ECHTZEITIMPLEMENTIERUNG EINES CELP-CODEC MIT UNGLEICHGEWICHTIGEM FEHLERSCHUTZ

Kai Clüver, Thomas Gries, Hui Li

Institut für Fernmeldetechnik, Technische Universität Berlin, Einsteinufer 25, D-1000 Berlin 10

1. Einführung

Das hier beschriebene Vorhaben diente der Entwicklung eines digitalen Sprachcodecs für den Einsatz in satellitengestützten Sprechfunkverbindungen zwischen Zivilflugzeugen und Bodenstationen.

Heutige Sprechfunkverbindungen sind bei transkontinentalen Flügen unsicher und häufig von sehr schlechter Qualität, da sie mit analoger Kurzwellenübertragung arbeiten. Mit Hilfe rechenintensiver Verfahren der Quellencodierung (für Sprachsignale) und Fehlerkorrekturverfahren kann eine grundlegende Verbesserung der Übertragungsqualität erreicht werden. Aus Gründen der Frequenzökonomie und der Realisierung von Sendern und Empfängern mit sehr kleinen Antennen ergibt sich allerdings eine Beschränkung der Netto-Übertragungsrate auf Werte um etwa 3..6 kbit/s.

Auch bei diesen geringen Bitraten wird eine transparente Sprachübertragung gefordert; mit den heute bekannten in Echtzeit arbeitenden Verfahren kann Transparenz aber nur angenähert erreicht werden. Wir entwickelten ausgehend von der Simulation bekannter niederratiger Sprachcodierverfahren ein für die Echtzeitanwendung geeignetes Verfahren, das als CELP-Coder (code-excited linear predictive) zu den Analysis-by-synthesis-Verfahren gezählt wird.

Die störenden Einflüsse des Satellitenkanals (Bitfehler durch Signaleinbrüche aufgrund des Fadingverhaltens) werden dabei mit den Verfahren der Vorwärtsfehlerkorrektur und mit angepaßten Modulationsverfahren auf ein Minimum reduziert.

2. CELP-Codierung des Sprachsignals

2.1. Beschreibung des CELP-Codierers

Die Struktur des CELP-Codierers ist in Bild 1 dargestellt. Das Sprachsignal wird zunächst in Segmente von 20 ms Länge aufgeteilt, für die dann das LPC-Synthesefilter und das Anregungssignal bestimmt werden. Die LPC-Analyse (12.Ordnung) erfolgt mit Hilfe des Autokorrelationsverfahrens.

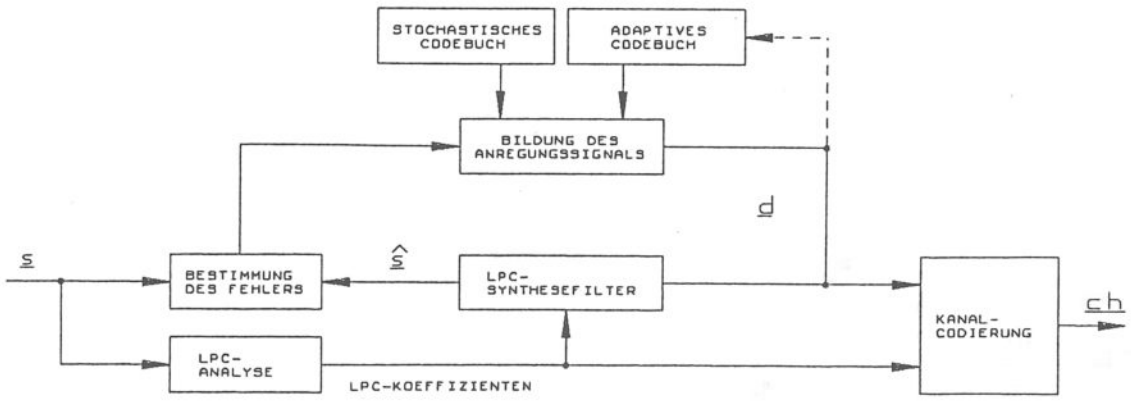

Bild 1: CELP-Coder

Das Anregungssignal wird für Teilsegmente von 5 ms Länge bestimmt. Es besteht aus einer Linearkombination je eines Vektors des festen stochastischen Codebuches und des adaptiven Codebuches, das als Langzeitprädiktor dient. Zur Optimierung werden die Codebücher nacheinander vollständig durchsucht; für jede Anregung wird im Codierer eine Synthese des Sprachsignals durchgeführt und ein mittlerer quadratischer Fehler zum Eingangssignal ermittelt. Das Anregungssignal, das den geringsten Fehler ergibt, wird nach der Codebuchsuche dazu verwendet, den Zustandsspeicher des Synthesefilters und das adaptive Codebuch zu aktualisieren und damit den Anfangszustand für die Bearbeitung des folgenden Teilsegmentes herzustellen.

Die Parameter des optimalen Anregungssignals und die LPC-Koeffizienten werden quantisiert und zum Decodierer (Bild 2) übertragen. Hier wird das Anregungssignal aus den Anteilen beider Codebücher gebildet, das adaptive Codebuch aktualisiert und die Synthese des Sprachsignals durchgeführt. Zur Verbesserung der subjektiven Sprachqualität durchläuft das synthetisierte Signal ein adaptives Postfilter, das die relativen Maxima des Spektrums überhöht und dadurch zur Verdeckung von Störgeräuschen beiträgt.

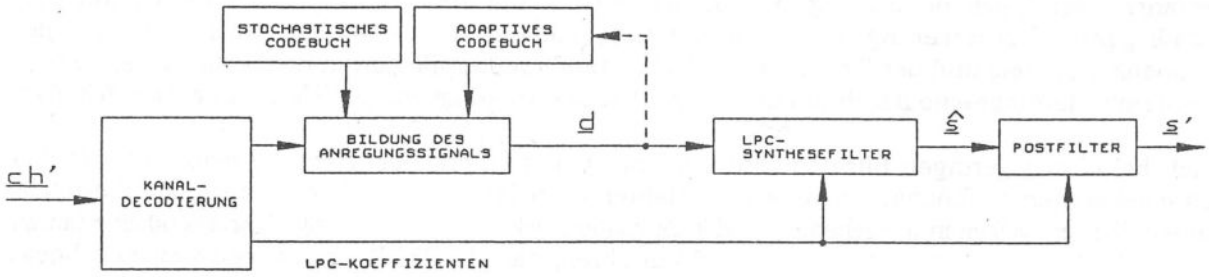

Bild 2: CELP-Decoder

2.2. Codebuchsuche

Als Gütekriterium bei der Codebuchsuche wird ein spektral gewichteter mittlerer quadratischer Fehler verwendet. Dabei fallen Abweichungen in Spektralbereichen geringer Energie stärker ins Gewicht als in Bereichen, in denen die Störung durch das Nutzsignal verdeckt wird (z.B. Formanten). Die Gewichtung geschieht derart, daß zum einen das Sprachsignal ein Gewichtungsfilter mit der Übertragungsfunktion

$$W(z) = A(z) \;/\; A(z/\alpha)$$

durchläuft und zum anderen die Synthese im Codierer durch ein gewichtetes Rekonstruktionsfilter

$$S(z) = 1 \;/\; A(z/\alpha)$$

durchgeführt wird. Dabei ist

$$A(z) = 1 - \sum_{i=1}^{12} a_i \cdot z^{-i} \quad (a_i\text{: Prädiktorkoeffizienten})$$

und der Gewichtungskoeffizient $\alpha = 0{,}8$.

Das Blockschema der Codebuchsuche ist in Bild 3 dargestellt: Das spektral gewichtete Sprachsignal $\underline{w}$ dient zunächst für die Codebuchsuche als Referenz. Das Ausgangssignal des Synthesefilters läßt sich nun in zwei Komponenten aufspalten: das Nachklingen, das von der Anregung der vorhergehenden Teilsegmente herrührt, (zero-input response) und die Antwort des leeren Filters auf das aktuelle Anregungssignal (zero-state response). Das Nachklingen ist vor der Bestimmung der Anregung bereits bekannt und wird von $\underline{w}$ subtrahiert; das Filtern eines Vektors kann nun durch eine Matrixmultiplikation ersetzt werden /ATA86/. Die Codebuchsuche wird nacheinander im adaptiven und im stochastischen Codebuch durchgeführt.

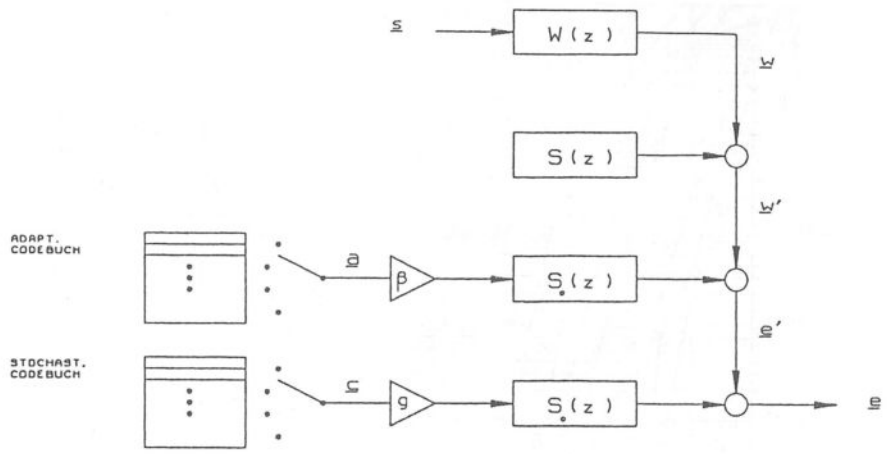

Bild 3: Codebuchsuche

Das adaptive Codebuch enthält das Anregungssignal der vorhergehenden Teilsegmente und dient in erster Linie dazu, die Periodizität stimmhafter Sprachsignale nachzubilden. Bei der Suche wird für einen vorgegebenen Bereich von Perioden jeder entsprechende Teilbereich (= Vektor) des Codebuches untersucht. Somit unterscheiden sich benachbarte Vektoren nur durch eine Verschiebung um einen Abtastwert, was eine deutliche Verringerung des Rechenaufwandes ermöglicht. Für Perioden unter 5 ms sind die Vektoren nur unvollständig bekannt, da sie teilweise Abtastwerte des noch zu bestimmenden Teilsegmentes enthalten müßten. In diesem Fall werden durch periodische Wiederholung der bekannten Anteile virtuelle Vektoren gebildet /KLE88/.

Das feste stochastische Codebuch enthält eine Pulsfolge, deren Abtastwerte zu ca. 82 % gleich Null sind ("sparse codebook"); es konnte festgestellt werden, daß ein derartiges Codebuch subjektiv wie objektiv zu einer Erhöhung der Sprachqualität gegenüber Gauß-Codebüchern führte. Zur Verringerung der Rechenzeit wurden wie beim adaptiven Codebuch überlappende Vektoren vorgesehen.

3. Kanalcodierung mit ungleichförmigem Fehlerschutz

3.1. Ermittlung der Bitfehlerempfindlichkeit

Der Quellencodierer sendet alle 20 ms einen Rahmen von 130 bit aus. Ein effektiver Schutz des codierten Sprachsignals vor Kanalfehlern erfordert einen der unterschiedlichen Empfindlichkeit der Bits entsprechenden ungleichförmigen Fehlerschutz. Die relative Fehlerempfindlichkeit der einzelnen Bits wurde durch informelle Hörvergleiche ermittelt und durch Simulation entsprechend gewichteter Bitfehlerraten nachgeprüft. Diese Vorgehensweise ergab eine deutlich stärkere Aussagekraft der Ergebnisse als eine Bestimmung der Empfindlichkeit nach objektiven Kriterien.

Die Bits des 20-ms-Rahmens werden im Codierer nach aufsteigender Fehlerempfindlichkeit sortiert und so für die Kanalcodierung durch einen RCPC-Code /HAG88/ vorbereitet.

3.2. Kanalcodierung und Modulation

Die in unserem System eingesetzten RCPC-Codes ermöglichen es, aufgrund ihrer unterschiedlichen Coderaten einen abgestuften Fehlerschutz zu realisieren. RCPC-Codes gehören zur Gruppe der Faltungscodes. Ausgehend von einem Basiscode (z.B. der Rate 1/3) erhalten wir durch verschiedene Punktierungsvorschriften Codes variabler Rate, die aber alle durch denselben Coder bzw. Decoder bearbeitet werden können. Es kommen Codes der Raten 8/9, 8/10 und 8/11 zur Anwendung, wobei zur Decodierung der Viterbi-Algorithmus benutzt wird.

Die durch die informellen Hörtests ermittelten noch zulässigen Restfehlerwahrscheinlichkeiten der einzelnen Bits werden durch die genannten punktierten Coderaten angenähert (s. Bild 4); ist der Kanal-SNR E_s/N_0 besser als ca. 6,7 dB, werden sie deutlich unterschritten.

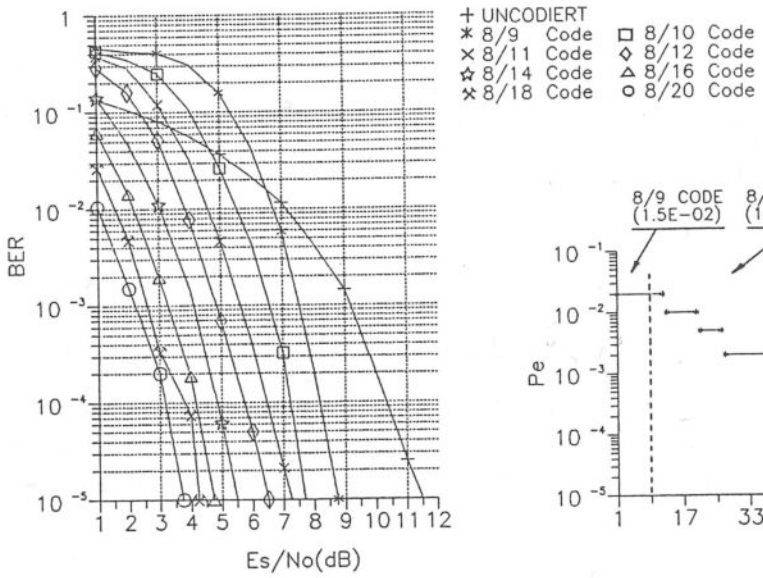

Bild 4a: BitfehlerwahrscheinlichkeitenBER ausgewählter RCPC-Codes (R=8/(8+l), M=4) bei QPSK und AWGN-Kanal; Simulation mit 'Soft-decision'.

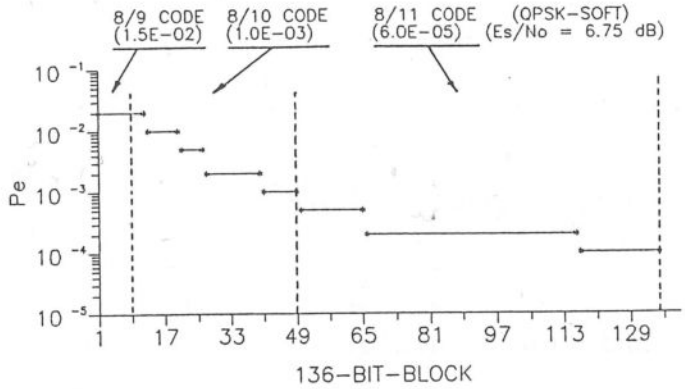

Bild 4b: Restfehlerwahrscheinlichkeit p_e einzelner CELP-Bitgruppen eines Blocks bei ausgewähltem E_s/N_0.

Die im Bild 4b genannte Blocklänge von 136 Bit ergibt sich implementationsbedingt durch den mit der Periode 8 arbeitenden Coder und beinhaltet am Rahmenende zusätzliche Null-Bits zum Speicherlöschen ('geblockte Faltungscodierung'). Die Decodierung wird um 24 Bit verzögert. Mit größeren Verzögerungen wären geringere Restfehlerwahrscheinlichkeiten erreichbar.

Als Modulationsverfahren wurde die DQPSK ausgewählt. Der Aufbau des Modulators und Demodulators wird vom Projektpartner (DLR Oberpfaffenhofen) vorgenommen.

Empfängerseitig übergibt der DQPSK-Demodulator die demodulierten Inphasen- und Quadraturwerte in grob quantisierter Form dem Kanaldecoder, der mit diesen Werten eine Soft-Decision durchführen kann.

4. Aufbau der Echtzeit-Hardware

Das Codec wurde auf einer Doppelprozessorkarte implementiert, die im Rahmen dieses Vorhabens als PC/AT-Einsteckkarte entwickelt wurde ("TU2DSP", siehe Bild 5). Jeder der beiden DSP32C-Signalprozessoren kann auf maximal 128 KWorte RAM in 40ns zugreifen.

Für die DSP-DSP-Kommunikation wurde eine einfache und schnelle PAL-Logik entworfen. Die Kommunikation nach außen erfolgt (mit Ausnahme des PC/AT-Anschlusses) über die seriellen Schnittstellen zum Kanal-Interface (Anschluß des Modulators / Demodulators über RS-422) bzw. über das auf der Karte befindliche PCM-Codec (8-kHz-Abtastfrequenz). Für den Einsatz ohne PC/AT-Steuerrechner kann die Prozessorkarte um eine Stand-alone-Bootlogik erweitert werden.

Ein komfortables Bedienprogramm ("TU1") lädt die Programme; es unterstützte die schnelle Optimierung des Verfahrens durch die Möglichkeit der Parametervariation während des laufenden Betriebs.

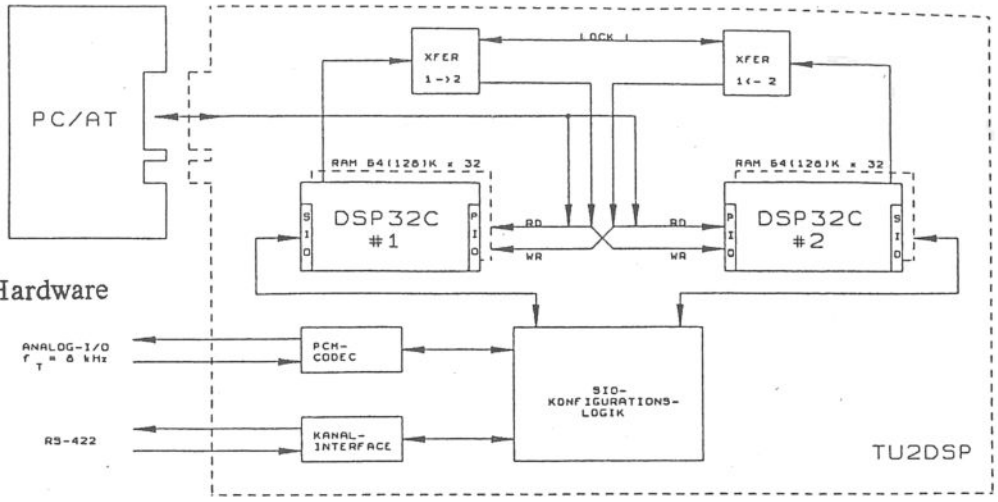

Bild 5: Echtzeit-Hardware

5. Zusammenfassung der Ergebnisse

Bei der Umsetzung der in der Simulation optimierten Verfahren erwies sich die Rechenkapazität des als Coder benutzten Signalprozessors als zu gering. Da die Codebuchsuche den überwiegenden Teil der Rechenzeit beansprucht, mußte nach der Optimierung des Suchalgorithmus der Codebuchumfang verringert werden. Eine Verkleinerung des adaptiven Codebuches schied aus, da dieses den wichtigeren Beitrag zur Qualität liefert. Daher wurde das stochastische Codebuch nach der zur Verfügung stehenden Rechenzeit bemessen.

Die erreichte Qualität unserer Echtzeitimplementation entspricht derjenigen vergleichbarer CELP-Codecs. Das Verfahren erwies sich als robust gegenüber Hintergrundgeräuschen. Da Eingangssignale mit tieffrequenten Anteilen zu größeren Codierverzerrungen führen, wurde das Codec mit einem eingangsseitigen 200-Hz-Hochpaß versehen.

Bitraten des implementierten Codierverfahrens:

Parameter	bit / s
Bitrate des Quellencodierers	6500
Bitrate des Kanalcodierers	2500
Synchroninformation	600

Bruttorate	9.600

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung und Technologie (Förderungskennzeichen 01 YH 88090) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren. Die Autoren danken den Herren E.Baumann und S.Wimmer für ihre unterstützende Mitarbeit.

Literatur:

/SCH85/ M. R. Schroeder, B. S. Atal: Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates, Proc. ICASSP, pp.937-940, 1985.

/ATA86/ B. S. Atal: High-Quality Speech at Low Bit Rates: Multi-Pulse and Stochastically Excited Linear Predictive Coders, Proc. ICASSP, pp.1681-1684, 1986.

/KLE88/ W. B. Kleijn, D. J. Krasinski, R. H. Ketchum: An Efficient Stochastically Excited Linear Predictive Coding Algorithm for High Quality Low Bit Rate Transmission of Speech, Speech Communication 7, pp.305-316, 1988.

/HAG88/ J. Hagenauer: Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications, IEEE Trans. COM-36, No. 4, pp.389-400, 1988.

/HAG89/ J. Hagenauer: Unequal Error Protection (UEP) for Statistically Time-Varying Channels, ITG-Fachbericht 107, pp.253-261, 1989.