

# VORSCHLAG FÜR EINEN GSM-VOLLRATEN-SPRACH- UND KANALCODEC MIT VERBESSERTER SPRACHQUALITÄT

Norbert Görtz

Lehrstuhl für Netzwerk- und Systemtheorie

Christian-Albrechts-Universität Kiel

Kaiserstraße 2

24143 Kiel

Tel.: 0431/77572-406

ng@techfak.uni-kiel.de

## ÜBERSICHT

Zunächst wird ein Schmalband-Sprachcodierer mit einer Bitrate von 10.7 kbit/s und deutlich verbesserter Sprachqualität gegenüber dem derzeit im GSM-Netz verwendeten RPE-LTP-Codec [1] beschrieben. Er basiert auf dem CELP<sup>1</sup>-Codierverfahren mit LSF<sup>2</sup>-Vektorquantisierung, adaptivem Codebuch und strukturierten Codebüchern, ähnlich dem "Algebraic CELP"-Ansatz [6]. Anschließend werden Maßnahmen zur Erkennung und Behandlung von Kanalfehlern erläutert. Die besonders empfindlichen Bits (im Sinne der subjektiven Sprachqualität) werden durch mehrere CR-Checks überwacht, so daß im Fehlerfall nicht alle Parameter, sondern nur solche mit Fehlern in "wichtigen" Bits durch Extrapolationsverfahren ersetzt werden. Als Kanalcodierer wird ein Faltungscoder der Rate 1/2 eingesetzt. Das Gesamtsystem ist so angelegt, daß es möglichst einfach in bestehende GSM-Systeme integriert und durch Festkomma-Signalprozessoren mit einer Rechenleistung von etwa 20 MIPS realisiert werden kann.

## 1. SPRACHCODEC

Der Sprachcoder ist in Abbildung 1 dargestellt. Er verarbeitet Rahmen von 160 Sprachabstastwerten bzw. 20 ms Dauer (bei einer Abtastrate von 8kHz), die wiederum in 4 Unterrahmen mit je 40 Abstastwerten unterteilt werden. Die Nummern in Klammern an den Blöcken in Abbildung 1 kennzeichnen die zeitliche Reihenfolge der Verarbeitungsschritte. Die Unterrahmenverarbeitung ist für Unterrahmen 3 dargestellt.

### 1.1. Rahmenverarbeitung

Zunächst wird die mittlere Leistung des Eingangssprachblocks bestimmt, die danach logarithmiert und mit Hilfe einer Tabelle mit 5 Bit quantisiert wird. Die quantisierte Leistung wird für die

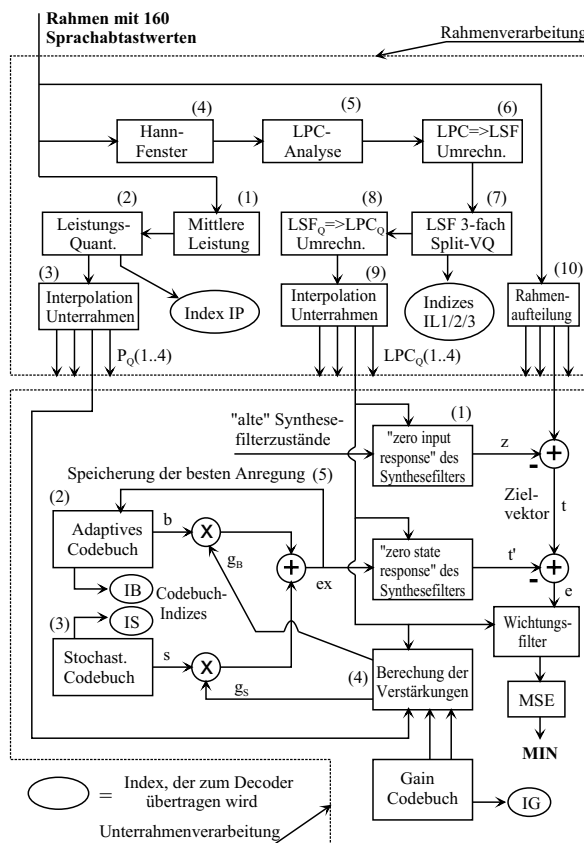


Abbildung 1: Sprachcoder

spätere Verwendung bei der Verstärkungsquantisierung in den Unterrahmen interpoliert. Weiterhin wird für jeden Rahmen eine LPC-Analyse der Ordnung 10 mit Hilfe der Autokorrelationsmethode durchgeführt. Anschließend werden die LPC-Koeffizienten mit Hilfe des Algorithmus in [4] in LSF (Line Spectrum Frequencies) umgerechnet. Danach werden die 10 LSF-Werte mit 25 Bit in 3 Gruppen mit 9,8 und 8 Bit, wie in [5], vektorquantisiert. Die erforderlichen drei Codebücher wurden mit dem LBG-Algorithmus [8] trainiert. Die quantisierten LSF-Werte werden in LPC-Koeffizienten "zurückgerechnet" und für die Verwendung in den Unterrahmen wie in [3] interpoliert.

### 1.2. Unterrahmenverarbeitung

In der Unterrahmenverarbeitung wird ein "synthetisiertes" Signal gesucht, das eine möglichst

<sup>1</sup>Code Excited Linear Prediction

<sup>2</sup>Line Spectrum Frequencies

gute Rekonstruktion des Sprachsignals sein soll. Als Kriterium hierfür wird der mittlere quadratische Fehler (**M**ean **S**quared **E**rror) verwandt. Aus einer beschränkten Menge möglicher Anregungssignale, den Anregungscodebüchern, wird für ein Allpol-Synthesefilter das Signal ausgesucht, das nach der Filterung den MSE minimiert. Die Koeffizienten des Synthesefilters werden der LPC-Analyse aus der Rahmenverarbeitung entnommen. Um die subjektive Sprachqualität zu erhöhen, wird der Fehler "e" vor der Berechnung des MSE durch ein Wichtungsfiler bewertet, das in Spektralbereichen mit hoher Sprachsignalleistung den Fehler geringer gewichtet und umgekehrt. Als Nullstellen des Filters werden die Pole des Synthesefilters gewählt, während die Pole gegenüber denen des Synthesefilters durch Multiplikation mit einem Faktor  $\gamma (=0.8)$  radial in Richtung des Ursprungs der z-Ebene verschoben werden. Diese Wahl ermöglicht die Vereinigung von Synthesefilter und Wichtungsfiler zu einem "gewichteten Synthesefilter": das Wichtungsfiler wird über die Summationspunkte geschoben. So entsteht kein zusätzlicher Rechenaufwand, bis auf die einmalige Wichtungsfilerung des Eingangssignals. Eine weitere Maßnahme zur Vereinfachung der Suche in den Anregungscodebüchern besteht in der Subtraktion des Ausschwinganteils, (zero-input-response) des gewichteten Synthesefilters vom wichtungsgefilterten Eingangssignal. Dadurch kann die Filterung während der Suche in den Codebüchern immer mit "zero-states" erfolgen. Der Ausschwinganteil wird aus den Eingangs- und Ausgangswerten des gewichteten Synthesefilters nach der Filterung des besten Anregungssignals im vorigen Rahmen und den Filterkoeffizienten des aktuellen Rahmens bestimmt. Es entsteht so ein modifizierter Zielvektor  $t'$  für die Suche in den Codebüchern.

Das Anregungssignal "ex" für das Synthesefilter ist die Summe eines adaptiven Anteils "b" und eines stochastischen Anteils "s", die Codebüchern entnommen und durch die Verstärkungsfaktoren  $g_B, g_S$  gewichtet werden. Im adaptiven Codebuch werden die "besten" Anregungssignale vergangener Unterrahmen gespeichert, um Ähnlichkeiten, die vor allem in periodischen Signalabschnitten (stimmhafte Sprache) auftreten, zur Bildung zukünftiger Anregungen auszunutzen. Der Codebuchindex IB codiert hier den Startpunkt des besten adaptiven Anregungsanteils in der Signalvergangenheit, der auch als "lag" bezeichnet wird. Die Suche im adaptiven Codebuch wird zuerst durchgeführt. Hierbei werden die für jeden dort entnommenen Codevektor opti-

malen Verstärkungsfaktoren verwandt, d.h. das Verstärkungscodebuch wird nicht benötigt, um den besten Codebucheintrag zu finden. Anschließend wird der Fehler, der nach der Verwendung des optimal verstärkten adaptiven Anregungsanteils verbleibt, durch die Wahl eines Codevektors aus dem festen, stochastischen Codebuch weiter verringert. Auch hier wird bei der Suche für jeden Codevektor optimale Verstärkungsfaktor benutzt. Anschließend werden durch eine Suche im Gaincodebuch die Verstärkungsfaktoren vektorquantisiert. Danach wird das "beste" Gesamt-Anregungssignal durch das Synthesefilter gefiltert, um die Zustände für den nächsten Rahmen zur Berechnung des Ausschwinganteils zu bestimmen. Abschließend wird das beste Anregungssignal im adaptiven Codebuch gespeichert, und die "ältesten" 40 Abtastwerte werden gelöscht. Die beschriebene Unterrahmenverarbeitung ist eine in CELP-Codern gebräuchliche Variante. Die Unterschiede bei den Codierern liegen vor allem in der Art der Anregungscodebücher, die für den vorgeschlagenen Codec im folgenden beschrieben werden.

### 1.2.1. *Adaptives Codebuch*

Das adaptive Codebuch wird über einen Versatz (lag) von 20..141 Abtastwerten in die Signalvergangenheit abgesucht. Für den Bereich 20..39 ist der entstehende Signalvektor kürzer als die Unterrahmenlänge von 40 Abtastwerten. Für diesen Fall wird mit  $2 \cdot \text{lag}$  zurückliegenden Abtastwerten aufgefüllt ("virtual lags" [3]). Das adaptive Codebuch enthält neben zurückliegenden Samples, die im Abtastraster liegen, auch (je nach lag-Wert) bis zu 5 interpolierte Zwischenwerte ("fractional lags"), um bei Signalen, deren Periode kein ganzzahliges Vielfaches einer Abtastperiode ist, die Signalperiode besser zu "treffen". Im Gegensatz zu anderen Codec-Designs wird hier das adaptive Codebuch in jedem Unterrahmen vollständig abgesucht. Die normalerweise entstehenden Komplexitätsprobleme konnten durch Näherungen bei den Korrelationsberechnungen, die zur Codevektorauswahl notwendig sind, ohne großen Qualitätsverlust verringert werden. Der lag-Index IB wird mit 8 Bit codiert.

### 1.2.2. *Stochastisches Codebuch*

Als stochastisches Anregungssignal werden 10 systematisch plazierte +1/-1-Pulse innerhalb des 40 Abtastwerte langen Codevektors verwendet. Dadurch ist es unnötig, ein Codebuch abzuspeichern. Das stochastische Anregungssignal wird durch den Index IS mit insgesamt 30 Bit codiert, für je-

Tabelle 1: Bitverteilung

Parameter: Index	Bit/Unter-Rahmen	Bit/Rahmen
LPC-Koeffizienten: IL1, IL2, IL3		9+8+8
mittl. Leistung: IP		5
adapt. Anregung: IB	8	32
Stoch. Anregung :		
Pulspositionen: $IS_p$	$2 * 10$	80
Pulszeichen: $IS_s$	$1 * 10$	40
Verstärkungen: IG	8	32
Summe		214

den Puls 2 Bit für 4 mögliche Positionen und ein Bit für das Vorzeichen. Durch die Vorgabe von zu testenden Puls/Vorzeichenkombinationen wird der Aufwand gegenüber der Absuche der theoretisch  $2^{30}$  möglichen drastisch reduziert. Das stochastische Codebuch ist ähnlich aufgebaut wie im ACELP-Ansatz [6].

### 1.2.3. Verstärkungsquantisierung

Die Anteile des Anregungssignals werden durch die Verstärkungen  $g_B$  und  $g_S$  skaliert. Sie werden gemeinsam innerhalb der Analyse-durch-Synthese-Schleife ("closed-loop") durch ein 8-Bit-Codebuch quantisiert, das mit Hilfe des LBG-Algorithmus trainiert wurde [8]. Das Codebuch enthält zwei Komponenten, den leistungsnormalisierten Beitrag  $P_0$  der adaptiven Anregung an der Gesamtanregung und den Korrekturfaktor  $G_S$ , der Näherungen bei der Berechnung der normalisierten Beiträge kompensiert. Der zugehörige Index wird mit  $IG$  bezeichnet. Die Verstärkungsquantisierung ist aus der im 8kbit/s-VSELP-Codec [3] abgeleitet.

### 1.3. Bitverteilung

Tabelle 1 stellt die Verteilung der Bits auf die Parameter des Sprachcoders zusammen. Der Codierer liefert 214 Bits für jeden Rahmen von 160 Abtastwerten. Dadurch ergibt sich eine Bitrate von **10.7 kbit/s** für die Sprachcodierung.

### 1.4. Decoder und Nachfilter

Abbildung 2 zeigt das Blockdiagramm des Decoders. Wie bei den meisten Codierern, die auf dem Analyse-durch-Synthese Prinzip beruhen, ist der Decoder Bestandteil des Coders, so daß viele Verarbeitungsschritte übernommen werden können. Die empfangenen Indizes werden durch Auslesen der korrespondierenden Codebucheinträge decodiert und zunächst zum Summenanregungssignal "ex" zusammengesetzt. Die Synthesefilterkoeffizienten werden den LSF-Codebüchern ent-

nommen und in LPC-Koeffizienten umgerechnet. Sie werden allerdings, anders als im Coder, nicht gewichtet, weil das Wichtungsfiler lediglich der Auswahl der im hörakustischen Sinn besten Codevektoren diene. Das Summenanregungssignal wird durch das Synthesefilter und abschließend durch das nur im Decoder vorhandene Nachfilter verarbeitet. Der subjektive Qualitätseindruck der Ausgangssprache soll durch das Nachfilter verbessert werden, indem zunächst im Spektrum die Linien im Abstand der Sprachgrundfrequenz stärker betont werden. Anschließend werden ebenfalls die Formanten durch ein entsprechendes Filter verstärkt, dessen Koeffizienten aus den quantisierten LPC-Koeffizienten bestimmt werden. Es handelt sich um eine angepaßte Version des Nachfilters aus dem LD-CELP-Algorithmus [2], wobei aber, im Gegensatz dazu, die Koeffizienten nicht rückwärtsadaptiv berechnet zu werden brauchen, weil sie dem Decoder übermittelt werden.

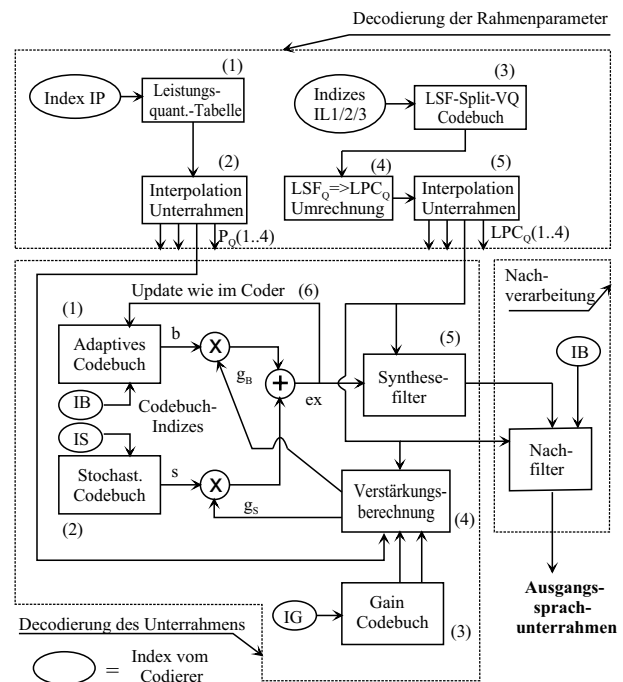


Abbildung 2: Sprachdecoder

### 1.5. Empfindlichkeit gegen Bitfehler

Für die Anwendung in Mobilfunknetzen ist, neben der "Clear-Channel"-Sprachqualität, vor allem die Robustheit des Coders bei Kanalfehlern von Bedeutung. In den folgenden Untersuchungen wurden einzelne Bits im Bitstrom vom Coder zum Decoder systematisch alle 10 Rahmen gestört. Die Störung der Parameter wurde 1,2,4 und 8 Rahmen hintereinander wiederholt, um die Beeinträchtigung der resultierenden Sprachqualität auch bei Bündelfehlern zu ermitteln. Die Qualitätsbeeinträchtigung, die durch die Zahlen

Tabelle 2: Qualitätsbeeinträchtigung

Qualitätseindruck	Bewertung
kaum hörbar	1
wenig störend	2
störend	3
sehr störend	4
katastrophal	5

in Tabelle 2 quantifiziert wird, wurde durch informelle Hörtests bestimmt. Die Ergebnisse für 1,2,4 und 8 Störungen hintereinander wurden gemittelt und in Abbildung 3 dargestellt. Bitnummer 1 entspricht der Stelle des Index im Binärcode mit der geringsten Wertigkeit. Für die systematische Störung eines Pulses aus dem stochastischen Codebuch ergibt sich für alle Bits von Position und Vorzeichen eine wenig störende Qualitätsbeeinträchtigung, die mit 2 bewertet wird. Die entsprechenden Diagramme wurden aus Platzgründen weggelassen. Zusammenfassend läßt sich feststel-

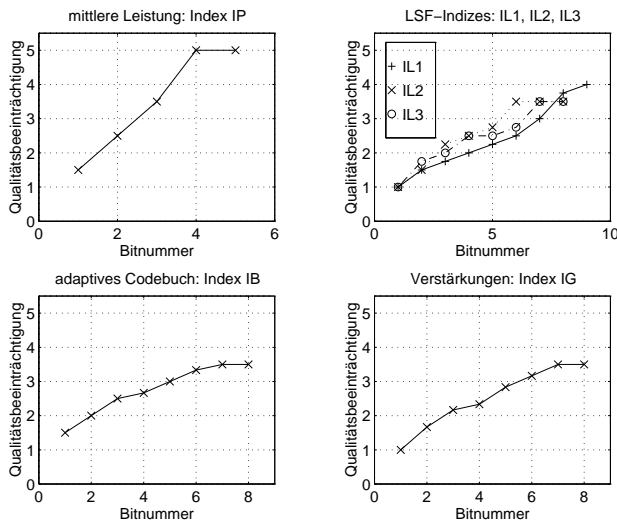


Abbildung 3: Qualitätsbeeinträchtigung durch systematische Störung einzelner Bits in Sprachcoderparametern.

len, daß die Bits 5,4 des Index IP der mittleren Leistung und die Bits 9,8 des Index IL1, der die ersten 3 LSF-Frequenzen quantisiert, die empfindlichsten sind. Die Empfindlichkeit der Bits steigt monoton mit der Position des Bits im Index.

### 1.6. Parameterabhängigkeiten

Für den Fall von Bitfehlern in Parameterindizes durch Störungen auf dem Übertragungskanal müssen Strategien zur Erkennung und Behandlung der Fehler im Sprachdecoder entwickelt werden. Hier sollen nun zunächst redundanzfreie Verfahren diskutiert werden, für die statistische Untersuchungen durchgeführt werden mußten, um zeitliche Abhängigkeiten der quantisierten Parameter zu messen. Dazu wurden die Verteilungs-

funktionen der Differenzen aufeinanderfolgender Werte eines Parameters berechnet. Beispielhaft sind die Verteilungsfunktionen der logarithmierten mittleren Leistung  $\log P$ , des ersten LSF-Koeffizienten  $L1$ , des "Verstärkungsfaktors"  $P0$  sowie des "lags" in Abbildung 4 dargestellt. Die

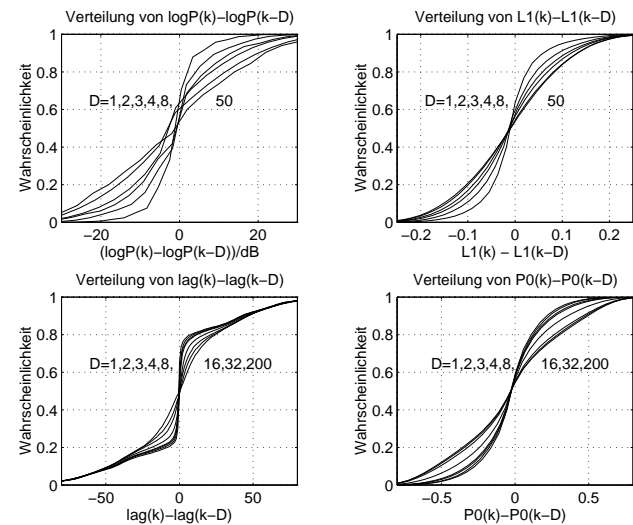


Abbildung 4: Verteilungsfunktionen der Differenzen aufeinanderfolgender Parameter mit D Rahmen / Unterrahmen Abstand

Sprachpausen wurden aus den Darstellungen, außer bei der mittleren Leistung, entfernt. Man erkennt, daß zeitlich aufeinanderfolgende Werte von  $\log P$  mit höherer Wahrscheinlichkeit "ähnliche" Werte besitzen, wenn die Werte einen kleinen zeitlichen Abstand  $D$  haben. Entsprechendes gilt auch für die Verstärkungsfaktoren  $GS$  und  $P0$ , für das  $lag$  und für die LSF-Koeffizienten. Bei der stochastischen Anregung konnte keine Korrelation mit anderen oder zeitlich zurückliegenden Parametern festgestellt werden. Neben den "Autokorrelationen" der Parameter wurden auch deren Abhängigkeiten von anderen Parametern untersucht. Zum einen ergeben sich mit hoher Wahrscheinlichkeit Wertebereiche für die Differenzen aufeinanderfolgender LSF-Indizes innerhalb eines Rahmens, die deutlich stärker eingeschränkt sind als durch die bekannte Stabilitätsbedingung [4], daß die LSF-Werte monoton mit ihrem Index steigen müssen. Außerdem gibt es eine Abhängigkeit zwischen  $P0$ , dem relativen Anteil der adaptiven Anregung an der Gesamtanregung des Synthesefilters, und dem "lag". Für Werte  $P0 > 0.7$  ist der  $lag$ -Änderungsbereich, nachdem eine Korrektur der "lags" von ganzzahligen Teilen und Vielfachen auf die Grundperiode erfolgt ist, sehr viel stärker eingeschränkt, als durch Abbildung 4 gegeben. Darüberhinaus ergaben sich aber keine zur Erkennung oder Behandlung von Kanalfehlern ausnutzbaren Abhängigkeiten.

### 1.7. Erkennung und Behandlung von Kanalfehlern ohne zusätzliche Redundanz

Aus den Verteilungsfunktionen der Parameterdifferenzen können Wertebereiche ermittelt werden, innerhalb derer der Parameter in Folgerahmen mit einer wählbaren Wahrscheinlichkeit liegen wird. Eine Verletzung dieser "Änderungsgrenzen" ist dann, mit der gewählten Wahrscheinlichkeit, durch einen Kanalfehler verursacht und kann zur Fehlererkennung verwandt werden. Eine solche redundanzfreie Fehlererkennung ist aber nur sinnvoll, wenn der Wertebereich einer Parameteränderung sich mit sehr hoher Wahrscheinlichkeit stark beschränken läßt. Das ist aber bei keinem der Parameter zu jedem Zeitpunkt der Fall, weil es z.B. starke Leistungssprünge im Signal gibt, die fälschlich als Fehler erkannt würden. Diese Tatsache schlägt sich darin nieder, daß Wahrscheinlichkeiten von 0.9 und mehr gewählt werden müssen, wenn die redundanzfreie Fehlererkennung ohne starke Beeinträchtigung der Clear-Channel-Qualität erfolgen soll. In diesem Fall werden aber längst nicht alle störenden Kanalfehler erkannt. Aus diesem Grund bietet sich eine Kombination der redundanzfreien Fehlererkennung mit redundanzbehafteten Verfahren als Ausweg an. Die redundanzfreien Verfahren können dann dazu benutzt werden, aus einer Gruppe von Parametern, für die eine redundante Fehlerdetektion vorliegt, die Parameter zu lokalisieren, welche wirklich gestört sind. Die praktische Wahl einer Grenze, ab der eine Parameteränderung in diesem Fall als zu groß und damit durch einen Kanalfehler verursacht angesehen werden kann, muß durch Hörtests ermittelt werden.

Als fehlerhaft erkannte Parameter können aufgrund ihrer zeitlichen Abhängigkeit durch ihre alten Werte ersetzt werden. Dieses Verfahren wird aber mit wachsendem Abstand  $D$  immer schlechter funktionieren. Hier ist die Reduktion der Signalenergie nach mehrfach hintereinander gestörten Parametern sowie die Verflachung spektraler Maxima in der LPC-Hüllkurve sinnvoll [7], um unangenehme Störgeräusche, allerdings um den Preis einer Signaldämpfung, zu vermeiden. Bei der Ersetzung der Unterrahmenparameter ist es möglich, zukünftige Werte zur Ersetzung eines Parameters heranzuziehen, ohne dabei das "Delay" der Codierung zu vergrößern. So wird, wenn z.B. der lag-Index des ersten Unterrahmens gestört ist und dies bereits mehrere Unterrahmen in Folge der Fall war, ein eventuell nicht mehr gestörter Wert in zukünftigen Unterrahmen des Rahmens anstelle des letzten ungestörten verwandt, wenn der zeitliche Abstand zu ihm kleiner

Tabelle 3: Grenzbits für die Fehlererkennung

Index	IP	IL1	IL2	IL3	IB	IG
Bit	3	3	3	3	2	4

ist. Entsprechend wird bei den Verstärkungsfaktoren verfahren, die zwischen altem und zukünftigem Wert noch zusätzlich interpoliert werden.

### 1.8. Grenzbits für die redundante Fehlererkennung

Die in Kapitel 1.7 angegebenen Verfahren zur Ersetzung als gestört erkannter Parameter wurden verwandt, um Parameter mit systematisch gestörten Bits (wie in Kapitel 1.5) im Decoder zu ersetzen. Als Grenzbits für eine redundante Fehlererkennung (CR-Checks) werden die Bits im Index eines Parameters angesehen, die, mit dem eingefügten Kanalfehler decodiert, einen etwas schlechteren Höreindruck liefern als die in Kapitel 1.7 angegebenen Verfahren zur Ersetzung gestörter Parameter. Es ergaben sich die in Tabelle 3 zusammengestellten Resultate. Die Bitnummer bezieht sich auf die Binärstelle im Index und ist ab 1 gezählt. Da für die stochastische Anregung keine Möglichkeiten der Parameterersetzung angegeben werden konnten, wurden hierfür auch keine Grenzbits bestimmt.

## 2. VORWÄRTS-FEHLERERKENNUNG UND KORREKTUR

Bei dem vorgestellten Sprachcodec mit einer Bitrate von 10.7 kbit/s verbleibt eine Bitrate von 12.1 kbit/s für Vorwärts-Fehlerkorrekturverfahren, um auf die Summenbitrate von 22.8 kbit/s im GSM-Vollratenkanal zu gelangen. Als Kanalcode wird ein Faltungscodes der Rate 1/2 verwandt, dessen Generatorpolynome vom derzeitigen GSM-Standard [1] übernommen wurden. Er wird mit Hilfe des Viterbi-Algorithmus mit "Soft Decision" decodiert. Die empfindlichsten 198 Bits des Sprachcoders werden durch den Faltungscodes geschützt, während die übrigen  $214-198=16$  Quellcodelbits ohne Fehlerkorrekturverfahren an den Ausgangsbitstrom des Faltungscodes angehängt werden. Es handelt sich hierbei um die 3 Bits eines Pulses der stochastischen Anregung und das unwichtigste Bit des gain-Index IG in jedem Unterrahmen. Innerhalb des Faltungscodes werden sechs 3-Bit-CR-Checks verwandt, um Kanalstörungen, die nach der Faltungsdecodierung verbleiben, zu erkennen. Ihre Generatorpolynome wurden vom GSM-Standard übernommen. Durch die Verwendung mehrerer CR-Checks können Kanalfehler im Bitstrom besser lokalisiert werden, so daß Parameter-

Tabelle 4: Vorwärts-Fehlerkorrektur/Erkennung

Parameter: Index	faltungs- codierte S-MSB	CRC-ge- schützte S-MSB	CRC- Nummer
mittl. Leistung: IP	5 (alle)	3	1
LSF 1..3: IL1	9 (alle)	7	2
LSF 4..6: IL2	8 (alle)	6	2
LSF 7..10: IL3	8 (alle)	6	2
lag-index: IB	8 (alle)	7	3(..6)
gain-index: IG	7 von 8	5	3(..6)
stoch. Codebuch			
Positionen: $IS_p$	18 von 20	-	-
Vorzeichen: $IS_s$	9 von 10	-	-

Ersetzungsverfahren nur dann erfolgen, wenn wichtige Bits von Parametern, siehe Tabelle 3, gestört sind. Wird ein Kanalfehler erkannt, so werden die Parameter-Ersetzungsverfahren aus Kapitel 1.7 im Sprachdecoder durchgeführt. Tabelle 4 gibt an, welche Bits des Sprachcoders vom Faltungscodierverfahren und von CR-Checks geschützt werden. Die Bitangaben beziehen sich immer auf die im subjektiven Sinn wichtigsten Bits (S-MSB) eines Parameters nach Abbildung 3. Zusammengefaßt werden 198 Quellcorderbits, 18 CRC-Bits und 4 Tail-Bits (zur Löschung des Faltungscodiergedächtnisses), d.h. insgesamt 220 Bits, auf den Eingang des 1/2-Faltungscoders gegeben. Dieser liefert 440 Ausgangsbits, an die noch die 16 ungeschützten Bits angehängt werden. Insgesamt ergeben sich damit 456 Bits pro 20 ms am Kanalcoderausgang bzw. eine Bitrate 22,8 kbit/s. Die weitere Verarbeitung, wie z.B. "Re-ordering" und "Interleaving", wurden vom GSM-Standard [1] übernommen.

### 2.1. Kombination der Fehlererkennungsverfahren mit und ohne Redundanz

Der Kanaldecodierer liefert 6 CRC-"Flags", die Kanalfehler innerhalb der Parametergruppen aus Tabelle 4 anzeigen. Falls z.B. der CRC mit der Nummer 3, der lag- und gain-Index des ersten Unterrahmens überwacht, Kanalfehler anzeigt, dann werden die Fehlererkennungsverfahren aus Kapitel 1.7 verwandt, um zu entscheiden, welche der Parameter gestört sind. Entsprechend wird für die übrigen Unterrahmen und für die LSF-Indizes verfahren. Für einen als gestört erkannten Parameter wird das zugehörige Ersetzungsverfahren aus Kapitel 1.7 benutzt. Falls mehr als 3 CR-Checks Kanalfehler detektieren, wird der Rahmen als "Lost-Frame" markiert. In diesem Fall werden die redundanzfreien Fehlererkennungsverfahren für alle Parameter, unabhängig vom CR-Check, der dem Parameter zugeordnet ist, durchgeführt, und bei Fehlerdetektion erfolgt die Parameterextrapolati-

on. Auf diese Weise kann die redundanzfreie Fehlererkennung, auch ohne parametereigene CRC-Detektion, verwandt werden, ohne daß die Clear-Channel-Qualität des Sprachcodierers verringert wird.

Die Wertebereiche der "erlaubten" Parameteränderungen für die redundanzfreie Fehlererkennung wurden durch Hörtests gemeinsam für alle Parameter, die durch einen CR-Check überwacht werden, optimiert, so daß sich die beste Sprachqualität ergab. Das Problem besteht darin, daß ein schmaler Wertebereich zu einer guten Erkennung führt, wenn ein Parameter wirklich gestört ist, allerdings treten dann im ungestörten Fall häufige Fehldetektionen auf, die eine Lokalisation von Fehlern bei anderen, in Wahrheit gestörten, Parametern erschweren. Das angegebene Verfahren führt zu einem brauchbaren Kompromiß.

### 3. SPRACHQUALITÄT

Die Sprachqualität ist deutlich besser als die des derzeit im GSM-System verwendeten Codec. Das gilt sowohl für den "Clear-Channel"-Fall als auch für Situationen mit stark gestörten Kanälen. Insbesondere die lauten "Knackser" und starken Lautstärkesprünge konnten durch die verbesserte Erkennung und Behandlung von Kanalfehlern stark reduziert werden. Die Clear-Channel-Qualität ist etwas besser als die des LD-CELP [2] und etwas schlechter als bei 64-kbit/s-log-PCM.

### 4. LITERATURVERZEICHNIS

- [1] ETSI/GSM, Rec. 06.10, "Full-rate speech transcoding", March 1992 und ETSI/GSM, Rec. 05.03, Channel Coding, Feb. 1992
- [2] ITU-T Rec. G.728, "Coding of Speech at 16kbit/s using Low-Delay Code Excited Linear Prediction"
- [3] Electronic Industries Association, Cellular Systems Dual-Mode Subscriber Equipment - Network Equipment Compatibility Specification, Dec. 1989
- [4] F.K. Soong, B.H. Juang, "Line Spectrum Pair (LSP) And Speech Data Compression", Proc. ICASSP, pp. 1.10.1-1.10.4, 1984
- [5] K.K. Paliwal, B.S. Atal, "Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame", IEEE Trans. on Speech and Audio Processing, Vol. 1, No. 1, pp. 3-14, January 1993
- [6] J-P. Adoul, P. Mabillean, M. Delprat and S. Morissette "Fast CELP coding based on algebraic codes", Proc. ICASSP, pp. 1957-1960, 1987
- [7] Kondo, "Digital Speech Coding for Low Bitrate Communications Systems", Wiley & Sons, 1994
- [8] Y. Linde, A. Buzo, R.M. Gray, "An Algorithm for Vector Quantizer Design", IEEE Trans. on Communications, Vol. COM-28, No. 1, January 1980