie die Anzahl der Bildpunkte, deren Prädiktionsfehler übertragen werden, durch eine Auswahl reduziert wird, so daß die verfügbare Bitrate effektiver verwendet werden kann. Zur Auswahl der zu codierenden Prädiktionsfehler und der anderen Steuerparameter des Coders wird hier eine sogenannte Vorwärtsanalyse eingesetzt. Durch die Vorwärtsanalyse können die unterschiedlichen subjektiven Bedeutungen verschiedener Prädiktionsfehler besser berücksichtigt werden. Außerdem ermöglicht die Vorwärtsanalyse eine optimale Einstellung der Bildfrequenz und Bandbreite des zu übertragenden Bildsignals.

In diesem Kapitel wird nach einer Analyse der Ansätze zur Vorwärtsanalyse auf die Klassifizierung der Prädiktionsfehler gemäß deren subjektiven Bedeutungen eingegangen. Anschließend werden die Algorithmen zur Vorwärtsanalyse im Detail beschrieben.

3.1 Ansätze und Aufgaben der Vorwärtsanalyse

In der konventionellen Teilbandcodierung von Bewegtbildsequenzen werden die Prädiktionsfehler des Gesamtbildes, wie in Bild 2.1 gezeigt, zuerst in Teilbänder zerlegt und die Teilbandabtwerte dann quantisiert. In Bildgebieten, in denen die entsprechenden Abtwerte nach der Quantisierung nicht in allen Teilbändern gleich Null sind, werden diese Abtwerte zusammen mit der Form und Position dieser Gebiete codiert und übertragen. Solche Gebiete werden in dieser Arbeit als codierte Gebiete bezeichnet. Im Decoder wird das Bild aus dem Prädiktionsbild und den übertragenen Teilbandabtwerten des Prädiktionsfehlerbildes wieder rekonstruiert. Zur Regulierung der Datenrate wird im Coder die Quantisiererschrittweite variiert. Je niedriger die verfügbare Datenrate ist, desto größer sind die Quantisierungs- und demzufolge auch die Rekonstruktionsfehler. Zur Veranschaulichung der Codierungsergebnisse werden Computersimulationen solchen Teilbandcoders durchgeführt. In Tabelle 3.1 werden die Größe der codierten Bildgebiete und die mittlere Datenrate gemessen in bit/Bildpunkt aus der Teilbandcodierung dargestellt. Diese Werte ergaben sich aus Simulationen mit den Testsequenzen 'Miss America' und 'Claire' (siehe Abschnitt 6.1) bei einer Übertragungsbitrate von 16 kbit/s. Bei der Berechnung der Datenrate wird nur die Datenmenge aus der Codierung der Teilbandabtwerte, nicht aber die aus der Codierung der Seiteninformation berücksichtigt. Tabelle 3.1 zeigt deutlich, daß bei sehr niedrigen Übertragungsbitraten nur eine extrem geringe Datenrate (ca. 0,1bit/Bildpunkt) für die Codierung der Teilbandabtwerte verwendet wird. Dadurch entstehen sehr große

Quantisierungsfehler in den Teilbändern. Nach der Teilbandsynthese erscheinen sie als subjektiv sehr störende Codierungsartefakte, wie in Bild 3.1 gezeigt, in den Rekonstruktionsbildern. Die durch die Quantisierung der Teilbandsignale verursachten Überschwinger wirken subjektiv an manchen Stellen störender als die ursprünglichen Prädiktionsfehler.

Tabelle 3.1 Größe der codierten Gebiete und die mittlere Datenrate der Teilbandcodierung von den Testsequenzen 'Miss America' und 'Claire' bei 16 kbit/s aus einem konventionellen Teilbandcoder

Testsequenz	Größe der codierten Gebiete [%]	Datenrate [bit/Bildpunkt]
Miss America	34,1	0,12
Claire	35,0	0,13

Aus diesen Ergebnissen ist zu erkennen, daß in diesem konventionellen Teilbandcoder die Größe der codierten Bildgebiete im Mißverhältnis zur verfügbaren Bitrate steht. Mit einer so niedrigen Datenrate ist eine akzeptable Qualität bei der Codierung der Prädiktionsfehler nicht mehr zu erreichen. Da die Rekonstruktionsfehler in manchen Gebieten subjektiv störender als die ursprünglichen Prädiktionsfehler sind, empfiehlt es sich, solche Gebiete von der Teilbandcodierung auszuschließen. Dadurch läßt sich die Größe der zu codierenden Gebiete reduzieren und die durchschnittliche Datenrate in bit/Bildpunkt erhöhen. Demzufolge ist eine Verbesserung der Bildqualität zu erwarten. Die Idee solche Gebiete von der Teilbandcodierung auszuschließen ist der erste Ansatz der Vorwärtsanalyse.

Neben der Größe der zu codierenden Gebiete können in einem Teilbandcoder noch weitere Parameter: nämlich die Bildfrequenz und Bandbreite sowie der Signal-Rausch-Abstand, eingestellt werden, um die erforderliche Datenreduktion zu erreichen. Die gegenseitige Anpassung zwischen der Bildfrequenz und Bandbreite sowie dem Signal-Rausch-Abstand ist der zweite Ansatz der Vorwärtsanalyse. Die Bildfrequenz wird durch die Anzahl der zu codierenden Bilder pro Sekunde definiert. Die Reduktion der Bildfrequenz läßt sich durch Weglassen von Eingangsbildern bei der Codierung erzielen. Diese fehlenden Bilder werden im Decoder durch Wiederholung der übertragenen Bilder ersetzt. Die Bandbreite des Bildsignals wird durch die höchsten Ortsfrequenzen in seinem zweidimensionalen Spektrum bestimmt. Bei der Teilbandcodierung kann die Reduktion der Signalbandbreite durch Nicht-Übertragung von hoch-



Bild 3.1 Rekonstruktionsbilder eines konventionellen Teilbandcoders bei einer Übertragungsbitrate von 16 kbit/s

frequenten Teilbändern realisiert werden. Der Signal–Rausch–Abstand des Rekonstruktionssignales läßt sich durch Variieren der Quantisiererschrittweite einstellen. Für die gleiche Datenratekompression gibt es verschiedene Einstellungsmöglichkeiten dieser Parameter. Diese Einstellungen führen jedoch zu unterschiedlichen subjektiven Bildqualitäten. Im allgemeinen wirkt sich eine einseitige Reduzierung eines dieser Parameter negativ aus. Zum Beispiel können bei einer sehr niedrigen Bildfrequenz die Bewegungsvorgänge der Objekte trotz einer perfekten Rekonstruktion der einzelnen übertragenen Bilder nicht mehr genau wiedergegeben werden. Andererseits kann der Bildinhalt bei großen Rekonstruktionsfehlern trotz einer hohen Bildfrequenz nur schwer zu erkennen sein. Deshalb ist eine aufeinander abgestimmte Einstellung dieser Parameter erforderlich, um bei einer gegebenen Datenrate eine möglichst hohe Bildqualität zu erzielen.

Bei der Teilbandcodierung werden nach der Teilbandsynthese immer solche Bildgebiete rekonstruiert, die in nicht überlappende Blöcke gleicher Größe in einem festen Raster aufgeteilt werden können. Die Kantenlänge dieser Blöcke ist gleich der Anzahl der Teilbänder in der jeweiligen Richtung. Zu jedem solchen Block gehört örtlich ein einziger Abtastwert in jedem Teilband, und die Anzahl der Teilbänder ist gleich der Anzahl der Abtastwerte in diesem Block. Aus diesem Grund werden in diesem Coder nur Gebiete aus solchen Blöcken für die Teilbandcodierung ausgewählt. Die Auswahl der zu codierenden Gebiete wird auch direkt als die Auswahl der zu codierenden Blöcke bezeichnet.

Die Anzahl der Teilbänder bestimmt nicht nur die Blockgröße sondern auch in folgender Weise die Codierungseffizienz der blockbasierten Teilbandcodierung, die in Kapitel 4 noch genau beschrieben wird. Durch die Erhöhung der Anzahl der Teilbänder läßt sich einerseits der Codierungsgewinn verbessern. Damit kann die erforderliche Datenrate in bit/Bildpunkt beim gleichen Signal–Rausch–Abstand reduziert werden. Andererseits ergibt eine größere Anzahl der Teilbänder auch eine größere Blockgröße, so daß die ausgewählten Blöcke insgesamt mehr Bildpunkte als im Fall mit wenigen Teilbändern enthalten können. Wegen der erhöhten Anzahl der zu codierenden Bildpunkte kann die gesamte Datenrate in bit/Bild aus der Codierung mit mehr Teilbändern in ungünstigen Fällen größer sein als die mit wenigen Teilbändern. Um dies zu vermeiden, läßt sich in diesem Coder die Anzahl der Teilbänder von 64 auf 16 für jedes Bild umstellen. Die Auswahl der Anzahl der Teilbänder für einzelne Bilder ist die weitere Aufgabe der Vorwärtsanalyse.

Zusammengefaßt hat die vorgesehene Vorwärtsanalyse die folgenden Aufgaben:

- Auswahl der zu codierenden Blöcke innerhalb eines Bildes;

- Festlegung der Bildfrequenz bei der Codierung und Übertragung;
- Festlegung der Bandbreite des zu übertragenden Prädiktionsfehlersignals;
- Auswahl der Anzahl der Teilbänder.

Der Begriff Vorwärtsanalyse wurde bereits in [20] in Verbindung mit der Auswahl von zu codierenden Blöcken in einem Hybrid-DCT-Coder verwendet. In [20] wurden nach einer Unterdrückung des Rauschens die verbleibenden Blöcke in Reihenfolge fallender Signalleistung nacheinander ausgewählt und codiert, bis alle verfügbaren Bits erschöpft werden. Im Vergleich zu dem Verfahren in [20] ist das in dieser Arbeit entwickelte Verfahren wesentlich vollständiger und leistungsfähiger. Erstens, werden hier nicht nur die zu codierenden Blöcke innerhalb eines Bildes sondern auch die Bildfrequenz und Bandbreite des zu übertragenden Signals ausgewählt. Zweitens, wird bei der Auswahl der zu codierenden Blöcke nicht nur die lokale Signalleistung sondern auch die subjektiven Bedeutungen unterschiedlicher Prädiktionsfehler berücksichtigt.

3.2 Klassifizierung der Prädiktionsfehler

Die Diskussion im letzten Abschnitt zeigt, daß bei sehr niedrigen Übertragungsbitraten die Teilbandcodierung auf einen ausgewählten Teil der Prädiktionsfehler eingeschränkt werden soll, um eine bessere subjektive Bildqualität zu erreichen. In diesem Abschnitt wird weiter erklärt, welche Prädiktionsfehler dabei bevorzugt ausgewählt werden. Dazu werden die subjektiven Bedeutungen verschiedener Prädiktionsfehler analysiert.

In Bild 3.2 ist ein typisches Prädiktionsfehlerbild aus der Testsequenz ‚Miss Amerika‘ dargestellt. Anhand ihrer Form und Stärke lassen sich die Prädiktionsfehler grob in die folgenden vier Klassen aufteilen:

1. Stochastisches Rauschen, das dem ganzen Prädiktionsbild überlagert ist. Im unbewegten Hintergrundbereich ist das Rauschen der dominierende Signalanteil der Prädiktionsfehler.
2. Linienartige Prädiktionsfehler, die hauptsächlich in der Nähe von Luminanzsprüngen in bzw. um bewegte Objekte auftreten. Diese Prädiktionsfehler werden meistens durch eine ungenaue Bewegungskompensation hervorgerufen.



Bild 3.2 Typisches Prädiktionsfehlerbild aus der Testsequenz 'Miss Amerika'

3. Prädiktionsfehler mit kleinen Amplituden und/oder Ausdehnungen, die durch kleine Änderungen im Bild herbeigeführt werden.
4. Restliche Prädiktionsfehler mit großen Amplituden oder Ausdehnungen, die durch wesentliche Bildinhaltsänderungen verursacht werden. In typischen Bildtelefonsequenzen können diese Änderungen meistens auf Augen- und Lippenbewegungen zurückgeführt werden.

Diese vier Klassen von Prädiktionsfehler haben unterschiedliche subjektive Bedeutungen. Das Rauschen ist unabhängig vom Bildinhalt. Die Übertragung von Rauschen ist nicht sinnvoll. Die linienartigen Prädiktionsfehler enthalten die Information über die genauen Positionen bewegter Objekte. Aber eine Verschiebung dieser Objekte um wenige Bildpunkte fällt normalerweise nicht auf und ist für die meisten Empfänger unwesentlich. Deshalb kann auf eine Übertragung solcher Prädiktionsfehler verzichtet werden. In der dritten Klasse, Prädiktionsfehler mit kleinen Amplituden und/oder Ausdehnungen, verbirgt sich bereits Information über Bildinhaltsänderungen, die für manche Empfänger interessant sein können. Aber bei sehr niedrigen Übertragungsbitraten, wo nur ein kleiner Teil aller Prädiktionsfehler sich übertragen läßt, kann diese Klasse im Vergleich zu der nächsten Klasse jedoch eine niedrigere Priorität bei der Übertragung erhalten. Die vierte Klasse der Prädiktionsfehler ist die wichtigste. Nur durch Übertra-

gung dieser Prädiktionsfehler können die wesentlichen Änderungen des Bildinhaltes den Empfängern ermittelt werden.

Aus dieser Analyse ergibt sich die folgende Vorgehensweise für die Auswahl der zu codierenden Prädiktionsfehler: Die Prädiktionsfehler der vierten Klasse sollen mit der höchsten Priorität ausgewählt und übertragen werden. Danach kann, abhängig von der Anzahl der verbleibenden Bits, die Übertragung der Prädiktionsfehler der dritten Klasse und anschließend der zweiten Klasse folgen.

Im nächsten Abschnitt werden die Algorithmen zur Vorwärtsanalyse im Detail beschrieben.

3.3 Algorithmen zur Vorwärtsanalyse

In Bild 3.3 wird der Ablauf der Algorithmen zur Vorwärtsanalyse schematisch dargestellt.

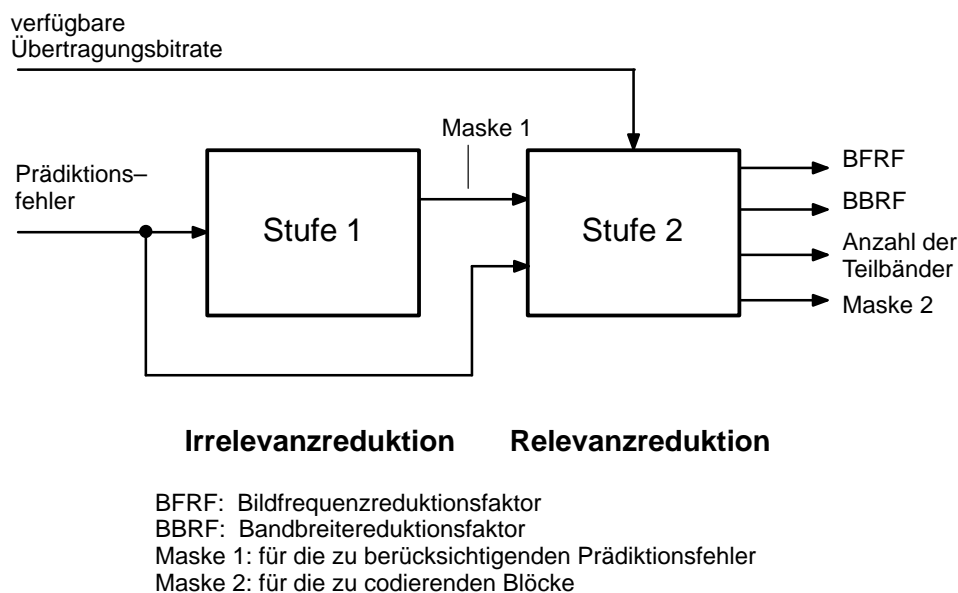


Bild 3.3 Ablauf der Algorithmen zur Vorwärtsanalyse

Als Eingangssignale der Vorwärtsanalyse dienen das aktuelle Prädiktionsfehlerbild und die verfügbare Übertragungsbitrate. Aus diesen Eingangssignalen werden die Maske für die zu codierenden Blöcke, der Bildfrequenzreduktionsfaktor (*BFRF*), der Bandbreitereduktionsfaktor (*BBRF*) und die Anzahl der Teilbänder als Ausgangssignale ermittelt. Die Maske der zu codierenden Blöcke gibt an, in welchen Blöcken die Prädik-

tionsfehler codiert und übertragen werden sollen. Der Bildfrequenz- und der Bandbreitereduktionsfaktor legen fest, in wie weit die Bildfrequenz bzw. Bandbreite der Eingangsbildsequenz reduziert werden sollen. $BFRF=1/3$ bedeutet z.B., daß nur jedes dritte Bild der Eingangsbildsequenz codiert werden soll. $BBRF=1/2$ legt fest, daß nur die Hälfte aller Teilbänder übertragen werden soll.

Die Algorithmen der Vorwärtsanalyse lassen sich in zwei Stufen unterteilen. In der ersten Stufe werden innerhalb des Bildes die Gebiete, die vorwiegend nur relevante Prädiktionsfehler (entsprechend Klasse 4) enthalten, ausgesucht. Dadurch werden die irrelevanten, oder unbedeutenden Prädiktionsfehler (entsprechend Klasse 1 bis 3) von der weiteren Teilbandcodierung ausgeschlossen. Als Ausgangssignal liefert diese Stufe eine Maske (Maske 1) für die in der nächsten Stufe weiter zu berücksichtigenden Prädiktionsfehler. Bei sehr niedrigen Übertragungsbitraten ist die Größe dieser Maske jedoch oft nicht klein genug, um die zugehörigen Prädiktionsfehler mit einer ausreichenden Genauigkeit codieren zu können. Deshalb wird in der zweiten Stufe, der aktuell verfügbaren Datenrate entsprechend, eine weitere Auswahl der zu codierenden Prädiktionsfehler nach ihrer lokalen Leistung blockweise durchgeführt, in der auch relevante Prädiktionsfehler von der Codierung ausgeschlossen werden. Gleichzeitig werden in dieser Stufe auch die zu verwendende Bildfrequenz und Bandbreite des zu übertragenden Signals sowie die Anzahl der Teilbänder festgelegt. Bei der Suche nach einer passenden Auswahl dieser Parameter wird eine Liste mit vorgegebenen Wertekombinationen verwendet. Da in der ersten Stufe hauptsächlich die irrelevante Information entfernt wird, wird diese Stufe in der folgenden Beschreibung als "Stufe zur Irrelevanzreduktion" bezeichnet. Demgegenüber wird die zweite Stufe "Stufe zur Relevanzreduktion" genannt. Im folgenden werden die Algorithmen in beiden Stufen im Detail beschrieben.

3.3.1 Stufe 1: Irrelevanzreduktion

In Bild 3.4 werden die einzelnen Verarbeitungsschritte dieser Stufe zusammen mit Beispielen ihrer Ein- und Ausgangssignale dargestellt.

Die erst beiden Schritte dienen zum Ausschließen derjenigen Gebiete von der weiteren Verarbeitung, die nur das stochastische Rauschen oder Prädiktionsfehler mit einer geringen Leistung enthalten. Dabei wird zuerst die lokale Leistung der Prädiktionsfehler gemessen, in dem für jeden Bildpunkt der quadratische Mittelwert des Prädiktionsfehlersignals innerhalb einer Nachbarschaft von 3×3 Bildpunkten berechnet wird. Anschließend wird eine binäre Maske durch Vergleich der lokalen Prädiktionsfehlerleistung mit einer vorgegebenen Schwelle T_1 gebildet. Diese Schwelle wird so eingestellt,

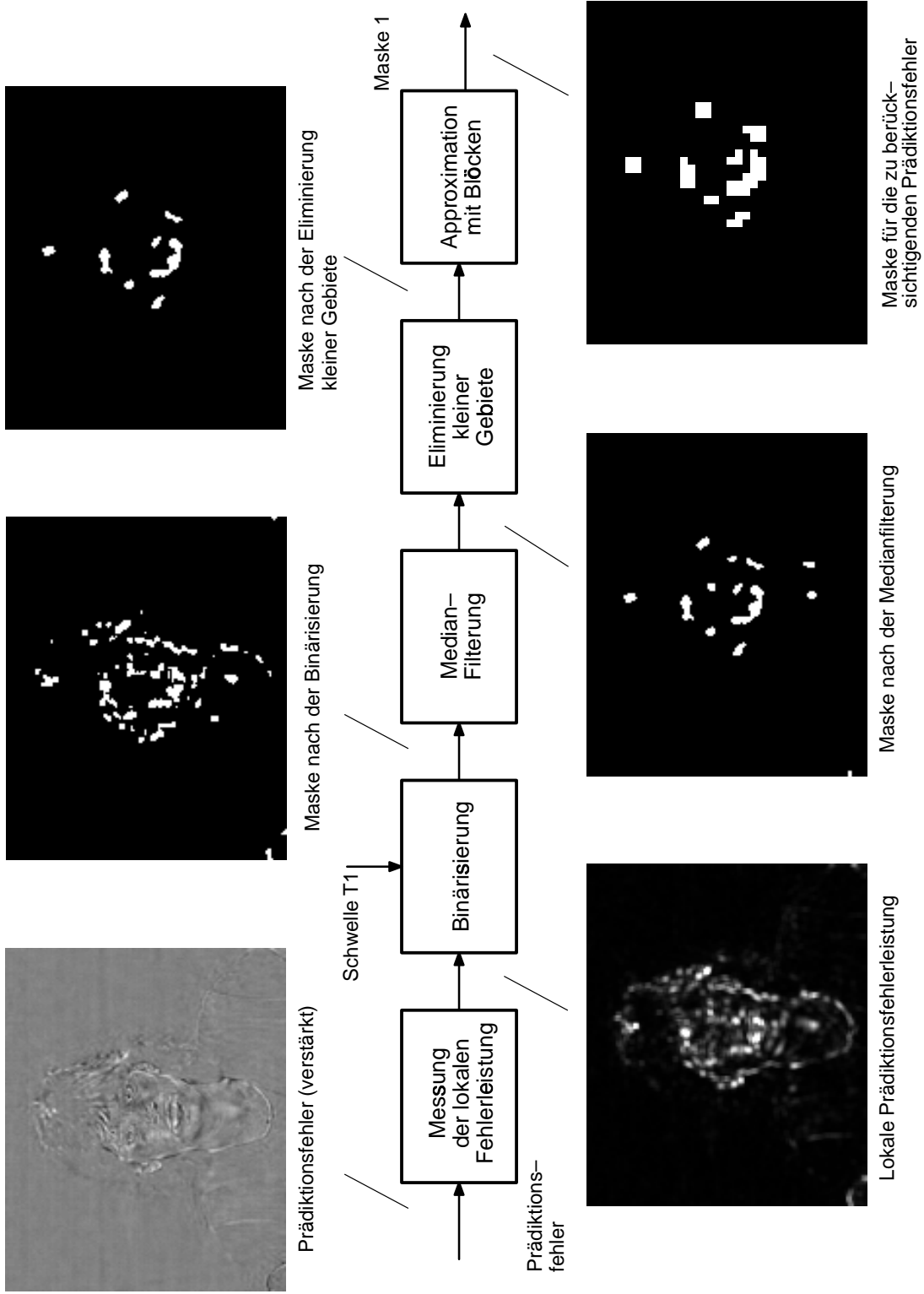


Bild 3.4 Aufbau der ersten Stufe der Vorwärtsanalyse

daß das Rauschen und die Prädiktionsfehler mit einer geringen Leistung unter die Schwelle fallen.

In den nächsten Schritten werden durch Medianfilterung und Eliminierung kleiner Gebiete die schmalen, linienförmigen Gebiete sowie die isolierten Gebiete mit einer kleinen Ausdehnung (Klassen 2 und 3) von der binären Maske entfernt. Dafür wird ein Medianfilter mit einer Fenstergröße von 5×5 Bildpunkte verwendet. Dadurch werden alle Linien entfernt, deren Breite kleiner als zwei Bildpunkte ist. Bei der Eliminierung kleiner Gebiete werden alle isolierten Gebiete mit einer Größe kleiner als 4×4 Bildpunkte entfernt.

Schließlich wird die beliebig geformte binäre Maske aus dem vorherigen Schritt mittels quadratischer Blöcke approximiert. Dabei wird die alte Maske in ein Raster aus Blöcken mit 4×4 Bildpunkten aufgeteilt. Die neue Maske enthält diejenigen Blöcke, in denen mindesten 50% der Bildpunkte zur alten Maske gehören. Der Grund dieser blockweisen Auswahl liegt darin, daß bei der Teilbandcodierung, wie in 3.1 erklärt, immer Bildbereiche aus rechteckigen Blöcken codiert werden. Da in diesem Coder die kleinste Anzahl der Teilbänder 16 beträgt, werden Blöcke mit 4×4 Bildpunkten bei der Approximation der Maske verwendet. Diese Maske aus quadratischen Blöcken stellt das Ergebnis dieser Stufe dar und wird zur weiteren Verarbeitung an die zweite Stufe übergeben.

3.3.2 Stufe 2: Relevanzreduktion

Der Aufbau der zweiten Stufe ist in Bild 3.5 dargestellt.

(i) Weitere Auswahl der zu codierenden Blöcke

Im ersten Schritt der zweiten Stufe werden die Prädiktionsfehler optional tiefpaßgefiltert, falls die Bandbreite der zu übertragenden Prädiktionsfehler voraussichtlich auf ein Viertel reduziert werden muß. Zur Schätzung dieser Bandbreite wurde der entsprechende Wert, der im vorangegangenen übertragenen Bild benutzt wird, verwendet. Diese Tiefpaßfilterung wird separat in horizontaler und vertikaler Richtung durchgeführt und unterdrückt dabei je Richtung die Hälfte des Spektrums der Prädiktionsfehler. Der Grund für diese Tiefpaßfilterung liegt darin, daß die hochfrequenten spektralen Anteile voraussichtlich nicht übertragen werden können und deshalb bei der weiteren Analyse nicht mehr berücksichtigt werden dürfen. Nach der Tiefpaßfilterung wird eine weitere Auswahl der zu codierenden Prädiktionsfehler blockweise nach der Fehlerleistung durchgeführt, wobei Blöcke mit einer höheren Fehlerleistung vorrangig ausgewählt werden. Dazu wird zuerst der quadratische Mittelwert der Prädiktionsfehler in jedem 4×4 Block berechnet. Dann wird diese Fehlerleistung optional örtlich gewichtet, um Prädiktionsfehler in manchen Gebieten bei der Codierung bevorzugt behandeln zu

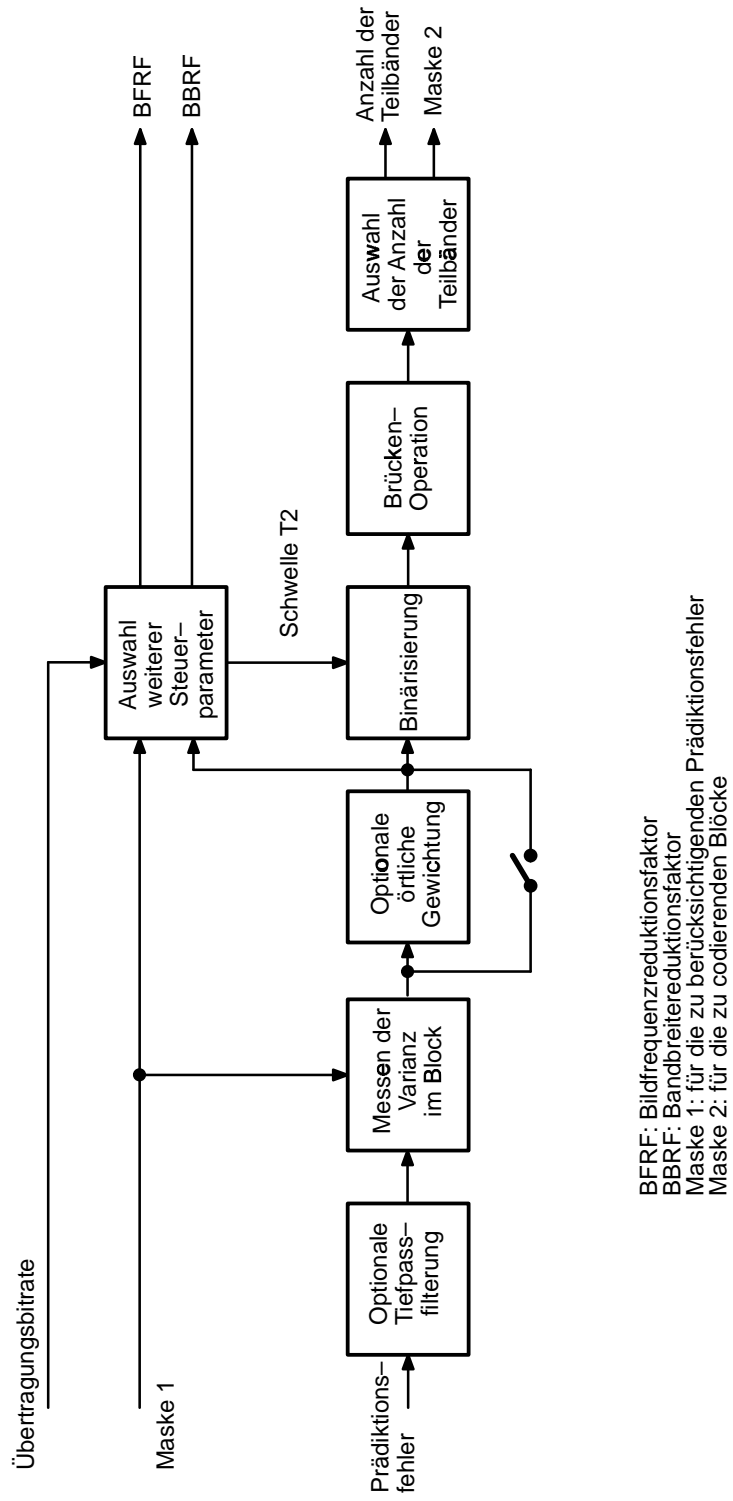


Bild 3.5 Aufbau der zweiten Stufe der Vorwärtsanalyse

können. Z.B. können Prädiktionsfehler in der Bildmitte mehr Gewicht als die am Rand bekommen. Anschließend wird eine neue Maske gebildet, indem die gewichtete Fehlerleistung in einzelnen Blöcken mit einer Schwelle T_2 verglichen wird. Die verwendete Schwelle wird vorher in einem Funktionsblock mit dem Name "Auswahl weiterer Steuerparameter" bestimmt. Die Funktionsweise dieses Blocks wird weiter unten genauer beschrieben. Um die Maske effizient codieren zu können, wird sie mit einer sogenannten 'Brücken'-Operation nachverarbeitet. Dabei werden einerseits einzelne, isolierte Blöcke aus der Maske entfernt, andererseits werden Blöcke zugefügt, um kleine Lücken innerhalb der Maske zu füllen.

(ii) Auswahl der Anzahl der Teilbänder

Die Auswahl der Anzahl der Teilbänder basiert auf einem Vergleich der Datenmengen, die von der Teilbandcodierung mit jeweils 16 und 64 Teilbändern zu erwarten sind. Diese Datenmenge ergibt sich aus der Codierung der Maske für die zu codierenden Blöcke sowie der Teilbandcodierung der Prädiktionsfehler in den Blöcken.

Zuerst wird die Differenz der Teildatenmengen aus der Codierung der jeweiligen Masken für 16 und 64 Teilbändern ermittelt. Für die Teilbandcodierung mit 16 Teilbändern kann die Maske nach der Brücken-Operation direkt verwendet werden, da diese Maske bereits aus Blöcken mit 4x4 Bildpunkten besteht. Für die Teilbandcodierung mit 64 Teilbändern muß eine neue Maske erzeugt werden. Diese neue Maske enthält alle solche Blöcke mit 8x8 Bildpunkten, in denen mindestens einer der 4 zugehörigen Blöcke mit 4x4 Bildpunkten bereits in der ursprünglichen Maske vorhanden ist. Wie in Kapitel 3.4 noch genau beschrieben wird, basiert die Codierung der Maske aus Blöcken mit 4x4 Bildpunkten auf der Codierung der Maske aus Blöcken mit 8x8 Bildpunkten. Dabei werden für jeden ausgewählten Block mit 8x8 Bildpunkten 4 zusätzliche Bits verwendet. Sie zeigen, welche der 4 zugehörigen Blöcke mit 4x4 Bildpunkten codiert werden. Dadurch wird insgesamt eine zusätzliche Datenmenge von

$$\Delta D_{mask} = 4 \cdot (N_{cod(64)}/64) = N_{cod(64)}/16 \quad (3.1)$$

Bits pro Bild für die Codierung der Maske aus Blöcken mit 4x4 Bildpunkten benötigt, wobei $N_{cod(64)}$ die Anzahl aller Bildpunkte in den zu codierenden Blöcken mit 8x8 Bildpunkten ist. $N_{cod(64)}/64$ ist dann die Anzahl der zu codierenden Blöcke mit 8x8 Bildpunkten.

Anschließend wird die Differenz der Teildatenmengen aus der Teilbandcodierung der Prädiktionsfehler in den jeweiligen Blöcken anhand

$$\Delta D_{SBC} = R_{(16)} \cdot N_{cod(16)} - R_{(64)} \cdot N_{cod(64)} \quad (3.2)$$

geschätzt. $R_{(16)}$ und $R_{(64)}$ sind die Schätzungen der bei der Teilbandcodierung mit 16 bzw. 64 Teilbändern benötigten Datenraten in bit/Bildpunkt. $N_{cod(16)}$ ist die Anzahl aller Bildpunkte in den zu codierenden Blöcken mit 4x4 Bildpunkten. Je nachdem, ob 16 oder 64 Teilbänder für das letzte codierte Bild benutzt wurden, erhält $R_{(16)}$ oder $R_{(64)}$ die sich im letzten Bild ergebene Datenrate der Teilbandcodierung. Der andere Wert wird aus der Beziehung $R_{(64)}/R_{(16)} = g$ berechnet, wobei $g=0.8$ experimentell eingestellt wurde.

Schließlich wird die Anzahl der Teilbänder für das aktuell zu codierende Bild auf 16 oder 64 festgelegt, je nachdem ob die Summe $\Delta D_{mask} + \Delta D_{SBC}$ negativ oder positiv ist. Entsprechend dieser Festlegung wird die passende Maske für die zu codierenden Blöcke der anschließenden blockbasierten Teilbandcodierung übergeben.

(iii) Auswahl weiterer Steuerparameter

Die Aufgabe dieses Funktionsblocks ist die Auswahl des Bildfrequenz- und Bandbreitereduktionsfaktors ($BFRF$ und $BBRF$) sowie der oben genannten Schwelle T_2 . Durch Veränderung der Schwelle T_2 wird die Anzahl der zu codierenden Blöcke variiert. Dementsprechend verändern sich auch die zu erwartenden Signal-Rausch-Abstände in diesen und den restlichen Blöcken mit den relevanten Prädiktionsfehlern. Durch Reduzierung der aktuellen Bildfrequenz läßt sich die verfügbare Datenmenge für die Codierung des aktuellen Bildes, demzufolge auch der Signal-Rausch-Abstand, bei der gleich bleibenden Übertragungsbitrate erhöhen. Durch Reduzierung der Bandbreite des zu übertragenden Prädiktionsfehlersignals sind mehr Bits pro Abtastwert für die Codierung der verbleibenden Teilbänder verfügbar.

Für die Auswahl dieser Parameter steht in diesem Coder eine Liste mit einstellbaren Wertekombinationen für $BFRF$, $BBRF$ und den zugehörigen Mindestwert des Signal-Rausch-Abstandes $PSNR_{min}$ zur Verfügung. Tabelle 3.2 zeigt die in diesem Coder verwendete Liste. Jede Zeile der Liste enthält eine dieser Wertekombinationen, die jeweils den angestrebten Kompromiß zwischen $BFRF$ und $BBRF$ darstellen. Die Zeilen der Liste sind in der Reihenfolge fallender Bildfrequenz und Bandbreite geordnet. Aus dieser Liste wird für jedes Bild die 1. passende Wertekombination von oben für $BFRF$ und $BBRF$ ausgewählt, bei welcher der zugehörige $PSNR_{min}$ erreicht wird. Dazu wird nacheinander, Zeile für Zeile, überprüft, ob dieses Kriterium erfüllt wird. Mit dieser Methode wird eine optimale Einstellung der Bildfrequenz und Bandbreite sowie des Signal-Rausch-Abstandes erzielt. Durch Anwendung von unterschiedlichen Listen

läßt sich diese Einstellung an den jeweiligen Bedarf verschiedener Anwendungen anpassen.

Tabelle 3.2 Liste mit den vorgegebenen Wertekombinationen für die auszuwählenden Steuerparameter

Schritt	Bildfrequenzreduktionsfaktor (<i>BFRF</i>)	Bandbreite- reduktions- faktor (<i>BBRF</i>)	$PSNR_{min}$ [dB]
1	1/1	1/1	34
2	1/1	1/2	33
3	1/2	1/2	33
4	1/2	1/4	32
5	1/3	1/4	32

In Bild 3.6 wird der Algorithmus zur Auswahl der Steuerparameter *BFRF*, *BBRF* und T_2 schematisch dargestellt. Zuerst wird die Verteilung der Prädiktionsfehlervarianz der einzelnen Blöcke mit den relevanten Prädiktionsfehlern in Form eines Histogramms ermittelt. Danach wird die erste Wertekombination aus der Liste eingelesen. Basierend auf den Werten von *BFRF* und *BBRF*, dem Histogramm und der Übertragungsbitrate in bit/s wird die Schwelle T_2 iterativ so eingestellt, daß von den zu codierenden und den nicht zu codierenden Blöcken ähnliche PSNR-Werte zu erwarten sind. Anschließend werden die geschätzten PSNR-Werte mit dem vorgegebenen Mindestwert $PSNR_{min}$ verglichen. Wird $PSNR_{min}$ erreicht, dann ist die passende Wertekombination gefunden. Anderenfalls wird dieser Vorgang mit der nächsten Wertekombination aus der Liste wiederholt. Der Auswahlvorgang wird so weit fortgesetzt, bis die passende Wertekombination gefunden, oder das Ende der Liste erreicht wird. In beiden Fällen wird die zuletzt geprüfte Wertekombination zusammen mit der zugehörigen Schwelle T_2 ausgewählt. Im folgenden wird gezeigt, wie die zu erwartenden PSNR-Werte geschätzt werden.

Durch die Schwelle T_2 werden die Blöcke mit den relevanten Prädiktionsfehlern in zwei Teile aufgeteilt. Zum ersten Teil (cod) gehören alle zu codierenden Blöcke, deren Prädiktionsfehlervarianz größer als die Schwelle T_2 ist. Der zweite Teil (noc) enthält die restlichen nicht zu codierenden Blöcke. Zur Schätzung der zu erwartenden PSNR-Werte in den beiden Teilen werden zuerst die Prädiktionsfehlervarianzen $\sigma_{pe,cod}^2$ und $\sigma_{pe,noc}^2$

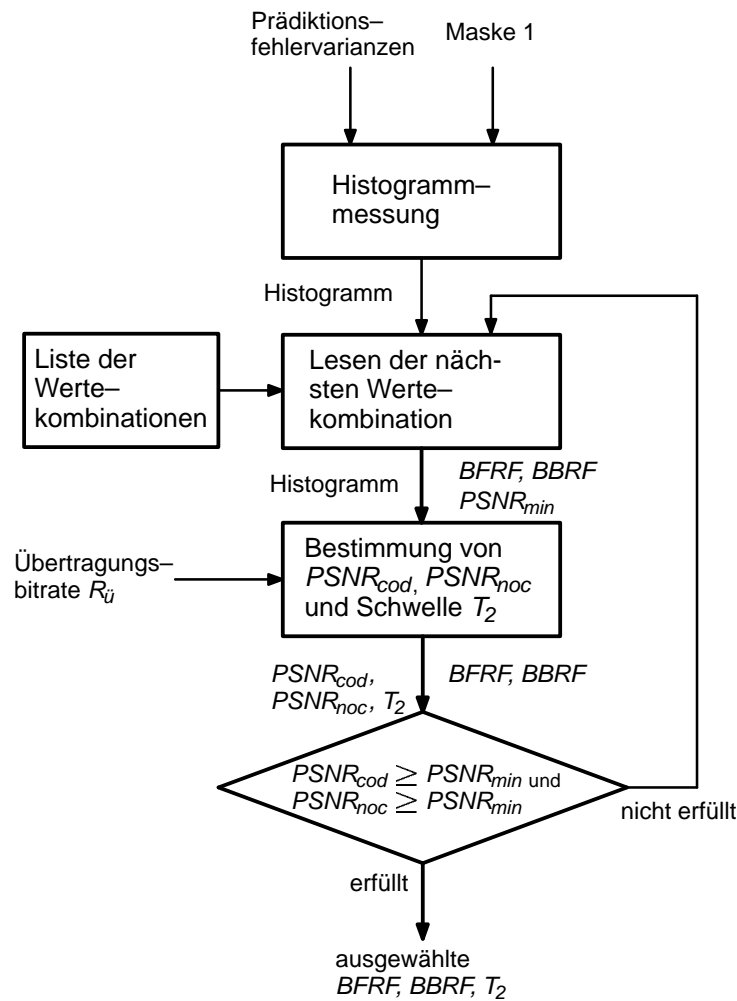


Bild 3.6 Algorithmen zur Auswahl der weiteren Steuerparameter

sowie die Anzahlen der Bildpunkte N_{cod} und N_{noc} in den entsprechenden Teilen anhand des Histogramms ermittelt.

Für die zu codierenden Blöcke wird die zu erwartende Rekonstruktionsfehlervarianz $\sigma_{re,cod}^2$ anhand

$$\sigma_{re,cod}^2 = \frac{1}{G_{SBC}} 2^{-2R} \sigma_{pe,cod}^2 \quad (3.3)$$

geschätzt, wobei G_{SBC} der Teilbandcodierungsgewinn [26] ist. Die aktuelle Datenrate der Teilbandcodierung R in bit/Bildpunkt läßt sich aus der gesamten Übertragungsbitrate R_{ii} in bit/s, der Eingangsbildfrequenz f_{Bild} in Bild/s, dem aktuellen $BFRF$, der Anzahl der zu codierenden Bildpunkte N_{cod} sowie dem Verhältnis p_{SBC} der Datenrate der Teilbandcodierung zur gesamten Datenrate des Codes gemäß

$$R = \frac{p_{SBC} \cdot R_{ii}}{f_{Bild} \cdot BFRF \cdot N_{cod}} \quad (3.4)$$

berechnen. Dabei werden für G_{SBC} und p_{SBC} die entsprechenden Messungen aus dem letzten codierten Bild verwendet. Der Einfluß von $BFRF$ geht in G_{SBC} ein.

In den nicht zu codierenden Blöcken ist die Rekonstruktionsfehlervarianz gleich der Prädiktionsfehlervarianz:

$$\sigma_{re,noc}^2 = \sigma_{pe,noc}^2 \quad (3.5)$$

Aus $\sigma_{re,cod}^2$ und $\sigma_{re,noc}^2$ berechnen sich die zu schätzenden PSNR–Werte in den beiden Teilen zu

$$PSNR_{cod} = 10 \log_{10} \frac{255^2}{\sigma_{re,cod}^2} \quad (3.6)$$

und

$$PSNR_{noc} = 10 \log_{10} \frac{255^2}{\sigma_{re,noc}^2} \quad (3.7)$$

3.4 Codierung der Ausgangssignale der Vorwärtsanalyse

Der Bildfrequenzreduktionsfaktor $BFRF$ wird anhand der zeitlichen Referenz des nächsten zu codierenden Bildes indirekt codiert. Dafür steht ein Codewort TR mit 8 bits im Bild–Header zur Verfügung.

Für den Bandbreitereduktionsfaktor $BBRF$ wird keine explizite Codierung benötigt, da das letzte codierte Teilband durch das Codewort EOB (End–Of–Block) bei der Lauflängencodierung der Teilbandkoeffizienten (siehe Kapitel 4.4) angegeben wird.

Zur Kennzeichnung der Anzahl der Teilbänder (16 oder 64), die für das aktuelle codierte Bild verwendet wird, ist ein Codewort NBAND (1 bit) im Bild–Header vorgesehen.

Die Information darüber, für welche Blöcke die Prädiktionsfehler codiert werden, wird, wie in H.263, Macroblock–Weise übermittelt. Für jeden Macroblock mit 16×16 Bildpunkten gibt zunächst das Codewort COD (1 bit) an, ob der Macroblock codiert wird. Ein nicht codierter Macroblock enthält weder den Displacementvektor noch die Teilbandkoeffizienten. Bei einem codierten Macroblock wird anhand zwei Codewörter CBPY und MCBPC festgelegt, welche zugehörigen Blöcke mit 8×8 Luminanz– bzw. Chrominanz–Abtastwerten übertragen werden. Falls für das aktuelle Bild 16 Teilbänder (entsprechend Blöcke mit 4×4 Bildpunkten) verwendet werden, zeigt für jeden codierten Block mit 8×8 Bildpunkten ein weiteres Codewort CSBP mit 4 bits, welche der 4 zugehörigen Blöcke mit 4×4 Bildpunkten übertragen werden.

In diesem Kapitel wurden die Ansätze und die Realisierung der Vorwärtsanalyse ausführlich beschrieben. Der durch die Vorwärtsanalyse erzielte Gewinn wird in Kapitel 6 anhand von Computersimulationen ermittelt und dargestellt.

4. Teilbandcodierung der ausgewählten Blöcke

Die durch die Vorwärtsanalyse ausgewählten Blöcke werden in diesem Coder mit einer Teilbandcodierung codiert und übertragen. In diesem Kapitel werden einige relevante Themen der Teilbandcodierung ausführlich behandelt. Im Abschnitt 4.1 wird das Prinzip einer blockbasierten Teilbandcodierung vorgestellt. Der Hauptunterschied zu einer konventionellen Teilbandcodierung liegt darin, daß die Teilbandzerlegung hier statt auf das Gesamtbild nur auf die ausgewählten Blöcke angewendet wird. Zur Vermeidung der Übertragung zusätzlicher Teilbandabstastwerte wird eine als symmetrische periodische Erweiterung bezeichnete Technik eingesetzt. Im Abschnitt 4.2 wird auf die Auswahl der Teilbandfilterbank eingegangen. Anhand der mittleren Entropien der Teilbandsignale, die sich unter einer subjektiv optimalen adaptiven Quantisierung ergeben, werden die unterschiedlichen Teilbandaufteilungen sowie die verschiedenen Teilbandfilter verglichen. Im Abschnitt 4.3 wird die Umschaltung zwischen einer Intraframe- und einer Interframe-Codierung der Teilbandsignale diskutiert. Schließlich wird in Abschnitt 4.4 die Codierung der quantisierten Teilbandsignale kurz beschrieben.

4.1 Blockbasierte Teilbandcodierung

In diesem Abschnitt werden zuerst die Probleme, die sich bei der konventionellen Teilbandcodierung mit sehr niedrigen Datenraten ergeben, analysiert. Anschließend wird das Prinzip der blockbasierten Teilbandcodierung vorgestellt.

4.1.1 Probleme bei der konventionellen Teilbandcodierung mit sehr niedrigen Datenraten

Bei der konventionellen bewegungskompensierenden Teilbandcodierung wird die Teilbandzerlegung auf das gesamte Prädiktionsfehlerbild angewendet. Diese Vorgehensweise wird in dieser Arbeit als bildbasierte Teilbandcodierung bezeichnet. Bei hohen Übertragungsdatenraten können die Prädiktionsfehler in der Regel für alle Bildbereiche übertragen werden. Deshalb ist die Anwendung der bildbasierten Teilbandcodierung in diesem Fall zweckmäßig. Bei sehr niedrigen Datenraten, bei denen die Prädiktionsfehler nur für sehr wenige Blöcke codiert und übertragen werden können, ergibt die bildbasierte Teilbandcodierung aber folgende zwei Probleme.

Das erste Problem ist, daß die Anzahl der insgesamt zu übertragenden Teilbandabstastwerte gegenüber der Anzahl der Bildpunkte in den zu codierenden Blöcken sich erheblich erhöht. Die Ursache ist, daß durch die Teilbandanalysefilterung auch außerhalb der zu codierenden Blöcke Teilbandabstastwerte ungleich Null erzeugt werden, da die lineare Faltung eines Eingangssignals der Länge N mit einem Filter der Länge L ein Aus-

gangssignal der Länge $N+L-1$ ergibt. Bei der Teilbandsynthese werden diese zusätzlichen Teilbandabtastwerte zur Rekonstruktion der zu codierenden Blöcke benötigt. Deswegen müssen sie zum Decoder übertragen werden. Andernfalls entstehen systematische Fehler. In Bild 4.1 wird dieses Problem anhand eines Beispiels veranschaulicht. Zur Rekonstruktion der mit schwarz markierten Blöcke im Originalbild müssen in den Teilbändern nicht nur für die schwarzen sondern auch für die grauen Bereiche Abtastwerte übertragen werden. Um einen quantitativen Eindruck zu vermitteln, wurden aus den Testsequenzen "Miss America" und "Claire" die Anzahl der Bildpunkte in den zu codierenden Blöcken und die Anzahl der insgesamt zu übertragenden Teilbandabtastwerte berechnet und in Tabelle 4.1 dargestellt.

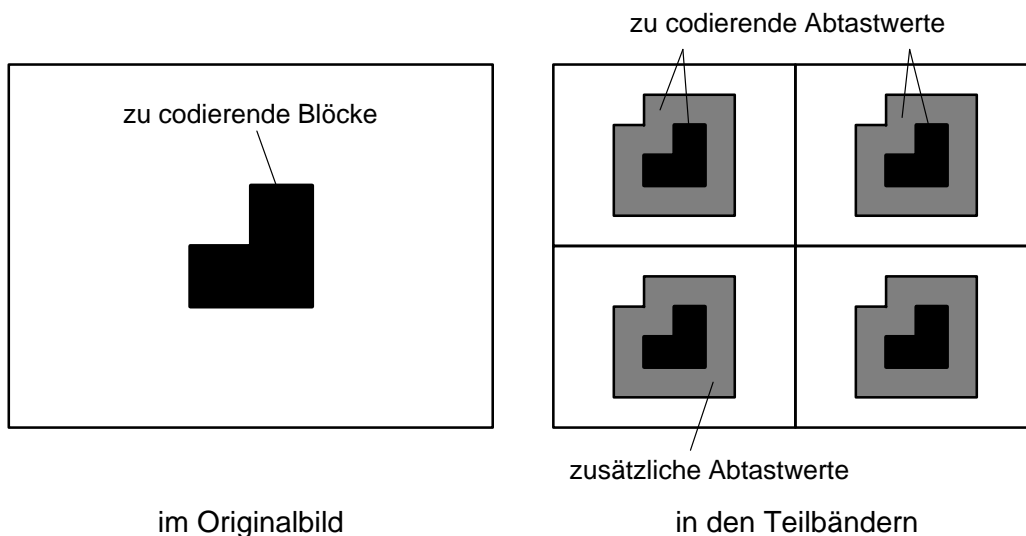


Bild 4.1 Beispiel zur Veranschaulichung des ersten Problems bei der bildbasierten Teilbandcodierung mit sehr niedrigen Datenraten

Das zweite Problem der bildbasierten Teilbandcodierung ist die örtliche Verbreiterung der Teilbandquantisierungsfehler durch die Teilbandsynthesefilterung. Wie Bild 4.2 zeigt, wird der Quantisierungsfehler eines Teilbandabtastwertes durch die Synthesefilterung zu ausgedehnten Rekonstruktionsfehlern in Form der Impulsantwortfunktion des Synthesefilters umgewandelt. Diese Rekonstruktionsfehler treten nicht nur in dem Block, zu dem der ursprüngliche Teilbandabtastwert örtlich gehört, sondern auch in den benachbarten Blöcken auf. In Rekonstruktionsbildern sind diese Fehler in Form von Ringing in der Nähe von Luminanzsprüngen besonders sichtbar.

Tabelle 4.1 Beispiele für die Anzahl der Bildpunkte in den zu codierenden Blöcken und der Anzahl der insgesamt zu übertragenden Teilbandabtastwerte aus der bildbasierten Teilbandcodierung

Testsequenz	Anzahl der zu codierenden Bildpunkte pro Bild	Anzahl der zu übertragenden Teilbandabtastwerte
Miss America	1014	6665
Claire	1039	5829

Zur Lösung dieser beiden Probleme wird die Teilbandzerlegung in diesem Coder nur auf die ausgewählten Blöcke angewendet. Diese Technik wird in dieser Arbeit als blockbasierte Teilbandcodierung bezeichnet. Das Wort "blockbasiert" betont den Unterschied zu "bildbasiert", bedeutet aber nicht, daß jeder Block einzeln in Teilbänder zerlegt wird. Durch die blockbasierte Teilbandcodierung wird die Verbreiterung der Quantisierungsfehler in die benachbarten Blöcke vermieden. In Verbindung mit einer symmetrischen periodischen Erweiterung der zu filternden Signale, die im nächsten Abschnitt beschrieben wird, vermeidet diese blockbasierte Teilbandcodierung ebenfalls die Übertragung der zusätzlichen Teilbandabtastwerte.

4.1.2 Teilbandzerlegung von Blockbündeln

Bei der blockbasierten Teilbandcodierung werden die Zeilen und Spalten von Bündeln der ausgewählten Blöcke separat durch kaskadierte eindimensionale Filterungen und

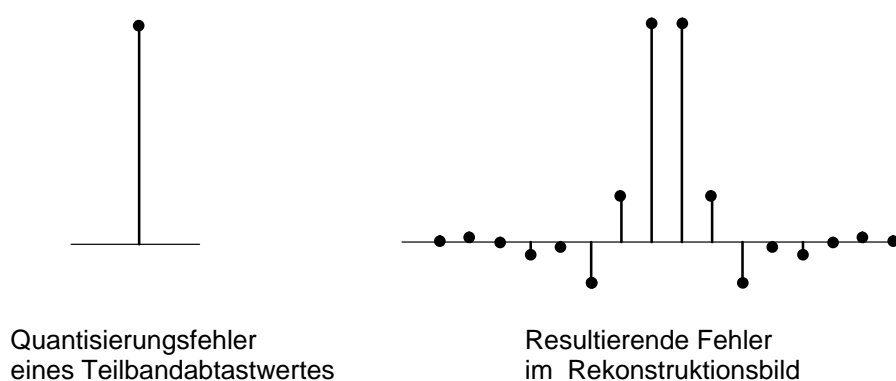


Bild 4.2 Beispiel der örtlichen Verbreiterung eines einzelnen Quantisierungsfehlers aus einem Teilband durch die Synthesefilterung

Unterabtastung in Teilbänder zerlegt. Als ein Blockbündel wird eine Menge von zusammenhängenden Blöcken bezeichnet, in der jeder Block mindestens eine gemeinsame Blockkante mit den restlichen Blöcken teilt. Die Filterung von gemeinsamen Zeilen oder Spalten eines Blockbündels wird über die Grenzen der einzelnen benachbarten Blöcke hinaus kontinuierlich durchgeführt. Eine Besonderheit bei der Filterung ist, daß das zu filternde Signal kürzer als das Filter sein kann. Im folgenden wird untersucht, wie das zu filternde Signal erweitert werden soll, damit das gefilterte Signal genau so viele Abtastwerte wie das zu filternde Signal aufweist. Dazu werden zunächst die bekannten Techniken betrachtet.

Bekannte Techniken zur Erweiterung eines Signals

Aus der Literatur [50] sind mehrere Techniken zur Erweiterung eines zu filternden Signals begrenzter Länge bekannt. Die einfachste ist die in Bild 4.3 dargestellte *periodische Erweiterung*.

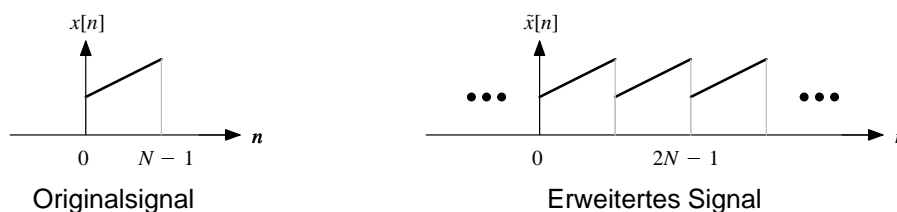


Bild 4.3 Periodische Erweiterung eines Signals mit einer begrenzten Länge

Dabei wird das zu filternde Signal periodisch wiederholt. Die Länge der Periode ist gleich der Länge des zu filternden Signals. Die Teilbandfilterung dieses erweiterten Signals ergibt wiederum periodische Signale mit der gleichen Periodenlänge. Nach der Reduktion der Abtastfrequenz ist die Anzahl der Abtastwerte aus allen Teilbändern gleich der Anzahl der Abtastwerte im Originalsignal. Diese periodische Erweiterung ist einfach zu realisieren, verursacht aber Diskontinuitäten an den Grenzen einzelner Perioden. An diesen Stellen werden der Anfang und das Ende des Originalsignals, die unterschiedliche Amplituden besitzen können, zusammengefügt. Wegen dieser Diskontinuitäten erhöht sich die lokale Signalleistung der hochfrequenten Anteile an den Rändern der Teilbänder. Demzufolge nimmt auch die zur Codierung der Teilbandsignale erforderliche Datenmenge zu.

Die andere Möglichkeit zur Erweiterung des zu filternden Signals ist die in Bild 4.4 veranschaulichte *symmetrische Erweiterung* [50].

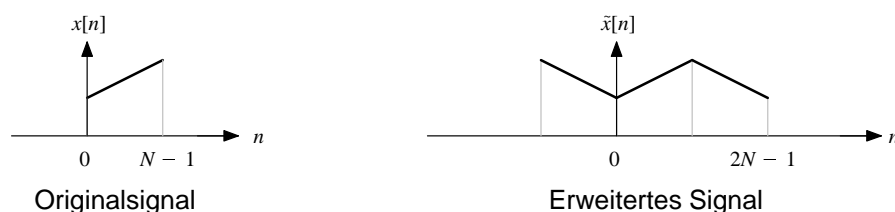


Bild 4.4 Symmetrische Erweiterung eines Signals mit einer begrenzten Länge

Diese Technik kann in Verbindung mit Filtern, die eine symmetrische Impulsantwortfunktion mit einer geraden Länge besitzen, für eine eindimensionale Teilbandzerlegung in zwei Teilbänder angewendet werden. Dabei wird das Eingangssignal durch Hinzufügen von zusätzlichen Abtastwerten erweitert. Diese hinzugefügten Abtastwerte sind jeweils um den Anfang und das Ende des Eingangssignals spiegel-symmetrisch zu den ursprünglichen Abtastwerten. Diese Symmetrie wird nach der Teilbandanalyse in den Teilbandsignalen beibehalten. Durch Ausnutzung dieser Symmetrie-Eigenschaft lassen sich die Teilbandsignale mit genau so vielen Abtastwerten wie im Eingangssignal beschreiben. Im Vergleich zur periodischen Erweiterung werden hier keine Diskontinuitäten in dem erweiterten Signal zugefügt. Allerdings muß das Eingangssignal eine gerade Länge haben, die mindestens halb so lang wie die Filterlänge ist.

Symmetrische Periodische Erweiterung

Um kurze Signale mit der blockbasierten Teilbandcodierung behandeln zu können, wird in diesem Coder eine symmetrische periodische Erweiterung (Bild 4.5) verwendet. Diese Technik ist eine Kombination der oben beschriebenen zwei Techniken, indem das Eingangssignal nach einer symmetrischen Spiegelung periodisch wiederholt wird. Eine Periode des erweiterten Signals ist doppelt so lang wie das ursprüngliche Signal. Von den beiden Halb-Perioden ist die erste gleich dem ursprünglichen Signal und die zweite spiegel-symmetrisch zu der ersten. Mathematisch läßt sich die symmetrische Erweiterung des Signals $x[n]$ mit der Länge N , wobei N eine gerade Zahl ist, gemäß

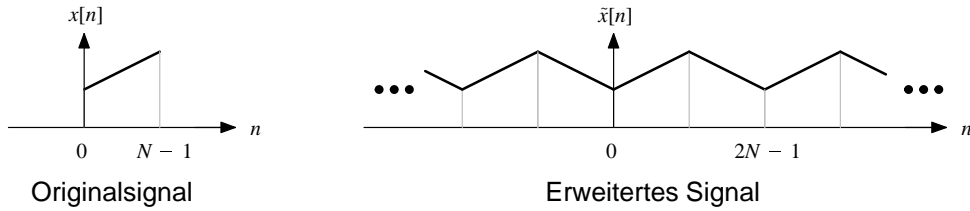


Bild 4.5 Symmetrische periodische Erweiterung eines Signals mit einer begrenzten Länge

$$\tilde{x}[n] = \begin{cases} x[n] & n = 0, 1, \dots, N-1 \\ x[2N-n-1] & n = N, N+1, \dots, 2N-1 \end{cases} \quad (4.1)$$

darstellen. Für die periodische Erweiterung gilt

$$\tilde{x}[n] = \tilde{x}[n + k \cdot 2N] \quad k = 1, 2, 3, \dots \quad (4.2)$$

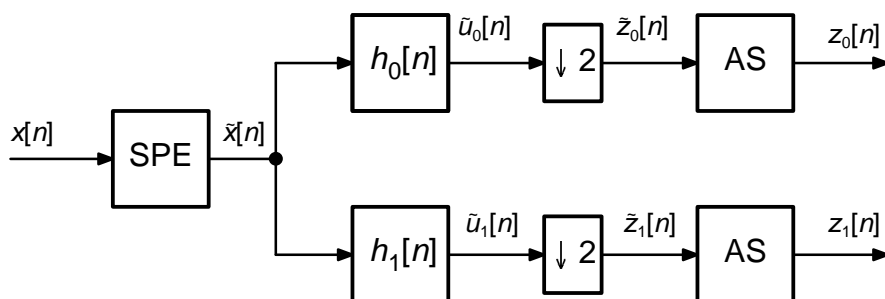
Im folgenden wird die Anwendung der symmetrischen periodischen Erweiterung bei der eindimensionalen Teilbandanalyse und –synthese mit zwei Teilbändern genau erläutert.

In Bild 4.6 wird an einem Beispiel die Anwendung der symmetrischen periodischen Erweiterung bei der Teilbandanalyse dargestellt. Das Eingangssignal $x[n]$ mit der geraden Länge N wird zuerst gemäß Gl. (4.1) zu einem symmetrischen periodischen Signal $\tilde{x}[n]$ erweitert. Das erweiterte Signal wird anschließend jeweils mit einem Tiefpaßfilter $h_0[n]$ und einem Hochpaßfilter $h_1[n]$ gefiltert. Wie in Bild 4.7 dargestellt, hat das Tiefpaßfilter eine symmetrische und das Hochpaßfilter eine antisymmetrische Impulsantwortfunktion mit einer geraden Länge L

$$h_0[n] = h_0[1-n], \quad n = -\frac{L}{2} + 1, \dots, -1, 0, 1, \dots, \frac{L}{2}, \quad (4.3)$$

$$h_1[n] = -h_1[1-n], \quad n = -\frac{L}{2} + 1, \dots, -1, 0, 1, \dots, \frac{L}{2}. \quad (4.4)$$

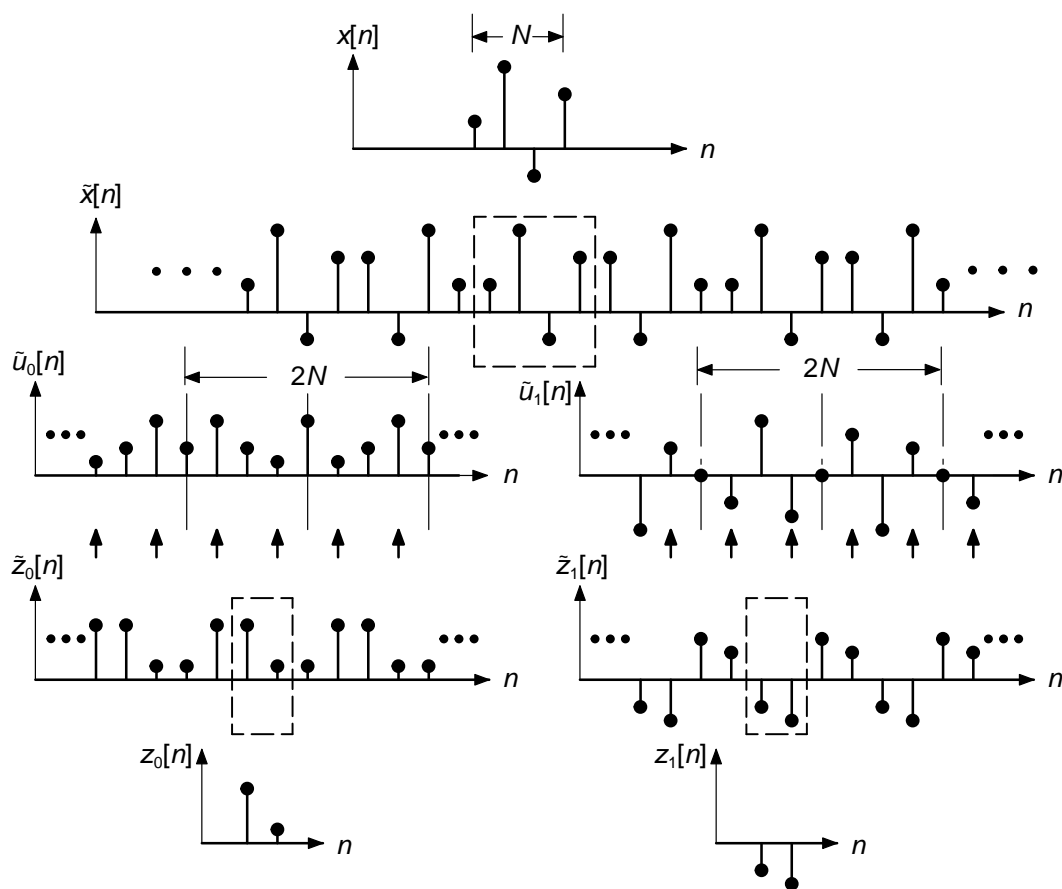
Da $\tilde{x}[n]$ ein periodisches Signal mit der Periodenlänge $2N$ ist, sind die gefilterten Signale $\tilde{u}_0[n]$ und $\tilde{u}_1[n]$ auch periodische Signale mit der gleichen Periodenlänge. Auf die gefilterten Signale $\tilde{u}_0[n]$ und $\tilde{u}_1[n]$ wird eine Reduktion der Abtastfrequenz um den



SPE: Symmetrische Periodische Erweiterung

AS: Abtastwert-Selektion

(a) Blockdiagramm



(b) Beispiel

Bild 4.6 Eindimensionale Teilbandanalyse mit zwei Teilbändern unter Anwendung der symmetrischen periodischen Erweiterung

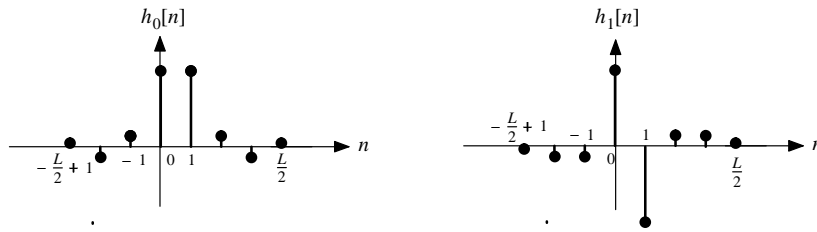


Bild 4.7 Impulsantwortfunktionen der Analysefilter in Bild 4.6

Faktor 2 angewendet, so daß die Abtastwerte mit ungeraden n erhalten bleiben. Nachfolgend wird gezeigt, daß für ungerade Werte von n die beiden Halbperioden von $\tilde{u}_0[n]$ symmetrisch sind. Die erste Halbperiode $\tilde{u}_0[1 + 2k]$ ($k = 0, 1, \dots, N/2 - 1$) läßt sich gemäß

$$\tilde{u}_0[1 + 2k] = \sum_{m=-\frac{L}{2}+1}^{\frac{L}{2}} \tilde{x}[1 + 2k - m] h_0[m] \quad (4.5)$$

darstellen. Unter Berücksichtigung der Symmetrieeigenschaft

$$\tilde{x}[n] = \tilde{x}[2N - 1 - n] \quad (4.6)$$

läßt sich die zweite Halbperiode $\tilde{u}_0[2N - (1 + 2k)]$ ($k = 0, 1, \dots, N/2 - 1$) wie folgt berechnen

$$\begin{aligned} \tilde{u}_0[2N - (1 + 2k)] &= \sum_{m=-\frac{L}{2}+1}^{\frac{L}{2}} \tilde{x}[2N - (1 + 2k) - m] h_0[m] \\ &= \sum_{m=-\frac{L}{2}+1}^{\frac{L}{2}} \tilde{x}[m + 2k] h_0[m] \end{aligned} \quad (4.7)$$

Indem der Index m durch $\lambda = 1 - m$ ersetzt wird, erhält man

$$\tilde{u}_0[2N - (1 + 2k)] = \sum_{\lambda=-\frac{L}{2}+1}^{\frac{L}{2}} \tilde{x}[1 - \lambda + 2k] h_0[1 - \lambda] \quad (4.8)$$

Setzt Gl. (4.3) in Gl. (4.8) ein, ergibt sich schließlich der Beweis

$$\begin{aligned}
\tilde{u}_0[2N - (1 + 2k)] &= \sum_{\lambda = -\frac{L}{2} + 1}^{\frac{L}{2}} \tilde{x}[1 - \lambda + 2k] h_0[\lambda] \\
&= \tilde{u}_0[1 + 2k] .
\end{aligned} \tag{4.9}$$

Auf gleiche Weise läßt sich zeigen, daß die beiden Halbperioden von $\tilde{u}_1[n]$ aufgrund von Gl. (4.4) antisymmetrisch sind, d.h.

$$\tilde{u}_1[2N - (1 + 2k)] = - \tilde{u}_1[1 + 2k] . \tag{4.10}$$

Werden bei der Reduktion der Abtastfrequenz um den Faktor 2 die Abtastwerte $\tilde{u}_0[1 + 2k]$ und $\tilde{u}_1[1 + 2k]$ übernommen, dann lassen sich $\tilde{z}_0[n]$ und $\tilde{z}_1[n]$ jeweils mit einer Halb-Periode mit der Länge $N/2$ vollständig beschreiben, so daß die beiden Teilbandsignale $z_0[n]$ und $z_1[n]$ insgesamt genau so viele Abtastwerte enthalten wie das ursprüngliche Eingangssignal.

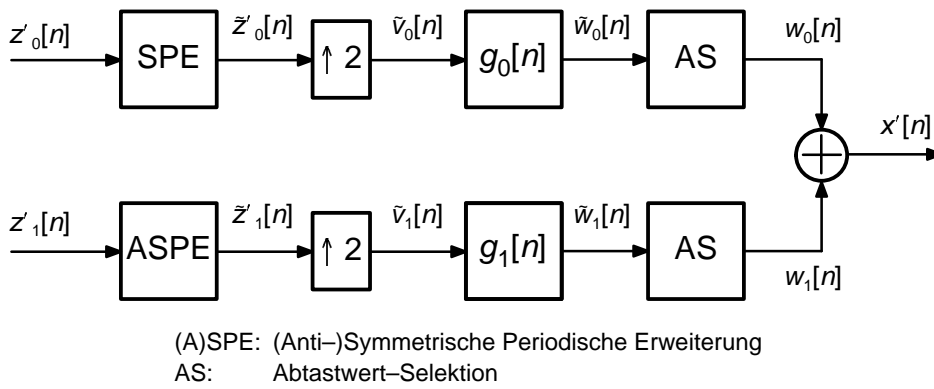
Zur Rekonstruktion der Eingangssignals werden die übertragenen Signale $z'_0[n]$ und $z'_1[n]$, wie Bild 4.8 zeigt, zunächst zu einem symmetrischen periodischen Signal $\tilde{z}'_0[n]$ bzw. antisymmetrischen periodischen Signal $\tilde{z}'_1[n]$ erweitert. Anschließend wird die Abtastfrequenz durch Einfügen von Nullen erhöht, so daß sich $\tilde{v}_0[n]$ und $\tilde{v}_1[n]$ ergeben. Die Synthesefilter $g_0[n]$ und $g_1[n]$ erfüllen die Bedingungen in Gl. (2.6) und (2.8) mit den Koeffizienten

$$g_0[n] = g_0[1 - n], \quad n = -\frac{L}{2} + 1, \dots, -1, 0, 1, \dots, \frac{L}{2}, \tag{4.11}$$

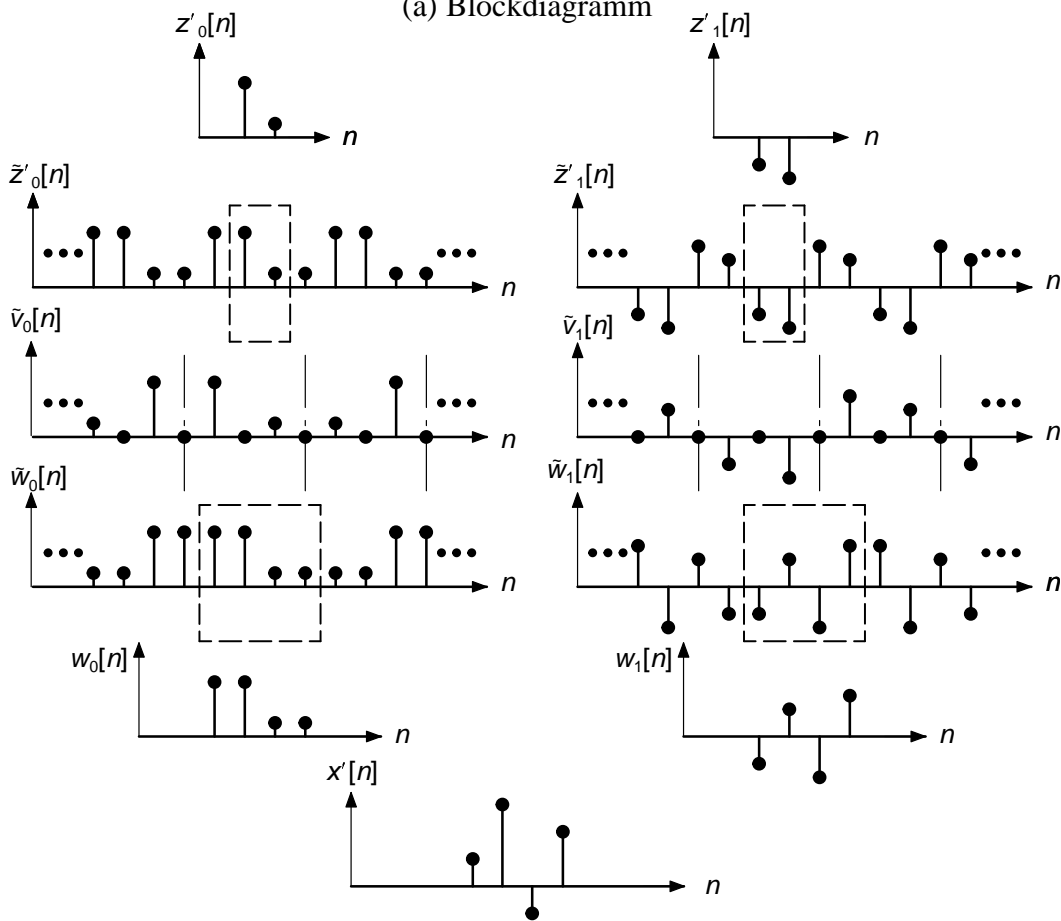
$$g_1[n] = -g_1[1 - n], \quad n = -\frac{L}{2} + 1, \dots, -1, 0, 1, \dots, \frac{L}{2}. \tag{4.12}$$

Durch Addition der richtig ausgewählten Halb-Perioden der gefilterten Signale $\tilde{w}_0[n]$ und $\tilde{w}_1[n]$ kann das Rekonstruktionssignal $x'_1[n]$ schließlich berechnet werden.

Eine allgemeine Teilbandanalyse und -synthese, bei der ein Eingangsbild zweidimensional in mehrere Teilbänder zerlegt wird, kann durch Kaskadierung der oben vorgestellten eindimensionalen Teilbandanalyse- bzw. -synthesestufen mit zwei Teilbändern realisiert werden. Dabei wird zunächst in der horizontalen und anschließend in der vertikalen Richtung gefiltert. Diese Stufen werden in dieser Arbeit jeweils als Basisfilterstufen für die Teilbandanalyse und -synthese bezeichnet. In Bild 4.8 wird als Beispiel eine Teilbandanalyse- und synthese mit 16 gleichförmigen Teilbändern dargestellt.



(a) Blockdiagramm



(b) Beispiel

Bild 4.8 Eindimensionale Teilbandsynthese mit zwei Teilbändern unter Anwendung der symmetrischen periodischen Erweiterung

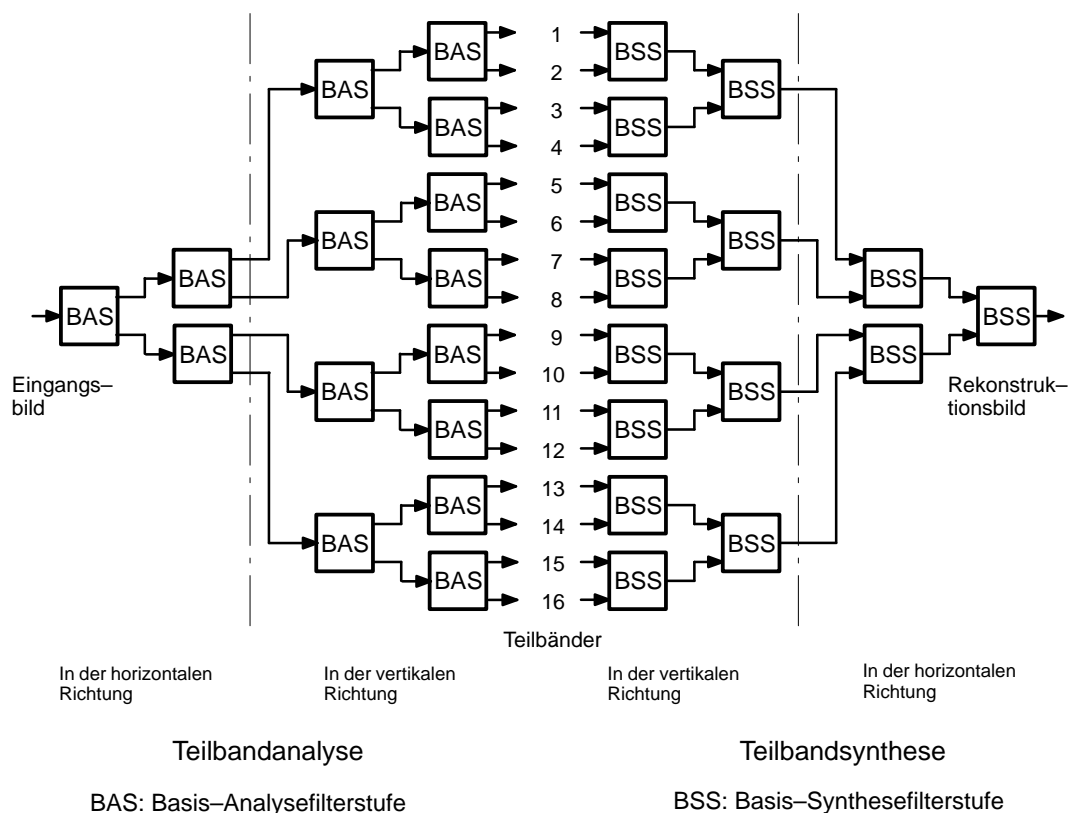


Bild 4.9 Teilbandanalyse und –synthese mit 16 Teilbändern durch Anwendung von Kaskaden aus Basisfilterstufen

4.2 Auswahl der Teilbandfilterbank

Eine Teilbandfilterbank besteht aus Teilbandfilterstufen und dient zur Teilbandanalyse und –synthese des zu codierenden Signals. Im folgenden werden die Teilbandfilterbank und die Teilbandfilterstufe einfach als Filterbank bzw. Filterstufe bezeichnet. In jeder Filterstufe wird neben der Filterung auch eine Umwandlung der Abtastfrequenz durchgeführt. Die Filterbankstruktur bestimmt die Bandaufteilung, nämlich die Anzahl und die Bandbreiten der Teilbänder, während die darin verwendeten Filter die spektrale Trennung zwischen den Teilbändern beeinflussen. Die Filterbank hat einen großen Einfluß auf den Gewinn einer Teilbandcodierung. Deshalb ist die Auswahl der Filterbank eine wichtige Aufgabe beim Entwurf eines Teilbandcoders.

Zur Reduzierung des Realisierungsaufwandes wird in dieser Arbeit die Auswahl auf Filterbänke, die durch eine Kaskadierung der in Kapitel 4.1.2 beschriebenen Basisfilterstufen entstehen, eingeschränkt. Die Aufgabe dieses Abschnittes ist es, Filterkaskaden einschließlich der darin verwendeten QM-Filter auszuwählen. Da der Einfluß einer Filterkaskade auf den Codierungsgewinn, der durch Ausnutzung der statistischen Eigenschaften des Signals sowie der menschlichen Wahrnehmung erzielt werden kann, sich nicht auf einfache Weise beschreiben läßt, wird die Auswahl in dieser Arbeit anhand experimenteller Untersuchungen durchgeführt. Dazu wird die mittlere Entropie von allen Teilbandsignalen, die anhand typischer Testsequenzen unter Anwendung der im nächsten Kapitel beschriebenen adaptiven Quantisierung bei der Sichtbarkeitschwelle ermittelt wird, als Auswahlkriterium verwendet. In 4.2.1 wird zunächst die Auswahl der Struktur der Filterkaskaden behandelt. In 4.2.2 folgt die Auswahl der QM-Filter für die einzelnen Basisfilterstufen.

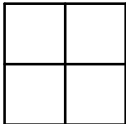
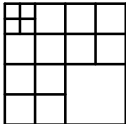
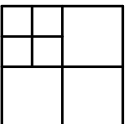
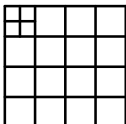
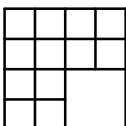
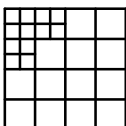
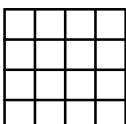
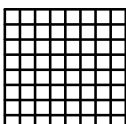
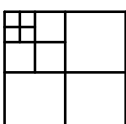
4.2.1 Auswahl der Struktur der Filterkaskaden

Die Struktur einer Filterkaskade wird durch die Anzahl der Filterstufen, die zur Erzeugung eines Teilbandes jeweils benutzt werden, festgelegt. Sie hat folgende Auswirkung auf den Codierungsgewinn. Je mehr Filterstufen verwendet werden, desto feiner ist die spektrale Zerlegung und um so größer ist der Codierungsgewinn. Mit der Anzahl der Filterstufen wächst jedoch die Filterlänge, so daß die Anpassung der Codierung an die lokalen Eigenschaften des Bildsignals eingeschränkt wird. Je nachdem, ob die Anzahl der Filterstufen für alle Teilbänder gleich oder ungleich ist, wird von einer gleichförmigen oder einer ungleichförmigen Bandaufteilung gesprochen. Bei einer ungleichförmigen Bandaufteilung benötigen einige Teilbänder wenige Filterstufen, so daß der Realisierungsaufwand reduziert werden kann. Im folgenden sollen zwei Fragen zur Auswahl der Filterkaskadestruktur geklärt werden: (1) Wie viele Filterstufen soll die Filterkaskade maximal besitzen? (2) Soll eine gleichförmige oder eine ungleichförmige Bandaufteilung eingesetzt werden?

Zur Erklärung dieser Fragen werden unterschiedliche Bandaufteilungen zuerst anhand der Codierungseffizienz der Luminanzkomponente miteinander verglichen. Dazu wird, wie oben beschrieben, die jeweilige mittlere Entropie der Teilbandsignale von den Original- und den Prädiktionsfehlerbildern ermittelt. In Tabelle 4.2 und 4.3 sind die Meßergebnisse aus der Testsequenz 'Claire' aufgeführt. Bei dieser Untersuchung werden Filterkaskaden aus QM-Filtern mit 16 Koeffizienten (QMF16A, siehe Angang A) verwendet.

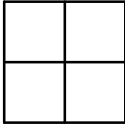
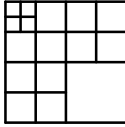
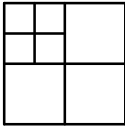
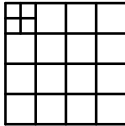
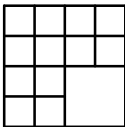
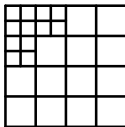
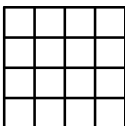
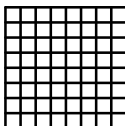
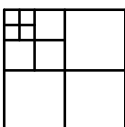
Diese Ergebnisse zeigen, daß unter den gleichförmigen Bandaufteilungen die Aufteilung mit 64 Teilbändern das Minimum der mittleren Entropie ergibt. Die Senkung der

Tabelle 4.2 Mittlere Entropie [bit/Abtastwert] der Teilbandsignale der Originalbilder bei verschiedenen Bandaufteilungen (Testsequenz: Claire)

	2,24		1,34
	1,57		1,31
	1,52		1,31
	1,48		1,33
	1,40	256 Teil- bänder	1,38

mittleren Entropie bei Erhöhung der Teilbandanzahl von 4 über 16 auf 64 läßt sich auf eine bessere spektrale Zerlegung zurückführen, während die Erhöhung der mittleren Entropie bei 256 Teilbändern auf eine schlechtere lokale Anpassung hindeutet. Bei den ungleichförmigen Bandaufteilungen, bei denen die weitere Zerlegung mancher hochfrequenten Teilbänder entfällt, ist die mittlere Entropie der Teilbandsignale in den meisten Fällen größer als die entsprechende Entropie aus derjenigen gleichförmigen Bandaufteilung, die die gleiche maximale Anzahl der Filterstufen hat. In den anderen Fällen ist diese Entropie nur geringfügig kleiner als die der entsprechenden gleichförmigen Aufteilung. Um die Teilbandabtastwerte anhand 2D-Lauflängen codieren zu können, wird in diesem Coder die gleichförmige Bandaufteilung eingesetzt. Angesichts des

Tabelle 4.3 Mittlere Entropie [bit/Abtastwert] der Teilbandsignale der Prädiktionsfehlerbilder bei verschiedenen Bandaufteilungen (Testsequenz: Claire)

	1,47		1,38
	1,44		1,35
	1,39		1,35
	1,35		1,33
	1,44	256 Teil- bänder	1,42

höheren Codierungsgewinns stellt eine solche Bandaufteilung mit 64 Teilbändern die beste Wahl dar.

Im nächsten Schritt werden die Formen der ursprünglich zu codierenden Bildgebiete bei der Auswahl der Teilbandanzahl berücksichtigt. Wie im letzten Kapitel erwähnt, besteht ein rekonstruiertes Bildgebiet der vorgeschlagenen blockbasierten Teilbandcodierung immer aus rechteckigen Blöcken der Größe 4x4 oder 8x8 Bildpunkte. Um ein beliebig geformtes Bildgebiet zu übertragen, wird immer ein aus solchen Blöcken zusammengesetztes, größeres Bildgebiet, das das ursprüngliche Gebiet umfaßt, codiert und übertragen. Demzufolge werden, außer den im ursprünglichen Gebiet vorhandenen Bildpunk-

ten, zusätzliche Bildpunkte übertragen. Um die Anzahl der zusätzlich zu codierenden Bildpunkte zu reduzieren, empfiehlt es sich in diesem Fall, kleinere Blöcke zu verwenden. Aus diesem Grund wird in diesem Coder neben der gleichförmigen Bandaufteilung mit 64 Teilbändern auch eine gleichförmige Bandaufteilung mit 16 Teilbändern verwendet. Die Auswahl zwischen den beiden wird im letzten Schritt der Vorwärtsanalyse durchgeführt.

Zusammenfassend werden zwei Filterkaskaden für diesen Coder ausgewählt: Die eine besteht aus jeweils drei Basisfilterstufen, und die andere aus jeweils zwei Basisfilterstufen für die horizontale und vertikale Teilbandzerlegung. Dadurch kann wahlweise eine Teilbandzerlegung mit 64 oder 16 gleichförmigen Teilbändern verwendet werden.

4.2.2 Auswahl der Teilbandfilter

Wie in Abschnitt 2.2 bereits erläutert, besitzen die Quadraturspiegelfilter (QMF) im Vergleich zu den konjugierten Quadraturspiegelfiltern (CQF) die folgenden Vorteile. Erstens haben die QMF eine lineare Phase, so daß die Phasenfehler dadurch vermieden werden können. Zweitens ermöglicht die symmetrische Impulsantwortfunktion der QMF die Anwendung der symmetrischen periodischen Erweiterung in den Basisfilterstufen. Aus diesen Gründen wird in dieser Arbeit auf die Anwendung von QM-Filtern beschränkt.

In Kapitel 2 wurde dargestellt, daß die Impulsantwortfunktionen $h_1[n]$, $g_0[n]$ und $g_1[n]$ in den Basisfilterstufen sich aus der Impulsantwortfunktion $h_0[n]$ ableiten lassen. Deshalb genügt hier die Auswahl des Tiefpaßfilters $h_0[n]$. Zur Auswahl stehen in dieser Arbeit die von Johnston entwickelten QM-Filter [27], die hier mit den Namen QMF lt bezeichnet. Dabei gibt l die Länge und t den Typ des Filters an. Z.B. wird das QM-Filter mit der Länge 16 und der Typ A als QMF16A genannt. Neben den QM-Filtern von Johnston wird auch ein kurzes Filter QMF02 mit der Impulsantwortfunktion (0.5,0.5) verwendet.

Bei der Auswahl der Filter wurden mehrere Aspekte berücksichtigt. Der erste Aspekt ist der Teilbandcodierungsgewinn. Dazu bestehen im Frequenz- und im Ortsbereich aber widersprüchliche Anforderungen. Im Frequenzbereich sollen die Aliasing-Anteile der Teilbandsignale möglichst klein gehalten werden. Anderenfalls erhöht sich die Signalleistung in hochfrequenten Teilbandsignalen und demzufolge auch die Datenrate. Zur Reduzierung der Aliasing-Anteile wird eine Filterübertragungsfunktion mit einem schmalen Übergangsbereich und einer hohen Sperrdämpfung benötigt. Solche Filter besitzen im allgemeinen eine lange Impulsantwortfunktion. Im Ortsbereich sollen die Teilbandsignale das lokale Spektrum wiedergeben, um den lokalen spektralen Unter-

schied besser ausnutzen zu können. Danach sollen die Teilbandabtastwerte möglichst aus einem kleinen Bildgebiet berechnet werden. Dies läßt sich nur durch Anwendung von kurzen Filtern erzielen.

Um den Zusammenhang zwischen der Filterlänge und dem Codierungsgewinn zu ermitteln, wurden Computersimulationen mit QM-Filtern unterschiedlicher Längen durchgeführt. Dabei wurden die mittlere Entropie der quantisierten Teilbandsignale bei verschiedenen Bandaufteilungen gemessen. Die Meßergebnisse von den Originalbildern der Testsequenz 'Claire' werden in Tabelle 4.4 dargestellt. Diese Ergebnisse zeigen, daß mit dem Filter QMF16A die kleinste mittlere Entropie der Teilbandsignale und dementsprechend der größten Codierungsgewinn erzielt werden können.

Tabelle 4.4 Mittlere Entropie der Teilbandsignale bei Anwendung von QM-Filtern unterschiedlicher Länge

Filterkaskaden aus	Mittlere Entropie der Teilbandsignale [bit/Abtastwert]		
	4 Teilbänder	16 Teilbänder	64 Teilbänder
QMF02	2,53	1,84	1,71
QMF08A	2,31	1,58	1,44
QMF16A	2,24	1,48	1,33
QMF32C	2,24	1,52	1,39

Der zweite bei der Auswahl der Filter berücksichtigte Aspekt ist die Einschränkung der Ausdehnung von Rekonstruktionsfehlern. Bei sehr niedrigen Übertragungsbitraten kann die Quantisierung nicht immer so gestaltet werden, daß die Rekonstruktionsfehler ständig unter der Sichtbarkeitsschwelle bleiben. Bei sichtbaren Rekonstruktionsfehlern hängen die resultierenden subjektiven Störungen nicht nur von der Amplitude sondern auch von der Ausdehnung der Fehler ab. Je größer die Ausdehnung ist, desto auffälliger sind die Fehler. Die Ausdehnung der Rekonstruktionsfehler, die von dem Quantisierungsfehler eines einzigen Teilbandabtastwertes verursacht werden, ist in jeder Richtung zur Länge des äquivalenten Synthesefilters proportional. Im Bild 4.9 wird an einem Beispiel erklärt, wie das äquivalente Synthesefilter einer Filterkaskade definiert ist und sich berechnen läßt. Dabei werden der zweite und dritte Funktionsblock zur Erhöhung der Abtastfrequenz vor das Filter SF₁ verschoben. Entsprechend werden zwischen den ursprünglichen Filterkoeffizienten jeweils zwei Null-Koeffizienten im zweiten Filter

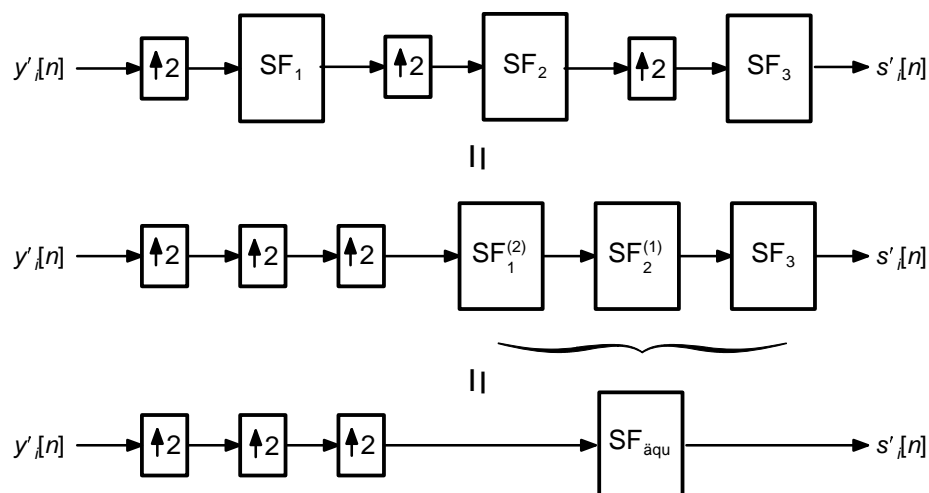


Bild 4.10 Das äquivalente Filter einer Filterkaskade für Teilbandsynthese

SF_2 und ein Null-Koeffizient im dritten Filter SF_3 eingefügt. Die Länge des äquivalenten Synthesefilters lässt sich gemäß Bild 4.9 mit der Gleichung

$$\begin{aligned}
 L_{SF_{\text{äqu}}} &= [(4 \cdot L_{SF_1} - 3) + (2 \cdot L_{SF_2} - 1) - 1] + L_{SF_3} - 1 \\
 &= 4 \cdot L_{SF_1} + 2 \cdot L_{SF_2} + L_{SF_3} - 6
 \end{aligned}
 \tag{4.13}$$

berechnen, wobei $L_{SF_{\text{äqu}}}$ und $L_{SF_{1,2,3}}$ die Längen der entsprechenden Filter sind.

Die Filterkaskade aus drei QMF16A-Filtern ergibt ein äquivalentes Filter mit 106 Koeffizienten. Hinsichtlich der Ausdehnung der Rekonstruktionsfehler ist diese Filterkaskade für die blockbasierte Teilbandcodierung bei sehr niedrigen Übertragungsbitraten nicht geeignet.

Zur Realisierung von Filterkaskaden, die einerseits einen hohen Codierungsgewinn aufweisen und andererseits eine akzeptable Länge der äquivalenten Filter besitzen, werden die Rollen der Filter in den verschiedenen Filterstufen betrachtet. In den Analysestufen wirken sich die Filter bezüglich der Aliasing unterschiedlich aus. Das Filter in der ersten Stufe verarbeitet das ganze Spektrum des Eingangssignals. Dabei kann ein Aliasing-Anteil bei einer Frequenz auftreten, die weit von seiner ursprünglichen Frequenz entfernt ist. Die in dieser Stufe eingefügten Aliasing-Anteile verbleiben nach weiteren Teilbandzerlegungen in den Teilbandsignalen. Im Gegensatz dazu wird in den

weiteren Filterstufen nur einen Teil des Spektrums verarbeitet. Um das Aliasing zu reduzieren, empfehlen sich also für die erste Analysestufe längere Filter mit einer besseren Bandtrennung. Für die weiteren Stufen genügen jedoch kürzere Filter. In den Synthesestufen leisten die Filter unterschiedliche Beiträge zur Länge des äquivalenten Synthesefilters. Wie in Gl. (4.13) gezeigt, kommt der größte Beitrag von der ersten Synthesestufe. Um die Länge der äquivalenten Filter zu reduzieren, empfehlen sich für diese Stufe kurze Filter. Da die Filter in der ersten Synthesestufe sich aus den Filtern in der letzten Analysestufe ableiten lassen, gilt diese Empfehlung also auch für die letzte Analysestufe. Aus dieser Betrachtung ergibt sich, daß in den verschiedenen Stufen der Filterkaskaden Filter mit unterschiedlichen Längen verwendet werden sollen. In der ersten Analysestufe sollen längere Filter und in der letzten kürze Filter eingesetzt werden. Basierend auf dieser Erkenntnis werden die Tiefpaßfilter für die einzelnen Analysestufen schließlich wie in Tabelle 4.5 dargestellt ausgewählt. Tabelle 4.6 zeigt, daß sich die mittlere Entropie der Teilbandsignale mit den ausgewählten Filterkombinationen im Vergleich zu den entsprechenden Filterkaskaden, die ausschließlich aus dem Filter QMF16A bestehen, nur jeweils geringfügig erhöht. Bei der dreistufigen Filterkaskade wird die Länge des äquivalenten Synthesefilters aber von 106 auf 34 reduziert.

In Anhang A sind die Impulsantwortfunktionen der in diesem Coder verwendeten QM-Filter aufgeführt.

Tabelle 4.5 Die ausgewählten Tiefpaßfilter für die einzelnen Analysestufen

Anzahl der Teilbänder	Filter in den einzelnen Stufen in einer Richtung		
	1	2	3
16	QMF16A	QMF08A	–
64	QMF16A	QMF08A	QMF02

Tabelle 4.6 Vergleich der Filterkaskaden in Tabelle 4.5 mit den entsprechenden Filterkaskaden aus nur QMF16A anhand der mittleren Entropie der Teilbandsignale

Filterkaskaden aus	Mittlere Entropie [bit/Abtastwert]	
	16 Teilbänder	64 Teilbänder
nur QMF16A	1,48	1,33
QMF16A, QMF08A, QMF02	1,53	1,39

4.3 Umschaltung zwischen einer Inter– und einer Intraframe–Codierung in jedem Teilband

In [16] wurde gezeigt, daß in höheren Ortsfrequenzen das Prädiktionsfehlersignal eine höhere Leistung als das Originalbildsignal besitzen kann. Aus der Informationstheorie ist bekannt, daß, bei Codierung zweier Signale mit dem gleichen Typ von Amplitudenverteilung, das Signal mit der kleineren Varianz weniger Bit pro Abtastwert bei gleicher Rekonstruktionsfehlervarianz benötigt. Da sowohl für das Originalbild als auch für das Prädiktionsfehlerbild die hochfrequenten Teilbandsignale sich gemäß [61] mit der gleichen Amplitudenverteilung beschreiben lassen, kann die mittlere Datenrate dadurch reduziert werden, daß für jedes Teilband immer jenes Teilbandsignal aus dem Originalbild oder dem entsprechenden Prädiktionsfehlerbild ausgewählt und codiert wird, das die geringere Varianz aufweist. Diese Umschaltung verhindert gleichzeitig die zeitliche Verbreitung der Prädiktionsfehler in den Teilbändern, die bei sehr niedrigen Übertragungsbitraten sich dann ergibt, wenn für diese Teilbänder die Prädiktionsfehler in mehreren nacheinander folgenden Bildern nicht codiert und übertragen werden können.

Um die Umschaltung zwischen der Inter– und der Intraframe–Codierung in jedem Teilband zu ermöglichen, wird die in Bild 2.9 gezeigte ursprüngliche Coderstruktur modifiziert. Die resultierende neue Coderstruktur wird in Bild 4.11 dargestellt. Die Hauptänderungen sind der Ersatz der Teilbandzerlegung des Prädiktionsfehlersignals durch zwei Teilbandzerlegungen, eine für das Originalbildsignal und eine für das Prädiktionsignal, sowie die Berechnung der Prädiktionsfehlersignale im Teilbandbereich statt im Bildbereich. Die Funktionsweise dieser Umschaltung läßt sich wie folgt beschreiben.

Zunächst wird die Teilbandanalyse sowohl auf das aktuelle Eingangsbild als auch auf das entsprechende Prädiktionsbild angewendet. Für jedes einzelne Teilband wird anschließend durch Vergleich der Signalvarianzen entschieden, ob das Teilbandsignal des Originalbildes oder das sogenannte Teilband–Prädiktionsfehlersignal codiert werden soll. Entsprechend wird die Schaltung S individuell für jedes Teilband umgeschaltet. In den Teilbändern, in denen das Teilbandsignal des Originalbildes übertragen werden soll, wird diese Schaltung auf die Position 1 geschaltet. In den anderen Teilbändern wird sie auf 2 geschaltet. Diese Einstellung wird einmal pro Bild vorgenommen. Vor der Teilbandsynthese wird in jedem Teilband, in dem das Teilband–Prädiktionsfehlersignal übertragen wurde, dieses Signal zunächst zum entsprechenden Teilbandsignal des Prädiktionsbildes addiert. Dann wird das Bildsignal in den ausgewählten und codierten Blöcken anhand der Teilbandsynthese rekonstruiert.

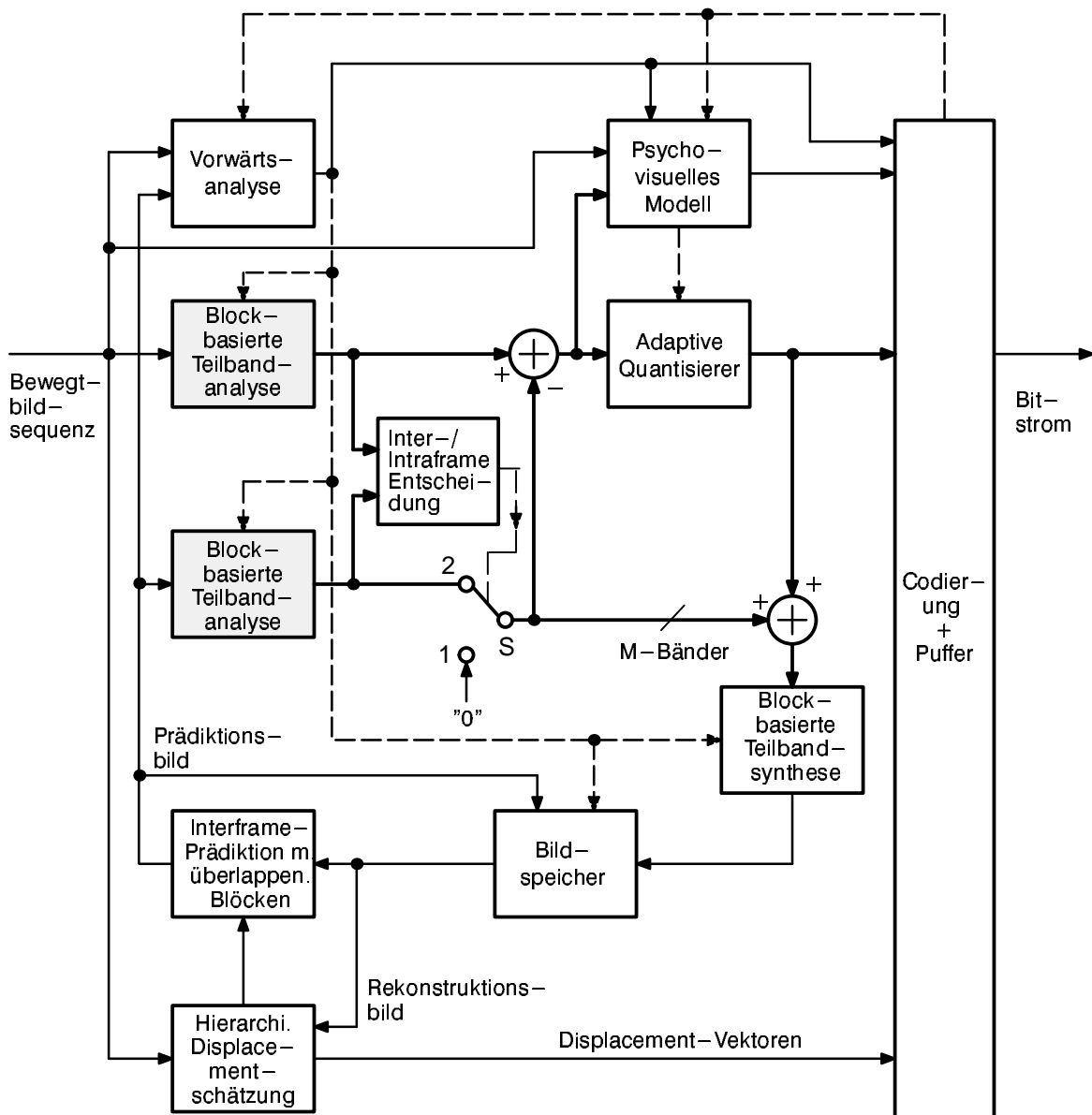


Bild 4.11 Blockschaltbild des bewegungskompensierenden Teilbandcoders mit der Umschaltung zwischen der Inter- und Intraframe-Codierung in jedem Teilband

4.4 Codierung der quantisierten Teilbandsignale

Die Codierung der quantisierten Teilbandsignale lehnt sich in diesem Coder eng an die Codierung der DCT-Koeffizienten in H.263 [25] an. Äquivalent zu den Blöcken der DCT-Koeffizienten in H.263 werden hier Folgen aus den Teilbandabstastwerten mit der gleicher Ortszugehörigkeit im Eingangsbild gebildet. Innerhalb einer Folge werden die Teilbandabstastwerte nach den steigenden zugehörigen Ortsfrequenzen geordnet.

Wie in Abschnitt 4.2.1 bereits erklärt, können in diesem Coder wahlweise 16 oder 64 Teilbänder verwendet werden. Dementsprechend ergeben sich Folgen mit 16 bzw. 64 Abstastwerten. Die Codierung der Folgen mit 64 Abstastwerten ist vergleichbar zur Codierung der 8x8 DCT-Koeffizienten in H.263. Dabei wird jede Folge durch Angabe der darin enthaltenen Null-Lauflängen und der Amplituden der Abstastwerte ungleich Null codiert. Eine Null-Lauflänge gibt die Anzahl der aufeinander folgenden Null-Abstastwerte an. Der Gewinn der Lauflängencodierung ergibt sich unter andere durch die häufig auftretenden langen Null-Lauflängen.

Aufgrund kürzerer Lauflängen ist von der direkten Lauflängencodierung der Folgen mit 16 Teilbandabstastwerten kein großer Gewinn zu erwarten. Um den Codierungsgewinn zu erhöhen, werden die zu einem 16x16 Makroblock zugehörigen 16 Folgen mit jeweils 16 Abstastwerten zu 4 neuen Folgen mit je 64 Abstastwerten zusammengefaßt und codiert. Bei der Zusammensetzung der Folgen mit jeweils 16 Abstastwerten werden deren Abstastwerte ineinander verkämmt angeordnet. d.h., zuerst wird nacheinander der 1. Abstastwert aus jedem der ursprünglichen Folgen in der neuen Folge angeordnet, dann folgt der 2. und der 3. Abstastwert usw..

In diesem Kapitel wurden verschiedene Themen der blockbasierten Teilbandcodierung ausführlich behandelt. In Kapitel 6 wird die blockbasierte Teilbandcodierung anhand Computersimulationen mit der bildbasierten Teilbandcodierung sowie der DCT-Codierung verglichen.

5. Adaptive Quantisierung unter Anwendung eines psychovisuellen Modells

Zur Ausnutzung der Irrelevanzreduktion ist die Anpassung der Quantisierung an die Eigenschaften der menschlichen visuellen Wahrnehmung erforderlich. Im Vergleich zur PCM oder DPCM kann eine solche Anpassung bei der Teilbandcodierung wegen der spektralen Zerlegung nicht nur im Orts– sondern auch im Frequenzbereich flexibler gestaltet werden. Im allgemeinen ergeben bei einer vorgegebenen Gesamtdatenrate unterschiedliche Verteilungen der Quantisierungsfehler auf die einzelnen Teilbänder unterschiedliche subjektive Qualitäten der Rekonstruktionsbilder. Hier entsteht die Frage, welche Verteilung der Quantisierungsfehler die optimale ist, bei der die höchste subjektive Bildqualität für eine vorgegebene Datenrate resultiert. In bisherigen Veröffentlichungen wird die Modulationsübertragungsfunktion (MTF) des menschlichen visuellen Systems zur Bestimmung der Verteilung der Quantisierungsfehler auf die Teilbänder verwendet. Dafür wird die Summe der mit der MTF gewichteten Quantisierungsfehlervarianzen der Teilbandsignale als Fehlerkriterium benutzt. Bei diesem Kriterium wird implizit angenommen, daß die subjektive Bewertung der Rekonstruktionsfehler proportional dieser Summe ist. Diese Annahme kann jedoch nicht aus der MTF direkt hergeleitet werden, weil die MTF lediglich die Abhängigkeit der Sichtbarkeitschwelle von der Ortsfrequenz beschreibt.

In diesem Kapitel soll auf die oben genannte Fragestellung durch Verwendung eines psychovisuellen Modells für das Luminanzsignal eingegangen werden. Das verwendete Modell berücksichtigt nicht nur die MTF sondern darüberhinaus auch die Maskierungseffekte. Nach einer kurzen Beschreibung des psychovisuellen Modells und der im Modell definierten "Potentiellen Sichtbarkeitsfunktion" wird der Zusammenhang zwischen den Quantisierungsfehlervarianzen der einzelnen Teilbandsignale und dem Wert der Potentiellen Sichtbarkeitsfunktion hergeleitet. Daraus wird die optimale Aufteilung der Datenrate in den Teilbändern bestimmt. Anschließend wird die Realisierung der adaptiven Quantisierung in diesem Coder beschrieben.

5.1 Ein Psychovisuelles Modell der menschlichen visuellen Wahrnehmung

Bei der Entwicklung der adaptiven Quantisierung der Teilbandsignale wurde in dieser Arbeit das Girodsche Modell der menschlichen visuellen Wahrnehmung zur Irrelevanzreduktion von Fernseh luminanzsignalen [17] verwendet. Dieses Modell ist eine Weiterentwicklung des Modells von Lukas und Budrikis [32] und stellt den Stand der Entwick-

lung in diesem Gebiet dar. Als Schwellenmodell liefert es die Aussage, ob ein additives, rauschartiges Störsignal lokal sichtbar oder nicht sichtbar ist. Wie in [17] beschrieben, stimmt für verschiedene visuelle Wahrnehmungsphänomene die Modellaussage mit dem Ergebnis der entsprechenden subjektiven Tests überein. Bei der Bestimmung der Sichtbarkeit des Störsignals berücksichtigt das Modell die folgenden visuellen Wahrnehmungsphänomene:

- Einfluß der örtlichen Ausdehnung und Dauer eines Störsignals
- Einfluß der Hintergrundhelligkeit
- Einfluß der Abtastfrequenzen und des Betrachtungsabstandes
- Örtlich–zeitliche Modulationsübertragungsfunktion
- Maskierung durch örtliche und zeitliche Helligkeitsprünge

Um das Zusammenwirken dieser Phänomene durch ein lineares Modell beschreiben zu können, wird angenommen, daß die durch die Quantisierung erzeugten Störsignale klein sind. In Bild 5.1 ist ein derartiges sogenanntes w -Modell dargestellt. Es besteht aus zwei Zweigen, dem Maskierungs- und dem Störsignalzweig, in denen die einzelnen Signalverarbeitungsschritte mit entsprechenden Teilmodellen abgebildet werden. Im Maskierungszweig werden die sogenannten w -Parameter aus dem ungestörten Originalbild berechnet. Sie stellen den Einfluß des Bildinhaltes auf die lokale Sichtbarkeit des Störsignals als Gewichtungen dar. Mit w_1 wird der Einfluß des Monitors, mit w_2 die Abbildung auf die Retina erfaßt; w_3 beschreibt die laterale Inhibition in der Fovea, auf welche sich die verschiedenen Maskierungseffekte zurückführen lassen. Im Störsignalzweig wird eine lokale Potentielle Sichtbarkeit (PSB) aus dem gewichteten und gefilterten Störsignal berechnet. Ein Vergleich der PSB-Werte mit einer Schwelle ergibt schließlich die Aussage, ob das Störsignal sichtbar oder nicht sichtbar ist. Die Potentielle Sichtbarkeit ist eine Funktion der Bildschirm-Ortskoordinaten.

In dieser Arbeit wird das vereinfachte w -Modell für breitbandige oder hochfrequente Störsignale verwendet. Da das in der Teilbandcodierung resultierende Rekonstruktionsfehlersignal als ein breitbandiges Störsignal angesehen werden kann, ist die Anwendung dieses Modells gerechtfertigt. In diesem Fall werden nur die w -Parameter w_1 und w_2 benutzt. In Bild 5.2 wird das Blockdiagramm des vereinfachten w -Modells gezeigt. Das Blockdiagramm enthält neben einigen Verknüpfungsoperationen mehrere Filter. In Anhang B sind die Impulsantwortfunktionen dieser Filter und die zugehörigen Parameter zusammengestellt. Eine ausführliche Beschreibung des w -Modells ist [17] zu entnehmen.

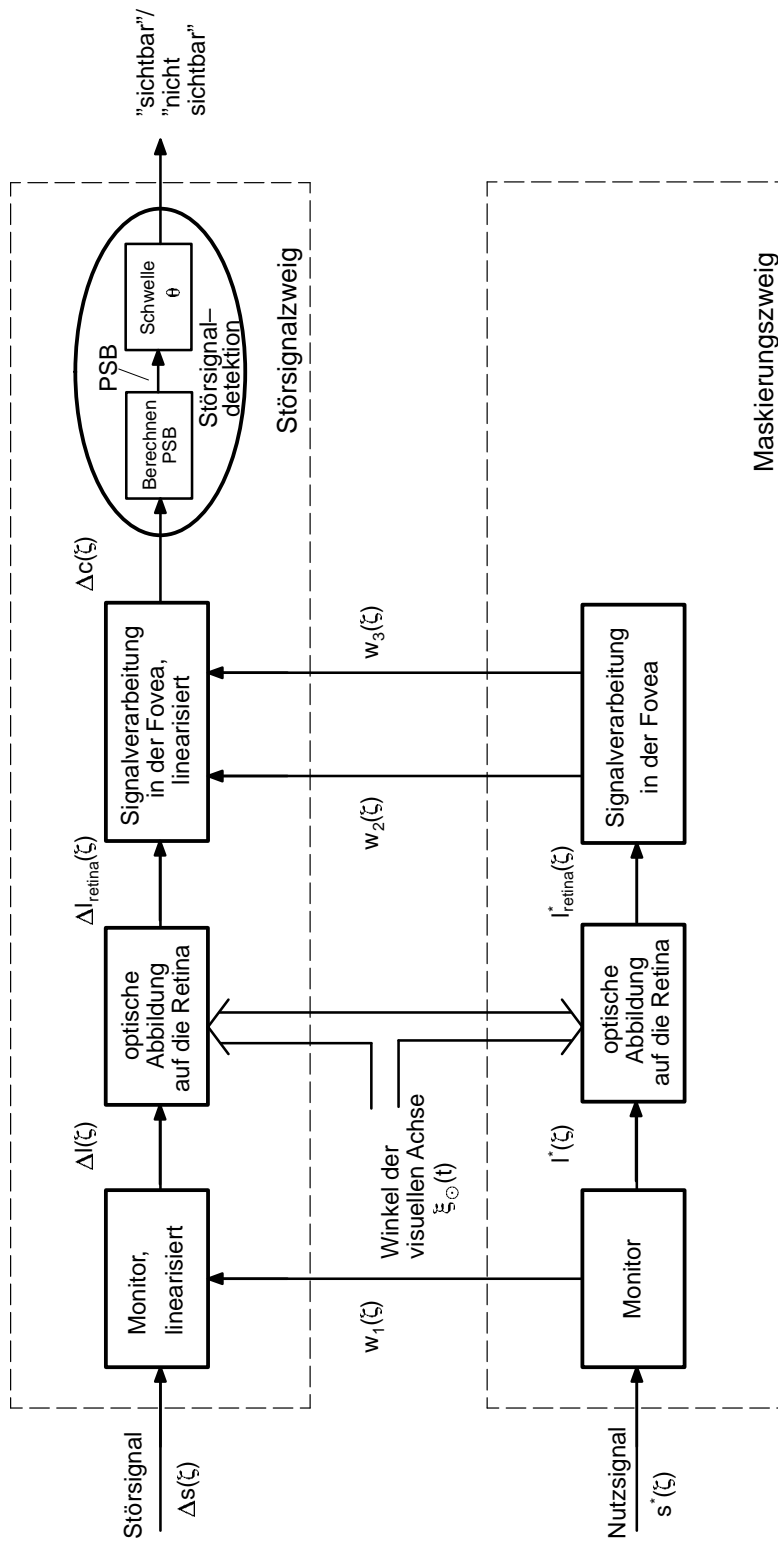
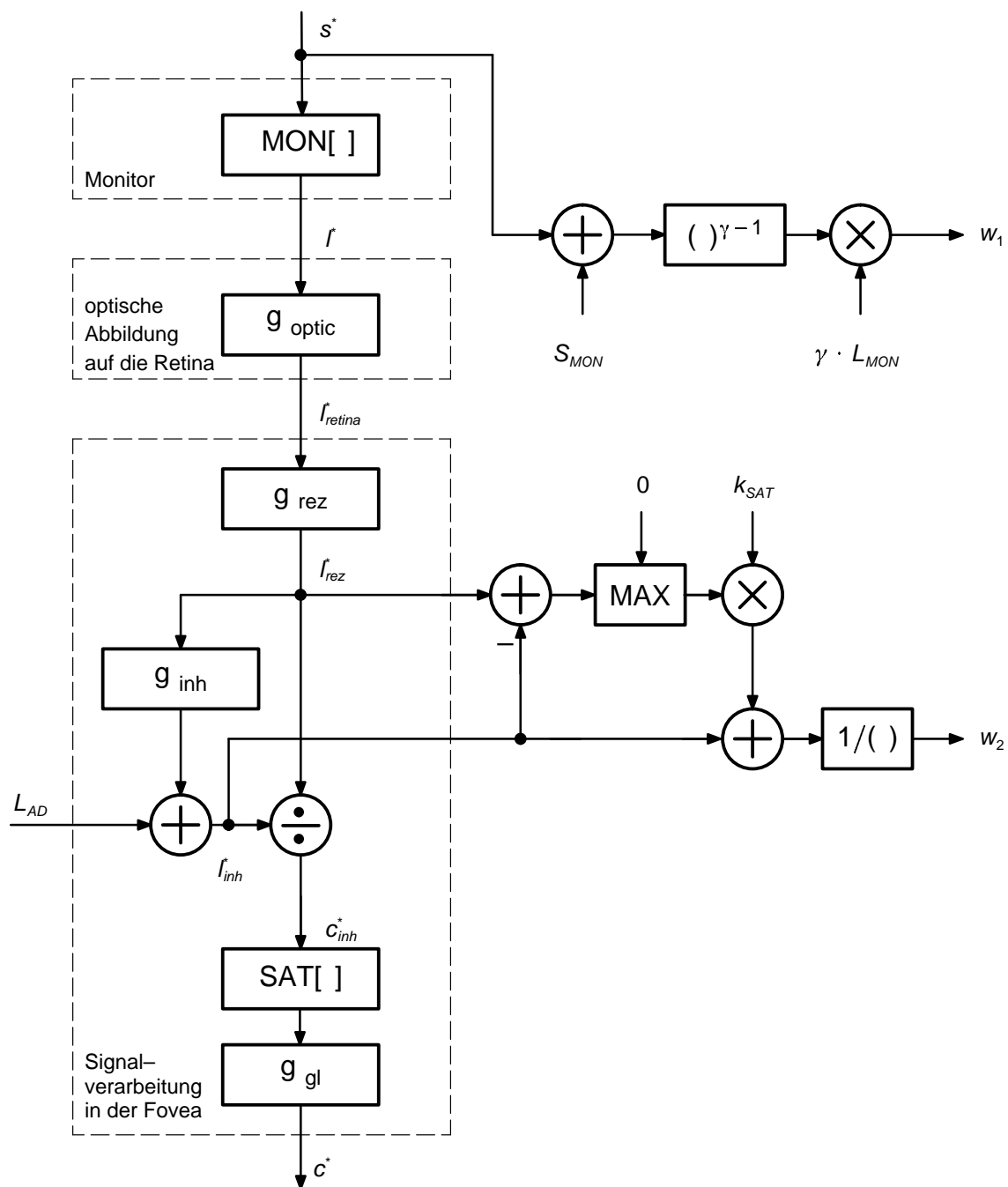
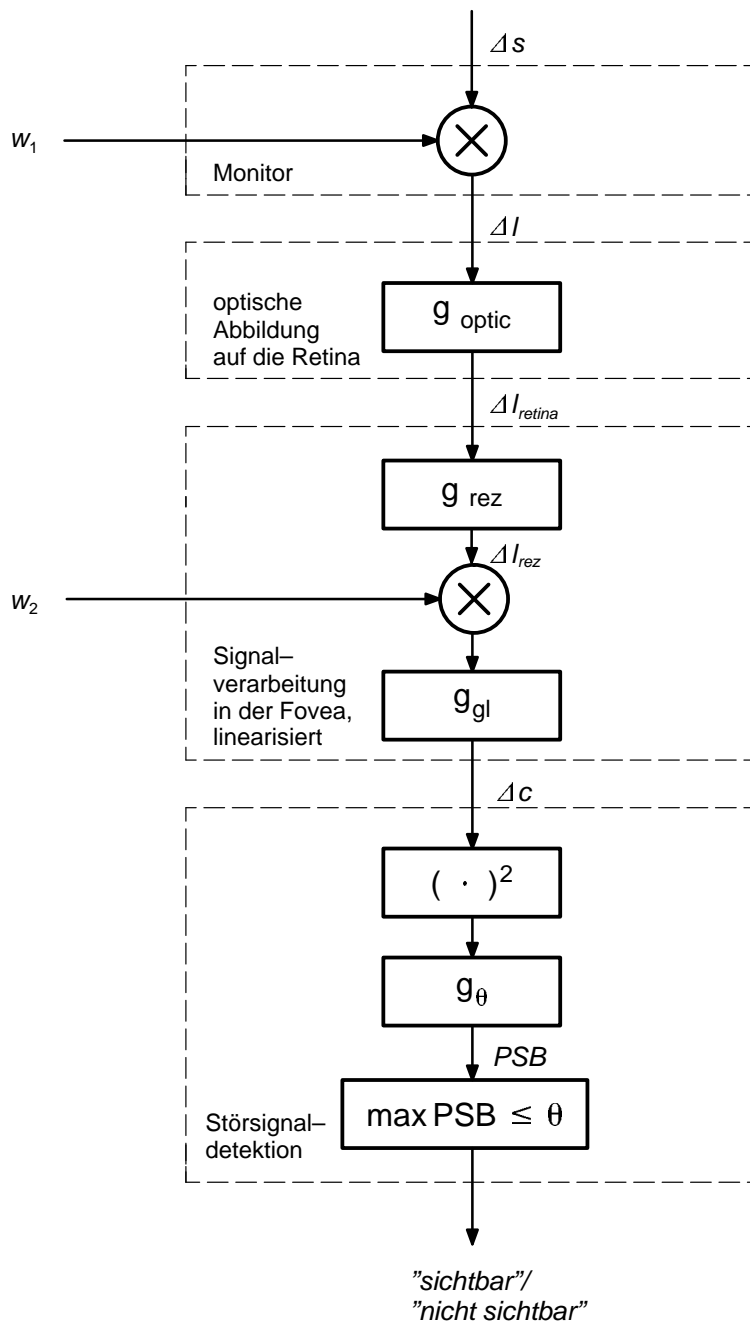


Bild 5.1 Struktur des linearisierten visuellen Modells (w-Modell)

Bild 5.2 (a) Maskierungsweig des vereinfachten w -Modells

Bild 5.2 (b) Störsignalzweig des vereinfachten w -Modells

Die folgenden Abschnitte konzentrieren sich auf die Anwendung des w -Modells zur Steuerung der Quantisierung. Zum Verständnis des nachfolgenden Kapitels genügt die Kenntnis der beiden Blockdiagramme in Bild 5.2(a) und (b) sowie der Definitionen der zugehörigen Filter.

5.2 Potentielle Sichtbarkeit der Rekonstruktionsfehler bei der Teilbandcodierung

Die analytische Herleitung der optimalen Verteilung der Quantisierungsfehler auf die Teilbänder setzt eine mathematische Beschreibung des Zusammenhangs zwischen den Teilbandquantisierungsfehlern und der resultierenden subjektiven Störung voraus. Zur quantitativen Messung der subjektiven Störung wird in dieser Arbeit die im w -Modell definierte Potentielle Sichtbarkeitsfunktion (PSB) des Störsignals benutzt. Diese Vorgehensweise ist gerechtfertigt, solange die durch die Quantisierungsfehler verursachten Störsignale die Sichtbarkeitsschwelle nicht wesentlich überschreiten. Um den Zusammenhang zwischen den Quantisierungsfehlern der Teilbandsignale und der Potentiellen Sichtbarkeitsfunktion zu beschreiben, wird das w -Modell gemäß Bild 5.3(a) zur Bewertung des Störsignals der Teilbandcodierung verwendet. Dabei werden das Originalbildsignal und die Quantisierungsfehlersignale als quasi-stationär angenommen. Außerdem wird die zeitliche Auswirkung der Quantisierungsfehler auf die nachfolgenden Bilder vernachlässigt, da die Bildfrequenz bei sehr niedrigen Übertragungsbitraten stark reduziert wird.

In [17] wird die Potentielle Sichtbarkeitsfunktion eines Störsignals als Funktion der Bildschirm-Koordinaten $\zeta = (x, y, t)$ folgenderweise definiert:

$$PSB(\zeta) = \Delta c^2(\zeta) * g_\theta(\zeta) \quad (5.1)$$

wobei Δc das gewichtete und gefilterte Störsignal am Eingang der Störsignaldetektion und g_θ die Impulsantwort eines Tiefpaßfilters ist. Mit g_θ wird die unterschwellige Summation erfaßt. Diese Definition zeigt, daß die Potentielle Sichtbarkeit ein lokal gewichteter Wert von $\Delta c^2(\zeta)$ ist. Die Gewichtungsfaktoren und die Größe des Gebietes, aus dem der Wert sich berechnen läßt, wird von dem Filter $g_\theta(\zeta)$ bestimmt. Um die Anwendung des Modells zu vereinfachen, werden nachfolgend mehrere Näherungen und Vereinfachungen der Gl. (5.1) eingefügt.

Bei einer Teilbandcodierung werden die Störsignale durch die Quantisierungsfehler der einzelnen Teilbandsignale hervorgerufen. Da sich im Störsignalzweig des w -Modells die Signalverarbeitung vom Monitor bis zur Fovea im Auge mit einem linearen System

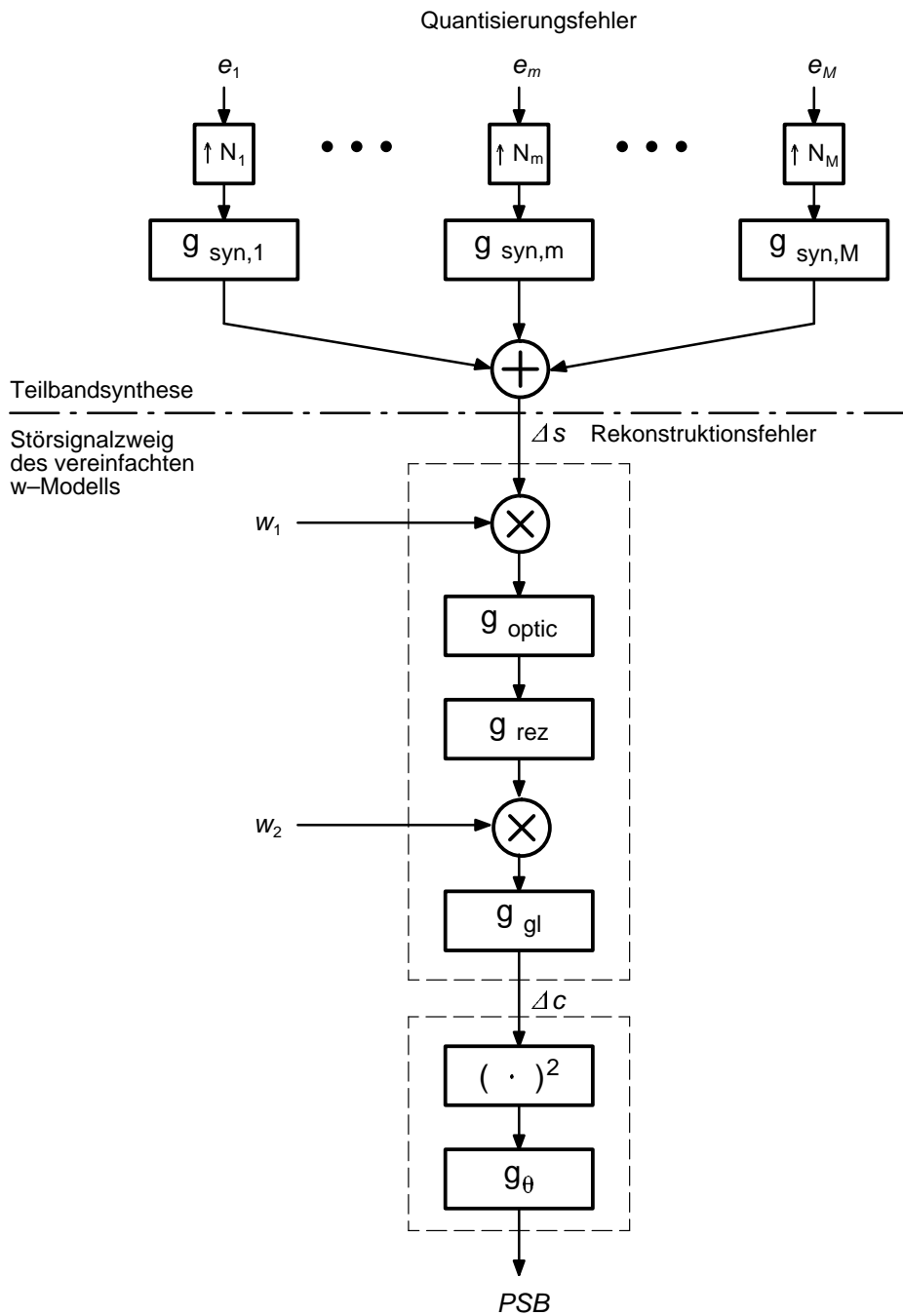


Bild 5.3 (a) Blockdiagramm zur Analyse der Potentiellen Sichtbarkeitsfunktion der Rekonstruktionsfehler bei der Teilbandcodierung

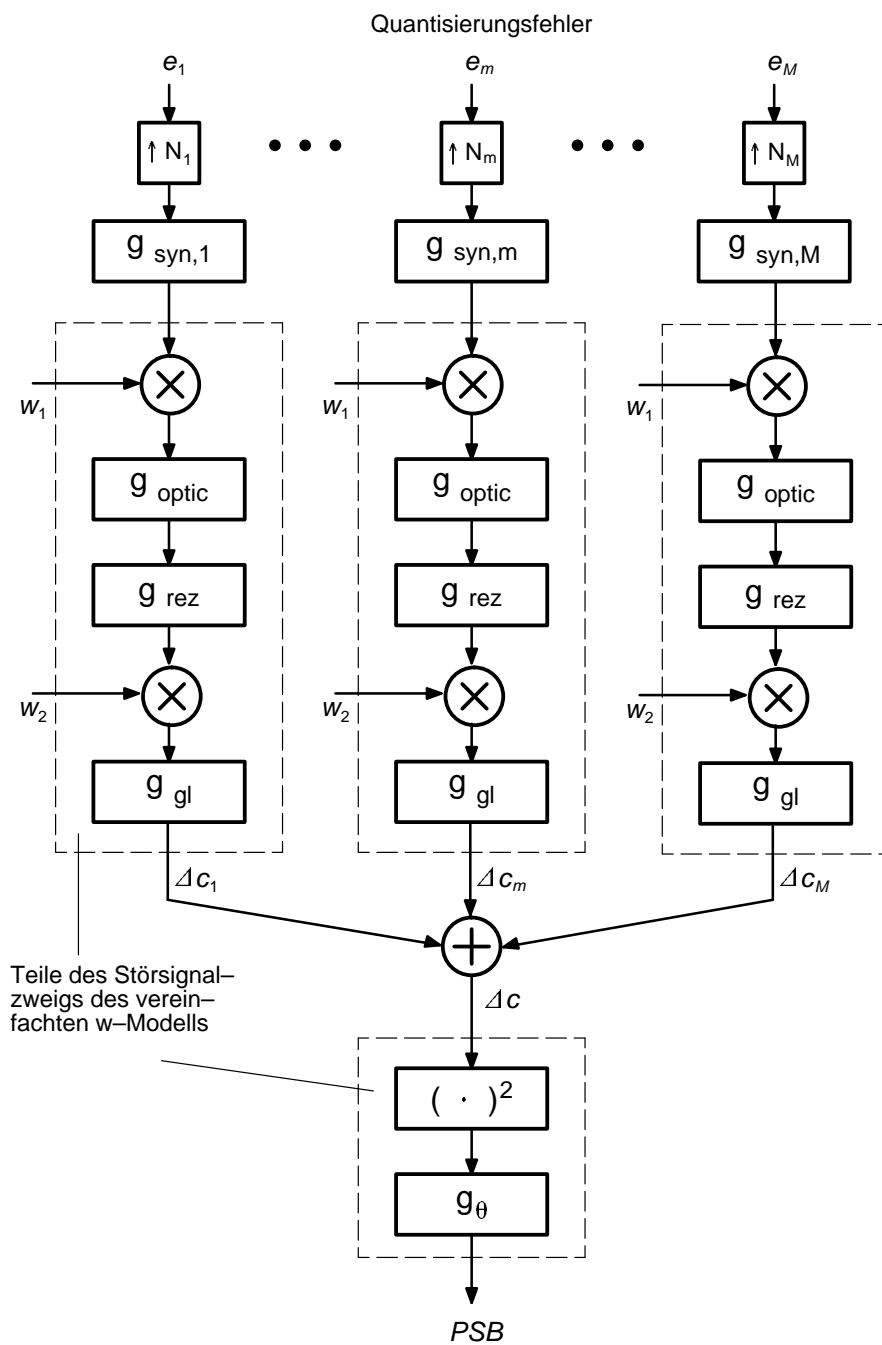


Bild 5.3 (b) Das funktional äquivalente Blockdiagramm zu Bild 5.3(a)

beschreiben läßt, kann statt Bild 5.3(a) Bild 5.3(b) zur Berechnung von Δc verwendet werden. Damit kann Δc wie in Gl. (5.2) durch dessen einzelne Bestandteile Δc_m beschrieben werden, wobei Δc_m von Quantisierungsfehlern des Teilbandes m hervorgerufen wird. M ist die Anzahl der Teilbänder.

$$\Delta c(\xi) = \sum_{m=1}^M \Delta c_m(\xi) \quad (5.2)$$

Daraus resultiert

$$\begin{aligned} \Delta c^2(\xi) &= \left[\sum_{m=1}^M \Delta c_m(\xi) \right]^2 \\ &= \sum_{m=1}^M \Delta c_m^2(\xi) + 2 \sum_{i=1}^M \sum_{j=1}^M \Delta c_i(\xi) \Delta c_j(\xi), \quad i \neq j. \end{aligned} \quad (5.3)$$

Da die Quantisierungen in den einzelnen Teilbändern unabhängig voneinander durchgeführt werden, kann angenommen werden, daß die Quantisierungsfehler e_m sowie die Signale Δc_m verschiedener Teilbandsignale unkorreliert und mittelwertfrei sind, so das

$$\sum_{i=1}^M \sum_{j=1}^M \Delta c_i(\xi) \Delta c_j(\xi) \approx 0, \quad i \neq j, \quad (5.4)$$

gilt, wenn M groß ist.

Unter Anwendung von Gl. (5.3) und (5.4) läßt sich die Potentielle Sichtbarkeitsfunktion der Teilbandquantisierungsfehler annähernd mit Gl. (5.5) berechnen [17]:

$$PSB(\xi) \approx \sum_{m=1}^M \Delta c_m^2(\xi) * g_\theta(\xi). \quad (5.5)$$

Zur Vereinfachung wird das Filter $g_\theta(\xi)$ in dieser Arbeit durch ein Mittelwertfilter

$$\hat{g}_\theta(\xi) = \begin{cases} 1/N_{pb} & \xi \in G_{pb} \\ 0 & \text{sonst} \end{cases} \quad (5.6)$$

ersetzt. G_{pb} umfaßt 3x3 Bildpunkte und $N_{pb} = 9$. Nach dieser Näherung ergibt sich

$$\begin{aligned}
 PSB(\xi) &\approx \sum_{m=1}^M [\Delta c_m^2(\xi) * \hat{g}_\theta(\xi)] \\
 &\approx \sum_{m=1}^M \overline{\Delta c_m^2(\xi)}
 \end{aligned} \tag{5.7}$$

wobei $\overline{\Delta c_m^2(\xi)}$ der lokale Mittelwert von $\Delta c_m^2(\xi)$ ist.

Nach Girod [17] kann jeder Summand der Gl. (5.7) dargestellt werden durch

$$\overline{\Delta c_m^2(\xi)} = \frac{1}{(2\pi)^3} \int_{B_m} S_{\Delta c_m}(\Omega) d\Omega, \tag{5.8}$$

wobei $S_{\Delta c_m}(\Omega)$ das Leistungsdichtespektrum von Δc_m ist. $\Omega = (\omega_x, \omega_y, \omega_t)$ ist der örtlich-zeitliche Kreisfrequenzvektor im Bildschirmkoordinatensystem. B_m ist die Bandbreite des Signals Δc_m .

$S_{\Delta c_m}(\Omega)$ ergibt sich aus dem Leistungsdichtespektrum $S_{e_m}(\Omega)$ des Quantisierungsfehlersignals e_m und der Leistungsübertragungsfunktion $S_{h_m}(\Omega)$ des Teilsystems h_m zwischen e_m und Δc_m nach Bild 5.3(b):

$$S_{\Delta c_m}(\Omega) = S_{e_m}(\Omega) \cdot S_{h_m}(\Omega). \tag{5.9}$$

Da die Bildfrequenz bei sehr niedrigen Übertragungsbiraten stark reduziert wird, kann die zeitliche Auswirkung der Quantisierungsfehler auf die nachfolgenden Bilder vernachlässigt werden. Dementsprechend läßt sich die Potentielle Sichtbarkeitsfunktion eines Bildes allein aus den Quantisierungsfehlern eines Bildes berechnen. Unter der Annahme, daß jedes Quantisierungsfehlersignal ein konstantes Leistungsdichtespektrum mit der Varianz $\sigma_{q,m}^2$ aufweist, läßt sich das Leistungsdichtespektrum $S_{e_m}(\Omega)$ des Quantisierungsfehlersignals e_m im Teilband m gemäß

$$S_{e_m}(\Omega) = \frac{(2\pi)^3}{B_m} \sigma_{q,m}^2 \tag{5.10}$$

darstellen.

Um den Zusammenhang zwischen den Quantisierungsfehlern und der Potentiellen Sichtbarkeitsfunktion auf einfache Weise darstellen zu können, wird zunächst ein Gebiet mit konstanten w -Parametern betrachtet. Die Leistungsübertragungsfunktion $S_{h_m}(\Omega)$ ergibt sich damit zu

$$S_{h_m}(\Omega) = w_1^2 w_2^2 \frac{N_m}{N} |G_{syn,m}(\Omega)|^2 \cdot |G_{opt}(\Omega)|^2 \cdot |G_{rez}(\Omega)|^2 \cdot |G_{gl}(\Omega)|^2, \quad (5.11)$$

wobei:

- w_1 und w_2 die w-Parameter sind
- $G_{syn,m}$ die Übertragungsfunktion des Synthesefilters für das Teilband m ist
- G_{opt} , G_{rez} und G_{gl} die Übertragungsfunktionen der entsprechenden Filter im w-Modell sind
- N_m/N der Decimation-Faktor des Teilbandes m ist.

Bei der Berechnung von $S_{h_m}(\Omega)$ geht die Interpolation des Teilbandsignals entsprechend seinem Decimation-Faktor N_m/N ein.

Durch Einsetzen von Gl. (5.9), (5.10) und (5.11) in Gl. (5.8) erhält man

$$\overline{\Delta c_m^2(\xi)} = \frac{\sigma_{q,m}^2}{B_m} \int_{B_m} S_{h_m}(\Omega) d\Omega = W_{SB,m}(\xi) \cdot \sigma_{q,m}^2, \quad (5.12)$$

wobei

$$W_{SB,m} = w_1^2 w_2^2 \cdot w_{SB,m} \quad (5.13)$$

und

$$w_{SB,m} = \frac{1}{B_m} \frac{N_m}{N} \int_{B_m} |G_{syn,m}(\Omega)|^2 \cdot |G_{opt}(\Omega)|^2 \cdot |G_{rez}(\Omega)|^2 \cdot |G_{gl}(\Omega)|^2 d\Omega \quad (5.14)$$

gelten.

Aus Gl. (5.7) und (5.12) ergibt sich der gesuchte Zusammenhang zwischen den Quantisierungsfehlervarianzen $\sigma_{q,m}^2$ der Teilbandsignale und dem resultierenden Wert der PSB-Funktion:

$$PSB(\xi) \approx \sum_{m=1}^M [W_{SB,m}(\xi) \cdot \sigma_{q,m}^2]. \quad (5.15)$$

Gleichung (5.15) zeigt, daß sich in diesem Fall die Potentielle Sichtbarkeitsfunktion als eine gewichtete Summe der Quantisierungsfehlervarianzen aller Teilbandsignale be-

rechnen läßt. Die Gewichtungsfaktoren $W_{SB,m}$ setzen sich, wie in Gl. (5.13) dargestellt, aus zwei Faktoren zusammen. Der erste Faktor $w_1^2 w_2^2$ ist ortsabhängig und erfaßt den lokalen Einfluß des Eingangssignals auf die PSB-Funktion. Der zweite Faktor $w_{SB,m}$ ist frequenzabhängig und beinhaltet die Wirkung des m -ten Teilbandsynthesefilters sowie der Filter des w -Modells.

0.34	0.13
0.12	0.044

(a) 4 Teilbänder

0.45	0.34	0.21	0.094
0.33	0.25	0.15	0.069
0.19	0.14	0.086	0.039
0.077	0.058	0.035	0.016

(b) 16 Teilbänder

0.48	0.45	0.39	0.32	0.25	0.18	0.12	0.077
0.45	0.42	0.36	0.30	0.23	0.17	0.11	0.072
0.38	0.36	0.31	0.25	0.20	0.14	0.093	0.061
0.30	0.28	0.25	0.20	0.16	0.11	0.074	0.049
0.23	0.21	0.19	0.15	0.12	0.085	0.056	0.037
0.16	0.15	0.13	0.10	0.082	0.059	0.039	0.025
0.098	0.091	0.079	0.065	0.050	0.036	0.024	0.016
0.061	0.057	0.049	0.040	0.031	0.023	0.015	0.010

(c) 64 Teilbänder

Tabelle 5.1 Ermittelte frequenzabhängige Gewichtungsfaktoren $w_{SB,m}$ für die gleichförmige Teilbandzerlegung mit 4, 16 und 64 Teilbändern. Die Gewichtungsfaktoren sind in Zeilen und Spalten in der Reihenfolge zunehmender mittlerer Ortsfrequenz der entsprechenden Teilbänder angeordnet.

Das Ergebnis in Gl. (5.15), das für Gebiete mit konstanten w -Parametern gilt, kann zur Schätzung der PSB-Funktion eines Rekonstruktionsbildes verwendet werden, indem man das Bild in Gebiete mit annähernd konstanten w -Parametern aufteilt und Gl. (5.15) in jedem Gebiet getrennt auswertet. Bei dieser Vorgehensweise können lokale Schätzfehler an den Rändern dieser Gebiete auftreten.

In Tabelle 5.1 sind die ermittelten frequenzabhängigen Gewichtungsfaktoren $w_{SB,m}$ für die gleichförmige Teilbandzerlegung mit 4, 16 und 64 Teilbändern angegeben. Diese Daten basieren auf der in Kapitel 4.2 ausgewählten Filterbank.

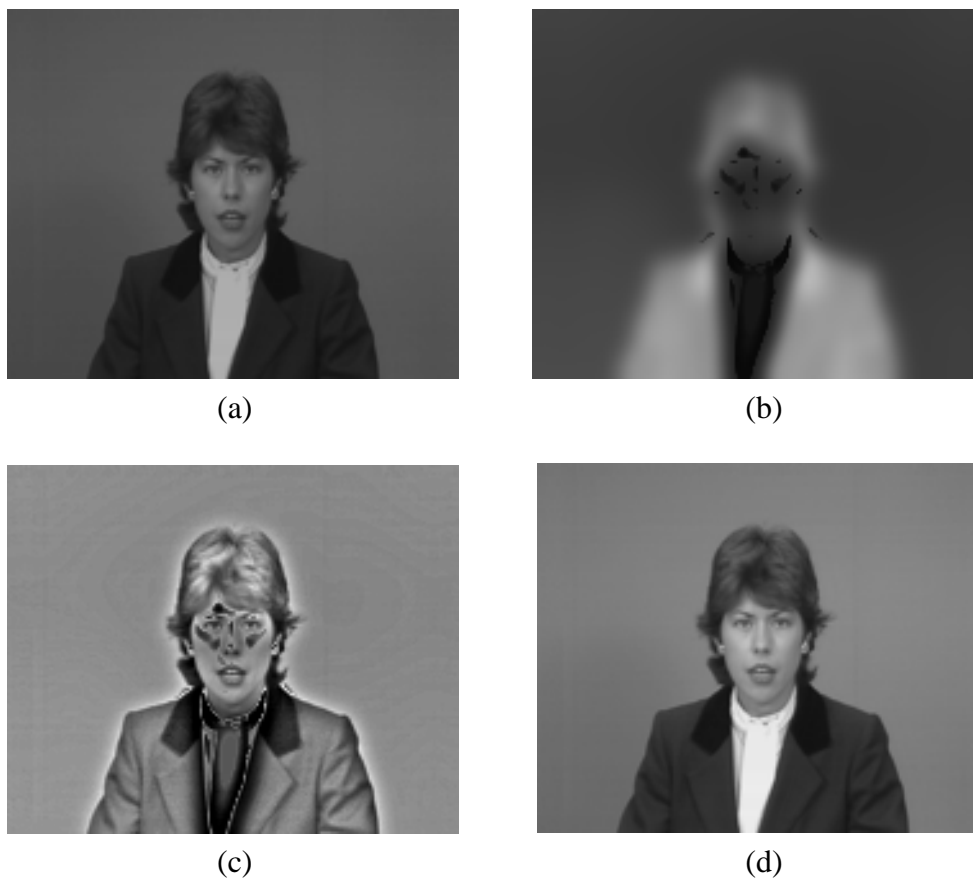


Bild 5.4 (a) w_1 , (b) w_2 und (c) $w_1 w_2$ für ein typisches Videotelefon-Bild (d). In dieser Darstellung sind die größeren Werte hell und die kleineren Werte dunkel wiedergegeben.

Die Tabellen zeigen, daß der Wert des Gewichtungsfaktors sich mit der wachsenden Ortsfrequenz verkleinert. Bei gleicher Quantisierungsfehlerleistung in den Teilbändern sind demzufolge die PSB-Anteile, die jeweils aus den Quantisierungsfehlern eines Teilbandsignals resultieren, unterschiedlich groß.

In Bild 5.4 sind als Beispiel die Werte der w -Parameter w_1 und w_2 sowie das zugehörige Produkt $w_1 w_2$ aus einem typischen Videotelefon-Bild dargestellt. Gleichzeitig wird auch das Originalbild gezeigt. Das Bild 5.4(c) veranschaulicht die lokale Veränderung und die örtliche Abhängigkeit des Produktes $w_1 w_2$ von dem Bildinhalt. Helle Gebiete in diesem Bild besitzen große Werte von $w_1 w_2$ und demzufolge eine höhere Wahrnehmungsempfindlichkeit für Quantisierungsfehler. Wie im nächsten Abschnitt dargestellt wird, müssen in diesen Gebieten die Teilbandsignale feiner quantisiert werden.

5.3 Optimale Aufteilung der Datenrate in den Teilbändern

Im letzten Abschnitt wurde gezeigt, wie aus den Quantisierungsfehlervarianzen einzelner Teilbandsignale der Wert der PSB-Funktion sich bestimmen läßt. In diesem Abschnitt wird analysiert, welche Datenrate den Teilbändern zugewiesen werden muß, so daß der resultierende PSB-Wert bei einer vorgegebenen gesamten Datenrate minimiert wird. Diese optimale Aufteilung der Datenrate ist die Grundlage der im nächsten Abschnitt zu beschreibenden adaptiven Quantisierung der Teilbandsignale.

Die Analyse wird zunächst für Bildgebiete mit jeweils konstanten w -Parametern in zwei Schritten durchgeführt. Im ersten Schritt wird die optimale Bitaufteilung ermittelt und daraus im zweiten Schritt die Quantisierungsfehlervarianzen der einzelnen Teilbandsignale bestimmt.

Optimale Bitaufteilung in Gebieten mit konstanten w -Parametern

Um eine direkte mathematische Beziehung zwischen den Datenraten R_m der Teilbandsignale y_m ($m = 1, \dots, M$) und dem Wert der PSB-Funktion herzustellen, wird hier der sogenannte Quantisierer-Performance-Faktor ϵ_m^{*2} [26] benutzt. ϵ_m^{*2} beschreibt für das Teilband m den Zusammenhang zwischen der Varianz der Quantisierungsfehler $\sigma_{q,m}^2$, der Datenrate R_m und der Varianz des zu quantisierenden Teilbandsignals $\sigma_{y,m}^2$ gemäß

$$\sigma_{q,m}^2 = \epsilon_m^{*2} 2^{-2R_m} \sigma_{y,m}^2. \quad (5.16)$$

Dabei ist R_m die Datenrate in der Einheit bit je übertragenen Abtastwert.

Im allgemeinen ist der Wert von ε_m^{*2} abhängig sowohl von den statistischen Eigenschaften des zu quantisierenden Signals y_m als auch von der verwendeten Quantisierung und Codierung.

Aus Gl. (5.15) und (5.16) ergibt sich der Zusammenhang zwischen den Datenraten der Teilbandsignale und dem entsprechenden Wert der PSB-Funktion:

$$PSB \approx \sum_{m=1}^M [W_{SB,m} \cdot \varepsilon_m^{*2} 2^{-2R_m} \sigma_{y,m}^2]. \quad (5.17)$$

Ausgehend von Gl. (5.17) kann die optimale Bitaufteilung, $R_1, \dots, R_m, \dots, R_M$ auf die Teilbandsignale bestimmt werden, in dem der Wert der PSB-Funktion in Gl. (5.17) unter der Nebenbedingung

$$\sum_{m=1}^M \frac{N_m}{N} R_m = R_{soll}, \quad (5.18)$$

minimiert wird. R_{soll} ist die vorgegebene Bitrate, bezogen auf einen übertragenen Bildpunkt. Diese Minimierung kann durch Anwendung des Lagrangeschen Verfahrens bestimmt werden. Sie ist in Anhang C beschrieben.

Zusammenfassend ergibt sich, daß bei der vorgegebenen Datenrate R_{soll} die optimale Bitaufteilung in Gebieten mit konstanten w -Parametern sich anhand von Gl. (C.16)

$$\underline{R}_m = R_{soll} + \frac{1}{2} \log_2 \frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{N_m/N \prod_{i=1}^M \left(\frac{W_{SB,i} \varepsilon_i^{*2} \sigma_{y,i}^2}{N_i/N} \right)^{N_i/N}}. \quad (5.19)$$

berechnen lassen. Diese Bitaufteilung führt zum Minimum der PSB-Funktion, nach Gl. (C.18):

$$PSB_{\min} = 2^{-2R_{soll}} \prod_{m=1}^M \left[\frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{\frac{N_m}{N}} \right]^{\frac{N_m}{N}}. \quad (5.20)$$

Bei einer gleichförmigen Bandaufteilung gilt $\frac{N_m}{N} = \frac{1}{M}$ für alle m , wobei M die Anzahl der Teilbänder ist. Aus Gl. (5.20) folgt in diesem Fall

$$PSB_{\min} = M \cdot 2^{-2R_{soll}} \prod_{m=1}^M [W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2]^{\frac{1}{M}}. \quad (5.21)$$

Optimale Verteilung der Quantisierungsfehler in einem Gebiet mit konstanten w -Parametern

Durch Einsetzen von Gl. (5.19) in Gl. (5.16) lassen sich die gesuchten optimalen Quantisierungsfehlervarianzen $\underline{\sigma}_{q,m}^2$ der einzelnen Teilbandsignale folgenderweise berechnen:

$$\begin{aligned}
 \underline{\sigma}_{q,m}^2 &= \varepsilon_m^{*2} 2^{-2R_m} \sigma_{y,m}^2 \\
 &= \varepsilon_m^{*2} \sigma_{y,m}^2 2^{-2 \left[R_{soll} + \frac{1}{2} \log_2 \frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{N_m/N \prod_{i=1}^M \left(\frac{W_{SB,i} \varepsilon_i^{*2} \sigma_{y,i}^2}{N_i/N} \right)^{N_i/N}} \right]} \\
 &= \varepsilon_m^{*2} \sigma_{y,m}^2 2^{-2R_{soll}} \frac{\frac{N_m}{N} \prod_{i=1}^M \left(\frac{W_{SB,i} \varepsilon_i^{*2} \sigma_{y,i}^2}{\frac{N_i}{N}} \right)^{\frac{N_i}{N}}}{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2} \\
 &= \frac{1}{W_{SB,m}} \frac{N_m}{N} 2^{-2R_{soll}} \prod_{i=1}^M \left(\frac{W_{SB,i} \varepsilon_i^{*2} \sigma_{y,i}^2}{\frac{N_i}{N}} \right)^{\frac{N_i}{N}}. \tag{5.22}
 \end{aligned}$$

Unter Berücksichtigung von Gl. (5.20) kann $\sigma_{q,m}^2$ schließlich mittels

$$\underline{\sigma}_{q,m}^2 = \frac{1}{W_{SB,m}} \frac{N_m}{N} PSB_{\min} \tag{5.23}$$

dargestellt werden.

Gleichung (5.23) zeigt, daß der aus den Quantisierungsfehlern im Teilband m resultierende PSB-Anteil $W_{SB,m} \underline{\sigma}_{q,m}^2$ (siehe Gl. (5.15)) proportional dem Decimation-Faktor N_m/N ist. Dieser Faktor beschreibt die Relation zwischen der Bandbreite eines Teilbandsignals und der Bandbreite des Originalsignals. Bei einer gleichförmigen Bandaufteilung sind die PSB-Anteile $W_{SB,m} \underline{\sigma}_{q,m}^2$ aus allen Teilbändern gleich groß, d.h.

$$\underline{\sigma}_{q,m}^2 = \frac{1}{W_{SB,m}} \frac{PSB_{\min}}{M}. \tag{5.24}$$

Gleichung (5.23) zeigt auch, daß die optimale Quantisierungsfehlervarianz eines Teilbandsignals umgekehrt proportional zum Gewichtungsfaktors $W_{SB,m}$ ist. Da

$W_{SB,m} = w_1^2 w_2^2 \cdot w_{SB,m}$ und der Wert $w_{SB,m}$ für ein tieffrequentes Teilband wesentlich größer als der für ein hochfrequentes Teilband ist, muß ein niederfrequentes Teilband feiner als ein hochfrequentes Teilband quantisiert werden.

Bis jetzt wurde die optimale Verteilung der Quantisierungsfehler in Bildgebieten mit konstanten w -Parametern behandelt. Im folgenden wird die optimale Verteilung der Quantisierungsfehler im gesamten Bild betrachtet.

Optimale Verteilung der Quantisierungsfehler im gesamten Bild

Der Ansatz zu dieser optimalen Verteilung ist, innerhalb eines Bildes überall den gleichen PSB-Wert zu erreichen, so daß die Bildqualität über dem gesamten Bild konstant bleibt. Anderenfalls könnte der PSB-Wert an einer Stelle größer als dieser konstante Wert sein und dort eine schlechtere Bildqualität erzeugen.

Zur Bestimmung dieses konstanten PSB-Wertes bei einer vorgegebenen Datenrate R_{soll} wird der zu codierende Bildbereich in K Gebiete mit jeweils konstanten w -Parametern unterteilt. In den einzelnen Gebieten gelten bei einer gleichförmigen Bandaufteilung mit M Teilbändern

$$PSB_{\min} = M \cdot 2^{-2R_k} \prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}} \quad (5.25)$$

und entsprechend

$$R_k = -\frac{1}{2} \log_2 \frac{PSB_{\min}}{M} + \frac{1}{2} \log_2 \prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}}. \quad (5.26)$$

Dabei sind R_k und $\left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k$ auf das Gebiet k bezogen.

Aus den Datenraten R_k ($k = 1, \dots, K$) in den einzelnen Gebieten berechnet sich die mittlere Datenrate für das gesamte Bild zu:

$$R = \sum_{k=1}^K \frac{P_k}{P} R_k, \quad (5.27)$$

wobei P_k ($k = 1, \dots, K$) die Anzahl der Bildpunkte in den einzelnen Gebieten und P die Anzahl der gesamten Bildpunkte sind.

Durch Einsetzen von Gl. (5.26) in Gl. (5.27) läßt sich die Datenrate R anhand PSB_{\min} darstellen:

$$R = \sum_{k=1}^K \frac{P_k}{P} \left\{ -\frac{1}{2} \log_2 \frac{PSB_{\min}}{M} + \frac{1}{2} \log_2 \prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}} \right\} \quad (5.28)$$

$$R = -\frac{1}{2} \log_2 \frac{\prod_{k=1}^K \left[\frac{PSB_{\min}}{M} \right]^{\frac{P_k}{P}}}{\prod_{k=1}^K \left[\prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}} \right]^{\frac{P_k}{P}}} \quad (5.29)$$

$$= -\frac{1}{2} \log_2 \frac{\frac{PSB_{\min}}{M}}{\prod_{k=1}^K \left[\prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}} \right]^{\frac{P_k}{P}}}$$

Aus (5.29) folgt die Beziehung zwischen dem Wert PSB_{\min} und der Datenrate R :

$$\begin{aligned} PSB_{\min} &= M \cdot 2^{-2R} \prod_{k=1}^K \left[\prod_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{1}{M}} \right]^{\frac{P_k}{P}} \\ &= M \cdot 2^{-2R} \left[\prod_{k=1}^K [w_1^2 w_2^2]_k^{\frac{P_k}{P}} \right] \cdot \left[\prod_{m=1}^M w_{SB,m}^{\frac{1}{M}} \right] \cdot \\ &\quad \prod_{m=1}^M \left[\prod_{k=1}^K \left[\varepsilon_m^{*2} \sigma_{y,m}^2 \right]_k^{\frac{P_k}{P}} \right]^{\frac{1}{M}}. \end{aligned} \quad (5.30)$$

Unter Anwendung der Ungleichung

$$\prod_{k=1}^K \left[\sigma_{y,m}^2 \right]_k^{\frac{P_k}{P}} \leq \sum_{k=1}^K \frac{P_k}{P} \left[\sigma_{y,m}^2 \right]_k = \sigma_{y,m}^2 \quad (5.31)$$

und der Näherung $[\varepsilon_m^{*2}]_k \approx \varepsilon_m^{*2}$ ergibt sich die obere Grenze für PSB_{\min} in Gl (5.30) gemäß

$$PSB_{\min}^{\hat{}} = M \cdot 2^{-2R} \left[\prod_{k=1}^K [w_1^2 w_2^2]_k^{\frac{p_k}{P}} \right] \cdot \left[\prod_{m=1}^M w_{SB,m}^{\frac{1}{M}} \right] \cdot \left[\prod_{m=1}^M [\varepsilon_m^{*2} \sigma_{y,m}^2]^{\frac{1}{M}} \right] . \quad (5.32)$$

Zur Vereinfachung der Berechnung wird jeder zu codierende Block als ein Gebiet mit annähernd konstanten w -Parametern betrachtet. Dann kann Gl. (5.32) in der Form

$$PSB_{\min}^{\hat{}} = M \cdot 2^{-2R} \left[\prod_{j_b=1}^{J_b} [\bar{w}_1(j_b) \bar{w}_2(j_b)]^{j_b} \right] \cdot \left[\prod_{m=1}^M w_{SB,m}^{\frac{1}{M}} \right] \cdot \left[\prod_{m=1}^M [\varepsilon_m^{*2} \sigma_{y,m}^2]^{\frac{1}{M}} \right] \quad (5.33)$$

geschrieben werden, wobei j_b der Index, und J_b die Anzahl der zu codierenden Blöcke ist. $\bar{w}_1(j_b)$ und $\bar{w}_2(j_b)$ sind die innerhalb eines zu codierenden Blocks j_b gebildeten Mittelwerte von w_1 und w_2 . Je nachdem ob 16 oder 64 Teilbänder verwendet werden, enthält ein Block 4x4 oder 8x8 Bildpunkten.

Um überall den gleichen Wert PSB_{\min} zu erzielen, sollen in den einzelnen Gebieten die Quantisierungsfehlervarianzen der Teilbandsignale, wie in Gl. (5.24) gezeigt, gemäß

$$[\sigma_{q,m}^2]_k = \frac{1}{[w_1^2 w_2^2]_k \cdot w_{SB,m}} \frac{PSB_{\min}}{M} \quad (5.34)$$

gesteuert werden. Diese Quantisierungsfehlervarianzen sind nicht nur von $w_{SB,m}$ frequenz– sondern auch von $[w_1^2 w_2^2]_k$ ortsabhängig. Wird Gl. (5.34) auf die einzelne zu codierende Blöcke angewendet, so ergeben sich die optimalen Quantisierungsfehlervarianzen $\sigma_{q,m}^2(j_b)$ für einzelne Teilbandabtastwerte:

$$\sigma_{q,m}^2(j_b) = \frac{1}{[\bar{w}_1^2(j_b) \bar{w}_2^2(j_b)] \cdot w_{SB,m}} \frac{PSB_{\min}^{\hat{}}}{M} , \quad j_b = 1, \dots, J_b . \quad (5.35)$$

Vergleich mit einer Quantisierung unter Anwendung des mittleren quadratischen Rekonstruktionsfehlers als Gütekriterium

In [26] wird eine Quantisierung der Teilbandsignale beschrieben, in der die Varianz der Rekonstruktionsfehler σ_{re}^2 , verursacht durch Quantisierungsfehler, bei einer vorgegebenen Datenrate R minimiert wird. Das erzielte Minimum $[\sigma_{re}^2]_{\min}$ und die entsprechende

Varianz der Quantisierungsfehler in den Teilbändern $[\sigma_q^2]_{obj}$ lassen sich bei einer gleichförmigen Teilbandzerlegung mit M Teilbändern durch

$$[\sigma_{re}^2]_{\min} = M \cdot 2^{-2R} \cdot \prod_{m=1}^M \left[\varepsilon_m^{*2} \sigma_{y,m}^2 \right]^{\frac{1}{M}} \quad (5.36)$$

und

$$[\sigma_{q,m}^2]_{obj} = [\sigma_q^2]_{obj} = \frac{[\sigma_{re}^2]_{\min}}{M} \quad (5.37)$$

darstellen. In diesem Fall ist die Varianz der Quantisierungsfehler $[\sigma_q^2]_{obj}$ in allen Teilbändern gleich groß.

Ein Vergleich der Gl (5.36) mit Gl. (5.33) zeigt den Zusammenhang zwischen \hat{PSB}_{\min} und $[\sigma_{re}^2]_{\min}$

$$\hat{PSB}_{\min} = \left[\prod_{j_b=1}^{J_b} [\bar{w}_1(j_b) \bar{w}_2(j_b)]^{j_b} \right]^{\frac{2}{J_b}} \cdot \left[\prod_{m=1}^M w_{SB,m}^{\frac{1}{M}} \right] \cdot [\sigma_{re}^2]_{\min} \quad (5.38)$$

Aus Gl. (5.34), (5.37) und (5.38) folgt für die Quantisierungsfehlervarianzen $[\sigma_{q,m}^2]_k$ und $[\sigma_q^2]_{obj}$ im Bildgebiet k die Beziehung:

$$[\sigma_{q,m}^2]_k = \frac{\prod_{j_b=1}^{J_b} [\bar{w}_1(j_b) \bar{w}_2(j_b)]^{j_b}}{[w_1^2 w_2^2]_k} \cdot \frac{\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}}}{w_{SB,m}} \cdot [\sigma_q^2]_{obj} \quad (5.39)$$

Zur Veranschaulichung der Ortsfrequenzabhängigkeit des Verhältnisses $[\sigma_{q,m}^2]_k / [\sigma_q^2]_{obj}$ wird Gl. (5.39) bei $[w_1^2 w_2^2]_k = \prod_{j_b=1}^{J_b} [\bar{w}_1(j_b) \bar{w}_2(j_b)]^{j_b}$ betrachtet. In diesem Fall gilt mit Gl. (5.37) und (5.39)

$$[\sigma_{q,m}^2]_k = \frac{\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}}}{w_{SB,m}} \cdot [\sigma_q^2]_{obj}, \quad (5.40)$$

und

$$[\sigma_{re}^2]_k = \sum_{m=1}^M [\sigma_{q,m}^2]_k = \frac{\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}} \cdot \sum_{m=1}^M \frac{1}{w_{SB,m}}}{M} \cdot [\sigma_{re}^2]_{\min}, \quad (5.41)$$

wobei $[\sigma_{re}^2]_k$ die Varianz der Rekonstruktionsfehler ist, die sich bei der subjektiv optimalen Quantisierung der Teilbandsignale im Bildgebiet k ergibt.

Um einen quantitativen Eindruck zu vermitteln, werden die Gl. (5.40) und (5.41) anhand der in Tabelle 5.1 angegebenen Gewichtungsfaktoren $w_{SB,m}$ für die Teilbandcodierung mit 4 Teilbändern ausgewertet:

$$\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}} = 0,12, \quad (5.42)$$

$$\frac{[\sigma_{q,m}^2]_k}{[\sigma_{q}^2]_{obj}} = \frac{\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}}}{w_{SB,m}} = \begin{bmatrix} 0,37 & 0,95 \\ 1,03 & 2,82 \end{bmatrix}, \quad (5.43)$$

$$\frac{[\sigma_{re}^2]_k}{[\sigma_{re}^2]_{\min}} = \frac{\prod_{i=1}^M w_{SB,i}^{\frac{1}{M}} \cdot \sum_{m=1}^M \frac{1}{w_{SB,m}}}{M} = 1,29. \quad (5.44)$$

Gl. (5.43) zeigt, daß in Folge der Quantisierung mit dem psychovisuellen Modell die Varianz der Quantisierungsfehler im Teilband mit höchsten Ortsfrequenzen $[\sigma_{q,4}^2]_k$ wesentlich größer als $[\sigma_{q}^2]_{obj}$ ist und demzufolge weniger Datenrate zur Codierung dieses Teilbandes benötigt wird. Von dieser Datenratensparnis profitiert das Teilband mit tiefsten Ortsfrequenzen, so daß $[\sigma_{q,1}^2]_k$ deutlich kleiner als $[\sigma_{q}^2]_{obj}$ eingestellt werden kann.

Bei der subjektiv optimalen Quantisierung der Teilbandsignale ist die Varianz der Rekonstruktionsfehler $[\sigma_{re}^2]_k$ erwartungsgemäß größer als das Minimum $[\sigma_{re}^2]_{\min}$, das durch die objektiv optimale Quantisierung erzielbar ist.

5.4 Realisierung der adaptiven Quantisierung der Teilbandsignale

Um die Quantisierungsfehlervarianz gemäß Gl. (5.35) zu realisieren, werden die Teilbandsignale gleichförmig quantisiert. Die Wahl der gleichförmigen Quantisierung ist

nicht nur durch ihre einfache Realisierbarkeit sondern auch durch ihre Effizienz begründet. In [26] wird gezeigt, daß bei einer gleichförmigen Quantisierung in Verbindung mit einer Entropiecodierung, die auch hier angewendet werden soll, die Datenrate minimal wird.

Bei einer feinen gleichförmigen Quantisierung des Teilbandsignals y_m weist das Quantisierungsfehlersignal eine Gleichverteilung auf. Deshalb besteht zwischen der Quantisierungsfehlervarianz $\sigma_{q,m}^2$ und der Quantisiererschrittweite $\underline{\Delta}_{q,m}^2$ die folgende Beziehung:

$$\sigma_{q,m}^2 = \frac{\underline{\Delta}_{q,m}^2}{12}. \quad (5.45)$$

Aus Gl. (5.35) und (5.45) berechnen sich die optimalen Quantisiererschrittweiten $\underline{\Delta}_{q,m}(j_b, n_p)$ für einen Abtastwert $y_m(j_b, n_p)$ in dem Teilband m des Bildes n_p zu

$$\underline{\Delta}_{q,m}(j_b, n_p) = 12 \cdot \underline{\sigma}_{q,m}(j_b, n_p) = \frac{12}{\bar{w}_1^2(j_b, n_p) \bar{w}_2^2(j_b, n_p) \cdot w_{SB,m}} \frac{\hat{PSB}_{\min}}{M}. \quad (5.46)$$

Dabei müssen die Mittelwerte $\bar{w}_1(j_b, n_p)$ und $\bar{w}_2(j_b, n_p)$ gemessen werden.

Aus Gl. (5.46) erhält man für die Quantisiererschrittweiten der Teilbandabtastwerte des Blocks j_b des Bildes n_p die folgende Beziehung

$$\underline{\Delta}_{q,m}(j_b, n_p) = \left[\frac{w_{SB,1}}{w_{SB,m}} \right]^{\frac{1}{2}} \cdot \underline{\Delta}_{q,1}(j_b, n_p), \quad (5.47)$$

wobei die Gewichtungsfaktoren $w_{SB,1}$ und $w_{SB,m}$ feste Größe sind, die der Tab. 5.1 entnommen werden können. Die Bezugsgröße $\underline{\Delta}_{q,1}(j_b, n_p)$ ist die Quantisiererschrittweite des Teilbandes $m = 1$ mit den tiefsten Ortsfrequenzen.

Für die Quantisierungsschrittweiten der Teilbandabtastwerte zweier Blöcke i_b und j_b des Bildes n_p gilt mit Gl. (5.46)

$$\underline{\Delta}_{q,m}(j_b, n_p) = \frac{\bar{w}_1(i_b, n_p) \bar{w}_2(i_b, n_p)}{\bar{w}_1(j_b, n_p) \bar{w}_2(j_b, n_p)} \cdot \underline{\Delta}_{q,m}(i_b, n_p). \quad (5.48)$$

In Fall zweier aufeinanderfolgenden Bilder n_p und $n_p - 1$ gilt entsprechend

$$\underline{\Delta}_{q,m}(j_b, n_p) = \underline{\Delta}_{q,m}(i_b, n_p - 1) \cdot \frac{\bar{w}_1(i_b, n_p - 1) \bar{w}_2(i_b, n_p - 1)}{\bar{w}_1(j_b, n_p) \bar{w}_2(j_b, n_p)} \cdot \left[\frac{\hat{PSB}_{\min}(n_p)}{\hat{PSB}_{\min}(n_p - 1)} \right]^{\frac{1}{2}}. \quad (5.49)$$

Gl. (5.47) und (5.48) zeigen, daß aus nur einer Quantisiererschrittweite $\underline{\Delta}_{q,1}(1, n_b)$ für den ersten Block ($j_b = 1$) und das erste Teilband ($m = 1$) die Quantisiererschrittweiten für alle Teilbandabtastwerte im Bild n_p gemäß

$$\underline{\Delta}_{q,1}(j_b, n_p) = \frac{\bar{w}_1(1, n_p) \bar{w}_2(1, n_p)}{\bar{w}_1(j_b, n_p) \bar{w}_2(j_b, n_p)} \cdot \underline{\Delta}_{q,1}(1, n_p) \quad (j_b = 2, \dots, J_b) \quad (5.50)$$

und

$$\underline{\Delta}_{q,m}(j_b, n_p) = \left[\frac{w_{SB,1}}{w_{SB,m}} \right]^{\frac{1}{2}} \cdot \underline{\Delta}_{q,1}(j_b, n_p) \quad (j_b = 1, \dots, J_b, m = 2, \dots, M). \quad (5.51)$$

berechnet werden können. Die Beziehung (5.49) bietet eine einfache Möglichkeit zur rekursiven Bestimmung der Quantisiererschrittweiten $\underline{\Delta}_{q,1}(1, n_p)$ von Bild zu Bild

$$\underline{\Delta}_{q,1}(1, n_p) = \underline{\Delta}_{q,1}(1, n_p - 1) \cdot \frac{\bar{w}_1(1, n_p - 1) \bar{w}_2(1, n_p - 1)}{\bar{w}_1(1, n_p) \bar{w}_2(1, n_p)} \cdot \left[\frac{\hat{PSB}_{\min}(n_p)}{\hat{PSB}_{\min}(n_p - 1)} \right]^{\frac{1}{2}}. \quad (5.52)$$

Dazu wird der Quotient $\hat{PSB}_{\min}(n_p)/\hat{PSB}_{\min}(n_p - 1)$ unter Verwendung von Gl. (5.33) gemäß

$$\frac{\hat{PSB}_{\min}(n_p)}{\hat{PSB}_{\min}(n_p - 1)} \approx 2^{-2[R_{\text{sol}}(n_p) - R_{\text{ist}}(n_p - 1)]} \cdot \frac{\prod_{j_b=1}^{J_b} [\bar{w}_1(j_b, n_p) \bar{w}_2(j_b, n_p)]^{\frac{2}{J_b}}}{\prod_{i_b=1}^{I_b} [\bar{w}_1(i_b, n_p - 1) \bar{w}_2(i_b, n_p - 1)]^{\frac{2}{I_b}}} \cdot \prod_{m=1}^M \left[\frac{\sigma_{y,m}^2(n_p)}{\sigma_{y,m}^2(n_p - 1)} \right]^{\frac{1}{M}}. \quad (5.53)$$

berechnet, wobei i_b der Index, und I_b die Anzahl der zu codierenden Blöcke des Bildes $n_p - 1$ ist. $R_{ist}(n_p - 1)$ ist die tatsächliche Datenrate des Bildes $n_p - 1$ und $R_{soll}(n_p)$ die Soll-Datenrate des Bildes n_p . Es wird angenommen, daß zwischen zwei Bildern die Werte von ε_m^{*2} sich wenig verändert haben, so daß

$$\prod_{m=1}^M \left[\frac{\varepsilon_m^{*2}(n_p)}{\varepsilon_m^{*2}(n_p - 1)} \right]^{\frac{1}{M}} \approx 1 \quad (5.54)$$

gilt.

Zur Vereinfachung wird ein Hilfsparameter $HP(n_p)$ gemäß

$$HP(n_p) = \underline{\Delta}_{q,1}(1, n_p) \cdot \bar{w}_1(1, n_p) \cdot \bar{w}_2(1, n_p) \quad (5.55)$$

eingeführt. Für das Bild $n_p - 1$ gilt entsprechend

$$HP(n_p - 1) = \underline{\Delta}_{q,1}(1, n_p - 1) \cdot \bar{w}_1(1, n_p - 1) \cdot \bar{w}_2(1, n_p - 1) \quad (5.56)$$

Aus Gl. (5.50) und (5.55) ergibt sich eine einfache Formel zur Bestimmung der optimalen Quantisiererschrittweiten für die einzelnen Abtastwerte im Teilband $m = 1$ des Bildes n_p :

$$\underline{\Delta}_{q,1}(j_b, n_p) = \frac{HP(n_p)}{\bar{w}_1(j_b, n_p) \bar{w}_2(j_b, n_p)} \quad (j_b = 1, \dots, J_b) . \quad (5.57)$$

Durch Einsetzen von $HP(n_p)$ und $HP(n_p - 1)$ gemäß Gl. (5.55) bzw. (5.56) in Gl. (5.52) erhält man

$$HP(n_p) = HP(n_p - 1) \cdot \left[\frac{\hat{PSB}_{\min}(n_p)}{\hat{PSB}_{\min}(n_p - 1)} \right]^{\frac{1}{2}} . \quad (5.58)$$

Ausgehend von dem Anfangswert

$$HP(1) = \underline{\Delta}_{q,1}(1, 1) \cdot \bar{w}_1(1, 1) \bar{w}_2(1, 1) \quad (5.59)$$

für $j_b = 1$ und $n_p = 1$ lassen sich die Werte $HP(n_p)$ für Bild $n_p = 2, 3, \dots$ und mit Gl. (5.57) die Quantisiererschrittweiten $\underline{\Delta}_{q,1}(j_b, n_p)$ des ersten Teilbandes bestimmen. Für die übrigen Teilbänder mit $m > 1$ lassen sich die Quantisiererschrittweiten $\underline{\Delta}_{q,m}(j_b, n_p)$ schließlich gemäß Gl. (5.51) berechnen.

Die Quantisiererschrittweiten $\underline{\Delta}_{q,1}(j_b, n_p)$ ($j_b = 1, \dots, J_b$) für die Abtastwerte im ersten Teilband müssen dem Decoder in Form von Seiteninformation für jedes Bild n_p mitgeteilt werden. Um diese Seiteninformation zu beschränken, wird jede errechnete Quantisiererschrittweite $\underline{\Delta}_{q,1}(j_b, n_p)$ aus Gl. (5.57) vorher auf einen der folgenden 31 Werte

$$\Delta_Q = \frac{Q}{\sqrt{M}}, \quad (Q=1, 2, \dots, 31) \quad (5.60)$$

gerundet, wobei M die Anzahl der Teilbänder ist. Als Seiteninformation wird der Quantisierungsparameter Q zum Decoder übertragen.

Eine weitere Möglichkeit der Einsparung von Seiteninformation besteht darin, nicht für jeden einzelnen, sondern nur für jeweils 4 bzw. 16 Abtastwerte des Teilbandes $m = 1$ einen Quantisierungsparameter Q zu übertragen. In diesem Fall wird zunächst aus den optimalen Quantisiererschrittweiten für diese Abtastwerte das Minimum

$$\Delta_{\min} = \min[\underline{\Delta}_{q,1}(j_b, n_p) \mid \text{aus 4 bzw. 16 Abtastwerten}] \quad (5.61)$$

ermittelt. Δ_{\min} bestimmt dann über Gl. (5.60) den zu übertragenden Quantisierungsparameter Q .

Aus Gl. (5.51) und (5.60) ergeben sich die Quantisiererschrittweiten der anderen Teilbänder gemäß

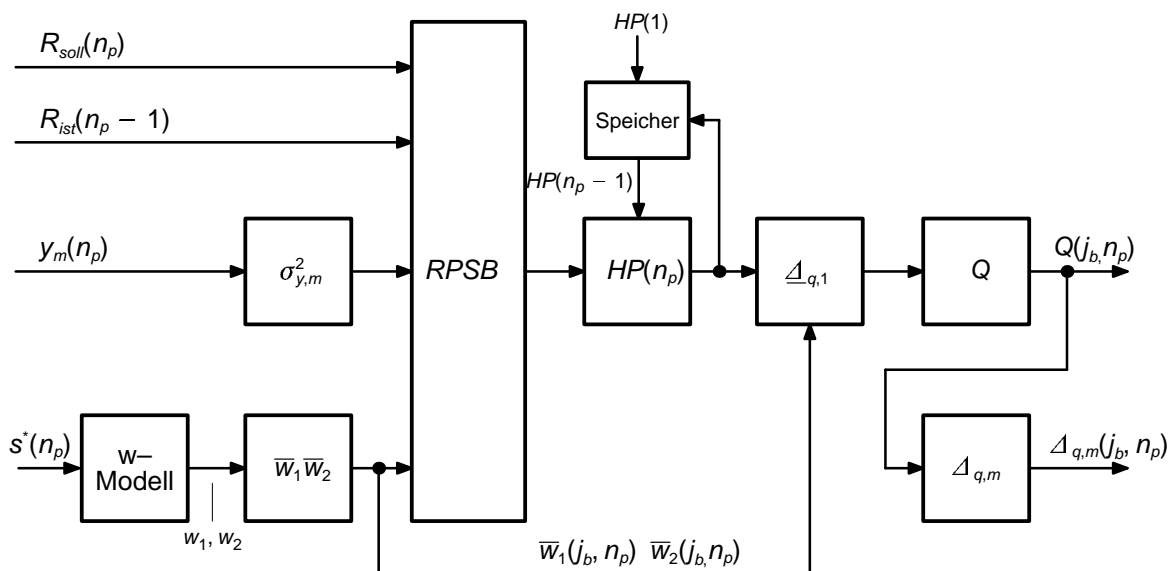
$$\Delta_{q,m}(j_b, n_p) = \left[\frac{w_{SB,1}}{w_{SB,m}} \right]^{\frac{1}{2}} \cdot \frac{Q(j_b, n_p)}{\sqrt{M}}. \quad (5.62)$$

Zusammenfassend ergibt sich folgender Ablauf zur Bestimmung der adaptiven Quantisiererschrittweiten für die einzelnen Teilbandabtastwerte:

1. Schätzen von $\hat{P}\hat{S}B_{\min}(n_p) / \hat{P}\hat{S}B_{\min}(n_p - 1)$ anhand Gl. (5.53),
2. Ermitteln des Hilfsparameters $HP(n_p)$ mit Gl. (5.58) und (5.59),
3. Berechnen der optimalen Quantisiererschrittweiten $\underline{\Delta}_{q,1}(j_b, n_p)$ für die einzelnen Abtastwerte des Teilbandes $m = 1$ gemäß Gl. (5.57),
4. Bestimmen der Quantisierungsparameter $Q(j_b, n_p)$ mittels (5.61) und (5.60),

5. Festlegen der zu verwendenden Quantisiererschrittweiten $\Delta_{q,m}(j_b, n_p)$ gemäß Gl. (5.62).

Dieser Ablauf wird in Bild 5.5 anhand eines Blockdiagramms veranschaulicht.



$$RPSB = \frac{\hat{PSB}_{\min}(n_p)}{\hat{PSB}_{\min}(n_p - 1)}$$

Bild 5.5 Bestimmung der adaptiven Quantisiererschrittweiten für die Teilbandsignale der Luminanzkomponente

In Tabelle 5.2 werden die Gewichtungsfaktoren $\left[\frac{w_{SB,1}}{w_{SB,m}} \right]^{\frac{1}{2}}$ in Gl. (5.62) für die gleichförmige Teilbandzerlegung mit 4, 16 und 64 Teilbändern aufgeführt. Diese Werte basieren auf den Daten in Tabelle 5.1.

1.0	1.6
1.7	2.8

(a) 4 Teilbänder

1.0	1.2	1.5	2.2
1.2	1.3	1.7	2.6
1.5	1.8	2.3	3.4
2.4	2.8	3.6	5.3

(b) 16 Teilbänder

1.0	1.0	1.1	1.2	1.4	1.6	2.0	2.5
1.0	1.1	1.2	1.3	1.4	1.7	2.1	2.6
1.1	1.2	1.2	1.4	1.5	1.9	2.3	2.8
1.3	1.3	1.4	1.5	1.7	2.1	2.5	3.1
1.4	1.5	1.6	1.8	2.0	2.4	2.9	3.6
1.7	1.8	1.9	2.2	2.4	2.9	3.5	4.4
2.2	2.3	2.5	2.7	3.1	3.7	4.5	5.4
2.8	2.9	3.1	3.5	3.9	4.6	5.7	6.9

(c) 64 Teilbänder

Tabelle 5.2 Gewichtungsfaktoren $\left[\frac{w_{SB,1}}{w_{SB,m}} \right]^{\frac{1}{2}}$ in Gl. (5.62) für die gleichförmige Teilbandzerlegung mit 4, 16 und 64 Teilbändern. Die Werte sind in Zeilen und Spalten in der Reihenfolge zunehmender mittleren Ortsfrequenz der entsprechenden Teilbänder angeordnet.

Die bisherigen Beschreibungen in diesem Kapitel handelten sich ausschließlich um die subjektiv optimale Quantisierung der Teilbandsignale der *Luminanzkomponente*, da das

dabei verwendete psychovisuelle Modell nur die menschliche visuelle Wahrnehmung des Störsignals in der Luminanzkomponente beschreibt. Für die *Chrominanzkomponenten* existieren noch keine vergleichbaren Modelle. Deshalb wird in diesem Coder die Quantisierung der Teilbandsignale der Chrominanzkomponenten durch Minimierung des mittleren quadratischen Rekonstruktionsfehlers [26] optimiert. Für die gleichförmige Quantisierung läßt sich die optimale Quantisierschrittweite $\underline{\Delta}_{q,C}(n_p)$ für das Bild n_p anhand

$$\underline{\Delta}_{q,C}(n_p) = \underline{\Delta}_{q,C}(n_p - 1) \cdot \left[\frac{\underline{\sigma}_{re,C}^2(n_p)}{\underline{\sigma}_{re,C}^2(n_p - 1)} \right]^{\frac{1}{2}} \quad (5.63)$$

rekursiv bestimmen, wobei

$$\underline{\sigma}_{re,C}^2 = M \cdot 2^{-2R_C} \cdot \prod_{m=1}^M \left[\varepsilon_{m_C}^{*2} \sigma_{y_{C,m}}^2 \right]^{\frac{1}{M}} \quad (5.64)$$

das Minimum der Varianz der Rekonstruktionsfehler bei der vorgegebenen Datenrate R_C für die Chrominanzkomponenten ist.

Der Faktor $\underline{\sigma}_{re,C}^2(n_p)/\underline{\sigma}_{re,C}^2(n_p - 1)$ in Gl. (5.63) wird aus Gl. (5.64) mit

$$\frac{\underline{\sigma}_{re,C}^2(n_p)}{\underline{\sigma}_{re,C}^2(n_p - 1)} = 2^{-2[R_{C, sol}(n_p) - R_{C, ist}(n_p - 1)]} \cdot \prod_{m=1}^M \left[\frac{\sigma_{y_{C,m}}^2(n_p)}{\sigma_{y_{C,m}}^2(n_p - 1)} \right]^{\frac{1}{M}} \quad (5.65)$$

berechnet. Aus $\underline{\Delta}_{q,C}(n_p)$ bestimmt über

$$\Delta_{Q_C} = \frac{Q_C}{\sqrt{M}}, \quad (Q_C = 1, 2, \dots, 31) \quad (5.66)$$

den Quantisierungsparameter Q_C für die Chrominanzkomponenten.

5.5 Codierung der Quantisierschrittweiten

Gl. (5.62) und (5.66) zeigen, daß die adaptiven Quantisierschrittweiten der einzelnen Teilbandabtwerte eindeutig aus $Q(j_b, n_p)$ und $Q_C(n_p)$ sich berechnen lassen. Um dem Decoder die verwendeten Quantisierschrittweiten mitzuteilen, werden diese Werte zum Decoder übertragen.

Der Quantisierungsparameter $Q(1, n_p)$ für den ersten Teilbandabstastwert der Luminanzkomponente sowie der Quantisierungsparameter $Q_C(n_p)$ für alle Teilbandabstastwerte der Chrominanzkomponenten im Bild n_p werden anhand zwei Codewörter PQUANTY und PQUANTC mit je 5 bit im Bild-Header übertragen.

Die weiteren Quantisierungsparameter $Q(j_b, n_p)$ ($j_b > 1$) für die Luminanzkomponente werden makroblockweise unter Anwendung einer Differenzcodierung übermittelt. Dazu werden die zu codierenden Makroblöcke in folgenden drei Typen aufgeteilt:

- Typ 1: hat keinen neuen Quantisierungsparameter. Der letzt verwendete Wert gilt für den aktuellen Makroblock weiter .
- Typ 2: hat einen neuen Quantisierungsparameter für alle 4 Blöcke mit 8x8 Bildpunkten im Makroblock.
- Typ 3: hat max. vier neue Quantisierungsparameter, jeweils für einen der 4 Blöcke im Makroblock.

Zur Angabe der Makroblocktyp wird für jeden Makroblock ein Codewort MCBPC verwendet. Für die Makroblöcke der Typ 3 wird ein weiteres Codewort QSBP mit 4 bit benutzt. Sie zeigen, welche Blöcke mit jeweils 8x8 Bildpunkten im Makroblock einen neuen Quantisierungsparameter erhalten sollen. Für jeden neuen Quantisierungsparameter wird dann die Differenz zwischen dem neuen und dem alten Wert übertragen.

In diesem Kapitel wurde die subjektiv optimale adaptive Quantisierung der Teilbandsignale unter Anwendung eines psychovisuellen Modells entwickelt. In Kapitel 6 wird der dadurch erzielte Gewinn experimentell ermittelt.

6. Experimentelle Untersuchungen

Für eine Bewertung des Konzepts des in dieser Arbeit beschriebenen bewegungskompensierenden Teilbandcoders wurden seine Algorithmen als Computerprogramme implementiert und mit Hilfe typischer Testsequenzen auf einem Computer simuliert. In diesem Kapitel werden die Ergebnisse aus diesen experimentellen Untersuchungen zusammenfassend dargestellt. Dazu werden in Abschnitt 6.1 zunächst die verwendeten Testsequenzen beschrieben. Anschließend werden in Abschnitt 6.2 die gemessenen Codierungsgewinne der einzelnen Komponenten angegeben, die in den letzten Kapiteln beschrieben wurden. In Abschnitt 6.3 wird dieser bewegungskompensierende Teilbandcoder mit dem bereits in Abschnitt 2.3 vorgestellten Hybrid-DCT-coder TMN5 anhand objektiver und subjektiver Gütekriterien verglichen.

6.1 Beschreibung der verwendeten Testsequenzen

Für die Untersuchungen in dieser Arbeit wurden die Testsequenzen *Miss America*, *Claire*, *Car Phone* und *Foreman* verwendet, die von internationalen Expertengruppen der "European COST 211bis Simulation Subgroup" und "ITU-T Study Group 15" zum Testen von Codern mit sehr niedrigen Übertragungsbitraten ausgewählt wurden [24]. Die Sequenz *Miss America* wurde von den British Telecom Research Laboratories (BTRL), Großbritannien, zur Verfügung gestellt, die Sequenz *Claire* stammt vom Centre de National D' Etudes Des Telecommunications (CNET), Frankreich. Die anderen zwei Sequenzen, *Car Phone* und *Foreman*, wurden von der Firma Siemens vorgeschlagen. In Bild 6.1 ist je ein Bild der vier Testsequenzen gezeigt. In dieser Arbeit wurden alle Sequenzen im "Quarter Common Intermediate Format" (QCIF) verwendet, das als Bildformat für Codierungssysteme mit sehr niedrigen Übertragungsbitraten in den Gremien des ITU-T vereinbart wurde. Jedes Bild besteht dabei aus 144 Zeilen und 176 Bildpunkten pro Zeile für die Luminanzkomponente und je aus 72 Zeilen und 88 Bildpunkten pro Zeile für die beiden Chrominanzkomponenten. Für alle Komponenten werden die Amplituden der Abtastwerte mit 8 bit codiert. Die Sequenzen wurden in nicht-teilbildverschränktem Abtastraster mit einer Bildfrequenz von 8,33 Hz verwendet.

Wie in Bild 6.1 gezeigt wird, repräsentieren die Sequenzen *Miss America* und *Claire* die typischen Szenen aus Bildtelefonanwendungen. Dabei stellt die Kopf-Schulter-Partie der sprechenden Person mit Moderatbewegungen das Hauptobjekt der Szene dar. In solchen Anwendungen steht die Kamera fest. In den meisten Fällen sind keine Bildinhaltsänderungen im Hintergrund vorhanden. Im Gegensatz dazu stellen die anderen zwei Sequenzen, *Car Phone* und *Foreman*, typische Szenen bei mobiler Kamera dar.



Bild 6.1 Originalbilder der verwendeten Testsequenzen
(a) Miss America, (b) Claire, (c) Car Phone, (d) Foreman

Dabei wurde die Sequenz *Car Phone* im Innenraum eines fahrenden Personenwagens und die Sequenz *Foreman* auf einer Baustelle aufgenommen. In solchen Fällen wurde die Kamera typischerweise mit freien Händen geführt und demzufolge wackeln die Bilder. Hinzu kommen noch Zoom und Schwenkung der Kamera. Die Szeneninhalte sind auch komplexer als die der ersten beiden Sequenzen. Neben dem kräftiger bewegten Hauptobjekt sind auch Bildinhaltsänderungen im Hintergrund zu sehen. Dadurch sind die Bildgebiete, deren Prädiktionsfehler codiert werden müssen, wesentlich größer.

Anhand der ersten beiden Testsequenzen sollen die mit den vorgestellten Techniken erziehbaren Gewinne ermittelt sowie die erreichbare Bildqualität bei sehr niedrigen Übertragungsbitraten dargestellt werden. Mit den anderen zwei Sequenzen soll hingegen die Grenze des entwickelten Codierungsverfahrens gezeigt werden.

6.2 Gewinne der einzelnen Komponenten des Coders

In diesem Abschnitt soll untersucht werden, ob die in den letzten Kapiteln untersuchten und entwickelten Komponenten des Coders tatsächlich die Erwartungen erfüllen können. Dafür wird zunächst ein Teilbandcoder mit den entwickelten Komponenten, der Vorwärtsanalyse, der blockbasierten Teilbandcodierung und der adaptiven Quantisierung, entsprechend Bild 4.11 auf einem Computer simuliert. Anschließend werden Vergleichssimulationen durchgeführt, in denen jeweils eine dieser Komponenten einzeln ausgeschaltet oder mit einer Standardkomponente ausgetauscht wird. Durch Vergleich der Simulationsergebnisse mit und ohne der jeweiligen Komponente wird dann der Gewinn ermittelt, der durch Anwendung dieser Komponente erzielbar ist. Als Testmaterial werden die typischen Bildtelefon-Sequenzen *Miss America* und *Claire* verwendet. Die gesamte Datenrate für die Luminanz und Chrominanzkomponenten sowie für die Seiteninformation wird jeweils auf 16 kbit/s eingestellt.

Zur Bewertung der Bildqualität wird in diesen Untersuchungen der konventionelle Signal-Rausch-Abstand (Peak Signal to Noise Ratio, kurz PSNR) sowie ein orts- und frequenzgewichteter Signal-Rausch-Abstand des Rekonstruktionsbildes verwendet.

Der konventionelle Signal-Rausch-Abstand wird gemäß,

$$PSNR = 10 \log_{10} \frac{255^2}{\sigma_r^2} \quad (6.1)$$

definiert, wobei σ_r^2 die Varianz der Rekonstruktionsfehler eines Bildes und 255 die maximale Signalamplitude ist. Zur Beurteilung der Codierungsergebnisse der gesamten Bewegtbildsequenz wird der PSNR-Mittelwert über alle Bilder dieser Sequenz gebildet. Das PSNR eines Rekonstruktionsbildes ist eine Messung der objektiven Abweichung des Rekonstruktionsbildes zu seinem ursprünglichen Eingangsbild. Dies entspricht jedoch nicht immer der subjektiven Bewertung der Bildqualität.

Zur Messung des Gewinns bei Anwendung des w-Modells wird zur Berücksichtigung der subjektiven Bildqualität ein gemäß Gl. (6.2) definierter orts- und frequenzgewichteter Signal-Rausch-Abstand (Spatial and Frequential Weighted PSNR, kurz SFWPSNR) verwendet.

$$SFWPSNR = 10 \log_{10} \frac{255^2}{\sum_{m=1}^M E[W_{SB,m} \cdot e_m^2]} \quad (6.2)$$

wobei e_m die lokalen Quantisierungsfehler im Teilband m und $W_{SB,m}$ die Gewichtungsfaktoren aus Gl. (5.13) sind. In Gl. (6.2) ist der Term $\sum_{m=1}^M E[W_{SB,m} \cdot e_m^2]$ eine Schätzung des PSB–Mittelwertes der Rekonstruktionsfehler. Wie in Abschnitt 5.2 bereits erwähnt, liefert der PSB–Wert eine quantitative Schätzung der Sichtbarkeit der Rekonstruktionsfehler an der Sichtbarkeitsschwelle. Deshalb kann der SFWPSNR als ein Gütekriterium zur Schätzung der subjektiven Bildqualität angewendet werden.

Alle nachfolgend angegebene PSNR– und SFWPSNR–Werte beziehen sich nur auf die Luminanzkomponente.

6.2.1 Vorwärtsanalyse

Zur Ermittlung des durch die Vorwärtsanalyse erzielten Gewinns werden Simulationen mit und ohne Vorwärtsanalyse durchgeführt. Im Fall ohne Vorwärtsanalyse werden die Teilbandanalyse und –synthese auf das gesamte Prädiktionsfehlerbild angewendet. Die Teilbandsignale werden wie in Kapitel 4 und 5 beschrieben lokal–adaptiv quantisiert und codiert.

Zum Vergleich werden die Größe der codierten Bildgebiete, die Datenrate der Teilbandcodierung in bit pro Bildpunkt und der PSNR–Mittelwert in den Rekonstruktionsbilder ermittelt und in Tabelle 6.1 zusammengetragen. Bei der Berechnung der Datenrate wird nur die Datenmenge aus der Codierung der Teilbandabtastwerte, nicht aber die aus der Codierung der Seiteninformation berücksichtigt. Im Fall ohne Vorwärtsanalyse enthalten die codierten Bildgebiete die Bildbereiche, in denen die zugehörigen quantisierten Abtastwerte mindestens in einem Teilband ungleich Null sind.

Tabelle 6.1 Vergleich der Coder mit und ohne Vorwärtsanalyse

Testsequenz		Größe der codierten Gebiete [%]	Datenrate [bit/Bildpunkt]	PSNR [dB]	
				Codierte Gebiete	Gesamt
Miss America	mit Vorwärtsanalyse	4,0	1,31	36,1	38,2
	ohne Vorwärtsanalyse	34,1	0,12	33,4	36,6
Claire	mit Vorwärtsanalyse	4,1	1,43	34,4	38,0
	ohne Vorwärtsanalyse	35,0	0,13	31,9	35,2

Diese Ergebnisse zeigen, daß durch Anwendung der Vorwärtsanalyse die Größe der zu codierenden Gebiete sich von ca. 34% auf ca. 4% der Bildfläche reduzieren läßt. Unter Beibehaltung der gesamten Datenmenge pro Bild erhöht sich die auf einen Bildpunkt bezogene Datenrate für die Teilbandcodierung von ca. 0.1 auf rund 1.3 bit/Bildpunkte. Demzufolge verbessert sich der PSNR–Mittelwert in den codierten Gebieten um ca. 2,5 dB. Bezogen auf das gesamte Bild ergibt sich ein PSNR–Gewinn von rund 2 dB.

Wie in Bild 6.2 dargestellt, wird auch eine wahrnehmbare Verbesserung der subjektiven Bildqualität durch Anwendung der Vorwärtsanalyse erzielt. Die Rekonstruktionsbilder



Bild 6.2 Rekonstruktionsbilder aus den Teilbandcodern
(a) mit der Vorwärtsanalyse
(b) ohne die Vorwärtsanalyse

enthalten in diesem Fall deutlich wenige Codierungsartefakte. Mehr Bilddetails lassen sich nach der Decodierung rekonstruieren. Die Rekonstruktionsbilder sehen insgesamt natürlich aus. Dies deutet darauf hin, daß in vielen Bildgebieten die Prädiktionsfehler

selbst weniger störend als die von der Teilbandcodierung bei sehr niedrigen Datenraten erzeugten Rekonstruktionsfehler sind.

6.2.2 Blockbasierte Teilbandcodierung

(1) Vergleich mit der bildbasierten Teilbandcodierung

Zum Vergleich der entwickelten blockbasierten Teilbandcodierung mit der konventionellen Teilbandcodierung werden die durch Vorwärtsanalyse ausgewählten Bildgebiete mit den beiden Verfahren codiert. Bei der blockbasierten Teilbandcodierung werden diese Bildgebiete nach der symmetrischen periodischen Erweiterung in die Teilbänder zerlegt. Die gesamte Anzahl der Abtastwerte in den Teilbändern ist mit der Anzahl der Bildpunkte der ausgewählten Bildgebiete identisch. Bei der bildbasierten Teilbandcodierung wird hingegen das gesamte Bild in die Teilbänder zerlegt. Vorher werden diejenigen Bildpunkte, die außerhalb der zu codierenden Bildgebiete liegen, auf Null gesetzt. Wie bereits in Abschnitt 4.1.1. gezeigt, ist die Anzahl der zu codierenden Teilbandabtastwerte in diesem Fall größer als die Anzahl der Bildpunkte in den ausgewählten Bildgebieten, da hier infolge der Analysefilterung auch außerhalb dieser Gebiete Teilbandabtastwerte ungleich Null vorhanden sind.

Zur Untersuchung der Codierungseffizienz der konventionellen Teilbandcodierung werden zwei Fälle betrachtet. Im einen Fall werden alle Abtastwerte codiert und übertragen. Im anderen werden nur Abtastwerte innerhalb der ausgewählten Bildgebiete übertragen, wobei auf der Decoderseite die fehlenden Abtastwerte der Prädiktionsfehler durch Null ergänzt werden. Für diese beiden Fälle sowie für die blockbasierte Teilbandcodierung werden die PSNR-Werte bei einer gleich bleibenden gesamten Datenrate von 16 kbit/s ermittelt. Die Ergebnisse aus je einem Bild der beiden Testsequenzen werden in Tabelle 6.2 dargestellt.

Tabelle 6.2 Vergleich der blockbasierten TBC mit einer bildbasierten TBC, Ergebnisse aus je einem Bild der beiden Testsequenzen.

Testsequenz	Größe der codierten Bildgebiete [%]	PSNR [dB]		
		Blockbasierte TBC	Bildbasierte TBC	
			Fall 1	Fall 2
Miss America	2,6	35,9	35,3	30,1
Claire	2,5	34,5	30,4	28,1

Fall 1: Übertragung nur der Abtastwerte innerhalb der ausgewählten Bildgebiete

Fall 2: Übertragung aller Abtastwerte ungleich Null

Aus diesen Ergebnissen ist zu erkennen, daß in den beiden Fällen der bildbasierten Teilbandcodierung die erzielten PSNR–Werte deutlich kleiner als die aus der blockbasierten Teilbandcodierung sind. Dies läßt sich wie folgt erklären: Im ersten Fall werden große Rekonstruktionsfehler durch das Verzichten der zusätzlichen Teilbandabstastwerte eingefügt. Im zweiten Fall ist die Anzahl der gesamten zu übertragenden Abstastwerte wegen der zusätzlichen Abstastwerte wesentlich höher. Dadurch verringert sich deutlich die verfügbare Bits pro Abstastwert bei einer gleich bleibenden gesamten Datenrate.

(2) Vergleich mit der DCT

Von der blockbasierten Teilbandcodierung wird ein höherer Codierungsgewinn als von der DCT erwartet, weil in der blockbasierten Teilbandcodierung die Korrelation zwischen den benachbarten zu codierenden Blöcken zusätzlich ausgenutzt werden kann. Zur Feststellung der Größe des Gewinns wird der blockbasierte Teilbandcoder mit einem DCT–Coder verglichen. Die Struktur dieses DCT–Coders ist mit der Struktur des Teilbandcoders identisch, wobei anstelle der Teilbandzerlegung die DCT verwendet wird.

Der Vergleich der beiden Coder wird zuerst bei gleichen PSNR–Mittelwerten in den codierten Gebieten durchgeführt. Dazu wird bei den Simulationen die gesamte Datenrate für den DCT–Coder so eingestellt, daß sich in den codierten Blöcken der gleiche PSNR–Mittelwert wie beim Teilbandcoder ergibt. Die Ergebnisse in Tabelle 6.3 zeigen, daß der DCT–Coder 0,2–0,3 bit pro Bildpunkt mehr benötigt, um die gleichen PSNR–Mittelwerte wie beim Teilbandcoder zu erreichen.

Tabelle 6.3 Vergleich der Datenraten der DCT mit der Teilbandcodierung bei gleichem PSNR–Mittelwert

Testsequenz	PSNR [dB]	Datenrate [bit/Bildpunkt]	
		DCT	TBC
Miss America	36	1,6	1,3
Claire	34	1,6	1,4

In Tabelle 6.3 bezieht sich die Datenrate nur auf die Codierung der DCT–Koeffizienten bzw. Teilbandabstastwerte.

In der zweiten Untersuchung werden die beiden Coder bei einer gleichen Datenrate von 16 kbit/s verglichen. Gegenüber dem DCT–Coder ergibt der Teilbandcoder ein PSNR–Gewinn von 0,3 dB.

6.2.3 Adaptive Quantisierung

Zur Beurteilung der in dieser Arbeit entwickelten adaptiven Quantisierung wird der Unterschied zwischen den SFWPSNR–Werten aus dem ursprünglichen Coder und einem Vergleichscoder ohne die adaptive Quantisierung ermittelt. Im Vergleichscoder ohne die adaptive Quantisierung wird eine feste Quantisiererschrittweite einmal pro Bild festgelegt, die auf jeden Abtastwert in allen Teilbändern angewendet wird. Diese feste Quantisiererschrittweite wird so eingestellt, daß sich aus diesem Coder die gleiche Datenrate wie aus dem Coder mit der adaptiven Quantisierung ergibt. In Tabelle 6.4 werden die SFWPSNR–Gewinne, die durch Anwendung der adaptiven Quantisierung erzielt wurden, aufgeführt.

Tabelle 6.4 SFWPSNR–Gewinne der adaptiven Quantisierung

Testsequenz	SFWPSNR–Gewinn [dB]
Miss America	0,7
Claire	0,9

Ein direkter Vergleich der Rekonstruktionsbilder aus den beiden Codern hat die Verbesserung der subjektiven Bildqualität durch die adaptive Quantisierung bestätigt. Eine genaue Analyse der eingestellten Quantisiererschrittweiten zeigt, daß bei der adaptiven Quantisierung häufig eine kleinere Quantisiererschrittweite als bei der nicht adaptiven Quantisierung für die tieffrequenten Teilbänder verwendet. Dadurch lassen sich die Grundstrukturen des Bildinhaltes besser rekonstruieren.

Tabelle 6.5 Zusammenfassung der Gewinne der einzelnen Komponenten des Coders

Komponenten	Kriterium	Gewinn [dB]
Vorwärtsanalyse	PSNR	~ 1.7 dB
TBC (gegenüber DCT)	PSNR	~ 0.3 dB
Adaptive Quantisierung	SFWPSNR	~ 0.8 dB

In Tabelle 6.5 werden die Gewinne der einzelnen Komponenten des Coders zusammengetragen. Alle Komponenten haben die Erwartungen erfüllt. Der größte Gewinn wird

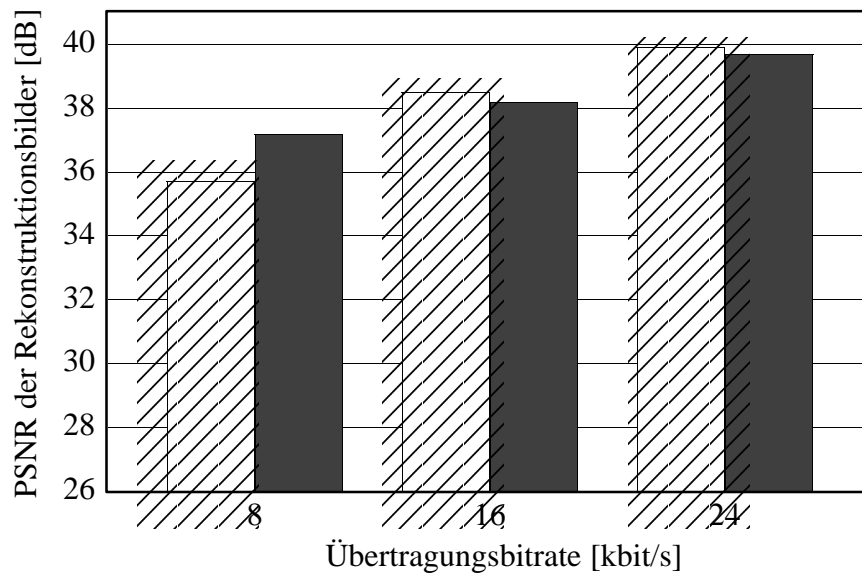
durch Anwendung der Vorwärtsanalyse erzielt. Danach folgt der Gewinn der adaptiven Quantisierung. Die Vorwärtsanalyse und die adaptive Quantisierung gehören beide zu denjenigen Techniken, die lediglich die Steuerung des Coders betreffen. Bei der Standardisierung der Codierverfahren wird die Steuerung aber nicht festgelegt. Diese Ergebnisse deuten darauf hin, daß die Codierungseffizienz auch bei standardisierten Codierungsverfahren durch Verbesserung der Codersteuerung, wie z.B. Vorwärtsanalyse, deutlich erhöht werden kann.

6.3 Vergleich des Teilbandcoders mit dem Hybrid-DCT-Coder TMN5

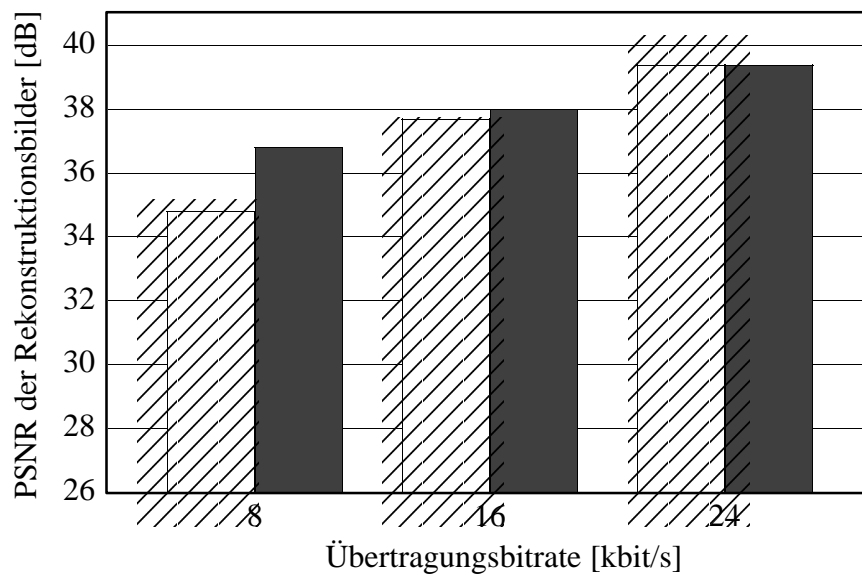
Zur Beurteilung der gesamten Codierungsergebnisse wird im folgenden der in dieser Arbeit entwickelte Teilbandcoder mit dem bereits in Abschnitt 2.3 vorgestellten Hybrid-DCT-Coder TMN5 anhand Computersimulationen verglichen. Wie bereits erwähnt, stellt der TMN5-Coder zur Zeit dieser Untersuchung den Stand der Technik dar.

In den nachfolgenden Untersuchungen werden die vier in Abschnitt 6.1 beschriebenen Testsequenzen nunmehr bei Gesamtdatenraten von 8, 16 und 24 kbit/s codiert. Die Gesamtdatenrate ergibt sich aus der Codierung der Luminanz- und Chrominanzkomponente sowie der Seiteninformation. Für die Simulationen des TMN5-Coders wird eine Software-Implementierung von Telenor Research benutzt. Von den vier optionalen Modi des Coders werden die uneingeschränkten Bewegungsvektoren sowie die fortgeschrittene Prädiktion mit der OBMC eingesetzt. Im Vergleich zum Teilbandcoder enthält der TMN5-Coder keine Vorwärtsanalyse und keine adaptive Quantisierung. Zur Regulierung der Bitrate wird ein einfacher Steuermechanismus implementiert. Bei den Simulationen wird jeweils die gleiche Bildfrequenz eingestellt.

In Bild 6.3 sind die PSNR-Mittelwerte der beiden Coder gegenüber dargestellt. Da bei 8 kbit/s die Testsequenzen *Car Phone* und *Foreman* nicht mit einer ausreichenden Qualität rekonstruiert werden konnten, werden die Ergebnisse nicht berücksichtigt. Von dem Teilbandcoder wurden eigentlich kleinere PSNR-Mittelwerte als von dem TMN5-Coder erwartet, da die in diesem Coder eingesetzte, mit dem w -Modell subjektiv optimierte Quantisierung größere Rekonstruktionsfehler als eine nach einem objektiven Kriterium optimierte Quantisierung erzeugt. Die Ergebnisse in Bild 6.3 zeigen jedoch, daß bei sehr niedrigen Datenraten der Teilbandcoder größere PSNR-Werte als der TMN5-Coder aufweist. Der Grund dieses PSNR-Gewinns liegt in der beim Teilbandcoder angewendeten Vorwärtsanalyse. Durch die Vorwärtsanalyse wird die Größe der zu codierenden Bildgebiete innerhalb eines Bildes deutlich reduziert. Demzufolge werden diese Gebiete mit kleineren Quantisierungsfehlern codiert. Einige restliche



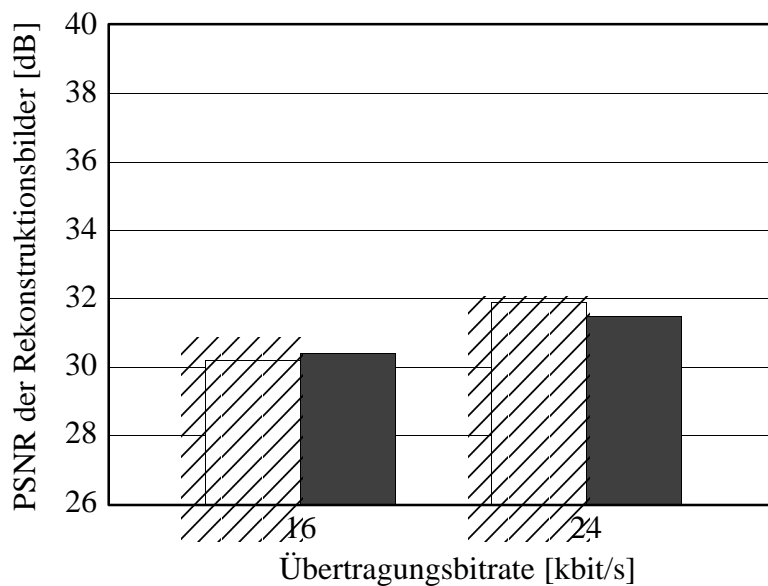
(a) Miss America



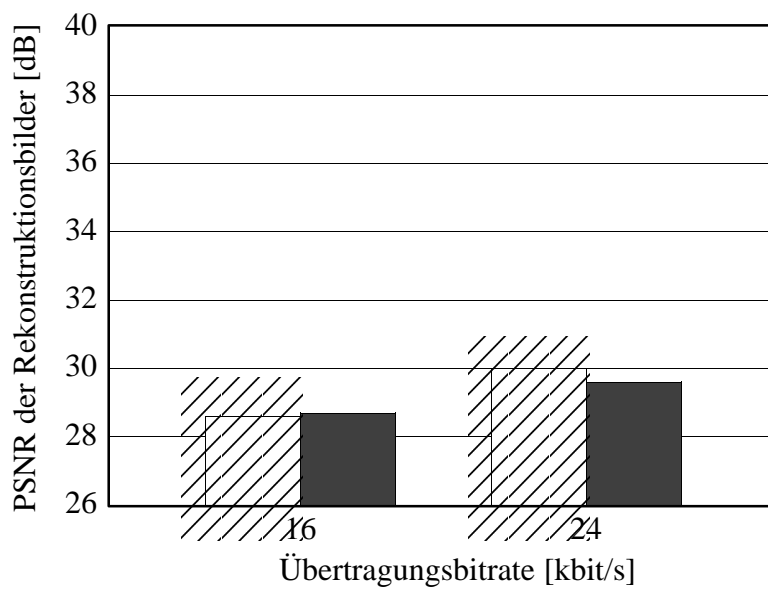
(b) Claire

 vom Hybrid-DCT-Coder TMN5
 vom blockbasierten TBC

Bild 6.3 Vergleich der PSNR-Mittelwerte aus dem blockbasierten Teilbandcoder und dem Hybrid-DCT-Coder TMN5 bei 8, 16 und 24 kbit/s



(c) Car Phone



(d) Foreman

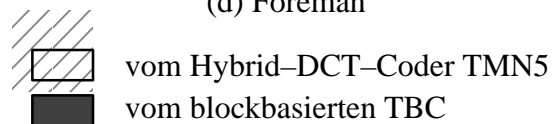


Bild 6.3 (Fortsetzung) Vergleich der PSNR-Mittelwerte aus dem blockbasierten Teilbandcoder und dem Hybrid-DCT-Coder TMN5 bei 16 und 24 kbit/s

Blöcke mit größeren Quantisierungsfehler werden nicht übertragen. Im TMN5-Coder dagegen wird eine große Anzahl von Blöcken codiert und übertragen, wobei der Prädiktionsfehler vieler Blöcke durch die Quantisierungsfehler verursacht wird, die erneut quantisiert und übertragen werden.

Zur Beurteilung der subjektiven Bildqualität werden in Bild 6.4 typische Rekonstruktionsbilder aus den beiden Codern zusammen mit den zugehörigen Originalbildern dargestellt. Bild 6.5 zeigt lokale Vergrößerungen der Gesichtspartie aus den Rekonstruktionsbildern der Testsequenzen *Miss America* und *Claire*. Ein Vergleich dieser Bilder zeigt, daß der Teilbandcoder in allen Fällen eine bessere subjektive Bildqualität gegenüber TMN5 aufweist. Die Rekonstruktionsbilder aus dem Teilbandcoder enthalten deutlich weniger störende Codierungsartefakte und mehr Details des Bildinhaltes.

Bei den Testsequenzen *Miss America* und *Claire* sind die Codierungsartefakte der Teilbandcodierung, wie Ringing-Effekte, bei einer Bildfrequenz von 8 Hz und einer Datenrate größer als 16 kbit/s kaum noch sichtbar. Bei 8 kbit/s lassen sich die einzelnen codierten Bilder zwar immer noch mit einer ausreichenden Qualität rekonstruieren, doch muß dabei die Bildfrequenz auf ca. 5 Hz reduziert werden.

Bei den Testsequenzen *Car Phone* und *Foreman* zeigt der Teilbandcoder gegenüber TMN5 bei 24 kbit/s zwar eine leichte Verbesserung, kann aber keine zufriedenstellende Bildqualität liefern. Die Rekonstruktionsbilder enthalten deutlich sichtbare Codierungsartefakte. Darüber hinaus verursacht die zwangsweise niedrig eingestellte Bildfrequenz eine ruckartige Bewegungswiedergabe. Dies zeigt die Grenze der mit dem Teilbandcoder erreichbaren Datenrateduktion. Bei sehr niedrigen Übertragungsbitraten können nicht alle Videosequenzen mit einer ausreichenden Bildqualität übertragen werden.

Zusammenfassend ergeben sich folgende Erkenntnisse: Mit der Teilbandcodierung kann eine ausreichende bis gute Bildqualität bei Bildtelefonanwendungen mit stationärer Kamera und einer Video-Datenrate von 8 bis 24 kbit/s erreicht werden. Um eine gute Bildqualität bei komplizierten Video-Sequenzen mit mobiler Kamera zu erzielen, wird eine höhere Übertragungsbitrate benötigt. Die Vorwärtsanalyse erweist sich als besonders wirkungsvoll und kann auch in anderen standardisierten Codierungsverfahren eingesetzt werden.



(1 a)



(2 a)



(1 b)



(2 b)



(1 c)



(2 c)

Miss America

Claire

Bild 6.4 Vergleich des Originalbildes (a) mit den Rekonstruktionsbildern aus dem Teilbandcoder (b) und dem Hybrid-DCT-Coder TMN5 (c) bei 16 kbit/s



(3 a)



(4 a)



(3 b)



(4 b)



(3 c)



(4 c)

Car Phone

Foreman

Bild 6.4 (Fortsetzung) Vergleich des Originalbildes (a) mit den Rekonstruktionsbildern aus dem Teilbandcoder (b) und dem Hybrid-DCT-Coder TMN5 (c) bei 24 kbit/s



(1a)



(1b)

(1) Miss America



(2a)



(2b)

(2) Claire

Bild 6.5 Lokale Vergrößerungen der Gesichtspartie der Rekonstruktionsbilder aus dem Teilbandcoder (a) und dem Hybrid-DCT-Coder TMN5 (b) bei 16 kbit/s

7. Zusammenfassung

In der vorliegenden Arbeit wird eine blockbasierte bewegungskompensierende Teilbandcodierung zur Übertragung von Bewegtbildsequenzen mit sehr niedrigen Übertragungsbitraten beschrieben. Dieses Codierungsverfahren stellt eine Weiterentwicklung bekannter Teilbandcodierungsverfahren dar. Das Ziel der Weiterentwicklung ist die Verbesserung der subjektiven Qualität der Rekonstruktionsbilder und die Ermöglichung einer automatischen Anpassung des Codierungsverfahrens an die vorgegebene Übertragungsbitrate und den Inhalt der zu codierenden Bildsequenz. Dabei werden die zu Beginn dieser Arbeit beschriebenen Mängel in den bekannten Teilbandcodern durch Anwendung von drei in dieser Arbeit entwickelten Techniken, nämlich der *Vorwärtsanalyse*, der *blockbasierten Teilbandcodierung* und der *adaptiven Quantisierung* gezielt behoben. Im folgenden werden zunächst für jede der drei Techniken der Ansatz und die daraus entwickelten Algorithmen kurz zusammengefaßt. Anschließend werden die durch Anwendung dieser Techniken erzielten Gewinne anhand experimenteller Ergebnisse dargestellt.

Um die sehr niedrige Übertragungsbitrate effizient zu nutzen, wird eine *Vorwärtsanalyse* der Prädiktionsfehler durchgeführt. Abhängig von der vorgegebenen Übertragungsbitrate und den Signaleigenschaften der zu übertragenden Bildsequenz werden in dieser Vorwärtsanalyse die Anzahl und die Auswahl der zu codierenden Blöcke eines Bildes sowie die momentane Bildfrequenz und Bandbreite des zu übertragenden Bildsignals festgelegt. Dazu werden die Prädiktionsfehler bezüglich ihrer subjektiven Bedeutung in zwei Stufen bewertet. In der ersten Stufe werden zunächst die *irrelevanten* Prädiktionsfehler bestimmt und von der weiteren Codierung ausgeschlossen. In der zweiten Stufe werden von den verbleibenden *relevanten* Prädiktionsfehlern blockweise diejenigen ausgewählt, deren lokale Leistung eine Schwelle überschreitet. Durch Veränderung der Schwelle wird die Anzahl der ausgewählten Blöcke variiert. Demzufolge verändert sich auch der zu erwartende Signal-Rausch-Abstand (PSNR) in den ausgewählten und den nicht-ausgewählten Blöcken im Rekonstruktionsbild. Die Schwelle wird adaptiv so eingestellt, daß der Signal-Rausch-Abstand in beiden Bildbereichen einen vorgegebenen Mindestwert überschreitet, so daß eine Mindestbildqualität in beiden Bildbereichen gewährleistet werden kann. Falls dies nicht möglich ist, werden dann die momentan zu erzielende Bildfrequenz und Bandbreite des zu übertragenden Signals schrittweise reduziert, bis eine Schwelle gefunden wird, bei der die Mindestanforderung an die oben genannten Signal-Rausch-Abstände erfüllt wird. Die Reduktion der Bildfrequenz und Bandbreite wird durch Auslassen von Eingangsbildern bzw. durch Vernachlässigung hochfrequenter Signalanteile bei der Codierung erzielt. Mit dieser Me-

thode wird eine abgestimmte Einstellung der Signal–Rausch–Abstände und der Bildfrequenz sowie Bandbreite erzielt.

Zur Codierung der ausgewählten Blöcke wird eine *blockbasierte Teilbandcodierung* entwickelt. Im Gegensatz zur konventionellen Teilbandcodierung wird hier die Teilbandanalyse und –synthese nur auf die ausgewählten Blöcke angewendet. Dabei werden benachbarte ausgewählte Blöcke zusammenhängend als eine Einheit verarbeitet. Die block–basierte Teilbandanalyse und –synthese besteht aus einer Kaskade von eindimensionalen Teilbandanalyse– bzw. –synthesestufen mit zwei Teilbändern. Um bei block–basierter Verarbeitung eine Vergrößerung der Anzahl der zu übertragenden Teilbandabstastwerte zu vermeiden, wird die aus der Literatur bekannte Technik der *„Symmetrischen Erweiterung“*, die ursprünglich für Behandlung von Bildrändern entwickelt wurde, so erweitert, daß sie auch zur Teilbandanalyse und –synthese von Blockbündeln angewendet werden kann. Die erweiterte Technik wird in dieser Arbeit als *„Symmetrische Periodische Erweiterung“* bezeichnet, da hier nach der symmetrischen Spiegelung zusätzlich eine periodische Wiederholung der Signale durchgeführt wird. Zur Erhöhung des Teilbandcodierungsgewinns werden außerdem adaptive Umschaltmechanismen implementiert. Zum einen läßt sich die Anzahl der Teilbänder bildweise zwischen 16 und 64 umschalten. Zum anderen wird für jedes einzelne Teilband entschieden, ob das Teilbandsignal des Originalbildes oder dessen Prädiktionsfehler übertragen werden soll. Durch diese Umschaltungen werden Anpassungen der Arbeitsweise der Teilbandcodierung an die Eigenschaften des zu codierenden Signals erzielt.

Zur Erhöhung der Irrelevanzreduktion wird eine *lokal–adaptive Quantisierung* unter Anwendung eines aus der Literatur bekannten psychovisuellen Modells entwickelt. Als Maß für die subjektive Wahrnehmung der Rekonstruktionsfehler in der Nähe der Sichtbarkeitsschwelle wird hier die in dem Modell definierte *Potentielle Sichtbarkeitsfunktion* angewendet. In dieser Arbeit wird erstmals gezeigt, daß die potentielle Sichtbarkeitsfunktion annähernd als eine gewichtete Summe der Quantisierungsfehlervarianzen einzelner Teilbänder dargestellt werden kann. Die Gewichtungsfaktoren sind zeit– und ortsvariant und können aus dem zu übertragenden Originalbildsignal, den Parametern des Modells und der verwendeten Teilbandfilterbank berechnet werden. Mit Hilfe dieser Erkenntnis werden die optimale Aufteilung der Datenrate in den Teilbändern und demzufolge die optimalen lokalen Quantisierschrittweiten der einzelnen Teilbandsignale berechnet. Im Gegensatz zu anderen bekannten Quantisierungsverfahren ermöglicht dieses Verfahren eine höhere Irrelevanzreduktion, weil hier neben der Modulationsübertragungsfunktion des menschlichen visuellen Systems auch andere

psychovisuelle Effekte, wie z.B. der Einfluß der Hintergrundhelligkeit und die Maskierung durch örtliche Helligkeitssprünge, mit ausgenutzt werden.

Zur Bewertung der drei entwickelten Konzepte werden zuerst die Gewinne der einzelnen in diesem Coder eingesetzten Techniken anhand von Computersimulationen mit den bekannten Bildtelefon-Testsequenzen *Miss America* und *Claire* bei einer Übertragungsbitrate von 16 kbit/s ermittelt. Aus diesen Computersimulationen ergeben sich folgende Ergebnisse:

- Mit Hilfe der Vorwärtsanalyse wird die Anzahl der zu codierenden Blöcke bei den verwendeten Bildtelefon-Testsequenzen von rund 30 % auf ca. 4 % reduziert. Dadurch läßt sich in den ausgewählten Blöcken die lokal zur Verfügung stehende Datenrate erheblich erhöhen, so daß diese Blöcke mit einer guten Bildqualität, entsprechend einem PSNR-Wert von 35,3 dB, codiert werden können. Durch den vorgegebenen Mindest-PSNR-Wert in der Vorwärtsanalyse ist auch in den anderen Bildbereichen eine Mindestqualität gewährleistet. Damit erzielt sich für das Gesamtbild eine Verbesserung des mittleren PSNR-Wertes von ca. 1,7 dB gegenüber einem Coder ohne Vorwärtsanalyse. Dabei wird eine Bildfrequenz von ca. 8 Hz erreicht.
- Der Einsatz der blockbasierten Teilbandcodierung wird als zweckmäßig bewiesen. Im Vergleich zur DCT eliminiert die block-basierte Teilbandcodierung die Block- und Moskitoeffekte und ermöglicht gleichzeitig die Ausnutzung der statistischen Bindungen zwischen den zusammenhängenden ausgewählten Blöcken. Gegenüber der DCT erzielt die blockbasierte Teilbandcodierung einen mittleren Gewinn des Signal-Rausch-Abstandes in den Rekonstruktionsbildern von rund 0.3 dB. Im Vergleich zu der konventionellen Teilbandcodierung vermeidet diese Technik die Erhöhung der Anzahl der zu übertragenden Teilbandabstastwerte und begrenzt die Verbreitung der Quantisierungsfehler im Rekonstruktionsbild über die ausgewählten und codierten Blöcke hinaus. Die Simulationsergebnisse zeigen, daß die Rekonstruktionsbilder weitgehend von den subjektiv störenden Ringing-Effekten befreit sind.
- Der erwartete Gewinn der lokal-adaptiven Quantisierung der Teilbandsignale wird durch die Simulationsergebnisse bestätigt. Bei der verwendeten Übertragungsbitrate sind die durch die Quantisierung verursachten Rekonstruktionsfehler entweder nicht mehr sichtbar oder subjektiv wesentlich weniger störend geworden. Der Signal-Rausch-Abstand, indem die Fehlersignale mittels eines Wahrneh-

mungsmodels gewichtet werden, verbessert sich in den Rekonstruktionsbildern um c.a. 0.8 dB.

Zur Beurteilung der Effizienz des Codierungsverfahrens wurde dies mit dem beim ITU-T entwickelten Testmodell TMN5 des internationalen Standards H.263 anhand von bekannten Testsequenzen *Miss America*, *Claire*, *Car Phone* und *Foreman* bei Übertragungsbitraten von 8, 16 und 24 kbit/s verglichen. Die Simulationen mit den Testsequenzen *Miss America* und *Claire* zeigen, daß bei Kopf-Schulter-Szenen mit Moderatbewegungen die subjektiv sehr störenden Codierungsartefakte, wie z.B. Block- und Moskitoeffekte, durch Anwendung der in dieser Arbeit beschriebenen Techniken eliminiert werden. Der Signal-Rausch-Abstand in den Rekonstruktionsbildern konnte bei 8 kbit/s um ca. 1 dB im Mittel erhöht werden. Bei den Testsequenzen *Car Phone* und *Foreman*, die stärkere Objekt- und Kamerabewegungen enthalten, wird auch eine subjektive Verbesserung der Bildqualität erzielt, obwohl der mittlere PSNR-Wert nicht verbessert wird. Der vorgestellte Coder kann seine Parameter automatisch an den Bildinhalt und die vorgegebene Übertragungsbitrate im Bereich von 8 bis 64 kbit/s anpassen, so daß eine manuelle Anpassung nicht erforderlich ist. Damit sind die zu Beginn der Arbeit gesetzten Ziele erreicht.

Die Effizienz und die selbständige Anpassungsfähigkeit des vorgestellten Coders erlaubt einen Einsatz in einem breiten Anwendungsbereich. Neben der Übertragung von Bewegtbildsequenzen mit sehr niedrigen Übertragungsbitraten ist dieser Coder auch geeignet für Speicherung von Bewegtbildsequenzen mit einer hohen Datenkompression. Die einzelnen, im Rahmen dieser Arbeit entwickelten Techniken sind auch in anderen Codern zur Erhöhung der Codierungseffizienz anwendbar. Zum einen läßt sich die Vorwärtsanalyse ohne Änderung der bestehenden Standards auch in andere blockbasierte Coder einzufügen. Zum anderen kann das psychovisuelle Modell in einer ähnlichen Weise zur Steuerung der Quantisierung von DCT-Koeffizienten dienen. Ferner ist auch möglich, die blockbasierte Teilbandcodierung auf Model-Failure-Gebiete [35] in einem objektbasierten Analyse-Synthese-Coder anzuwenden.

Andererseits sind weitere Verbesserungen des vorgestellten Codierungsverfahrens möglich. Dazu zählt zunächst die Verbesserung des Einschwingverhaltens nach einem Szenenwechsel. In der vorliegenden Version des Coders werden ab dem zweiten Eingangsbild mehrere Bilder aus einer Szene ausgelassen, um eine ausreichende Datenmenge zur Codierung des ersten Bildes der Szene bereitzustellen. Beim Empfänger wird das erste Bild der neuen Szene erst nach einer entsprechenden Verzögerung dargestellt. Die Verzögerungszeit kann jedoch reduziert werden, wenn Techniken aus der progressi-

ven Übertragung von Standbildern einbezogen werden. Darüberhinaus kann die Effizienz der bewegungskompensierenden Prädiktion durch Einführung einer globalen Kompensation von Kamerabewegungen erhöht werden. Auf dieser Weise lassen sich die Probleme, die bei Codierung von Szenen mit Kameraschwenkungen aufgetreten sind, beseitigen. Weiterhin kann die in diesem Coder verwendete skalare Quantisierung durch eine Interband-Vektorquantisierung ersetzt werden, wobei zur Bildung eines Interband-Vektors von jedem Teilband ein Abtastwert an gleicher Position entnommen wird. Dadurch können noch verbleibende statistische Bindungen höherer Ordnung zwischen den Teilbändern ausgenutzt werden.

8. Literaturverzeichnis

- [1] S.O. Aase and T.A. Ramstad, "Ringing reduction in low bit-rate image subband coding using projection onto a space of paraboloids", *Signal Processing: Image Communication*, Vol. 5, No. 3, S. 273–282, May 1993.
- [2] K. Aizawa, H. Harashima, T. Saito, "Model-based analysis-synthesis image coding (MBASIC) system for a person's face", *Signal Processing: Image Communication*, Vol. 1, Nr. 2, S. –, Oktober 1989.
- [3] M.H. Ahmad Fadzil and T.J. Dennis, "Video subband VQ coding at 64 kbit/s using short-kernel filter banks with an improved motion estimation technique", *Signal Processing: Image Communication*, Vol. 3, No. 1, S. 3–21, February 1991.
- [4] C. Auyeung, J. Kosmach, M. Orchard and T. Kalafatis, "Overlapped block motion compensation", *Proc. SPEI*, Vol. 1818, Visual Communications and Image Processing, S. 561–571, 1992.
- [5] M. Bierling, "Displacement estimation by hierarchical blockmatching", *Proc. SPEI*, Vol. 1001, Visual Communications and Image Processing, S. 942–951, November 1988.
- [6] M. Bierling, *Hierarchische Displacementschätzung zur Bewegungskompensation in digitalen Fernsehbildsequenzen*, Dissertation, Universität Hannover, VDI-Fortschritt-Berichte, Reihe 10, Nr. 179, Düsseldorf: VDI-Verlag, 1991.
- [7] J. Biemond, B.P. Thieme, D.E. Boekee, "Subband-coding of moving video using hierarchical motion estimation and vector quantization", *International Workshop on 64 kbit/s Coding of Moving Video*, Hannover, Deutschland, Paper 3–2, Juni 1988.
- [8] A. v. Brandt, "Teilbandcodierung von Bewegtbildsequenzen mit 2 Mbit/s", *Frequenz*, Vol. 40, S. 190–197, 1986.
- [9] CCITT, "Draft revision of recommendation H.261: Video codec for audio visual services at p x 64 kbit/s", Study Group XV, WP/1/Q4, Specialist group on coding for visual telephony, Doc. No. 584, November 1989.
- [10] CCITT SG XV, Doc. #525, "Description of Ref. Model 8 (RM8)", 1989.
- [11] A. Croisier, D. Esteban and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques", *Proc. IEEE Int. Conf. Inform. Sci. Syst.*, Patras, Greece, S. 443–446, Aug. 1976.

-
- [12] D. Esteban and C. Galand, "Applications of quadrature mirror filters to band voice coding schemes", *Proc. Int. Conf. Acoust., Speech, Signal Processing*, S.191–195, 1977.
- [13] R. Forchheimer, O. Fahlander, T. Kronander, "A semantic approach to the transmission of face images", *Picture Coding Symposium (PCS'84)*, Cesson-Sevigne, Frankreich, Nr. 10.5, Juli 1984.
- [14] P. Gerken, "Objekt-based analysis-synthesis coding of image sequences at very low bit rates", *IEEE Trans. Circ. and Syst. for Video Technol.*, Vol. 4, No. 3, S. 228–235, June 1994.
- [15] H. Gharavi, A. Tabatabai, "Subband coding of digital images using two-dimensional quadrature mirror filtering", in *Visual Comm. Signal Processing*, Proc. SPIE, Vol. 707, S. 51–61, 1986.
- [16] B. Girod, "The efficiency of the motion-compensating prediction for hybrid coding of video sequences", *IEEE J. Select. Areas Comm.*, Vol. SAC-5, No. 7, S. 1140–1154, August 1987.
- [17] B. Girod, *Ein Modell der menschlichen visuellen Wahrnehmung zur Irrelevanzreduktion von Fernseh luminanzsignalen*, Dissertation, Universität Hannover, VDI-Fortschritt-Berichte, Reihe 10, Nr. 84, Düsseldorf: VDI-Verlag, 1988.
- [18] B. Girod, "Psychovisual aspects of image communication", *Signal Processing*, Vol. 28, S. 239–251, 1992.
- [19] R.M. Gray, "Vector quantization", *IEEE ASSP Mag.*, S. 4–29, April 1984.
- [20] K. Grotz, J. Mayer, G. Süßmeier, "A 64 kbit/s videophone codec with forward analysis and control", *Signal Processing: Image Communication*, Vol. 1, No. 2, S. 103–116, Oktober 1989.
- [21] M. Hötter, *Objektorientierte Analyse-Synthese-Codierung basierend auf dem Modell bewegter, zweidimensionaler Objekte*, Dissertation, Universität Hannover, VDI-Fortschritt-Berichte, Reihe 10, Nr. 217, Düsseldorf: VDI-Verlag, 1992.
- [22] M. Hötter, "Optimization and efficiency of an object-oriented analysis-synthesis coder", *IEEE Trans. Circ. and Syst. for Video Technol.*, Vol. 4, Nr. 2, S.181–194, April 1994.
- [23] ISO/IEC JTC1/SC29/WG11, "Proposal Package Description (PPD)", November 1994.

-
- [24] ITU–T, "Video Codec Test Model, TMN5", Study Group 15, Working Party 15/1/, Expert's group on very low bitrate visual telephony, Januar 1995.
- [25] ITU–T, "Draft ITU–T Recommendation H.263: Video coding for low bitrate communication", Study Group 15, Working Party 15/1/, Expert's group on very low bitrate visual telephony, November 1995.
- [26] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice–Hall, Englewood Cliffs, NJ, 1984.
- [27] J.D. Johnston, "A filter family designed for use in quadrature mirror filter banks", *Proc. Internat. Conf. Acoust. Speech Signal Process. (ICASSP)*, Denver, U.S.A., S. 291–294, April 1980.
- [28] G. Karlsson and M. Vetterli, "Extension of finite length signals for subband coding", *Signal Processing*, Vol. 17, Nr. 2, S. 161–166, Juni 1989.
- [29] T. Kronander, *Some aspects of perception based image coding*, Dissertation, Nr. 203, Dep. Elec. Eng. Linköping University, 1989.
- [30] D. LeGall, A. Tabatabai, "Sub–band coding of digital images using symmetric short kernel and arithmetic coding", *Proc. IEEE Internat. Conf. Acoust. Speech Signal Process.*, S. 761–764, April 1988.
- [31] Y. Linde, A. Buzo, R.M. Gray, "An algorithm for vector quantizer design", *IEEE Trans. Commun.*, Vol. COM–28, Nr. 1, S. 84–95, Januar 1980.
- [32] F. X. J. Lukas, "Picture quality prediction based on a visual model", *IEEE Trans. Commun.*, Vol. COM–30, Nr. 7, S. 1679–1692, Juli 1982.
- [33] F. May, "Algorithmen und Realisierung eines ISDN–Bildtelefon–Coders", *Nachrichtentechnische Zeitschrift (NTZ)*, Bd. 42, Heft 3, S.130–133, 1989.
- [34] H.G. Musmann, P. Pirsch, H.J. Grallert, "Advances in picture coding", *Proceedings of the IEEE*, Vol. 73, Nr. 4, S. 523–548, April 1985.
- [35] H.G. Musmann, M. Hötter and J. Ostermann, "Object–oriented analysis–synthesis coding of moving images", *Signal Processing: Image Communication*, Vol. 3, No. 2, S. 117–138, November 1989.
- [36] H.G. Musmann, "Source models for very low bit rate video coding", Eingereicht bei *Signal Processing: Image Communication*.
- [37] S. Nogaki and M. Ohta, "An overlapped block motion compensation for high quality picture coding", *1992 IEEE internat. Symposium on Circuits and Systems*, S. 184–187, May 1992.

-
- [38] J. Ostermann, "Objekt-based analysis-synthesis coding based on the source model of moving rigid 3D objects", *Signal Processing: Image Communication*, Vol. 6, No. 2, S. 143–161, May 1994.
- [39] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.
- [40] W.K. Pratt, *Digital Image Processing*, Wiley Interscience, New York, 1987.
- [41] D. Qian, "Block-based motion compensated subband coding at very low bit rates", *COST 211^{ter} European workshop on new techniques for coding of video signals at very low bitrates*, Hannover, Germany, Nr. 3.4, Dezember 1993.
- [42] D. Qian, "Block-based motion compensated subband coding at very low bit rates using a psychovisual model", *International Workshop on Coding Techniques for very low bit-rate video (VLBV 94)*, Colchester, Großbritannien, Nr. 6.8, April 1994.
- [43] D. Qian, "A motion compensated subband coder for very low bit rates", *Signal Processing: Image Communication*, Vol. 7, No. 6, S. 397–418, October 1995.
- [44] D. Qian and A. Hutter, "Block-based subband-coding", Proposal for MPEG4 standard, *ISO/IEC JTC1/SC29/WG11 Document*, October 1995.
- [45] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circ. and Syst. for Video Technol.*, Vol. 6, S. 243–250, June 1996.
- [46] H. Schiller and M. Hötter, "Investigation on colour coding in an object-oriented analysis-synthesis coder", *Signal Processing: Image Communication*, Vol. 5, No. 4, S. 319–326, October 1993.
- [47] B. Schmale, "Local adaptive quantization of HDVT subband signals", *Int. Workshop on HDTV'93*, Ottawa, October 1993.
- [48] J.M. Shapiro, "Embedded image coding using zero-trees of wavelet coefficients", *IEEE Trans. Signal Processing*, Vol. 41, S. 3445–3463, Dec. 1993.
- [49] M.T.J. Smith and T.P. Barnwell, "Exact reconstruction techniques for tree-structured subband coders", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-34, No. 3, S. 434–441, Juni 1986.
- [50] M.T.J. Smith and S.L. Eddins, "Sub-band coding of images with octave band tree structure", *Proc. IEEE Int. Conf. on Acoust. Speech, Signal Process.*, Dallas, S. 1382–1385, April 1987.

-
- [51] M.T.J. Smith and S.L. Eddins, "Analysis/synthesis techniques for subband image coding", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-38, Nr. 8, S. 1446–1456, Aug. 1990.
- [52] G. Sullivan, "Multi-Hypothesis motion compensation for low bit-rate video coding", *Proc. Internat. Conf. Acoust. Speech Signal Process. (ICASSP)*, Minneapolis, U.S.A., S. V437–V440, April 1993.
- [53] M. Temerinac and B. Edler, "LINC: A common theory of transform and subband coding", *IEEE Trans. Commun.*, vol. COM-41, No. 2, S.266–274, February 1993.
- [54] P.P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect-reconstruction property", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-35, Nr. 4, S. 476–492, Apr. 1987.
- [55] P.P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial", *Proc. IEEE*, Vol. 78, Nr. 1, S. 56–93, Jan. 1990.
- [56] L. Vandendorpe, "Optimized quantization for image subband coding", *Signal Processing: Image Communication*, Vol. 4, No. 1, S. 65–79, November 1991.
- [57] L. Vandendorpe, "CQF filter banks matched to signal statistics", *Signal Processing*, Vol. 29, S. 237–249, 1992.
- [58] M. Vetterli, "Multi-dimensional sub-band coding: some theory and algorithms", *Signal Processing*, Vol. 6, S. 97–112, 1984.
- [59] M. Vetterli, "Filter banks allowing perfect reconstruction", *Signal Processing*, Vol. 10, Nr. 3, S. 219–244, 1986.
- [60] P.H. Westerink, J. Biemond and D.E. Boekee, "An optimal bit allocation algorithm for subband coding", *Proc. Internat. Conf. Acoust. Speech Signal Process. (ICASSP)*, New York, S. 757–760, April 1988.
- [61] P.H. Westerink, *Subband Coding of Image*, Dissertation, Technische Universiteit Delft, Okt. 1989.
- [62] P.H. Westerink, J. Biemond and F. Muller, "Subband coding of image sequences at low bit rates", *Signal Processing: Image Communication*, Vol. 2, No. 4, S. 441–448, December 1990.
- [63] J.W. Woods and S. O'Neil, "Subband coding of images", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-34, No. 5, S. 1278–1288, October 1986.

- [64] Z. Xiong, K. Ramchandran, M.T. Orchard and Y.Q. Zhang, "A comparative study of DCT- and wavelet-based image coding", *IEEE Trans. Circ. and Syst. for Video Technol.*, Vol. 9, No. 5, S. 692-695, August 1999.

Anhang

A. Impulsantwortfunktionen der in dieser Arbeit verwendeten QM-Filter

Die Filter mit 8 bzw. 16 Koeffizienten wurden ursprünglich von Johnston in [27] vorgestellt. In dieser Arbeit werden dieser Filter als QMF16A und QMF08A bezeichnet.

Filter mit 2 Koeffizienten (QMF02):

0,5

0,5

Filter mit 8 Koeffizienten (QMF08A):

$0,94221178 \cdot 10^{-2}$

$-0,70915014 \cdot 10^{-1}$

$0,69686897 \cdot 10^{-1}$

0,49180600

0,49180600

$0,69686897 \cdot 10^{-1}$

$-0,70915014 \cdot 10^{-1}$

$0,94221178 \cdot 10^{-2}$

Filter mit 16 Koeffizienten (QMF16A):

$0,10501670 \cdot 10^{-2}$

$-0,50545260 \cdot 10^{-2}$

$-0,25897560 \cdot 10^{-2}$

$0,27641400 \cdot 10^{-1}$

$-0,96663760 \cdot 10^{-2}$

$-0,90392230 \cdot 10^{-1}$

$0,97798170 \cdot 10^{-1}$

0,48102840

0,48102840

$0,97798170 \cdot 10^{-1}$

$-0,90392230 \cdot 10^{-1}$

$$-0,96663760 \cdot 10^{-2}$$

$$0,27641400 \cdot 10^{-1}$$

$$-0,25897560 \cdot 10^{-2}$$

$$-0,50545260 \cdot 10^{-2}$$

$$0,10501670 \cdot 10^{-2}$$

B. Filterimpulsantwortfunktionen und Parameter des w-Modells

Dieser Anhang enthält eine Zusammenstellung der Filterimpulsantwortfunktionen sowie der Parameter des in [17] beschriebenen W-Modelles.

B.1 Filterimpulsantwortfunktionen

$$MON[s_u] = L_{MON} \cdot (s_u + s_{MON})^y \quad (\text{B.1})$$

$$g_{optic}(\xi) = \frac{1}{2\pi \cdot r_{optic}^2} \exp\left\{\frac{-x^2 - y^2}{2r_{optic}^2}\right\} \cdot \delta_1(t) \quad (\text{B.2})$$

$$g_{rez}(\xi) = \begin{cases} \delta_2(\xi) \cdot \frac{1}{\tau_{rez}} \exp\left\{\frac{-t}{\tau_{rez}}\right\} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad (\text{B.3})$$

$$g_{inh}(\xi) = g_{inh/xy}(\xi) \cdot g_{inh/t}(t) \quad (\text{B.4})$$

$$g_{inh/xy}(\xi) = \frac{1}{2\pi \cdot r_{inh}^2} \exp\left\{\frac{-x^2 - y^2}{2r_{inh}^2}\right\} \quad (\text{B.5})$$

$$g_{inh/t}(t) = \begin{cases} \frac{t}{\tau_{inh}^2} \exp\left\{\frac{-t}{\tau_{inh}}\right\} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad (\text{B.6})$$

$$SAT(c_{inh}) = \begin{cases} c_{inh} & \text{für } c_{inh} \leq 1 \\ 1 + \frac{1}{k_{sat}} \ln(k_{sat} \cdot (c_{inh}-1) + 1) & \text{für } c_{inh} > 1 \end{cases} \quad (\text{B.7})$$

$$g_{gl}(\xi) = \begin{cases} \delta_2(\xi) \cdot \frac{1}{\tau_{gl}} \exp\left\{\frac{-t}{\tau_{gl}}\right\} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad (\text{B.8})$$

$$g_{\theta}(\xi) = g_{\theta/xy}(\xi) \cdot g_{\theta/t}(t) \quad (\text{B.9})$$

$$g_{\theta/xy}(\xi) = \frac{1}{2\pi \cdot r_{\theta}^2} \exp\left\{-\frac{x^2 + y^2}{2r_{\theta}^2}\right\} \quad (\text{B.10})$$

$$g_{\theta/t}(t) = \begin{cases} \frac{1}{\tau_{\theta}} \exp\left\{-\frac{t}{\tau_{\theta}}\right\} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad (\text{B.11})$$

B.2 Parameter

Bezeichnung	Formelzeichen	Wert
γ -Charakteristik des Monitors	γ	2,2
	s_{MON}	15
	L_{MON}	$0,35 \cdot 10^{-3} \text{ cd/m}^2$
Streuungsparameter der Augenoptik	r_{optic}	1,0 min
Rezeptorzeitkonstante	τ_{rez}	18 ms
Adaptationsleuchtdichte	L_{ad}	$7,0 \text{ cd/m}^2$
Streuungsparameter der Inhibition	r_{inh}	8,0 min
Inhibitionszeitkonstante	τ_{inh}	35 ms
Sättigungsparameter	k_{sat}	8,0
Ganglienzell-Zeitkonstante	τ_{gl}	18 ms
Streuungsparameter der unterschwelligigen Summation	r_{θ}	13 min
Zeitkonstante der unterschwelligigen Summation	τ_{θ}	230 ms
Schwelle	θ	$1,25 \cdot 10^{-5}$

C. Herleitung der optimalen Bitaufteilung auf die Teilbandsignale

Die optimale Bitaufteilung \underline{R}_m auf die Teilbandsignale y_m ($m = 1, \dots, M$) wird in einem Gebiet mit konstanten w-Parametern dann erzielt, wenn die PSB-Funktion

$$PSB = \sum_{m=1}^M [W_{SB,m} \cdot \varepsilon_m^{*2} 2^{-2R_m} \sigma_{y,m}^2]. \quad (C.1)$$

ein Minimum erreicht, wobei die Gesamtbitrate:

$$R = \sum_{m=1}^M \frac{N_m}{N} R_m = R_{soll} \quad (C.2)$$

gleich der vorgegebenen Bitrate R_{soll} ist. Die Bitraten \underline{R}_m sowie R und R_{soll} werden hier in bit/Abtastwert gemessen. Die Bedeutungen der Symbole in Gl. (C.1) und (C.2) wurden bereits in Abschnitt 5.2 erläutert.

Zur Vereinfachung der Notierung werden Konstanten

$$a_m = W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 \text{ und } b_m = \frac{N_m}{N} \quad (C.3)$$

($m = 1, \dots, M$) eingefügt, so daß

$$PSB = \sum_{m=1}^M a_m 2^{-2R_m} \quad (C.4)$$

und

$$R_{soll} = \sum_{m=1}^M b_m R_m \quad (C.5)$$

gelten. Mit Hilfe eines Langrangeschen Multiplikators λ kann die gesuchte Bitaufteilung \underline{R}_m durch Auflösen der Gleichung

$$\frac{\partial}{\partial R_m} \left[\sum_{i=1}^M a_i 2^{-2R_i} + \lambda \left(\sum_{j=1}^M b_j R_j - R_{soll} \right) \right] = 0 \quad (C.6)$$

wie folgt bestimmt werden. Nach der partiellen Differenzierung der Gl. (C.6) ergeben sich

$$a_m 2^{-2R_m} \ln 2 (-2) + \lambda b_m = 0 \quad (\text{C.7})$$

und

$$R_m = -\frac{1}{2} \log_2 \frac{\lambda b_m}{2a_m \ln 2} = -\frac{1}{2} \log_2 \lambda c_m \quad (\text{C.8})$$

für $m = 1, \dots, M$, wobei $c_m = \frac{b_m}{2a_m \ln 2}$ ist.

Durch Einsetzen von Gl. (C.8) in Gl. (C.5) lässt sich der Wert von λ wie folgt berechnen:

$$R_{soll} = \sum_{m=1}^M b_m \left[-\frac{1}{2} \log_2 \lambda c_m \right] \quad (\text{C.9})$$

$$-2R_{soll} = \log_2 \lambda \sum_{m=1}^M b_m + \sum_{k=1}^M (b_k \log_2 c_k) \quad (\text{C.10})$$

$$\log_2 \lambda = -\frac{2R_{soll} + \sum_{k=1}^M (b_k \log_2 c_k)}{\sum_{m=1}^M b_m} \quad (\text{C.11})$$

$$\lambda = 2^{-\frac{2R_{soll} + \sum_{k=1}^M (b_k \log_2 c_k)}{\sum_{m=1}^M b_m}} \quad (\text{C.12})$$

Unter Berücksichtigung folgender Beziehungen

$$\sum_{m=1}^M b_m = \sum_{m=1}^M \frac{N_m}{N} = 1 \quad (\text{C.13})$$

und

$$2^{-\sum_{k=1}^M b_k \log_2 c_k} = 2^{-\log_2 \prod_{k=1}^M c_k^{b_k}} = \frac{1}{\prod_{k=1}^M c_k^{b_k}} \quad (\text{C.14})$$

folgt aus Gl. (C.12)

$$\lambda = \frac{2^{-2R_{soll}}}{\prod_{k=1}^M c_k^{b_k}}. \quad (C.15)$$

Aus Gl. (C.8) und (C.15) ergibt sich die gesuchte Bitaufteilung

$$\begin{aligned} R_m &= -\frac{1}{2} \log_2 \frac{2^{-2R_{soll}} c_m}{\prod_{k=1}^M c_k^{b_k}} \\ &= R_{soll} + \frac{1}{2} \log_2 \frac{a_m/b_m}{\prod_{k=1}^M (a_k/b_k)^{b_k}} \\ &= R_{soll} + \frac{1}{2} \log_2 \frac{\frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{N_m/N}}{\prod_{k=1}^M \left(\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{N_k/N} \right)^{N_k/N}} \\ &= R_{soll} + \frac{1}{2} \log_2 \frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{N_m/N \prod_{k=1}^M \left(\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{N_k/N} \right)^{N_k/N}}, \end{aligned} \quad (C.16)$$

wobei im vorletzten Schritt die Werte von a_m und b_m eingesetzt werden.

Schließlich läßt sich das bei dieser optimalen Bitaufteilung erzielte Minimum der PSB-Funktion aus Gl. (C.1) und (C.16) berechnen:

$$PSB_{\min} = \sum_{m=1}^M \left\{ W_{SB,m} \cdot \varepsilon_m^{*2} \sigma_{y,m}^2 \cdot 2^{-2 \left[R_{soll} + \frac{1}{2} \log_2 \frac{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2}{N_m/N \prod_{k=1}^M \left(\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{N_k/N} \right)^{N_k/N} \right]} \right\} \quad (C.17)$$

$$\begin{aligned}
PSB_{\min} &= \sum_{m=1}^M \left[W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2 2^{-2R_{soll}} \frac{\frac{N_m}{N} \prod_{k=1}^M \left(\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{\frac{N_k}{N}} \right)^{\frac{N_k}{N}}}{W_{SB,m} \varepsilon_m^{*2} \sigma_{y,m}^2} \right] \\
&= 2^{-2R_{soll}} \prod_{k=1}^M \left[\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{\frac{N_k}{N}} \right]^{\frac{N_k}{N}} \sum_{m=1}^M \frac{N_m}{N} . \\
&= 2^{-2R_{soll}} \prod_{k=1}^M \left[\frac{W_{SB,k} \varepsilon_k^{*2} \sigma_{y,k}^2}{\frac{N_k}{N}} \right]^{\frac{N_k}{N}} .
\end{aligned} \tag{C.18}$$

Lebenslauf

Deyu Qian

geboren am 31. 3. 1957 in Zhejiang, China
als Sohn des Bauingenieures Chongzhan Qian
und der Volkswirtin Pinxian Li

seit dem 22. 7. 1983 verheiratet mit Wei Ma, Ärztin
zwei Kinder

- Schulbildung 1964 – 1971 Grundschule in Peking und Hennan, China
1971 – 1977 Gymnasien in Hennan und Tianjin, China
Abschluß Abitur
- Studium 2/1978 – 1/1982 Bachelor–Studium der Elektrotechnik an der Uni-
versität Tianjin, China, mit dem Schwerpunkt Nachrichtentechnik.
Abschluß Bachelor
2/1982 – 12/1984 Master–Studium der Elektrotechnik an der Univer-
sität Tianjin, China, mit dem Schwerpunkt Nachrichtentechnik. Ab-
schluß Master
9/1985 – 7/1986 Deutsche Sprachkurse am Institut für Fremdspra-
chen Xi’an, China
4/1987 – 9/1987 Deutsche Sprachkurse am Goethe–Institut Mann-
heim mit einem Stipendium der Konrad–Adenauer–Stiftung
- Beruf 8/1977 – 1/1978 Mechaniker–Lehrling im Textilfabrik Tianjin, China
12/1984 – 9/1985 und 7/1986 – 3/1987 wissenschaftlicher Assistent
am Forschungslaboratorium für Bildinformation, Universität Tian-
jin, China
10/1987 – 12/1994 Stipendiat/wissenschaftlicher Mitarbeiter am In-
stitut für Theoretische Nachrichtentechnik und Informationsver-
arbeitung, Universität Hannover, mit einem Stipendium der Konrad–
Adenauer–Stiftung
1/1995 – 5/1996 wissenschaftlicher Mitarbeiter am Fraunhofer–Insti-
tut für Festkörpertechnologie, München
6/1996 – 6/2000 Konzeptingenieur bei Fa. Micronas, Freiburg
7/2000 – 9/2000 Entwicklungsingenieur bei Fa. Infineon Technolo-
gies, München
seit 10/2000 Entwicklungsingenieur bei Fa. Micronas, München