



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 601 21 201 T2 2007.05.31**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 330 818 B1**

(51) Int Cl.⁸: **G10L 19/00 (2006.01)**

(21) Deutsches Aktenzeichen: **601 21 201.0**

(86) PCT-Aktenzeichen: **PCT/IB01/02021**

(96) Europäisches Aktenzeichen: **01 983 716.0**

(87) PCT-Veröffentlichungs-Nr.: **WO 2002/037475**

(86) PCT-Anmeldetag: **29.10.2001**

(87) Veröffentlichungstag

der PCT-Anmeldung: **10.05.2002**

(97) Erstveröffentlichung durch das EPA: **30.07.2003**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **28.06.2006**

(47) Veröffentlichungstag im Patentblatt: **31.05.2007**

(30) Unionspriorität:

702540 31.10.2000 US

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE, TR**

(73) Patentinhaber:

Nokia Corp., Espoo, FI

(72) Erfinder:

**MÄKINEN, Jari, FIN-33100 Tampere, FI; MIKKOLA,
J., Hannu, FIN-33300 Tampere, FI; VAINIO, Janne,
FIN-33880 Lempäälä, FI; ROTOLA-PUKKILA, Jani,
FIN-33820 Tampere, FI**

(74) Vertreter:

Becker, Kurig, Straus, 80336 München

(54) Bezeichnung: **VERFAHREN UND VORRICHTUNG ZUR VERSCHLEIERUNG VON FEHLERHAFTEN RAHMEN
WÄHREND DER SPRACHDEKODIERUNG**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

GEBIET DER ERFINDUNG

[0001] Die vorliegende Erfindung bezieht sich allgemein auf das Dekodieren von Sprachsignalen aus einem kodierten Bitstrom und insbesondere auf das Verbergen beschädigter Sprachparameter, wenn Fehler in Sprachrahmen während der Sprachdekodierung erkannt werden.

HINTERGRUND DER ERFINDUNG

[0002] Sprach- und Audiokodieralgorithmen haben eine breite Vielzahl von Anwendungen in Kommunikations-, Multimedia- und Speichersystemen. Die Entwicklung der Kodieralgorithmen wird durch die Notwendigkeit angetrieben, Übertragungs- und Speicherkapazität zu sparen, während die hohe Qualität des synthetischen Signals aufrecht gehalten wird. Die Komplexität des Kodierers wird beispielweise durch die Verarbeitungsleistung der Anwendungsplattform beschränkt. Bei einigen Anwendungen, beispielsweise der Sprachspeicherung, kann der Kodierer ziemlich komplex sein, während der Dekodierer so einfach wie möglich sein sollte.

[0003] Moderne Sprach-Kodierer/Dekodierer arbeiten durch das Verarbeiten des Sprachsignals in kurzen Segmenten, die Rahmen genannt werden. Eine typische Rahmenlänge eines Sprach-Kodierers/Dekodierers beträgt 20 ms, was 160 Sprachabstastwerten entspricht, wenn man eine Abtastfrequenz von 8 kHz annimmt. In den Breitband-Kodierern/Dekodierern entspricht die typische Rahmenlänge von 20 ms 320 Sprachabstastwerten, wenn man eine Abtastfrequenz von 16 kHz annimmt. Der Rahmen kann weiter in eine Anzahl von Unterrahmen aufgeteilt werden. Für jeden Rahmen bestimmt der Kodierer eine parametrische Darstellung des Eingangssignals. Die Parameter werden quantisiert und durch einen Kommunikationskanal in digitaler Form übertragen (oder in einem Speichermedium gespeichert). Der Dekodierer erzeugt ein synthetisches Sprachsignal auf der Basis der empfangenen Parameter, wie das in [Fig. 1](#) gezeigt ist.

[0004] Ein typischer Satz extrahierter Kodierparameter umfasst Spektralparameter (wie Parameter der linearen Vorhersagekodierung (LPC)), die bei einer kurzzeitigen Vorhersage des Signals verwendet werden, Parameter, die für eine Langzeitvorhersage (LTP) des Signals verwendet werden, verschiedene Verstärkungsparameter und Ansteuerungsparameter. Der LTP-Parameter ist eng mit der Grundfrequenz des Sprachsignals verbunden. Dieser Parameter wird oft als ein sogenannter Abstands-Verzögerungs-Parameter (pitch-lag parameter) bezeichnet, der die Grundperiodizität in Begriffen von Sprachabstastwerten beschreibt. Auch einer der Ver-

stärkungsparameter steht in sehr enger Beziehung zur Grundperiodizität und wird als LTP-Verstärkung bezeichnet. Die LTP-Verstärkung ist ein sehr wichtiger Parameter, um die Sprache so natürlich wie möglich zu machen. Die Beschreibung der Kodierparameter oben passt allgemein ausgedrückt zu einer Vielzahl von Sprach-Kodierern/Dekodierern, die die sogenannten Code-Excited Linear Prediction (CELP) Kodierer/Dekodierer einschließen, die einige Zeit die erfolgreichsten Sprach-Kodierer/Dekodierer waren.

[0005] Sprachparameter werden durch einen Kommunikationskanal in einer digitalen Form übertragen. Manchmal ändert sich der Zustand des Kommunikationskanals, und das kann Fehler im Bitstrom verursachen. Dies wird Rahmenfehler (schlechte Rahmen) verursachen, das heißt einige der Parameter, die ein spezielles Sprachsegment (typischerweise 20 ms) beschreiben, sind beschädigt. Es gibt zwei Arten von Rahmenfehlern: total beschädigte Rahmen und teilweise beschädigte Rahmen. Diese Rahmen werden manchmal im Dekodierer überhaupt nicht empfangen. In paketbasierten Übertragungssystemen kann, wie in normalen Internetverbindungen, die Situation auftreten, dass das Datenpaket den Empfänger überhaupt nicht erreicht, oder dass das Datenpaket so spät ankommt, dass es aufgrund der Echtzeitnatur gesprochener Sprache nicht verwendet werden kann. Der teilweise beschädigte Rahmen ist ein Rahmen, der beim Empfänger ankommt und der noch einige Parameter enthalten kann, die nicht fehlerhaft sind. Dies ist gewöhnlicherweise die Situation in einer leitungsvermittelten Verbindung, wie in der existierenden GSM-Verbindung. Die Bitfehlerrate (BER) bei teilweise beschädigten Rahmen liegt typischerweise bei ungefähr 0,5–5%.

[0006] Aus der obigen Beschreibung kann man sehen, dass die zwei Fälle der schlechten oder beschädigten Rahmen unterschiedliche Lösungen erforderlich machen, um mit der Verschlechterung der rekonstruierten Sprache durch den Verlust der Sprachparameter fertig zu werden.

[0007] Die verlorenen oder fehlerhaften Sprachrahmen sind Folgen des schlechten Zustands des Kommunikationskanals, der Fehler im Bitstrom verursacht. Wenn ein Fehler im empfangenen Sprachrahmen erkannt wird, wird ein Fehlerkorrekturverfahren gestartet. Dieses Fehlerkorrekturverfahren umfasst gewöhnlicherweise ein Ersetzungsverfahren und ein Stummschaltungsverfahren (muting procedure). Im Stand der Technik werden die Sprachparameter des schlechten Rahmens durch gedämpfte oder modifizierte Werte des vorherigen guten Rahmens ersetzt. Einige Parameter (wie die Ansteuerung in CELP-Parametern) im beschädigten Rahmen können jedoch noch für die Dekodierung verwendet werden.

[0008] [Fig. 2](#) zeigt das Prinzip des Verfahrens des

Stands der Technik. Wie in [Fig. 2](#) gezeigt ist, wird ein Puffer, der als "Parameterverlauf" bezeichnet ist, verwendet, um die Sprachparameter des letzten guten Rahmens zu speichern. Wenn ein schlechter Rahmen erkannt wird, wird die Schlechtrahmenanzeige (BFI) auf 1 gesetzt, und das Fehlerverbergungsverfahren wird gestartet. Wenn die BFI nicht gesetzt ist (BFI = 0), wird der Parameterverlauf aktualisiert, und Sprachparameter werden für die Dekodierung ohne eine Fehlerverbergung verwendet. Im System des Stands der Technik verwendet das Fehlerverbergungsverfahren den Parameterverlauf für das Verbergen der verlorenen oder fehlerhaften Parameter in den beschädigten Rahmen. Einige Sprachparameter können vom empfangenen Rahmen verwendet werden, obwohl er als ein schlechter Rahmen (BFI = 1) klassifiziert ist. Beispielsweise wird in einem adaptiven GSM-Mehrraten-Sprach-Kodierer/Dekodierer (Adaptive Multi-Rate, AMR) (ETSI-Spezifikation 06.91) der Anregungsvektor vom Kanal immer verwendet. Wenn die Sprachrahmen vollständig verlorene Rahmen sind (beispielsweise in manchen IP-basierten Übertragungssystemen), werden keine Parameter vom empfangenen schlechten Rahmen verwendet. In einigen Fällen wird kein Rahmen empfangen oder der Rahmen wird so spät ankommen, dass er als ein verlorener Rahmen klassifiziert werden muss.

[0009] In einem System des Stands der Technik verwendet die LTP-Verzögerungs-Verbergung den letzten guten LTP-Verzögerungswert mit einem leicht modifizierten Bruchteil, und die Spektralparameter werden durch die letzten guten Parameter ersetzt, die leicht zu einer konstanten Mitte hin verschoben sind. Die Verstärkungen (LTP und festes Kodebuch) können gewöhnlicherweise durch den gedämpften letzten guten Wert oder durch den Mittelwert mehrerer letzter guter Werte ersetzt werden. Dieselben substituierten Sprachparameter werden für alle Unterrahmen mit einer leichten Modifikation bei einigen von ihnen verwendet.

[0010] Die LTP-Verbergung des Stands der Technik kann adäquat für stationäre Sprachsignale sein, beispielsweise für stimmhafte Sprache (voiced speech) oder stationäre Sprache (stationary speech). Für nicht stationäre Sprachsignale kann es jedoch sein, dass das Verfahren des Stands der Technik unangenehme und hörbare Artefakte erzeugt. Wenn beispielsweise das Sprachsignal stimmlos oder nicht stationär ist, so hat das einfache Ersetzen des Verzögerungswerts im schlechten Rahmen durch den letzten guten Verzögerungswert die Wirkung der Erzeugung eines kurzen stimmhaften Sprachsegments in der Mitte eines stimmlosen Sprachstoßes (siehe [Fig. 10](#)). Die Wirkung, die als "Bing-Artefakt" bekannt ist, kann nervend sein.

[0011] Es ist vorteilhaft und wünschenswert, ein

Verfahren und ein System für eine Fehlerverbergung bei einer Sprachdekodierung zu liefern, um die Sprachqualität zu verbessern.

[0012] Die US-6,188,980 beschreibt einen Dekodierer für synthetische Sprache aus einem kodierten Signal, das angeregte lineare Voraussageparameter und LSF-Vektoren umfasst. Wenn ein Fehler bei der Übertragung des Signals von einem Kodierer auftritt, kann die Sequenz der LSF-Werte im LSF-Vektor ein oder mehrere Paare von LSF-Werten, die gestört sind, aufweisen. Der Dekodierer führt ausgewählt ein Löschen, ein LSF-Verbergen oder ein Pair-Flipping aus, je nachdem wie viele Paare in der Sequenz gestört sind.

ZUSAMMENFASSUNG DER ERFINDUNG

[0013] Die vorliegende Erfindung zieht Vorteil aus der Tatsache, dass es eine erkennbare Beziehung zwischen den Langzeitvoraussage-(LTP)-Parametern in den Sprachsignalen gibt. Insbesondere weist die LTP-Verzögerung eine starke Korrelation mit der LTP-Verstärkung auf. Wenn die LTP-Verstärkung hoch und vernünftig stabil ist, so ist die LTP-Verzögerung typischerweise sehr stabil, und die Variation zwischen nebeneinander liegenden Verzögerungswerten ist klein. In diesem Fall zeigen die Sprachparameter eine stimmhafte Sprachsequenz an. Wenn die LTP-Verstärkung niedrig oder unstabil ist, so ist die LTP-Verzögerung typischerweise stimmlos, und die Sprachparameter zeigen eine stimmlose Sprachsequenz an. Wenn die Sprachsequenz als stationär (stimmhaft) oder nicht stationär (stimmlos) klassifiziert ist, kann der beschädigte oder schlechte Rahmen in der Sequenz unterschiedlich verarbeitet werden.

[0014] Gemäß der Erfindung wird ein Verfahren für das Verbergen von Fehlern in einem kodierten Bitstrom, der Sprachsignale repräsentiert, der in einem Sprachdekodierer empfangen wird, bereit gestellt, wobei der kodierte Bitstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind, und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen voraus gehen, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, wobei das Verfahren folgende Schritte umfasst: Vorsehen einer oberen Grenze und einer unteren Grenze auf der Basis der zweiten Langzeitvoraussage-Verzögerungswerte; Bestimmen, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb oder außerhalb der oberen und unteren Grenzen liegt; Ersetzen des ers-

ten Langzeitvoraussage-Verzögerungswerts im teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt, und Behalten des ersten Langzeitvoraussage-Verzögerungswert im teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenzen liegt.

[0015] Das Verfahren kann auch das Ersetzen des ersten Langzeitvoraussage-Verstärkungswerts im teilweise beschädigten Rahmen durch einen dritten Verstärkungswert umfassen, wenn der erste Langzeit-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt.

[0016] Gemäß der Erfindung wird auch ein Sprachsignal-Sender- und Empfängersystem für das Kodieren von Signalen in einen kodierten Bitstrom und das Dekodieren des kodierten Bitstroms in synthetische Sprache bereit gestellt, wobei der kodierte Bitstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind, und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen voraus gehen, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, und ein erstes Signal verwendet wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei das System umfasst: ein erstes Mittel, das auf das erste Signal reagiert, für das Bestimmen, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb einer oberen Grenze und unteren Grenzen liegt, und für das Liefern eines zweiten Signals, das diese Bestimmung anzeigt; ein zweites Mittel, das auf das zweite Signal reagiert, für das Ersetzen des ersten Langzeitvoraussage-Verzögerungswerts im teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt, und das Behalten des ersten Langzeitvoraussage-Verzögerungswert im teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenzen liegt.

[0017] Weiterhin wird gemäß der Erfindung auch ein Dekodieren für das Synthetisieren von Sprache aus einem kodierten Bitstrom geliefert, wobei der kodierte Bitstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind, und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen voraus gehen,

wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, und ein erstes Signal verwendet wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei der Dekodierer umfasst: ein erstes Mittel, das auf das erste Signal reagiert, für das Bestimmen, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb einer oberen Grenze und einer unteren Grenze liegt, und für das Liefern eines zweiten Signals, das diese Bestimmung anzeigt; ein zweites Mittel, das auf das zweite Signal reagiert, für das Ersetzen des ersten Langzeitvoraussage-Verzögerungswerts im teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt, und das Behalten des ersten Langzeitvoraussage-Verzögerungswert im teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenzen liegt.

[0018] Nochmals weiter wird gemäß der Erfindung eine Mobilstation bereit gestellt, die ausgelegt ist, um einen kodierten Bitstrom zu empfangen, der Sprachdaten enthält, die Sprachsignale repräsentieren, wobei der kodierte Bitstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind, und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen voraus gehen, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, und ein erstes Signal verwendet wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei die Mobilstation umfasst: ein erstes Mittel, das auf das erste Signal reagiert, für das Bestimmen, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb einer oberen Grenze und einer unteren Grenze liegt, und für das Liefern eines zweiten Signals, das diese Bestimmung anzeigt; ein zweites Mittel, das auf das zweite Signal reagiert, für das Ersetzen des ersten Langzeitvoraussage-Verzögerungswerts im teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt, und das Behalten des ersten Langzeitvoraussage-Verzögerungswert im teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenzen liegt.

[0019] Zusätzlich wird gemäß der Erfindung ein Element in einem Telekommunikationsnetz geliefert, das ausgelegt ist, um einen kodierten Bitstrom zu empfangen, der Sprachdaten von einer Mobilstation enthält, wobei die Sprachdaten eine Vielzahl von Sprachrahmen einschließen, die in Sprachsequenzen angeordnet sind, und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen voraus gehen, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langezeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, und ein erstes Signal verwendet wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei das Element umfasst:

ein erstes Mittel, das auf das erste Signal reagiert, für das Bestimmen, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb einer oberen Grenze und einer unteren Grenze liegt, und für das Liefern eines zweiten Signals, das diese Bestimmung anzeigt; ein zweites Mittel, das auf das zweite Signal reagiert, für das Ersetzen des ersten Langzeitvoraussage-Verzögerungswerts im teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt, und das Behalten des ersten Langzeitvoraussage-Verzögerungswert im teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenzen liegt.

[0020] Der dritte Verzögerungswert kann auf den zweiten Langzeitvoraussage-Verzögerungswerten und einer adaptiv begrenzten zufälligen Verzögerungsschwankung basieren.

[0021] Das zweite Mittel kann weiter den ersten Langzeit-Verstärkungswert im teilweise beschädigten Rahmen durch einen dritten Verstärkungswert ersetzen, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenzen liegt. Weiterhin kann der dritte Verzögerungswert auf der Basis der zweiten Langzeitvoraussage-Verstärkungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung bestimmt werden.

[0022] Die vorliegende Erfindung wird beim Lesen der Beschreibung in Verbindung mit den [Fig. 3](#) bis [Fig. 11c](#) deutlich.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0023] [Fig. 1](#) ist ein Blockdiagramm, das einen allgemeinen verteilten Sprach-Kodierer/Dekodierer

darstellt, wobei der kodierte Bitstrom, der Sprachdaten enthält, von einem Kodierer zu einem Dekodierer über einen Kommunikationskanal oder ein Speichermedium befördert wird.

[0024] [Fig. 2](#) ist ein Blockdiagramm, das eine Fehlerverbergungsvorrichtung des Stands der Technik in einem Empfänger darstellt.

[0025] [Fig. 3](#) ist ein Blockdiagramm, das die Fehlerverbergungsvorrichtung in einem Empfänger gemäß der vorliegenden Erfindung darstellt.

[0026] [Fig. 4](#) ist ein Flussdiagramm, das das Verfahren der Fehlerverbergung gemäß der vorliegenden Erfindung darstellt.

[0027] [Fig. 5](#) ist eine schematische Darstellung einer Mobilstation, die ein Fehlerverbergungsmodul enthält, gemäß der vorliegenden Erfindung.

[0028] [Fig. 6](#) ist eine schematische Darstellung eines Telekommunikationsnetzes, das einen Dekodierer gemäß der vorliegenden Erfindung verwendet.

[0029] [Fig. 7](#) ist ein Schaubild von LTP-Parametern, das die Verzögerungs- und Verstärkungsprofile in einer stimmhaften Sprachsequenz darstellt.

[0030] [Fig. 8](#) ist ein Schaubild von LTP-Parametern, das die Verzögerungs- und Verstärkungsprofile in einer stimmlosen Sprachsequenz darstellt.

[0031] [Fig. 9](#) ist ein Schaubild von LTP-Verzögerungswerten in einer Serie von Unterrahmen, das die Differenz zwischen der Lösung der Fehlerverbergung gemäß dem Stand der Technik und der Lösung gemäß der vorliegenden Erfindung darstellt.

[0032] [Fig. 10](#) ist ein anderes Schaubild von LTP-Verzögerungswerten in einer Serie von Unterrahmen, das die Differenz zwischen der Lösung der Fehlerverbergung gemäß dem Stand der Technik und der Lösung gemäß der vorliegenden Erfindung darstellt.

[0033] [Fig. 11a](#) ist ein Schaubild von Sprachsignalen, das eine fehlerfreie Sprachsequenz zeigt, wobei der Ort des schlechten Rahmens des Sprachkanals dort liegt, wo das die [Fig. 11b](#) und [Fig. 11c](#) zeigen.

[0034] [Fig. 11b](#) ist ein Schaubild von Sprachsignalen, das die Verbergung von Parametern in einem schlechten Rahmen gemäß der Lösung des Stands der Technik zeigt.

[0035] [Fig. 11c](#) ist ein Schaubild von Sprachsignalen, das die Verbergung von Parametern in einem schlechten Rahmen gemäß der vorliegenden Erfindung zeigt.

BESTE ART FÜR DAS AUSFÜHREN DER ERFINDUNG

[0036] [Fig. 3](#) zeigt einen Dekodierer **10**, der ein Dekodiermodul **20** und ein Fehlerverbergungsmodul **30** einschließt. Das Dekodiermodul **20** empfängt ein Signal **140**, das normalerweise Sprachparameter **102** darstellt, für eine Sprachsynthese. Das Dekodiermodul **20** ist aus dem Stand der Technik bekannt. Das Fehlerverbergungsmodul **30** ist ausgelegt, um einen kodierten Bitstrom **100** zu empfangen, der eine Vielzahl von Sprachströmen einschließt, die in Sprachsequenzen angeordnet sind. Eine Schlechtrahmen-Detektionsvorrichtung **32** wird verwendet, um beschädigte Rahmen in den Sprachsequenzen zu detektieren, und liefert ein Schlechtrahmenanzeigesignal (BFI-Signal) **110**, das ein BFI-Flag darstellt, wenn ein beschädigter Rahmen detektiert wird. Die BFI ist auch aus dem Stand der Technik bekannt. Das BFI-Signal **110** wird verwendet, um zwei Schalter **40** und **42** zu steuern. Normalerweise sind die Sprachrahmen nicht beschädigt, und das BFI-Flag ist 0. Der Anschluss S ist betriebsmäßig mit dem Anschluss 0 in den Schaltern **40** und **42** verbunden. Die Sprachparameter **102** werden in einem Pufferspeicher oder einen "Parameterverlaufsspeicher" **50** und das Dekodiermodul **20** für eine Sprachsynthese befördert. Wenn ein schlechter Rahmen durch die Schlechtrahmendetektionsvorrichtung **32** detektiert wird, wird das BFI-Flag auf 1 gesetzt. Der Anschluss S ist mit dem Anschluss 1 in den Schaltern **40** und **42** verbunden. Somit werden die Sprachparameter **102** einer Analysevorrichtung **70** geliefert, und die Sprachparameter, die für die Sprachsynthese benötigt werden, werden von einem Parameterverbergungsmodul **60** an das Dekodiermodul **20** geliefert. Die Sprachparameter **102** umfassen typischerweise LPC-Parameter für die Kurzzeit-Voraussage, Anregungsparameter, Langzeitvoraussage-(LTP)-Verzögerungsparameter, einen LTP-Verstärkungsparameter und andere Verstärkungsparameter. Der Parameterverlaufsspeicher **50** wird verwendet, um die LTP-Verzögerung und die LTP-Verstärkung einer Anzahl nicht beschädigter Sprachrahmen zu speichern. Die Inhalte des Parameterverlaufsspeichers **50** werden konstant aktualisiert, so dass der letzte LTP-Verstärkungsparameter und der letzte LTP-Verzögerungsparameter, die im Speicher **50** gespeichert sind, solche des letzten nicht beschädigten Sprachrahmens sind. Wenn ein beschädigter Rahmen in einer Sprachsequenz im Dekodierer **10** empfangen wird, wird das BFI-Flag auf 1 gesetzt, und die Sprachparameter **102** des beschädigten Rahmens werden zur Analysevorrichtung **70** durch den Schalter **40** befördert. Durch das Vergleichen des LTP-Verstärkungsparameters im beschädigten Rahmen und der LTP-Verstärkungsparameter, die im Speicher **50** gespeichert sind, ist es für die Analysevorrichtung **70** möglich, zu bestimmen, ob die Sprachsequenz stationär oder nicht stationär ist, auf der Basis der Größe und ihrer Variation in den

LTP-Verstärkungsparametern in benachbarten Rahmen. Typischerweise sind in einer stationären Sequenz die LTP-Verstärkungsparameter hoch und vernünftig stabil, der LTP-Verzögerungswert ist stabil und die Variation in benachbarten LTP-Verzögerungswerten ist klein, wie das in [Fig. 7](#) gezeigt ist. Im Gegensatz dazu sind in einer nicht stationären Sequenz die LTP-Verstärkungsparameter niedrig und unstabil, und die LTP-Verzögerung ist auch unstabil, wie das in [Fig. 8](#) gezeigt ist. Die LTP-Verzögerungswerte ändern sich mehr oder weniger zufällig. [Fig. 7](#) zeigt die Sprachsequenz für das Wort "viiniä". [Fig. 8](#) zeigt die Sprachsequenz für das Wort "exhibition".

[0037] Wenn die Sprachsequenz, die den beschädigten Rahmen einschließt, stimmhaft oder stationär ist, wird die letzte gute LTP-Verzögerung aus dem Speicher **50** abgerufen und an das Parameterverbergungsmodul **60** befördert. Die abgerufene gute LTP-Verzögerung wird verwendet, um die LTP-Verzögerung des beschädigten Rahmens zu ersetzen. Da die LTP-Verzögerung in einer stationären Sprachsequenz stabil und ihre Variation klein ist, ist es vernünftig, eine vorherige LTP-Verzögerung mit einer kleinen Modifikation zu verwenden, um den entsprechenden Parameter im beschädigten Rahmen zu verbergen. Nachfolgend verursacht ein RX-Signal **104**, dass die Ersetzungsparameter, wie sie mit der Bezugszahl **134** bezeichnet sind, zum Dekodiermodul **20** durch den Schalter **42** befördert werden.

[0038] Wenn die Sprachsequenz, die den beschädigten Rahmen einschließt, stimmlos oder nicht stationär ist, berechnet die Analysevorrichtung **70** einen Ersetzungs-LTP-Verzögerungswert und einen Ersetzungs-LTP-Verstärkungswert für das Parameterverbergen. Da die LTP-Verzögerung in einer nicht stationären Sprachsequenz unstabil und ihre Variation in benachbarten Rahmen typischerweise sehr groß ist, sollte die Parameterverbergen es erlauben, dass die LTP-Verzögerung in einer fehlerverborgenen nicht stationären Sequenz in zufälliger Weise fluktuiert. Wenn die Parameter im beschädigten Rahmen total beschädigt sind, wie bei einem verlorenen Rahmen, wird die Ersetzungs-LTP-Verzögerung berechnet unter Verwendung eines gewichteten Mittelwerts der vorherigen guten LTP-Verzögerungswerte zusammen mit einer adaptiv begrenzten zufälligen Schwankung. Die adaptiv begrenzte zufällige Schwankung kann innerhalb Grenzen variieren, die aus dem Verlauf der LTP-Werte berechnet werden, so dass die Parameterfluktuation in einem fehlerverborgenen Segment ähnlich dem vorherigen guten Abschnitt derselben Sprachsequenz ist.

[0039] Eine beispielhafte Regel für die LTP-Verzögerungs-Verbergen wird durch ein Satz von Bedingungen folgendermaßen bestimmt:

Wenn
 $\text{minGain} > 0,5 \text{ UND } \text{LadDif} < 10; \text{ ODER}$

lastGain > 0,5 UND secondLastGain > 0,5, dann wird die zuletzt empfangene gute LTP-Verzögerung für den gesamten beschädigten Rahmen verwendet. Ansonsten wird Update_lag ein gewichteter Mittelwert des LTP-Verzögerungspuffers mit einer Zufallsanordnung für den gesamten beschädigten Rahmen verwendet. Update_lag wird in der folgenden Weise berechnet:

Der LTP-Verzögerungspuffer wird sortiert und die drei größten Pufferwerte werden abgerufen. Der Mittelwert dieser drei größten Werte wird als die gewichtete mittlere Verzögerung (WAL) bezeichnet, und die Differenz von diesen größten Werten wird als gewichtete Verzögerungsdifferenz (WLD) bezeichnet.

[0040] Es sei RAND die Zufallsanordnung mit der Skalierung von $(-WLD/2, WLD/2)$, dann $Update_lag = WAL + RAND(-WLD/2, WLD/2)$, wobei minGain der kleinste Wert des LTP-Verstärkungspuffers ist; LagDif die Differenz zwischen den kleinsten und den größten LTP-Verzögerungswerten ist; lastGain die letzte empfangene gute LTP-Verstärkung ist; und secondLastGain die zweitletzte empfangene gute LTP-Verstärkung ist.

[0041] Wenn die Parameter des beschädigten Rahmens teilweise beschädigt sind, dann wird der LTP-Verzögerungswert im beschädigten Rahmen entsprechend ersetzt. Dass der Rahmen teilweise beschädigt ist, wird durch einen Satz beispielhafter LTP-Merkmalsskriterien bestimmt, die nachfolgend angegeben sind:

Wenn

- (1) $LagDif < 10$ UND $(minLag - 5) < T_{bf} < (maxLag + 5)$; ODER
- (2) $lastGain > 0,5$ UND $secondLastGain > 0,5$ UND $(lastLag - 10) < T_{bf} < (lastLag + 10)$; ODER
- (3) $minGain < 0,4$ UND $lastGain = minGain$ UND $minLag < T_{bf} < maxLag$; ODER
- (4) $LagDif < 70$ UND $minLag < T_{bf} < maxLag$; ODER
- (5) $meatilag < T_{bf} < maxLag$

wahr ist, dann wird T_{bf} verwendet, um die LTP-Verzögerung im beschädigten Rahmen zu ersetzen. Ansonsten wird der beschädigte Rahmen als ein total beschädigter Rahmen behandelt, wie das oben beschrieben ist. In den obigen Bedingungen:

maxLag ist der größte Wert des LTP-Verzögerungspuffers;

meanLag ist der Mittelwert des LTP-Verzögerungspuffers;

minLag ist der kleinste Wert des LTP-Verzögerungspuffers;

lastLag ist der letzte empfangene gute LTP-Verzögerungswert; und

T_{bf} ist eine dekodierte LTP-Verzögerung, die gesucht wird, wenn die BFI gesetzt ist, aus dem adaptiven Kodbuch, wie wenn BFI nicht gesetzt ist.

[0042] Zwei Beispiele einer Parameterverbergung sind in den [Fig. 9](#) und [Fig. 10](#) gezeigt. Wie gezeigt ist, ist das Profil der Ersetzungs-LTP-Verzögerungswerte im schlechten Rahmen gemäß dem Stand der Technik ziemlich flach, aber das Profil der Ersetzung gemäß der vorliegenden Erfindung ermöglicht einige Fluktuation ähnlich dem fehlerfreien Profil. Der Unterschied zwischen der Lösung des Stands der Technik und der vorliegenden Erfindung wird weiter in den [Fig. 11b](#) beziehungsweise [Fig. 11c](#) auf der Basis von Sprachsignalen in einem fehlerfreien Kanal dargestellt, wie sie in [Fig. 11a](#) gezeigt sind.

[0043] Wenn die Parameter im beschädigten Rahmen teilweise beschädigt sind, kann die Parameterverbergung weiter optimiert werden. In teilweise beschädigten Rahmen können die LTP-Verzögerungen in den beschädigten Rahmen dennoch ein akzeptables synthetisiertes Sprachsegment ergeben. Gemäß den GSM-Spezifikationen wird das BFI-Flag durch einen Mechanismus einer zyklischen Redundanzprüfung (CRC) oder einen anderen Fehlerdetektionsmechanismus gesetzt. Diese Fehlerdetektionsmechanismen detektieren Fehler in den signifikantesten Bits im Kanalkodierverfahren. Somit kann, sogar wenn nur einige Bits fehlerhaft sind, der Fehler detektiert werden, und das BFI-Flag wird entsprechend gesetzt. In der Lösung der Parameterverbergung des Stands der Technik wird der gesamte Rahmen verworfen. Somit wird Information, die in den korrekten Bits enthalten ist, weggeworfen.

[0044] Typischerweise ist im Kanalkodierverfahren die BER pro Rahmen ein guter Indikator für den Kanalzustand. Wenn der Kanalzustand gut ist, so ist die BER pro Rahmen klein, und ein hoher Prozentsatz der LTP-Verzögerungswerte in den fehlerhaften Rahmen sind korrekt. Wenn beispielsweise die Rahmenfehlerrate (FER) 0,2% beträgt, so sind über 70% der LTP-Verzögerungswerte korrekt. Sogar wenn die FER 3% erreicht, so sind ungefähr 60% der LTP-Verzögerungswerte noch korrekt. Die CRC kann genau einen schlechten Rahmen detektieren und das BFI-Flag entsprechend setzen. Die CRC liefert jedoch keine Schätzung der BER im Rahmen. Wenn das BFI-Flag als einziges Kriterium für die Parameterverbergung verwendet wird, dann kann es sein, dass ein hoher Prozentsatz der korrekten LTP-Verzögerungswerte verschwendet wird. Um zu verhindern, dass eine große Menge korrekter LTP-Verzögerungen weggeworfen wird, ist es möglich ein Entscheidungskriterium für die Parameterverbergung auf der Basis des BFI-Verlaufs auszubilden. Es ist auch möglich, beispielsweise die FER als das Entscheidungskriterium zu verwenden. Wenn die LTP-Verzögerung das Entscheidungskriterium erfüllt, so ist kein Parameterverbergen notwendig. In diesem Fall, befördert die Analysevorrichtung **70** die Sprachparameter **102**, so wie sie sie empfangen hat, durch den Schalter **40** zum Parameterverbergungsmodul **60**,

das diese dann zum Dekodiermodul **20** durch den Schalter **42** befördert. Wenn die LTP-Verzögerung das Entscheidungskriterium nicht erfüllt, wird der beschädigte Rahmen unter Verwendung der LTP-Merkmalsskriterien für eine Parameterverbergung weiter untersucht, wie das hier oben beschrieben wurde.

[0045] Bei stationären Sprachsequenzen ist die LTP-Verzögerung sehr stabil. Ob die meisten der LTP-Verzögerungswerte in einem beschädigten Rahmen korrekt oder fehlerhaft sind, kann mit hoher Wahrscheinlichkeit korrekt vorhergesagt werden. Somit ist es möglich, ein sehr strenges Kriterium für die Parameterverbergung zu verwenden. Bei nicht stationären Sprachsequenzen kann es durch die instabile Natur der LTP-Parameter schwierig sein, vorherzusagen, ob der LTP-Verzögerungswert in einem beschädigten Rahmen korrekt ist. Ob die Vorhersage korrekt oder falsch ist, ist jedoch bei nicht stationärer Sprache weniger wichtig als bei stationärer Sprache. Wenn man es erlaubt, dass fehlerhafte LTP-Verzögerungswerte bei der Dekodierung stationärer Sprache verwendet werden, so kann das bewirken, dass die synthetisierte Sprache unerkennbar ist, während die Erlaubnis der Verwendung fehlerhafter LTP-Verzögerungswerte bei der Dekodierung nicht stationärer Sprache, gewöhnlicherweise nur die hörbaren Artefakte erhöht. Somit kann das Entscheidungskriterium für die Parameterverbergung bei nicht stationärer Sprache relativ locker sein.

[0046] Wie früher erwähnt wurde, fluktuiert die LTP-Verstärkung bei nicht stationärer Sprache stark. Wenn derselbe LTP-Verstärkungswert vom letzten guten Rahmen wiederholt verwendet wird, um den LTP-Verstärkungswert eines oder mehrerer beschädigter Rahmen in einer Sprachsequenz zu ersetzen, wird das LTP-Verstärkungsprofil im Verstärkungsverbergungssegment flach sein (ähnlich der LTP-Verzögerungsersetzung des Stands der Technik, wie sie in den [Fig. 7](#) und [Fig. 8](#) gezeigt ist), in starkem Kontrast zum fluktuierenden Profil der nicht beschädigten Rahmen. Die plötzliche Änderung im LTP-Verstärkungsprofil kann unangenehme hörbare Artefakte verursachen. Um diese hörbaren Artefakte zu minimieren, ist es möglich, es dem Ersetzungs-LTP-Verstärkungswert zu ermöglichen, im Fehlerverbergungssegment zu fluktuieren. Für diesen Zweck kann die Analysevorrichtung **70** auch verwendet werden, um auf der Basis der Verstärkungswerte im LTP-Verlauf die Grenzen zu bestimmen, zwischen denen der Ersetzungs-LTP-Verstärkungswert fluktuieren kann.

[0047] Die LTP-Verstärkungs-Verbergung kann in der nachfolgend beschriebenen Weise ausgeführt werden. Wenn die BFI gesetzt ist, so wird ein Ersetzungs-LTP-Verstärkungswert gemäß einem Satz von LTP-Verstärkungsverbergungsregeln berechnet. Die Ersetzungs-LTP-Verstärkung wird als Updated_gain

bezeichnet.

- (1) Wenn $\text{gainDif} > 0,5$ UND $\text{lastGain} = \text{maxGain} > 0,9$ UND $\text{subBF} = 1$, dann $\text{Updated_gain} = (\text{secondLastGain} + \text{thirdLastGain})/2$;
- (2) Wenn $\text{gainDif} > 0,5$ UND $\text{lastGain} = \text{maxGain} > 0,9$ UND $\text{subBF} = 2$, dann $\text{Updated_gain} = \text{meanGain} + \text{randvar} \cdot (\text{maxGain} - \text{meanGain})$;
- (3) Wenn $\text{gainDif} > 0,5$ UND $\text{lastGain} = \text{maxGain} > 0,9$ UND $\text{subBF} = 3$, dann $\text{Updated_gain} = \text{meanGain} - \text{randvar} \cdot (\text{meanGain} - \text{minGain})$;
- (4) Wenn $\text{gainDif} > 0,5$ UND $\text{lastGain} = \text{maxGain} > 0,9$ UND $\text{subBF} = 4$, dann $\text{Updated_gain} = \text{meanGain} + \text{randvar} \cdot (\text{maxGain} - \text{nzeanGain})$;

[0048] In der vorhergehenden Bedingungen kann Updated_gain nicht größer als lastGain sein. Wenn die vorherigen Bedingungen nicht erfüllt werden können, werden die folgenden Bedingungen verwendet:

- (5) Wenn $\text{gainDif} > 0,5$, dann $\text{Updated_gain} = \text{lastGain}$;
- (6) Wenn $\text{gainDif} < 0,5$ UND $\text{lastGain} = \text{maxGain}$, dann $\text{Updated_gain} = \text{meanGain}$;
- (7) Wenn $\text{gainDif} < 0,5$, dann $\text{Updated_gain} = \text{lastGain}$,

Wobei

meanGain der Mittelwert des LTP-Verstärkungspuffers ist;

maxGain der größte Wert des LTP-Verstärkungspuffer ist;

minGain der kleinste Wert des LTP-Verstärkungspuffer ist;

randVar ein Zufallswert zwischen 0 und 1 ist,

gainDif die Differenz zwischen den kleinsten und den größten LTP-Verstärkungswerten im LTP-Verstärkungspuffer ist;

lastGain die letzte empfangene gute LTP-Verstärkung ist;

secondLastGain die zweitletzte empfangene gute LTP-Verstärkung ist;

thirdLastGain die drittletzte empfangene gute LTP-Verstärkung ist; und

subBF die Reihenfolge des Unterrahmens ist.

[0049] [Fig. 4](#) zeigt das Verfahren der Fehlerverbergung gemäß der vorliegenden Erfindung. Wenn der kodierte Bitstrom im Schritt **160** empfangen wird, wird der Rahmen im Schritt **162** geprüft, um zu sehen, ob er beschädigt ist. Wenn der Rahmen nicht beschädigt ist, so wird im Schritt **164** der Parameterverlauf der Sprachsequenz aktualisiert, und dann werden die Sprachparameter des aktuellen Rahmens im Schritt **166** dekodiert. Wenn der Rahmen schlecht oder beschädigt ist, werden die Parameter aus dem Parameterverlaufsspeicher im Schritt **170** abgerufen. Ob der beschädigte Rahmen ein Teil der stationären Sprachsequenz oder der nicht stationären Sprachsequenz ist, wird im Schritt **172** bestimmt. Wenn die Sprachsequenz stationär ist, wird die LTP-Verzögerung des

letzten guten Rahmens verwendet, um die LTP-Verzögerung im beschädigten Rahmen im Schritt **174** zu ersetzen. Wenn die Sprachsequenz nicht stationär ist, werden ein neuer Verzögerungswert und ein neuer Verstärkungswert auf der Basis des LTP-Verlaufs im Schritt **180** berechnet, und sie werden verwendet, um die entsprechenden Parameter im beschädigten Rahmen im Schritt **182** zu ersetzen.

[0050] **Fig. 5** zeigt ein Blockdiagramm einer Mobilstation **200** gemäß einer beispielhaften Ausführungsform der Erfindung. Die Mobilstation umfasst Teile, die typisch für die Vorrichtung sind, wie ein Mikrofon **201**, eine Tastatur **207**, eine Anzeige **206**, einen Kopfhörer **214**, einen Sende/Empfangs-Schalter **208**, eine Antenne **209** und eine Steuereinheit **205**. Zusätzlich zeigt die Figur Sender- und Empfängerblöcke **204**, **211**, die für eine Mobilstation typisch sind. Der Senderblock **204** umfasst einen Kodierer **221** für das Kodieren des Sprachsignals. Der Senderblock **204** umfasst auch Operationen, die für die Kanalkodierung, die Entschlüsselung und die Modulation erforderlich sind, als auch HF-Funktionen, die in **Fig. 5** aus Gründen der Klarheit nicht aufgezeichnet wurden. Der Empfängerblock **211** umfasst auch einen Dekodierblock **220** gemäß der Erfindung. Der Dekodierblock **220** umfasst ein Fehlerverbergungsmodul **222** wie das Parameterverbergungsmodul **30**, das in **Fig. 3** gezeigt ist. Das Signal, das vom Mikrofon **201** kommt, in der Verstärkungsstufe **202** verstärkt und im A/D-Wandler digitalisiert wird, wird zum Senderblock **204** genommen, typischerweise zur Sprachkodiervorrichtung, die im Senderblock enthalten ist. Das Sendesignal, das verarbeitet, moduliert und verstärkt wird durch den Senderblock wird über den Sende/Empfangsschalter **208** zur Antenne **209** geliefert. Das zu empfangende Signal wird von der Antenne über den Sende/Empfangsschalter **208** zum Empfängerblock **211** genommen, der das empfangene Signal demoduliert und die Entschlüsselung und die Kanalkodierung durchführt. Das sich ergebende Sprachsignal wird über den D/A-Wandler **212** zu einem Verstärker **213** und weiter zu einem Kopfhörer **214** geleitet. Die Steuereinheit **205** steuert den Betrieb der Mobilstation **200**, liest die Steuerbefehle, die vom Benutzer über die Tastatur **207** gegeben werden und gibt Nachrichten an den Benutzer mittels der Anzeige **206**.

[0051] Das Parameterverbergungsmodul **30** gemäß der Erfindung kann auch in einem Telekommunikationsnetz **300** verwendet werden, wie einem normalen Telefonnetz, oder einem Mobilnetz, wie dem GSM-Netz. **Fig. 6** zeigt ein Beispiel eines Blockdiagramms eines solchen Telekommunikationsnetzes. Beispielsweise kann das Telekommunikationsnetz **300** Telefonvermittlungsstellen oder entsprechende Vermittlungssysteme **360** umfassen, mit denen normale Telefone **370**, Basisstationen **340**, Basisstationssteuerungen **350** und andere zentrale Vorrichtun-

gen **355** des Telekommunikationsnetzes verbunden sind. Die Mobilstationen **330** können eine Verbindung mit dem Telekommunikationsnetz über die Basisstationen **340** errichten. Ein Dekodierblock **320**, der ein Fehlerverbergungsmodul **322** ähnlich dem Fehlerverbergungsmodul **30**, das in **Fig. 3** gezeigt ist, umfasst, kann am vorteilhaftesten beispielsweise in der Basisstation **340** platziert werden. Der Dekodierblock **320** kann aber auch in der Basisstationssteuerung **350** oder beispielsweise der Zentral- oder Schaltvorrichtung **355** platziert werden. Wenn das Mobilstationssystem getrennte Transkoder verwendet, beispielsweise zwischen den Basisstationen und den Basisstationssteuerungen, für das Transformieren des kodierten Signals, das über den Funkkanal übertragen wurde, in ein typisches Signal von 64 kBit/s, das in einem Telekommunikationssystem übertragen wird, und umgekehrt, so kann der Dekodierblock **320** auch in einem solchen Transkoder platziert werden. Im allgemeinen kann der Dekodierblock **320**, der das Parameterverbergungsmodul **322** einschließt, in jedem Element des Telekommunikationsnetzes **300** platziert werden, das den kodierten Datenstrom in einen unkodierten Datenstrom transformiert. Der Dekodierblock **320** dekodiert und filtert das kodierte Sprachsignal, das von der Mobilstation **330** kommt, wonach das Sprachsignal in der gewöhnlichen Weise als unkomprimiertes Signal im Telekommunikationsnetz **300** übertragen werden kann.

[0052] Es sollte angemerkt werden, dass das Fehlerverbergungsverfahren der vorliegenden Erfindung in Bezug auf stationäre und nicht stationäre Sprachsequenzen beschrieben wurde, und dass stationäre Sprachsequenzen gewöhnlicherweise stimmhaft sind, und dass nicht stationäre Sprachsequenzen gewöhnlicherweise stimmlos sind. Somit wird verständlich, dass das beschriebene Verfahren für eine Fehlerverbergung in stimmhaften und stimmlosen Sprachsequenzen verwendet werden kann.

[0053] Die vorliegende Erfindung ist auf Sprach-Kodierer/Dekodierer des CELP-Typs anwendbar und kann auch an andere Typen von Sprach-Kodierern/Dekodierern angepasst werden.

Patentansprüche

1. Verfahren zum Verbergen von Fehlern in einem kodierten Bitdatenstrom, der Sprachsignale repräsentiert, die in einem Sprachdecoder (**10**, **220**, **320**) empfangen werden, wobei der kodierte Bitdatenstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen vorausgegangen sind, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungs-

wert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen, wobei das Verfahren die folgende Schritte umfasst:

Bereitstellen einer oberen Grenze und einer unteren Grenze auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte; Ermitteln, ob der erste Langzeitvoraussage-Verzögerungswert innerhalb oder außerhalb der oberen und unteren Grenze ist; Ersetzen des ersten Langzeitvoraussage-Verzögerungswerts in dem teilweise beschädigten Rahmen durch einen dritten Verzögerungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze (**182**) ist; und Beibehalten des ersten Langzeitvoraussage-Verzögerungswerts in dem teilweise beschädigten Rahmen, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenze ist.

2. Verfahren nach Anspruch 1, weiter umfassend den Schritt des Ersetzen des ersten Langzeitvoraussage-Verstärkungswerts in dem teilweise beschädigten Rahmen durch einen dritten Verstärkungswert, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze (**182**) ist.

3. Verfahren nach Anspruch 1, wobei der dritte Verzögerungswert auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung berechnet wird, die auf weitere Grenzen beschränkt ist, die auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte (**180**) ermittelt werden.

4. Verfahren nach Anspruch 2, wobei der dritte Verstärkungswert auf Basis der zweiten Langzeitvoraussage-Verstärkungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung berechnet wird, die auf Grenzen beschränkt ist, die auf Basis der zweiten Langzeitvoraussage-Verstärkungswerte (**180**) ermittelt werden.

5. Sprachsignal-Sender- und Empfänger-System (**204, 211**) zum Kodieren von Signalen in einem kodierten Bitdatenstrom und Dekodieren des kodierten Bitdatenstroms in synthetische Sprache, wobei der kodierte Bitdatenstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen vorausgegangen sind, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstär-

kungswerte einschließen und ein erstes Signal (**110**) benutzt wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei das System umfasst: ein erstes Mittel (**70**), das auf das erste Signal (**110**) reagiert, um zu ermitteln, ob die erste Langzeitvoraussage-Verzögerung innerhalb einer oberen Grenze und einer unteren Grenze ist und ein zweites Signal (**130**) bereitzustellen, das die Ermittlung anzeigt; ein zweites Mittel (**60**), das auf das zweite Signal reagiert, um den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen durch einen dritten Verzögerungswert zu ersetzen, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist; und den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen beizubehalten, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenze ist.

6. System (**204, 211**) gemäß Anspruch 5, wobei der dritte Verzögerungswert auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung ermittelt wird.

7. System (**204, 211**) gemäß Anspruch 5, wobei das zweite Mittel weiter den ersten Langzeitvoraussage-Verstärkungswert in dem teilweise beschädigten Rahmen durch einen dritten Verstärkungswert ersetzt, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist.

8. System (**204, 211**) gemäß Anspruch 7, wobei der dritte Verstärkungswert auf Basis der zweiten Langzeitvoraussage-Verstärkungswerte und einer adaptiv begrenzten zufälligen Verstärkungsschwankung ermittelt wird.

9. Decoder (**10, 220, 320**) zum Synthetisieren von Sprache aus einem kodierten Bitdatenstrom, wobei der kodierte Bitdatenstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen vorausgegangen sind, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen und ein erstes Signal (**110**) benutzt wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei der Decoder umfasst: ein erstes Mittel (**70**), das auf das erste Signal (**110**) reagiert, um zu ermitteln, ob die erste Langzeitvoraussage-Verzögerung innerhalb einer oberen Grenze und einer unteren Grenze ist und ein zweites Sig-

nals **(130)** bereitzustellen, das die Ermittlung anzeigt; ein zweites Mittel **(60)**, das auf das zweite Signal reagiert, um den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen durch einen dritten Verzögerungswert zu ersetzen, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist; und den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen beizubehalten, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenze ist.

10. Decoder **(10, 220, 320)** gemäß Anspruch 9, wobei der dritte Verzögerungswert auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung ermittelt wird.

11. Decoder **(10, 220, 320)** gemäß Anspruch 9, wobei das zweite Mittel weiter den ersten Langzeit-Verstärkungswert in dem teilweise beschädigten Rahmen durch einen dritten Verstärkungswert ersetzt, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist.

12. Decoder **(10, 220, 320)** gemäß Anspruch 11, wobei der dritte Verstärkungswert auf Basis der zweiten Langzeitvoraussage-Verstärkungswert und einer adaptiv begrenzten zufälligen Verstärkungsschwankung ermittelt wird.

13. Mobilstation **(200)**, die eingerichtet ist, um einen kodierten Bitdatenstrom zu empfangen, der Sprachdaten enthält, die Sprachsignale repräsentiert, wobei der kodierte Bitdatenstrom eine Vielzahl von Sprachrahmen einschließt, die in Sprachsequenzen angeordnet sind und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen vorausgegangen sind, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen und ein erstes Signal **(110)** benutzt wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei die Mobilstation umfasst:

ein erstes Mittel **(70)**, das auf das erste Signal **(110)** reagiert, um zu ermitteln, ob die erste Langzeitvoraussage-Verzögerung innerhalb einer oberen Grenze und einer unteren Grenze ist und ein zweites Signal **(130)** bereitzustellen, das die Ermittlung anzeigt; ein zweites Mittel **(60)**, das auf das zweite Signal reagiert, um den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen durch einen dritten Verzögerungswert zu ersetzen,

wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist; und den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen beizubehalten, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und unteren Grenze ist.

14. Mobilstation **(200)** gemäß Anspruch 13, wobei der dritte Verzögerungswert auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung ermittelt wird.

15. Mobilstation **(200)** gemäß Anspruch 13, wobei das zweite Mittel weiter den ersten Langzeit-Verstärkungswert in dem teilweise beschädigten Rahmen durch einen dritten Verstärkungswert ersetzt, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist.

16. Mobilstation **(200)** gemäß Anspruch 15, wobei der dritte Verstärkungswert auf Basis der zweiten Langzeitvoraussage-Verstärkungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung ermittelt wird.

17. Element **(340)** in einem Telekommunikationsnetz, das eingerichtet ist, um einen kodierten Bitdatenstrom zu empfangen, der Sprachdaten enthält, von einer Mobilstation, wobei die Sprachdaten eine Vielzahl von Sprachrahmen einschließen, die in Sprachsequenzen angeordnet sind und die Sprachrahmen mindestens einen teilweise beschädigten Rahmen einschließen, dem ein oder mehrere nicht beschädigte Rahmen vorausgegangen sind, wobei der teilweise beschädigte Rahmen einen ersten Langzeitvoraussage-Verzögerungswert und einen ersten Langzeitvoraussage-Verstärkungswert einschließt, und die nicht beschädigten Rahmen zweite Langzeitvoraussage-Verzögerungswerte und zweite Langzeitvoraussage-Verstärkungswerte einschließen und ein erstes Signal **(110)** benutzt wird, um den teilweise beschädigten Rahmen anzuzeigen, wobei das Element umfasst:

ein erstes Mittel **(70)**, das auf das erste Signal **(110)** reagiert, um zu ermitteln, ob die erste Langzeitvoraussage-Verzögerung innerhalb einer oberen Grenze und einer unteren Grenze ist und ein zweites Signal **(130)** bereitzustellen, das die Ermittlung anzeigt; ein zweites Mittel **(60)**, das auf das zweite Signal **(110)** reagiert, um den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen durch einen dritten Verzögerungswert zu ersetzen, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist; und den ersten Langzeitvoraussage-Verzögerungswert in dem teilweise beschädigten Rahmen beizubehalten, wenn der erste Langzeitvoraussage-Verzögerungswert innerhalb der oberen und

unteren Grenze ist.

18. Element **(340)** gemäß Anspruch 17, wobei der dritte Verzögerungswert auf Basis der zweiten Langzeitvoraussage-Verzögerungswerte und einer adaptiv begrenzten zufälligen Verzögerungsschwankung ermittelt wird.

19. Element **(340)** gemäß Anspruch 17, wobei das zweite Mittel ferner den ersten Langzeit-Verstärkungswert in dem teilweise beschädigten Rahmen durch einen dritten Verstärkungswert ersetzt, wenn der erste Langzeitvoraussage-Verzögerungswert außerhalb der oberen und unteren Grenze ist.

20. Element **(340)** gemäß Anspruch 19, wobei der dritte Verstärkungswert auf Basis der zweiten Langzeitvoraussage-Verstärkungswerte und einer adaptiv begrenzten zufälligen Verstärkungsschwankung ermittelt wird.

Es folgen 12 Blatt Zeichnungen



FIG. 1

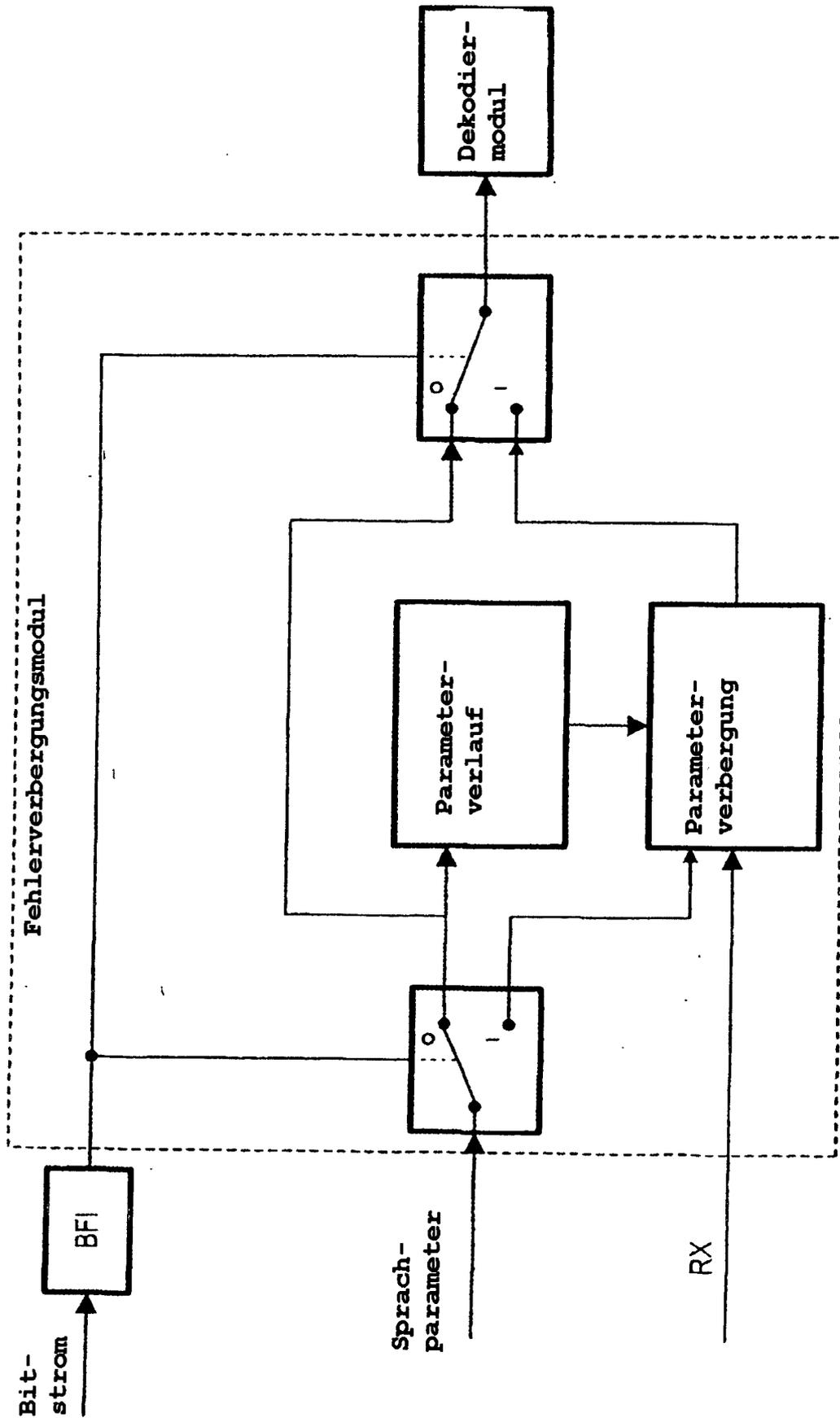


FIG. 2

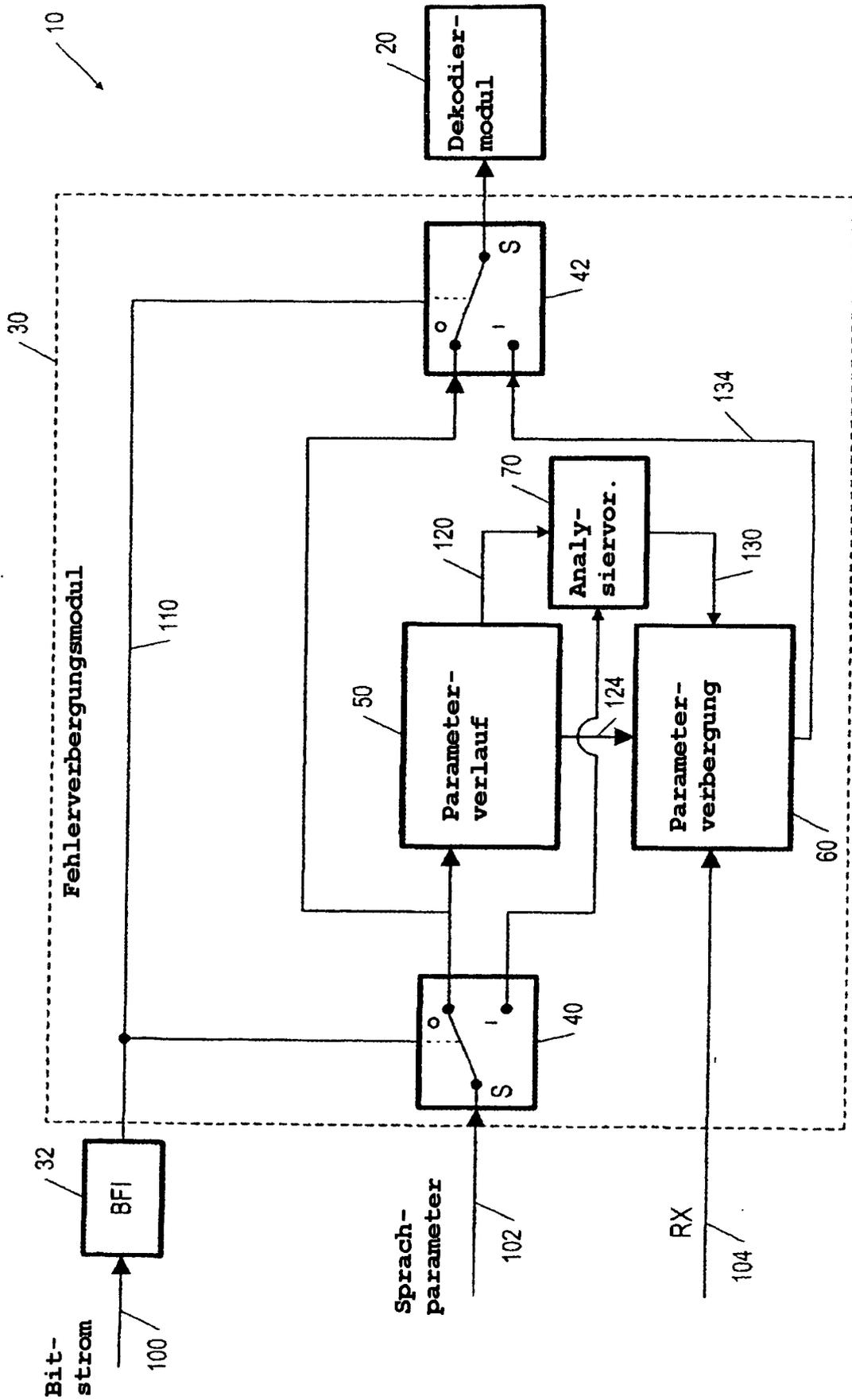


FIG. 3

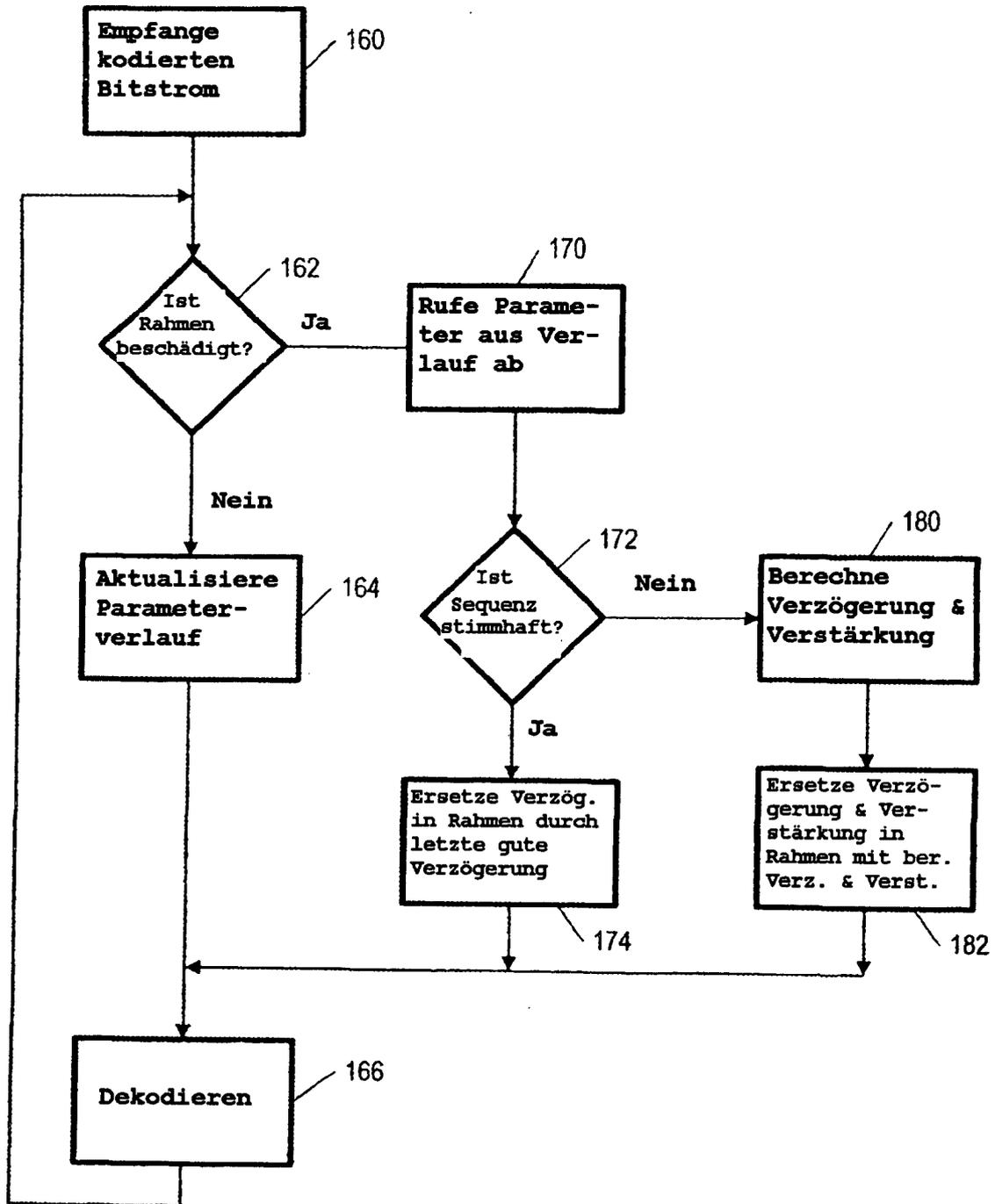


FIG. 4

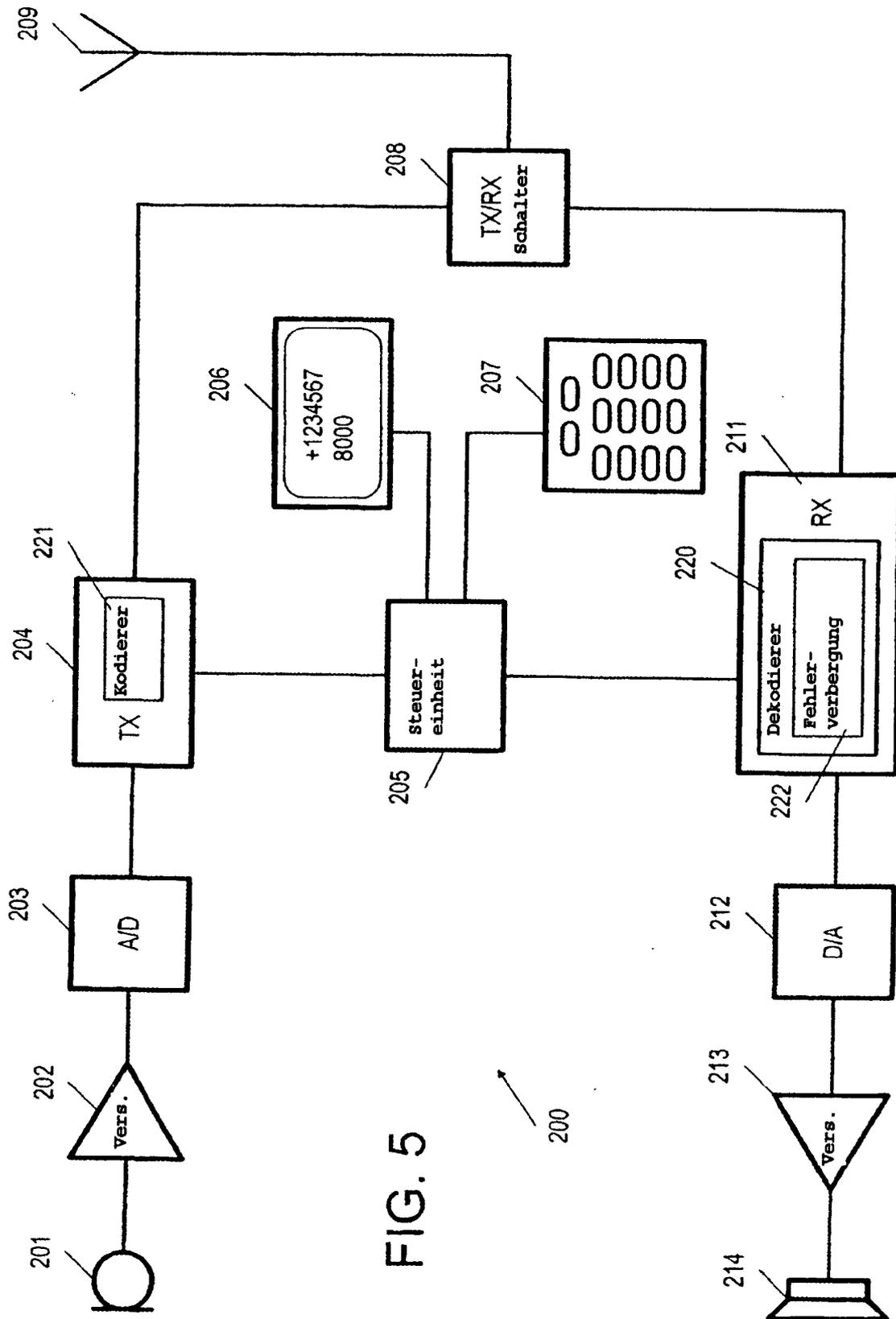


FIG. 5

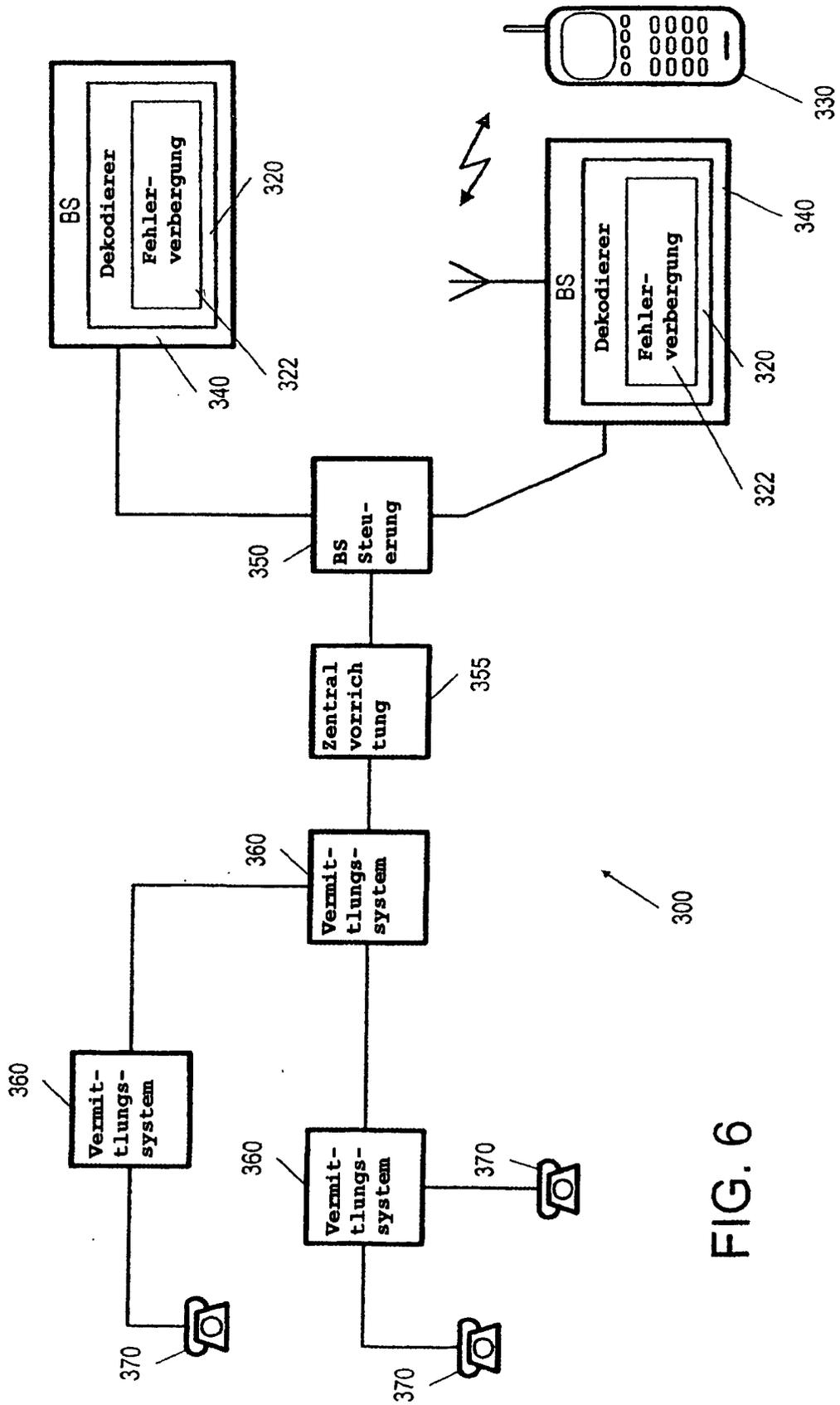


FIG. 6

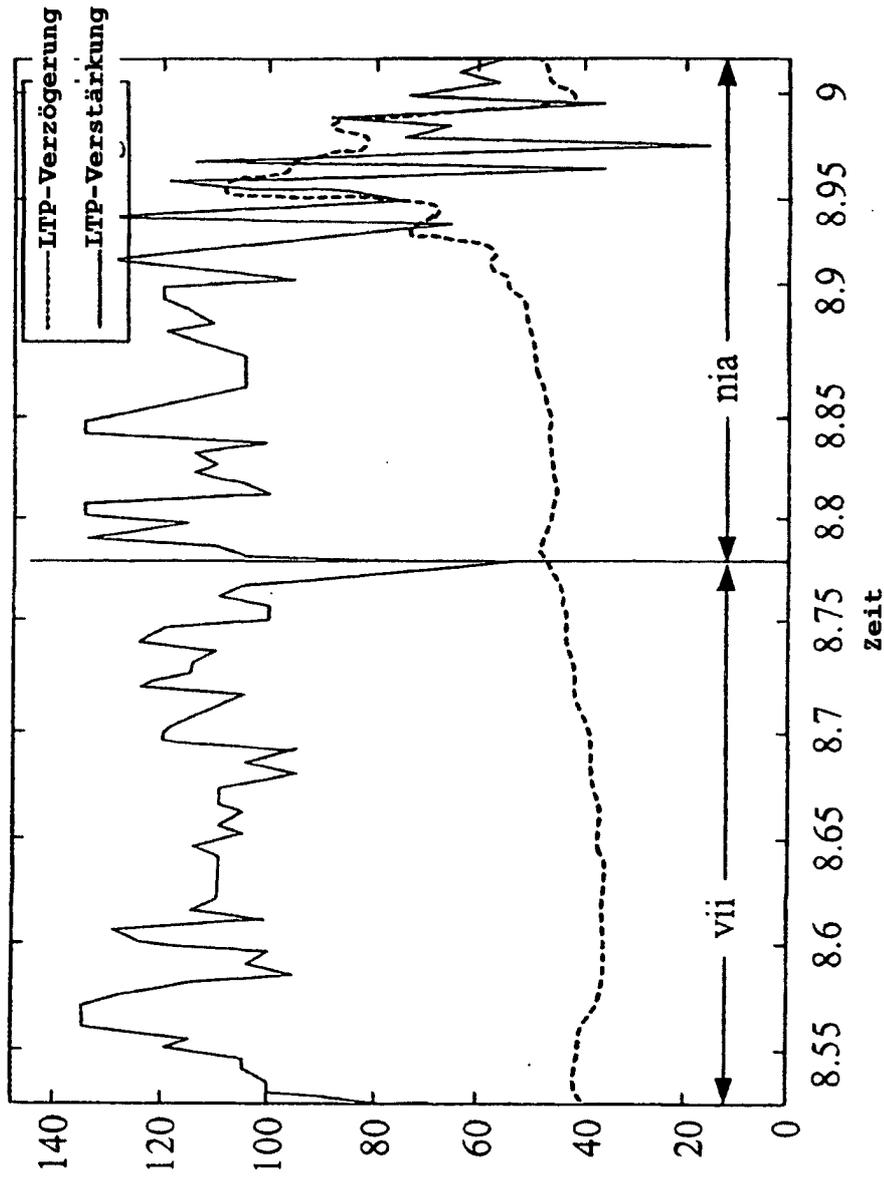


FIG. 7

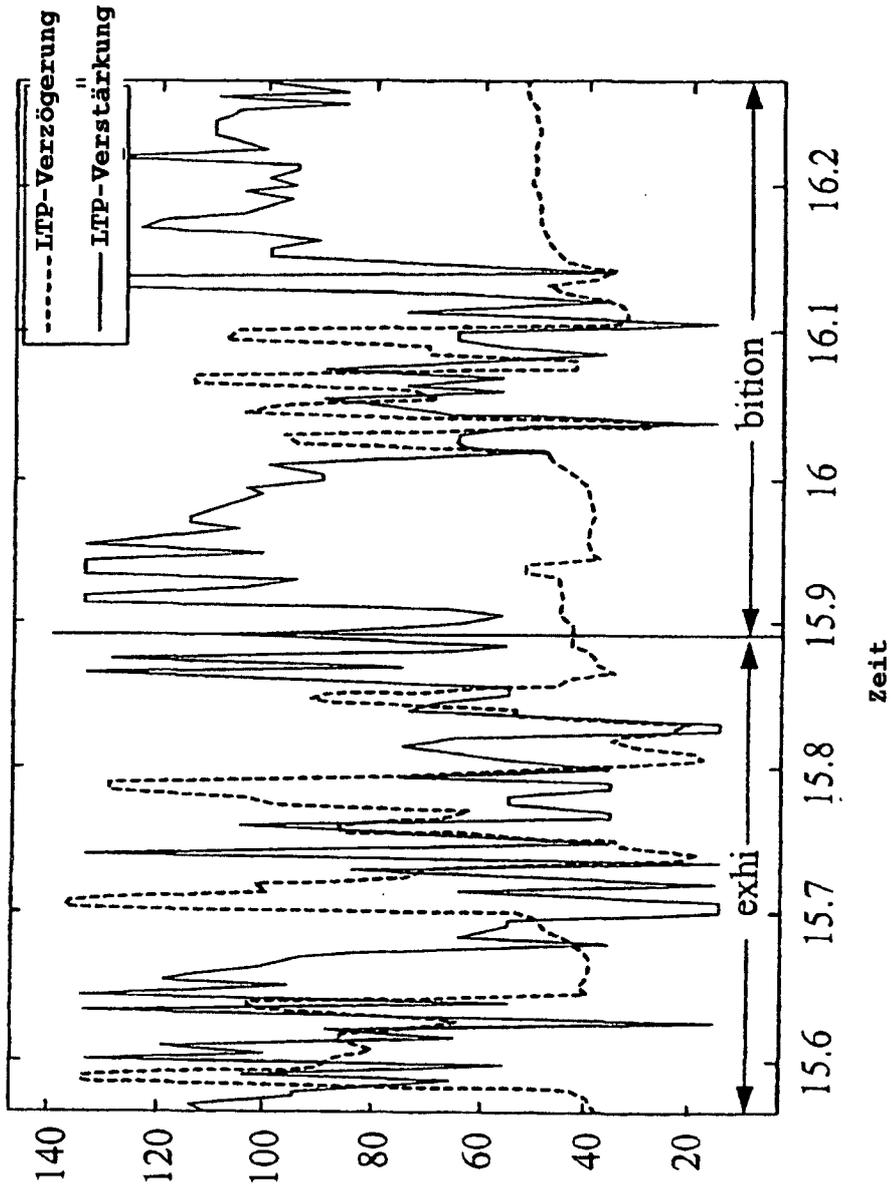


FIG. 8

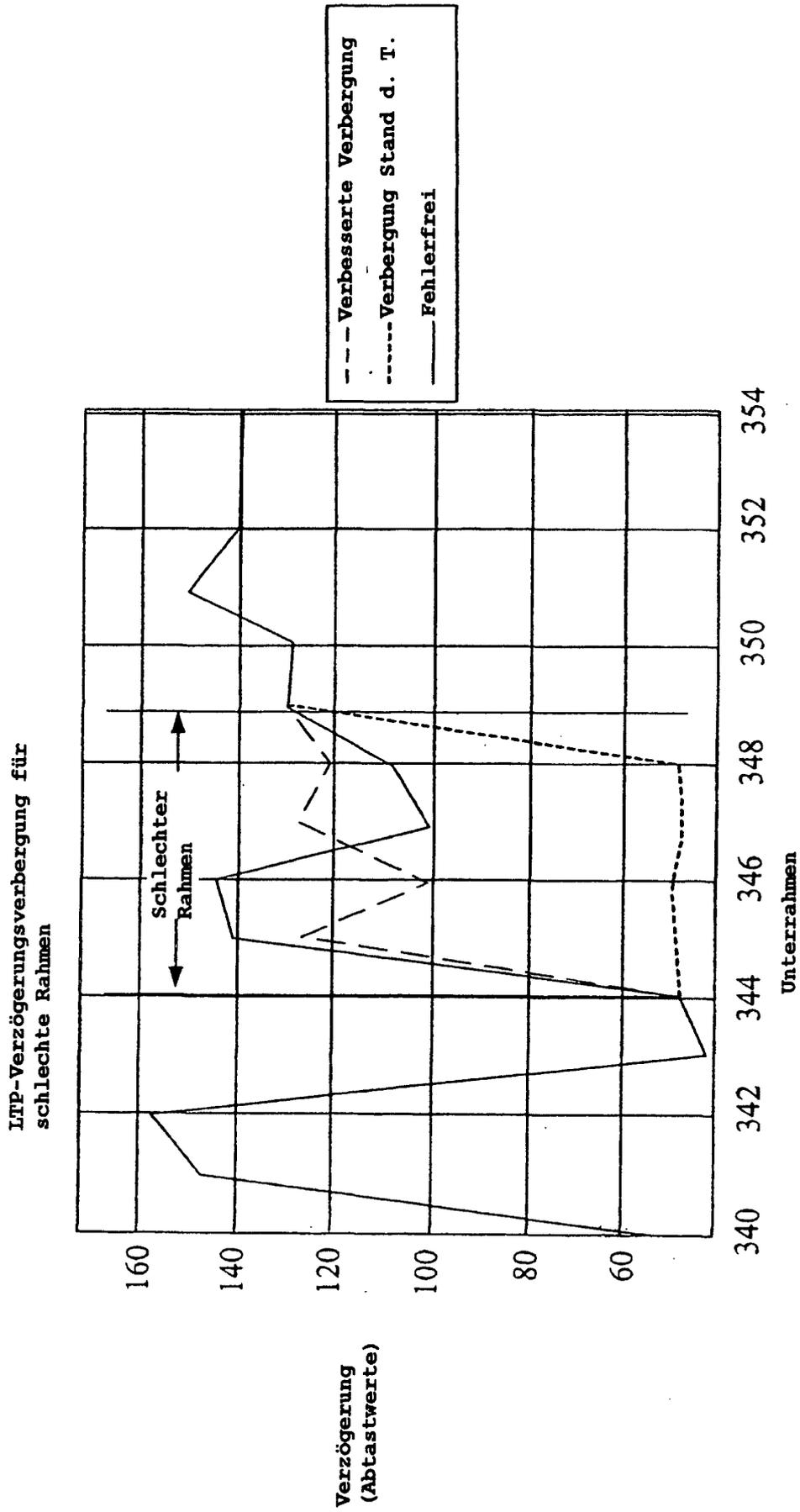


FIG. 9

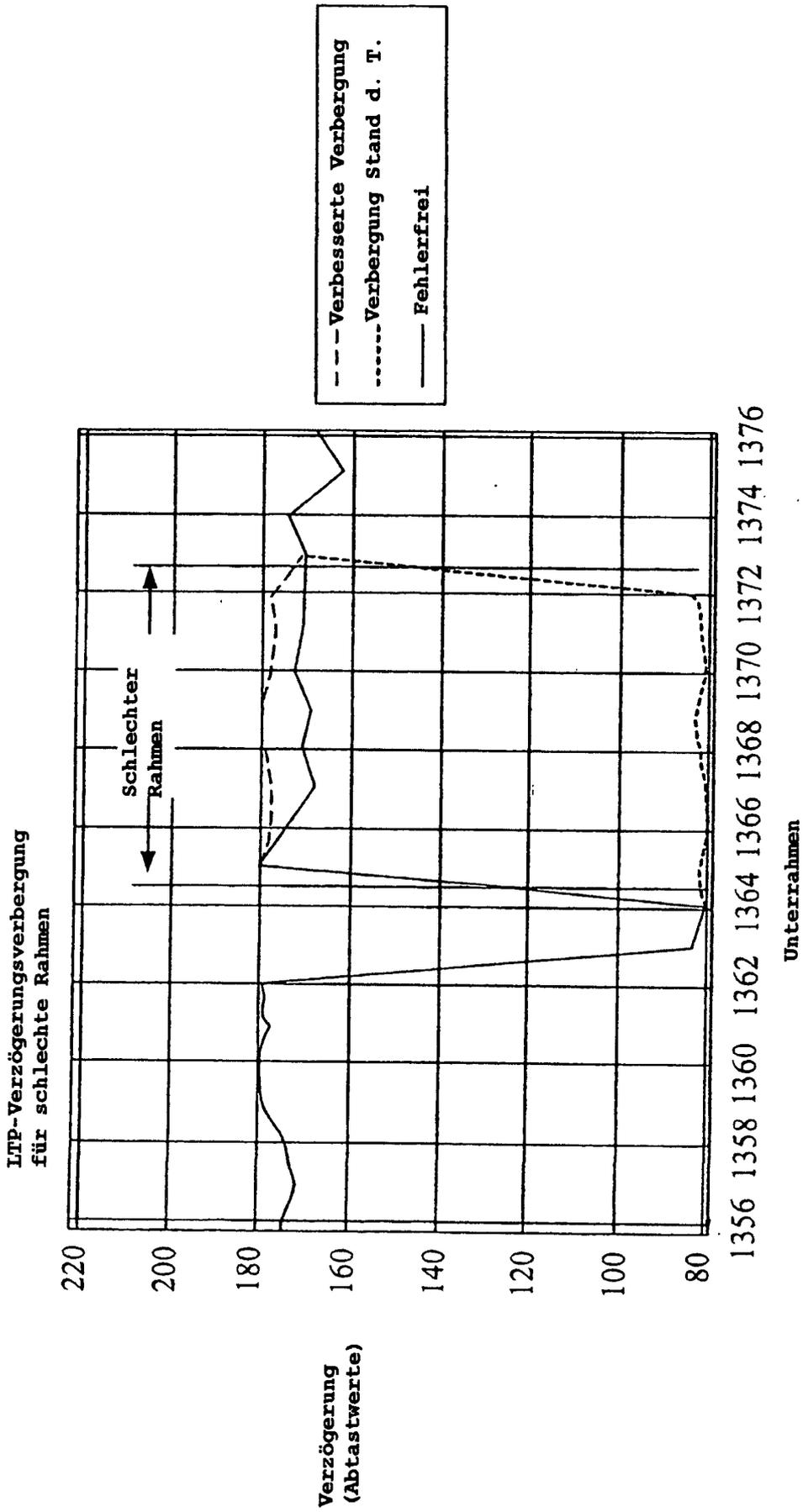


FIG. 10

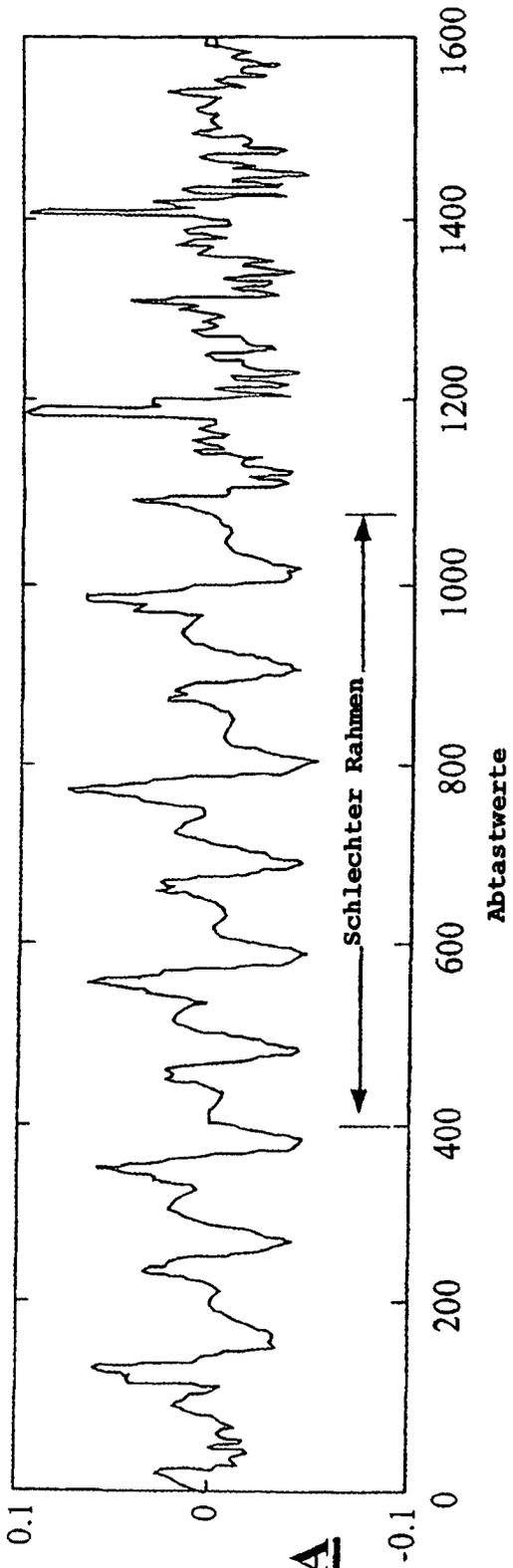


FIG. 11A

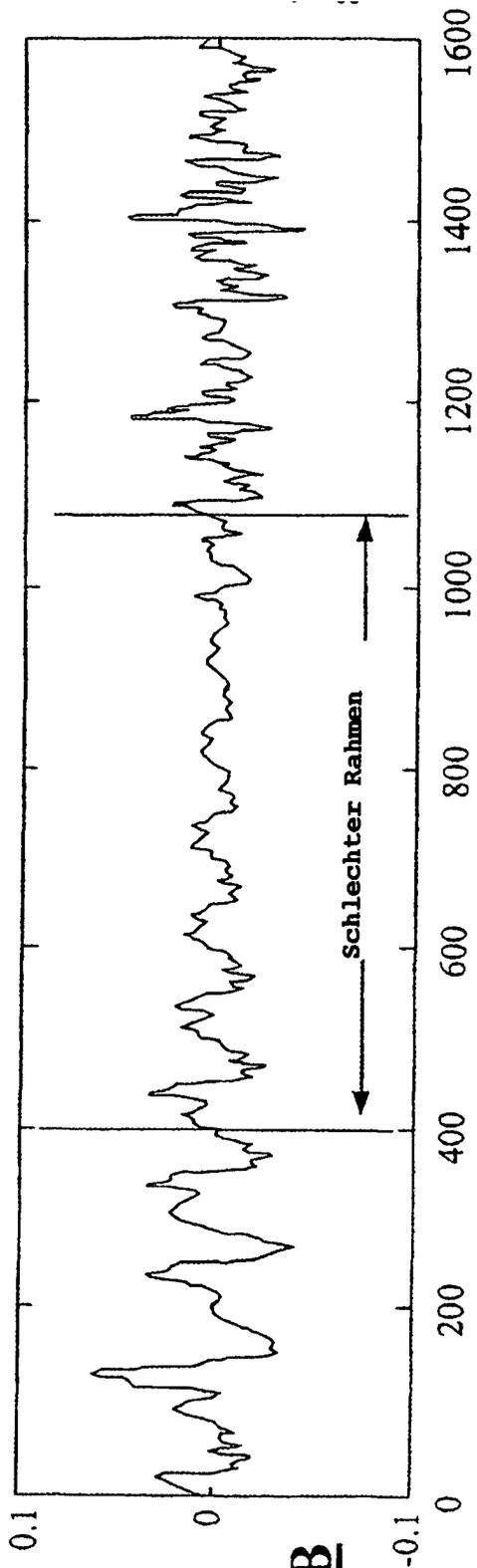


FIG. 11B

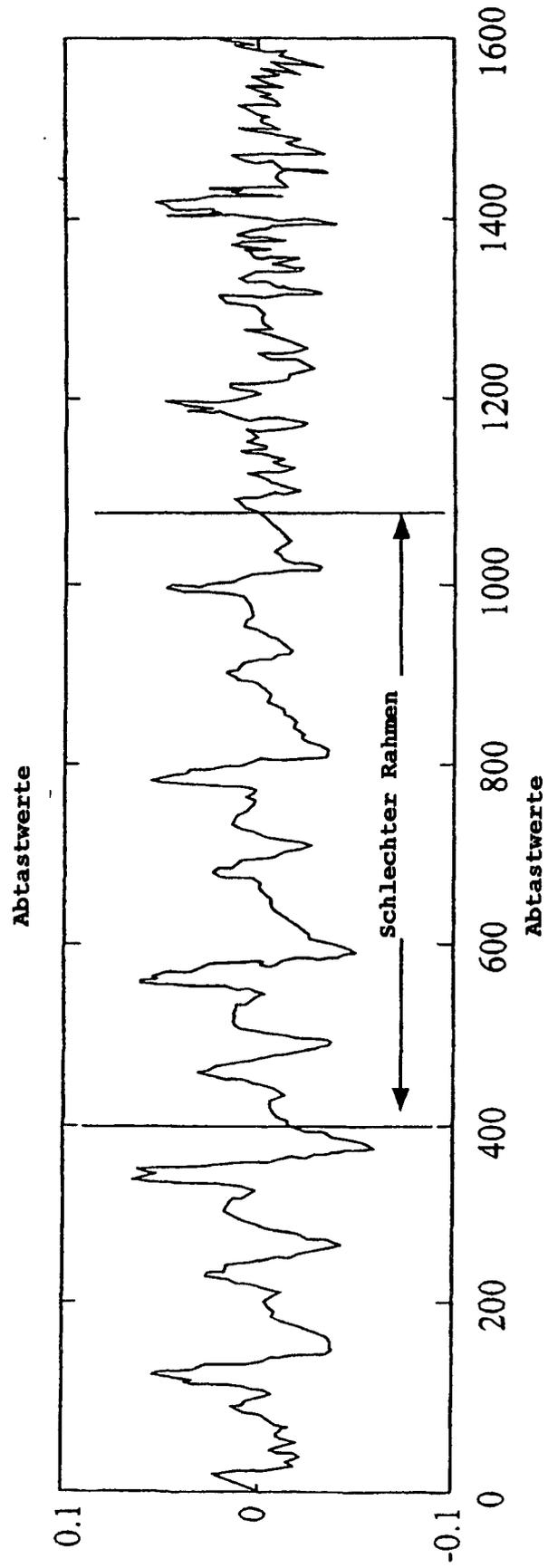


FIG. 11C