



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 695 32 879 T2** 2005.04.21

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 0 718 820 B1**

(21) Deutsches Aktenzeichen: **695 32 879.4**

(96) Europäisches Aktenzeichen: **95 118 647.7**

(96) Europäischer Anmeldetag: **27.11.1995**

(97) Erstveröffentlichung durch das EPA: **26.06.1996**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **14.04.2004**

(47) Veröffentlichungstag im Patentblatt: **21.04.2005**

(51) Int Cl.7: **G10L 19/14**
G10L 19/06

(30) Unionspriorität:

31448394 **19.12.1994** **JP**

32249494 **26.12.1994** **JP**

32249594 **26.12.1994** **JP**

17848495 **14.07.1995** **JP**

(73) Patentinhaber:

**Matsushita Electric Industrial Co., Ltd., Kadoma,
Osaka, JP**

(74) Vertreter:

**Tiedtke, Bühling, Kinne & Partner GbR, 80336
München**

(84) Benannte Vertragsstaaten:

DE, FR, GB

(72) Erfinder:

Morii, Toshiyuki, Inagi-shi, Tokyo, JP

(54) Bezeichnung: **Vorrichtung zur Sprachkodierung, zur linear-prädiktiven Analyse und zur Geräuschverminderung**

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

[0001] Die vorliegende Erfindung bezieht sich auf eine Vorrichtung zur Sprachcodierung, bei der Sprachinformationsstücke in digitale Signale mit einem geringen Informationsumfang codiert werden, und die digitalen Signale werden übertragen und decodiert, um eine effiziente Datenübertragung auszuführen. Die vorliegende Erfindung bezieht sich auch auf eine Vorrichtung zum Analysieren lineare Prädiktionskoeffizienten, wobei ein digitales Sprachsignal mit einer Analysierzeitlänge analysiert wird, um einen linearen Prädiktionskoeffizienten unter Verwendung der Vorrichtung zur Sprachcodierung zu gewinnen. Auch bezieht sich die vorliegende Erfindung auf eine Vorrichtung zur Geräuschverminderung, bei der ein in der Sprachinformation enthaltenes Geräusch auf einen moderaten Grad vermindert wird, bevor die Sprachinformation in der Vorrichtung zur Sprachcodierung codiert wird.

[0002] Auf dem Gebiet der digitalen mobilen Übertragung, wie einem Mobiltelefon, ist ein Kompressionscodierverfahren für mit einer geringen Bitrate gesendeter Sprachsignale erforderlich, weil es mehr Teilnehmer bei der digitalen mobilen Übertragung gibt, und die Forschungen und Entwicklungen bezüglich Kompressionscodierverfahren sind in verschiedenen Forschungseinrichtung ausgeführt worden. In Japan ist ein Codierverfahren, das vektorsummenstimulierte Linearprädiktion (VSELP) genannt wird, von Motorola company vorgeschlagen worden, wobei Signale mit einer Bitrate von 11,2 kbits pro Sekunde (kbps) als Normcodierverfahren für ein digitales Mobiltelefon angewandt, übertragen werden. Das digitale Mobiltelefon, das nach dem VSELP-Codierverfahren hergestellt wird, ist in Japan seit Herbst 1994 im Handel. Auch ein anderes Codierverfahren, das PSI-CELP, durch Tonhöhen synchronen Innovationscodes stimulierte Linearprädiktion, genannt wird, vorgeschlagen von NTT moving communication network Co., LTD., wobei Signale mit einer Bitrate von 5,6 kbps gesendet werden, wird in Japan als nächstes Normcodierverfahren für das nächste Mobiltelefon verwendet, und die Entwicklung des nächsten mobilen Telefons geht weiter. Die Normcodierverfahren werden durch Verbessern einer CELP gewonnen, die offenbart ist von M. R. Schroeder in "High Quality Speech at Low Bit Rates" Proc. ICASSP, 1985, Seiten 937 bis 940. In diesem CELP-Codierverfahren wird die aus eingegebener Sprache gewonnene Sprachinformation in Tonquelleninformationen separiert, die auf Vibrationstönen oder Stimmbändern der Sprachtraktinformation beruhen, basierend auf Gestalten des Sprachtraktes, der sich von den Stimmbändern bis zum Mund erstreckt. Die Tonquelleninformation wird gemäß einer Vielzahl von Tonquellenabtastungen codiert, die in einem Codebuch gespeichert sind, während die Sprachtraktinformation mit der eingegebenen Sprache berücksichtigt und verglichen wird, und die Sprachtraktinformation wird dann mit einem linearen Prädiktionskoeffizienten codiert. Das heißt, eine Analyse durch Synthese (A-b-S)-Verfahren wird beim CELP-Codierverfahren angewandt.

[0003] Ein grundlegender Algorithmus für das CELP-Codierverfahren wird beschrieben.

[0004] Fig. 1 ist ein grundlegendes Blockdiagramm einer herkömmlichen Vorrichtung zur Sprachcodierung nach dem CELP-Codierverfahren.

[0005] Wenn in Fig. 1 eine Stimme oder Sprache an eine Eingabesprachempfangseinheit **102** einer herkömmlichen Vorrichtung **101** zur Sprachcodierung als Stücke von Sprachdaten gegeben wird, werden eine Autokorrelationsanalyse und eine lineare Prädiktionskoeffizientenanalyse für alle Sprachdaten in einer LPC-Analysiereinheit **103** (Analysiereinheit mit linearem Prädiktionskoeffizienten) ausgeführt, um einen linearen Prädiktionskoeffizienten für alle Sprachdaten zu gewinnen. Danach werden in der Einheit **103** alle linearen Prädiktionskoeffizienten codiert, um einen LPC-Code zu erhalten, und der LPC-Code wird decodiert, um einen wiedergegebenen linearen Prädiktionskoeffizienten zu gewinnen.

[0006] Danach werden alle ersten Tonquellenabtastwerte, die in einem Adaptivcodebuch **104** gespeichert sind, und alle zweiten Tonquellenabtastwerte, die in einem probabilistischen Codebuch **105** gespeichert sind, herausgenommen und an eine Addiereinheit **106** gegeben. In der Addiereinheit **106** wird eine optimale Verstärkung für sowohl die ersten als auch die zweiten Tonquellenabtastwerte errechnet, die Tonquellenabtastwerte werden in ihrer Leistung justiert gemäß der optimalen Verstärkung, und eine Vielzahl von Synthesetonquellen werden gewonnen als Ergebnis aller Kombinationen von leistungsjustierten ersten Tonquellenabtastwerten und den leistungsjustierten zweiten Tonquellenabtastwerten. Das heißt, jede der Synthesetonquellen wird gewonnen durch Addieren entweder der ersten leistungsjustierten Tonquellenabtastwerte oder der zweiten leistungsjustierten Tonquellenabtastwerte.

[0007] In einer LPC-Synthetisierereinheit **107** werden danach die Synthesetonquellen gefiltert mit dem wiedergegebenen linearen Prädiktionskoeffizienten, den die LPC-Analysiereinheit **103** gewonnen hat, um eine Vielzahl von Synthesesprachen zu erhalten. In einer Vergleichseinheit **108** wird danach ein Abstand zwischen al-

len Sprachdaten errechnet, die die Sprachenempfangseinheit **102** aufgenommen hat, und jede der Synthesprachen wird errechnet, eine spezielle Synthesprache gemäß dem speziellen Abstand, der der Minimalwert unter den Abständen ist, wird ausgewählt aus den Synthesprachen, und ein spezieller erster Tonquellenabstastwert und ein spezieller zweiter Tonquellenabstastwert gemäß der speziellen Synthesprache wird gewonnen.

[0008] In einer Parametercodiereinheit **109** werden danach die optimalen Verstärkungen in der Codiereinheit **106** errechnet und codiert, um eine Vielzahl von Verstärkungscodes zu gewinnen. Der in der LPC-Analysiereinheit **103** gewonnene LPC-Code, die Indexcodes, die die speziellen Tonquellenabstastwerte aufzeigen, gewonnen in der Vergleichseinheit **108**, und die Verstärkungscodes werden auf eine Sendeleitung **110** in einer Gruppe übertragen. Auch eine Synthesetonquelle wird erzeugt aus dem Verstärkungscodes gemäß dem speziellen ersten Tonquellenabstastwert und dem speziellen ersten Tonquellenabstastwert in der Einheit **109**. Die Synthesetonquelle wird gespeichert im adaptiven Codebuch **104** als erster Tonquellenabstastwert, und der spezielle erste Tonquellenabstastwert wird aufgegeben.

[0009] In der LPC-Synthesiereinheit **107** wird darüber hinaus das akustische Gefühl für alle Sprachdaten mit dem linearen Prädiktionskoeffizienten, einem Frequenzbetonungsfilterkoeffizienten und einem durch Ausführen einer langfristigen Prädiktionsanalyse für alle Sprachdaten gewonnenen langfristige Prädiktionskoeffizienten, gewichtet. Die Tonquellenabstastwerte werden aus den Unterrahmen gewonnen, die man durch Einteilen aller Analysierblöcke im adaptiven Codebuch **104** und dem probabilistischen Codebuch **105** erhält.

[0010] Auch die lineare Prädiktionskoeffizientenanalyse, die die LPC-Analysiereinheit **103** ausführt, wird in verschiedenen Codierv Verfahren verwendet. Eine herkömmliche lineare Prädiktionskoeffizientenanalyse wird anhand **Fig. 2** beschrieben.

[0011] **Fig. 2** ist ein Blockdiagramm einer herkömmlichen linearen Prädiktionskoeffizienten-Analysier Vorrichtung.

[0012] Wenn, wie in **Fig. 2** gezeigt, eine Sprache in eine Eingabesprachaufnahmeinheit **112** einer herkömmlichen linearen Prädiktionskoeffizienten-Analysier Vorrichtung **111** gegeben wird, dann erfolgt das Umsetzen der Sprache in eine Vielzahl von Sprachsignalen X_i , die jeweils eine vorgeschriebene Analysierperiode aufweisen, und jedes der Sprachsignale X_i , das zeitsequentiell abgegeben wird, wird multipliziert mit einem Fensterkoeffizienten W_i in einer Fensterstelleinheit **113**. Ein Koeffizient in einem Hamming-Fenster, einem Hanning-Fenster, einem Blackman-Harris-Fenster oder dergleichen wird für den Fensterkoeffizienten W_i verwendet. Eine Fensterstellverarbeitung in der Einheit **113** wird folgendermaßen formuliert:

$$Y_i = W_i * X_i$$

[0013] Hier bedeutet i die Anzahl von Sprachsignalen ($i = 1$ bis L), L bedeutet die Anzahl von Sprachsignalen und Y_i bedeutet eine Vielzahl von fensterverarbeiteten Sprachsignalen.

[0014] Danach wird eine automatische Korrelationsanalyse für die fensterverarbeiteten Sprachsignale Y_i in einer Autokorrelationsanalysiereinheit **114** wie folgt ausgeführt:

$$V_j = \sum_{i=j+1}^L Y_i * Y_{i-j}$$

[0015] Hier bedeutet V_j eine Vielzahl von Autokorrelationsfunktionen, und j bedeutet die Anzahl von Autokorrelationsfunktionen.

[0016] Danach wird eine lineare Prädiktionsanalyse auf der Grundlage des Autokorrelationsverfahrens in einer linearen Prädiktionskoeffizienten-Analysiereinheit **115** ausgeführt, um einen linearen Prädiktionskoeffizienten für jedes der Sprachsignale zu erhalten. Die linearen Prädiktionsanalyse ist offenbart in verschiedenen Sprachinformationsverarbeitungsdokumenten, wie beispielsweise "The Autocorrelation Method", in der Literatur geschrieben von L. R. Labiner und R. W. Schaffer "Digital Processing of Speech Signals", Seiten 401 bis 403.

[0017] Da auch die Sprachinformation, gewonnen aus der eingegebenen Sprache, codiert wird gemäß einem der verbesserten CELP-Codierv Verfahren, kann eine Vielzahl von Sprachsignalen, die die Sprachinformation aufzeigen, mit einer sehr geringen Bitrate übertragen werden. Da jedoch die Sprachinformation komprimiert ist

gemäß einem Sprachvokalisiermodell, kann die Toninformation einschließlich der Sprachinformation nicht in geeigneter Weise verarbeitet werden gemäß einem der verbesserten CELP-Codierverfahren. Das heißt, im Falle, daß das Hintergrundgeräusch oder ein eingestelltes Geräusch mit den Sprachsignalen vorhanden ist, gibt es den Nachteil, daß die Sprachsignale nicht in effizienter Weise codierbar sind, und in der wiedergegebenen Sprach tritt Allophonie auf. Um diesen Nachteil zu beseitigen, ist ein Verfahren zum Reduzieren von Geräuschen vorgeschlagen worden, die in den Sprachsignalen vorhanden sind. Beispielsweise wird ein mit den Sprachsignalen vorhandenes Geräusch reduziert durch einen Geräuschbeseitiger im Standard-PSI-CELP-Codierverfahren, bevor die Sprachsignale codiert werden. Dieser Geräuschbeseitiger ist aufgebaut aus einem Kalman-Filter. Das heißt, das Vorhandensein von Sprache wird festgestellt, und die Sprache wird in adaptiver Weise vom Kalman-Filter gesteuert, um das in der Sprache vorhandene Geräusch zu reduzieren. Das Hintergrundgeräusch kann folglich zu einem gewissen Grad vom Geräuschbeseitiger entfernt werden. Ein Geräusch mit hohem Pegel oder ein Geräusch, das in der Sprache enthalten ist, kann jedoch nicht in effektiver Weise vermindert oder herausgezogen werden.

[0018] Ein effektiveres Geräuschreduzierverfahren ist als spektrales Subtraktionsverfahren in der Literatur offenbart, verfaßt von S. F. Boll "Suppression of Acoustic Noise in Speech Using Spectral Subtraction" IEEE, Trans. ASSP., Band 27, Nummer 2, Seiten 113 bis 120, 1979. Beim Spektralsubtraktionsverfahren wird eine diskrete Fourier-Transformation ausgeführt, um eine Vielzahl von eingegebenen Sprachsignalen in eine Vielzahl von Spektren umzusetzen, und eines oder mehrere Geräusche werden aus diesen Spektren subtrahiert. Dieses Verfahren wird hauptsächlich angewandt für eine Spracheingabeeinheit und eine Vorrichtung zur Spracherkennung. Eine herkömmliche Vorrichtung zur Geräuschverminderung, in der die Spektralsubtraktionsverfahren angewandt wird, um ein im Sprachsignal enthaltenes Geräusch aus dem Sprachsignal zu entfernen, wird anhand **Fig. 4** beschrieben.

[0019] Wie in **Fig. 4** gezeigt, wird ein Geräuschspektrum in einer ersten Prozedur angenommen, und ein Geräusch, dessen Spektrum angenommen wird, wird aus dem Sprachsignal in einer zweiten Prozedur subtrahiert. In der ersten Prozedur wird eine Vielzahl von Geräuschsignalen S_n , die ein Geräusch aufzeigen, seriell in einen Analog-Digital-Umsetzer (A/D-Umsetzer) **122** einer herkömmlichen Vorrichtung **121** zum Geräuschsubtrahieren eingegeben, und die Geräuschsignale S_n werden umgesetzt in eine Vielzahl digitaler Geräuschsignale. In diesem Falle ist irgendein Sprachsignal nicht in den Geräuschsignalen S_n enthalten. Danach wird eine diskrete Fourier-Transformation in einer Fourier-Transformationseinheit **123** für jeden Rahmen digitaler Geräuschsignale ausgeführt, und ein Geräuschspektrum wird aus jedem Rahmen gewonnen. Jeder Rahmen ist zusammengesetzt aus einer Serie digitaler Geräuschsignale mit einer konstanten Zeitlänge. Danach wird ein Durchschnittsgeräuschspektrum in einer Geräuschanalyseinheit **124** gewonnen durch Durchschnittsbildung einer Vielzahl von Geräuschspektren, und ein Durchschnittsgeräuschspektrum wird gespeichert in einer Geräuschspektrumspeichereinheit **125** als repräsentatives Geräuschspektrum des Geräuschs. Die erste Prozedur wird ausgeführt für verschiedene Geräuschsignale, die verschiedene Geräuscharten aufzeigen, und eine Vielzahl von spektral aufzeigenden repräsentativen Geräuschen, die verschiedene Arten von Geräuschen aufzeigen, wird in der Speichereinheit **125** gespeichert. In der zweiten Prozedur wird eine Vielzahl von Sprachsignalen S_s seriell in einen A/D-Umsetzer **126** eingegeben, die ein Sprache aufzeigen, die ein Geräusch enthält, und eine Vielzahl von digitalen Sprachsignalen wird gewonnen. Danach wird eine diskrete Fourier-Transformation in einer Fourier-Transformationseinheit **127** ausgeführt, und ein Sprachspektrum einschließlich eines aktuellen Geräuschspektrums wird gewonnen. Danach wird ein repräsentatives Geräuschspektrum, das zu dem aktuellen Geräuschspektrum paßt, aus der Speichereinheit **125** ausgelesen, und das ausgelesene repräsentative Geräuschspektrum wird aus dem Sprachspektrum in einer Geräuschsubtrahiereinheit **128** abgezogen, um das aktuelle Geräuschspektrum zu beseitigen. Eine inverse Fourier-Transformation wird danach für das Sprachspektrum in einer inversen Fourier-Transformationseinheit **129** ausgeführt, und ein Sprachausgangssignal S_o wird gewonnen.

[0020] Um sowohl das Geräusch als auch das Sprachspektrum zu gewinnen, wird ein Amplitudenspektrum für jedes der Geräusche und die Sprache ausgeführt, die darin enthalten sind. Das heißt, eine reale Komponente einer Norm, festgelegt in einer komplexen Ebene für die Amplitude eines Geräuschs oder einer Sprache, und eine Imaginärkomponente der Norm werden jeweils quadriert, die quadrierte Realkomponente und die quadrierte Imaginärkomponente werden addiert, um einen quadrierten Absolutwert zu gewinnen, und eine Quadratwurzel des quadrierten Absolutwerts wird errechnet als Amplitudenspektrum. Auch im Falle, bei dem die inverse Fourier-Transformation für das Amplitudenspektrum ausgeführt wurde, aus dem ein Geräuschspektrum subtrahiert wird, wird eine Phasenkomponente eines jeden Sprachsignals S_s als Phasenkomponente des Amplitudenspektrums verwendet.

[0021] (1) Um eine Sprachcodiervorrichtung und eine Decodiervorrichtung in kleiner Größe als Mobiltelefon

einzusetzen, ist es erforderlich, die Speicherkapazität eines Nurlesespeichers (ROM) zu reduzieren, in dem eine Vielzahl erster Tonquellenabstastwerte eines adaptiven Codebuchs und eine Vielzahl zweiter Tonquellenabstastwerte eines probabilistischen Codebuchs gespeichert sind. Da jedoch eine große Anzahl von Codevektoren erforderlich ist, um eine Vielzahl feststehender Tonquellen zu speichern, die die zweiten Tonquellenabstastwerte in der herkömmlichen Sprachcodier Vorrichtung **101** repräsentieren, ist es schwierig, die Vorrichtung **101** in kleiner Größe zu realisieren. Um die Anzahl von Codevektoren im ROM zu reduzieren, wird beispielsweise ein langer Vektor verschoben, um als Vielzahl von Codevektoren verwendet zu werden. Gleichermaßen werden ähnliche Codevektoren gewonnen durch Verschieben des langen Vektors, und dann kommt der Nachteil auf, daß die Qualität der wiedergegebenen Sprache verschlechtert ist, verglichen mit der, die wiedergegeben wird unter einer großen Anzahl von Codevektoren, die sich voneinander unterscheiden. Da es auch erforderlich ist, den Codevektor jedes Mal zu errechnen und zu erzeugen, gibt es den weiteren Nachteil, daß ein großer Umfang an Rechnungen erforderlich ist.

[0022] (2) Da auch das VSELP-Codierverfahren und das PSI-CELP-Codierverfahren erzielt werden durch Verbessern des CELP-Codierverfahrens, wird dieselbe Verarbeitung für beliebig eingegebene Stimme oder Sprache im VSELP-Codierverfahren und im PSI-CELP-Codierverfahren ausgeführt. Die eingegebene Stimme oder Sprache kann folglich nicht in effizienter Weise codiert werden.

[0023] Sprachinformationsstücke, die unter realen Umständen aufgezeichnet wurden, unterscheiden sich weitestgehend voneinander in einem Gesichtspunkt örtlicher Eigenschaften. Jede Sprachinformation ist zusammengesetzt aus einem oder mehreren Sprachabschnitten oder einem oder mehreren Ruheabschnitten. Sprache vom Sprachabschnitt ist zusammengesetzt aus einem oder mehreren Konsonanten oder aus einem oder mehreren Vokalen. Jeder Konsonant ist klassifiziert in einen stimmlosen Konsonanten und in einen stimmhaften Konsonanten. Jeder Vokal ist eingeteilt in einen Vokalstationärabschnitt und in einen Vokalübergangsabschnitt. Im Vokalstationärabschnitt ist eine Sprachtonhöhe und eine Bewegung des Mundes stabil. Im Vokalübergangsabschnitt ändern sich immer die Sprachtonhöhe und die Mundbewegung. Da der ruhige Abschnitt, der stimmlose Konsonant, der stimmhafte Konsonant, der Vokalstationärabschnitt und der Vokalübergangsabschnitt folglich unterschiedliche Eigenschaften haben, gibt es ein optimales Codierverfahren für diese.

[0024] In Fällen, bei denen das CELP-Codierverfahren verwendet wird, ist nachstehend beschrieben, wie die Sprachinformation codiert wird, während die örtlichen Eigenschaften berücksichtigt werden. Da es keine Sprache in einem Ruheabschnitt gibt, existiert nur ein Geräusch in den realen Umständen im Ruheabschnitt, und eine Zeitlänge des Ruheabschnitts ist erforderlich, um darüber zu informieren. Die Zeitinformation des Ruheabschnitts kann folglich codiert werden mit einer sehr geringen Bitrate unter Fortlassen der Tonquellenabstastwerte. Der stimmlose Konsonant wird klassifiziert in einen Verschlusslaut, wie beispielsweise ein Phonem /p/, /t/ oder dergleichen oder in einen Reibelaut, wie ein Phonem /s/, /h/ oder dergleichen. Da sich die Stimmleistung eines Verschlusslautes ganz kurzzeitig ändern kann und es wichtig ist, die kurzzeitige Änderung zu erkennen, ist es vorzuziehen, daß der Verschlusslaut codiert wird in einer Einheit kurzer Rahmenlänge. Die ersten Tonquellenabstastwerte, die im adaptiven Codebuch **104** gespeichert sind, sind nicht erforderlich, um den Verschlusslaut zu codieren. Ebenfalls sind, im Falle des Reibelauts, eine Ausbreitungseigenschaft und eine Zeitlänge wichtig. Die ersten Tonquellenabstastwerte, die im adaptiven Codebuch **104** gespeichert sind, sind folglich nicht erforderlich, um den Reibelaut zu codieren. Im Falle des stimmhaften Konsonanten sind eine kurzzeitige Stimmleistungsänderung, Sprachtraktinformation und Tonquelleninformation von Bedeutung. Folglich ist der größte Informationsumfang erforderlich zum Codieren des stimmhaften Konsonanten. Im Vokalstationärabschnitt wird eine Vielzahl von Wellen mit ähnlich gestalteten Wellenformen hintereinander gebildet. Der Vokalstationärabschnitt kann folglich codiert werden unter Verwendung eines geringen Informationsumfangs in Fällen, bei denen die ersten Tonquellenabstastwerte, gespeichert im adaptiven Codebuch **104**, Verwendung finden. Im Vokalübergangsabschnitt ist die Änderung der Sprachtraktinformation und der Tonquelleninformation größer als im Vokalstationärabschnitt, und eine Sprachleistung im Vokalübergangsabschnitt ist groß. Folglich kann die Verschlechterung der Tonqualität leicht bemerkbar sein. Ein großer Informationsumfang ist folglich in derselben Weise wie beim stimmhaften Konsonanten erforderlich, um den Vokalübergangsabschnitt zu codieren.

[0025] In Fällen, bei denen das Codierverfahren örtlich von sowohl dem Ruheabschnitt, dem Verschlusslaut, dem Reibelaut, dem stimmhaften Konsonanten, dem Vokalstationärabschnitt als auch dem Vokalübergangsabschnitt geändert wird zum adaptiven Verteilen von Informationsstücken, kann die eingegebene Sprache in effektiver Weise codiert werden. Das heißt, da sich die örtlichen Eigenschaften der Sprachinformation, aufgezeichnet unter realen Umständen, weitestgehend voneinander unterscheiden, in Fällen, bei denen die Sprachinformationen in adaptiver Weise codiert werden, während die örtlichen Eigenschaften in positiver Weise berücksichtigt werden, kann die Codiereffizienz verbessert werden und eine Vielzahl von Synthesprachen kann

vorzugsweise mit einer geringen Durchschnittsbitrate gewonnen werden. Basierend auf dieser Idee wird ein Codierverfahren vorgeschlagen, bei dem eine Vielzahl von Codiermodulen Verwendung findet. Beispielsweise wird ein Sprachcodierverfahren variabler Bitrate verwendet, wie es offenbart ist im Dokument 2-Q-23, gelesen in Spring research convention of Japanese Acoustic Society, und im QCELP-Verfahren, das die Qaucom Company vorgeschlagen hat. Das QCELP-Verfahren wird verwendet als Standardcodierverfahren (TIA-IS96) für ein digitales zelliges Telefon in Nordamerika.

[0026] Eines der Vielzahl von Codiermodulen wird ausgewählt entsprechend einer einfachen Regel beim Sprachcodierverfahren variabler Bitrate und beim QCELP-Verfahren. Es gibt folglich eine Wahrscheinlichkeit, daß ein Codiermodul nicht für einen Sprachinformationsteil eingerichtet ist und ausgewählt wird aufgrund eines Fehlers, und dann gibt es den Nachteil, daß eine kratzende Allophonie auftritt. Um diesen Nachteil bei der Vorrichtung zur Sprachcodierung zu beseitigen, die betrieben wird gemäß der Analyse nach dem Syntheseverfahren, wird die Sprachinformation codiert unter Verwendung aller Codiermodule, eine Vielzahl von Codierstörungen gemäß den Codiermodulen wird untereinander verglichen, und ein Codiermodul gemäß einer Codierverzerrung, die am geringsten unter den Codierverzerrungen ist, wird als am besten angepaßtes Codiermodul verwendet. In diesem Falle wird jedoch der Umfang der Rechnungen, die erforderlich sind zum Bestimmen des adaptiven Codiermoduls, extrem groß, und es ist schwierig, die Vorrichtung zum Sprachcodieren, die betrieben wird gemäß dem obigen Auswahlverfahren, in einer kleingebauten Kommunikationsvorrichtung, wie einem Mobiltelefon, unterzubringen. Es ist ebenfalls schwierig, eine komplizierte Regel zur korrekten Auswahl des passenden Codiermoduls zum Zwecke des Vermeidens vom Auftreten der Allophonie bereitzustellen.

[0027] (3) Auch in Fällen, bei denen die Fensterkoeffizienten W_i verwendet werden bei der herkömmlichen Analysiervorrichtung **111** für lineare Prädiktionskoeffizienten, weil ein Wert eines jeden Fensterkoeffizienten W_i ein Mittenabschnitt einer Analysierperiode mit hohem Wert ist und Werte jeden Fensterkoeffizienten W_i an beiden Endabschnitten der Analysierperiode sehr niedrig ist, gibt es den Nachteil, daß ein Informationsstück für jedes fensterverarbeitete Sprachsignal Y_i ein Informationsstück für jedes Sprachsignal X_i am Mittenabschnitt der Analysierperiode darstellt. Um diesen Nachteil zu umgehen, werden ein rückwärtiger Abschnitt des vorangehenden Sprachsignals X_{i-1} am hinteren Abschnitt einer vorangehenden Analysierperiode, ein aktuelles Sprachsignal X_i an einer aktuellen Analysierperiode und einem Vorderteil eines nachfolgenden Sprachsignals X_{i+1} an einem Vorderabschnitt einer nachfolgenden Analysierperiode, abgegeben aus der Eingabesprachempfangseinheit **112** in dieser Reihenfolge mit einem Fensterkoeffizienten W_i für das aktuelle Sprachsignal X_i nach einem normalen CELP-Codierverfahren, multipliziert. In diesem Falle kann ein Informationsstück über das gesamte aktuelle Sprachsignal X_i auf einem Informationsstück für ein aktuelles fensterverarbeitetes Sprachsignal Y_i wiedergegeben werden.

[0028] Da jedoch die Multiplikation zu erwarten ist, bis der Vorderabschnitt der nachfolgenden Analysierperiode passiert, wird ein in einem Codec ausgeführter Codierprozeß verzögert um eine Periode, die dem Vorderabschnitt der nachfolgenden Analysierperiode äquivalent ist. Um diese Codierprozeßverzögerung zu vermindern, wird der Vorderabschnitt der nachfolgenden Analysierperiode abgekürzt auf mehrere Millisekunden in einem Codec, der verwendet wird für eine digitale Mobilübertragung, wie ein Mobiltelefon. In diesem Falle ist es schwierig, daß die Information über das gesamte aktuelle Sprachsignal X_i auf der Information für das aktuell fensterverarbeitete Sprachsignal Y_i wiedergegeben wird. Wenn folglich ein Sprachstück, wie ein betonter Konsonant, in dem sich das Sprachspektrum weitgehend ändert, in die Eingabesprachempfangseinheit **112** eingegeben ist, gibt es den Nachteil, daß die Qualität der wiedergegebenen Sprache örtlich verschlechtert ist.

[0029] (4) Obwohl das Spektrumsabtraktionsverfahren, angewandt in der herkömmlichen Vorrichtung **121** zur Geräuschsubtraktion, effektiver ist, um ein Geräusch aus einer Sprache zu subtrahieren in den Fällen, bei denen das Verfahren bei der Vorrichtung zur Echtzeitverarbeitung angewandt wird, gibt es viele Nachteile bei dem Geräuschannahmeverfahren oder bei den Herstellkosten der Vorrichtung. Ein erster Nachteil mit der, daß die Annahme des Geräuschspektrums schwierig ist, weil eine Position des Sprachsignals in Datenstücken vorhanden ist, die sich nicht spezifizieren lassen. Ein zweiter Nachteil besteht darin, daß der Rechenumfang in der Vorrichtung groß ist. Ein dritter Nachteil besteht darin, daß die erforderliche Speicherkapazität zum Speichern des Geräuschspektrums in einem Direktzugriffsspeicher groß ist. Ein vierter Nachteil besteht darin, daß ein Sprachspektrum, aus dem ein Geräuschspektrum mit hoher Intensität subtrahiert wird, weitestgehend verzerrt ist, und die Qualität der wiedergegebenen Sprache verschlechtert ist.

[0030] Das Dokument WO 92 06470 A offenbart ein Verfahren und ein System zum Codieren analoger Signale, wie beispielsweise Sprachsignale, durch orthogonales Vergleichen der Signale mit Sequenzen, die in Tochtercodebüchern gespeichert sind, hergeleitet aus einem Muttercodebuch, um so passende Sequenzen für ein Ergebnis des Vergleichs zu bestimmen. Die beschriebene Technik ist anwendbar auf ein analoges

CELP-Signalcodiersystem.

[0031] Das Dokument Easton M. G. at al.: "A CELP Codebook and Search Technique Using a Hopfield Net", Speech Processing 1, Toronto, 14. bis 17. Mai 1991, Band 1, 14. Mai 1991, Institute of Electrical and Electronics Engineers, Seiten 685 bis 688, XP000245321, offenbart eine Suchtechnik unter Verwendung eines neutralen Hopfield-Netzes zum Suchen eines Codebuchs und zur Auswahl des nächstliegenden optimalen Codevektors. Das Codebuch ist ein nach Art eines Massennistplatzes stimulierbares Codebuch mit einer speziellen Struktur für das CELP-Codieren, wobei ein Hauptcodebuch eingeteilt wird in mehrere Untercodebücher.

[0032] Eine Aufgabe der vorliegenden Erfindung ist es, unter Berücksichtigung der Nachteile der oben beschriebenen herkömmlichen Sprachcodier Vorrichtung eine verbesserte Vorrichtung zur Sprachcodierung zu schaffen.

[0033] 26 Um diese Aufgabe zu lösen, sieht die vorliegende Erfindung eine Vorrichtung zur Sprachcodierung vor, wie sie in den anliegenden Patentansprüchen angegeben ist.

KURZE BESCHREIBUNG DER ZEICHNUNG

[0034] Aufgaben, Merkmale und Vorteile der vorliegenden Erfindung werden aus der nachstehenden Beschreibung in Verbindung mit der nachstehend aufgeführten beiliegenden Zeichnung deutlich.

[0035] **Fig. 1** ist ein funktionales Blockdiagramm einer herkömmlichen Vorrichtung zur Sprachcodierung nach dem CELP-Codierverfahren;

[0036] **Fig. 2** ist ein Blockdiagramm einer herkömmlichen linearen Prädiktionskoeffizienten-Analysier Vorrichtung, die für eine Vorrichtung zur Sprachcodierung verwendet wird;

[0037] **Fig. 3** zeigt einen Fensterkoeffizienten, durch den ein Abschnitt eines vorhergehenden Sprachsignals zu einer vorhergehenden Analysierperiode, ein gesamtes aktuelles Sprachsignal bei einer aktuellen Analysierperiode und ein Abschnitt eines nachfolgenden Sprachsignals zu einer nachfolgenden Analysierperiode in der in **Fig. 3** gezeigten herkömmlichen linearen Prädiktionskoeffizienten-Analysier Vorrichtung multipliziert werden;

[0038] **Fig. 4** ist ein Blockdiagramm einer herkömmlichen Geräuschverminderungsvorrichtung, bei der ein Spektralsubtraktionsverfahren zur Verminderung des Rauschens angewandt wird, das in einem Sprachsignal enthalten ist;

[0039] **Fig. 5** ist ein Blockdiagramm einer Vorrichtung zur Sprachcodierung nach einem ersten Ausführungsbeispiel der vorliegenden Erfindung;

[0040] **Fig. 6** ist ein Blockdiagramm einer Auswahlregel, die eine Vorrichtung zum Herrichten einer Codiermodulauswahlregel herrichtet, verwendet in einer in **Fig. 5** gezeigten Vorrichtung zur Sprachcodierung;

[0041] **Fig. 7** zeigt eine Vielzahl von Eigenschaftsparametern und Signal-/Geräuschverhältnissen (S/N-Verhältnisse), die in einer Codierverzerrungsspeichereinheit von der Vorrichtung zur in **Fig. 6** gezeigten Auswahlregelvorbereitung gespeichert sind;

[0042] **Fig. 8** zeigt ein neuronales Netz, das aufgebaut ist in einer statistischen Verarbeitungseinheit der Vorrichtung zur Auswahlregelvorbereitung, die in **Fig. 6** gezeigt ist, und in einer Codiermodulauswahleinheit der Vorrichtung zur Sprachcodierung eingesetzt ist, die in **Fig. 5** gezeigt ist;

[0043] **Fig. 9** zeigt eine Bitrate in jedem der fünf Codiermodule, die zum Abschätzen der Auswahl eines geeigneten Codiermoduls im in **Fig. 8** gezeigten neuronalen Netz verwendet wird;

[0044] **Fig. 10** zeigt die Wahrscheinlichkeit, daß ein geeignetes Codiermodul in korrekter Weise in der Codiermodulauswahleinheit ausgewählt wird;

[0045] **Fig. 11** ist ein Blockdiagramm einer Vorrichtung zur Sprachcodierung nach einem zweiten Ausführungsbeispiel, das nicht in den Schutzbereich der vorliegenden Erfindung fällt;

[0046] **Fig. 12** zeigt eine feststehende Tonquelle, die gewonnen wird durch Auslesen der feststehenden Ton-

quelle in einer Normalrichtung von einem Unterrahmen eines probabilistischen Codebuchs der Sprachcodier-
vorrichtung, die in **Fig. 11** gezeigt ist, und eine umkehrbefestigte Tonquelle, die gewonnen wird durch Auslesen
der feststehenden Tonquelle in umgekehrter Richtung vom Unterrahmen des probabilistischen Codebuchs;

[0047] **Fig. 13** ist ein Blockdiagramm einer Vorrichtung zur Klangwellenverbesserung nach dem zweiten Aus-
führungsbeispiel;

[0048] **Fig. 14** ist ein Blockdiagramm einer anderen Vorrichtung zur Tonquellenverbesserung nach dem zwei-
ten Ausführungsbeispiel;

[0049] **Fig. 15** ist ein Blockdiagramm einer Analysiervorrichtung für lineare Prädiktionskoeffizienten nach ei-
nem dritten Ausführungsbeispiel, das nicht in den Schutzbereich der vorliegenden Erfindung fällt;

[0050] **Fig. 16** zeigt eine Addition einer Vielzahl autokorrelierter Funktionen, die gewonnen werden in einer
Partialanalysierperiode, bei der ein Partialsprachsignal multipliziert wird mit einem Fensterkoeffizienten;

[0051] **Fig. 17** ist ein Blockdiagramm einer Geräuschreduziervorrichtung nach einem vierten Ausführungsbei-
spiel, das nicht in den Schutzbereich der vorliegenden Erfindung fällt; und

[0052] **Fig. 18** ist ein Blockdiagramm eines Geräuschreduziersystems, bei dem die in **Fig. 17** gezeigte Vor-
richtung zur Geräuschreduzierung vorgesehen ist.

DETAILLIERTE BESCHREIBUNG DER AUSFÜHRUNGSBEISPIELE

[0053] Bevorzugte Ausführungsbeispiele der Vorrichtung zur Sprachcodierung sind nachstehend anhand der
Zeichnung beschrieben.

[0054] Ein erstes Ausführungsbeispiel der vorliegenden Erfindung wird als erstes beschrieben.

[0055] **Fig. 5** ist ein Blockdiagramm einer Vorrichtung zur Sprachcodierung nach einem ersten Ausführungs-
beispiel der vorliegenden Erfindung.

[0056] Wie in **Fig. 5** gezeigt, ist die Vorrichtung **11** zur Sprachcodierung ausgestattet mit:
einer Spracheingabeempfangseinheit **12** zum Aufnehmen einer Stimme oder einer Sprache und zum Umset-
zen der Stimme oder Sprache in eine Vielzahl digitaler Sprechsignale mit derselben vorherbeschriebenen Zeitlän-
ge;
einer Autokorrelationsanalysiereinheit **13** zum Aufteilen eines jeden der digitalen Sprachsignale in Zeitmultip-
lexart, ausgegeben von der Einheit **12**, in eine Vielzahl digitaler Analysesignale zu Analysierperioden, die ei-
nem Ruheabschnitt von Sprache entsprechen, einem Verschlusslaut der Sprache, einem Reibelaut eines
stimmlosen Konsonanten, eines stimmhaften Konsonanten, eines stimmlosen Konsonanten, eines stationären
Vokalabschnitts von einem Vokal oder einem Vokalübergangsabschnitt eines Vokals, Analysieren einer Auto-
korrelation eines jeden Analysesignals und Erzielen eines Autokorrelationskoeffizienten eines Leistungsinforma-
tionsstückes für jedes digitale Analysesignal;
einer linearen Prädiktionskoeffizienten-Analysiereinheit **14** (LPC-Analysiereinheit) zum Erzielen eines linearen
Prädiktionskoeffizienten für jedes der digitalen Analysesignale durch Analysieren der Autokorrelationskoeffizi-
enten, die die Analysiereinheit **13** erzielt;
einer vorläufigen Tonhöhenauswahleinheit **15** zum inversen Filtern eines jeden der digitalen Analysesignale
mit dem zugehörigen linearen Prädiktionskoeffizienten, gewonnen in der LPC-Analysiereinheit **14**, um eine li-
neare Prädiktionsrestwellenform für jedes digitale Analysesignal zu gewinnen, Analysieren einer langfristigen
primären Korrelation der linearen Prädiktionsrestwellenform für jedes digitale Analysesignal und Auswählen ei-
nes Kandidaten einer Fundamentalfrequenz (oder Tonhöhe) der langfristigen primären Korrelation für jedes di-
gitale Analysesignal;
einer Auswahlregelspeichereinheit **16** zum Speichern einer Codiermodulauswahlregel;
einer Codiermodulspeichereinheit **17** zum Speichern einer Vielzahl von Codiermodulen, die jeweils einem Ru-
heabschnitt einer Sprache entsprechen, einem Verschlusslaut eines stimmlosen Konsonanten der Sprache, ei-
nem Reibelaut des stimmlosen Konsonanten der Stimme, einem stimmhaften Konsonant der Stimme, einem statio-
nären Vokalabschnitt eines Vokals der Stimme oder eines Vokalübergangsabschnitts eines Vokals der Stimme;
einer Codiermodulauswahleinheit **18** zum Auswählen eines passenden Codiermoduls, der am besten geeignet
ist für Codieren eines jeden digitalen Analysesignals aus den Codiermodulen, die in der Codiermodulspeiche-
reinheit **17** gespeichert sind, gemäß der Modulauswahlregel, die in der Auswahlregelspeichereinheit **16** ge-

speichert ist, unter Verwendung einer ersten Gruppe von Spracheigenschaften, die gewonnen sind in der Autokorrelationsanalysiereinheit **13**, wie der Autokorrelationskoeffizient, die Leistungsinformation bei einer zugehörigen Analysierperiode, ein Leistungsänderungsgrad zur zugehörigen Analysierperiode und eine Differenz in der Leistungsinformation des digitalen Analysesignals zwischen der zugehörigen Analysierperiode und einer Analysierperiode, die vor der zugehörigen Analysierperiode liegt, einer zweiten Gruppe von Spracheigenschaften, gewonnen von der LPC-Analysiereinheit **14**, wie der lineare Prädiktionskoeffizient, eine Spektralhülle und eine Differenz im Spektrum zwischen einer zugehörigen Analysierperiode und einer vor der zugehörigen Analysierperiode liegenden Analysierperiode, und eine dritte Gruppe von Spracheigenschaften, die gewonnen wird in der vorläufigen Tonhöhenauswahleinheit **15**, wie eine Leistung der linearen Prädiktionsrestwellenform bei einer zugehörigen Analysierperiode, eine Leistung einer langfristigen Prädiktionsrestwellenform zu der zugehörigen Analysierperiode, eine Leistungsdifferenz in der linearen Prädiktionsrestwellenform zwischen der zugehörigen Analysierperiode und einer vor der zugehörigen Analysierperiode liegenden Analysierperiode, eine Leistungsdifferenz in der langfristigen Prädiktionsrestwellenform zwischen der zugehörigen Analysierperiode und der vorigen Analysierperiode, ein Leistungsänderungsgrad der linearen Prädiktionsrestwellenform zur zugehörigen Analysierperiode, ein Leistungsänderungsgrad der langfristigen Prädiktionsrestwellenform zu der zugehörigen Analysierperiode und dem langfristigen Prädiktionskoeffizienten der linearen Prädiktionsrestwellenform;

einer Codiereinheit **19** zum Codieren eines jeden digitalen Analysesignals zu einem codierten Signal im passenden Codiermodul, ausgewählt in der Modulauswahleinheit **18**; und mit

einer Codiersignalvereinheitlichungseinheit **20** zum seriellen Sammeln einer Vielzahl codierter Signale, gewonnen durch Codieren der digitalen Analysesignale der digitalen Sprachsignale in der Codiereinheit **19**, Vereinheitlichen der codierten Signale zu einer Serie codierter Signale und Abgeben der Serie codierter Signale auf eine Sendeleitung **21**.

[0057] Der Kandidat der fundamentalen Frequenz (oder Tonhöhe), gewonnen in der Tonhöhenvorlaufauswahleinheit **15** für jedes digitale Analysesignal wird verwendet für eine Vorlaufauswahl eines adaptiven Codebuchs und wird verwendet für ein Wichtungsfiler für akustisches Gefühl. Die Autokorrelationsanalysiereinheit **13**, die LPC-Analysiereinheit **14** und die Tonhöhenvorlaufauswahleinheit **15** bauen eine Sprachanalysiereinheit **22** auf, in der Eigenschaftsparameter für jedes der digitalen Analysesignale gewonnen wird durch Analysieren von Spracheigenschaften der digitalen Analysesignale.

[0058] Als nächstes anhand der **Fig. 6** bis **8** beschrieben ist ein Aufbereitungsverfahren der Codiermodulauswahlregel, die im voraus in der Auswahlregelspeichereinheit **16** gespeichert ist.

[0059] **Fig. 6** ist ein Blockdiagramm einer Vorrichtung zur Aufbereitung einer Auswahlregel, um die Codiermodulauswahlregel aufzubereiten, die in der Vorrichtung **11** zur Sprachcodierung verwendet wird.

[0060] Wie in **Fig. 6** gezeigt, ist die Vorrichtung **31** zur Aufbereitung der Auswahlregel ausgestattet mit: einer Sprachaufzeichnungseinheit **32** zum Aufzeichnen einer Vielzahl von Abtastsprachen, die vokalisiert sind von verschiedenen Personen, und Umsetzen der Sprachabtastwerte in digitale Sprachabtastsignale; einer Sprachsignalspeichereinheit **33** zum Speichern der digitalen Sprachabtastsignale; einer Codieroperationsbefehlseinheit **34** zum Betreiben der in **Fig. 5** gezeigten Sprachanalysiereinheit **22**, um eine Vielzahl von Eigenschaftsparametern eines jeden der digitalen Abtastanalysiersignale zu errechnen, die gewonnen wurden aus jedem der digitalen Sprachabtastsignale, die in der Sprachsignalspeichereinheit **33** gespeichert sind, und Errechnen der Codierverzerrung (oder eines Signal/Geräuschverhältnisses) für jeden der Eigenschaftsparameter gemäß einem jeden der in der Codiermodulspeichereinheit **17** gespeicherten Codiermodule; und mit einer Codierverzerrungsspeichereinheit **35** zum Speichern der Codierverzerrungen, die jeweils von einer Analysierperiode spezifiziert werden, einer Gruppe von Eigenschaftsparametern und ein Codiermodul.

[0061] Als Beispiel werden die Codierverzerrungen für alle Codiermodule in den Eigenschaftsparametern gespeichert, und zwar in der Codierverzerrungsspeichereinheit **35**, für jede Analysierperiode, wie in **Fig. 7** gezeigt. Die Codierverzerrung und die Eigenschaftsparameter werden bestimmt auf dem Umstand, daß eine Sprache in "Ba", vokalisiert von einem Mann, zu Analysierperioden von 21 Millisekunden analysiert wird.

[0062] Wie in **Fig. 7** gezeigt, werden für jede Analysierperiode gespeichert: eine Sprachleistung, ein primärer Autokorrelationskoeffizient, ein sekundärer Autokorrelationskoeffizient, ein Sprachleistungsverhältnis einer Sprachleistung zu einem Kopfabschnitt einer Analysierperiode und demjenigen eines hinteren Abschnitts der Analysierperiode (das Verhältnis wird gewonnen durch Teilen einer kleineren Sprachleistung durch eine große Sprachleistung), ein langfristiger Prädiktionskoeffizient für einen Kandidaten einer Fundamentalfrequenz einer

langfristigen Primärkorrelation mit dem höchsten Wert in einer linearen Prädiktionsrestwellenform, einem Impulswert, der die Impulseigenschaft einer langfristigen Restwellenform der linearen Prädiktionsrestwellenform aufzeigt, eine normierte Leistung der langfristigen Restwellenform der linearen Prädiktionsrestwellenform und der Kandidat für die fundamentale Frequenz (oder Tonhöhe) der langfristigen Primärkorrelation mit dem höchsten Wert in der linearen Prädiktionsrestwellenform werden gespeichert als Eigenschaftsparameter für jede analysierte Periode. Auch werden für jede Analysierperiode gespeichert: eine Codierverzerrung, die angezeigt wird durch ein S/N-Verhältnis für jeden der fünf Codiermodule und die Anzahl passender Codiermodule, für die die Codierverzerrung am geringsten ist (oder der Störabstand am höchsten ist).

[0063] Auch die Vorrichtung **31** zur Aufbereitung der Auswahlregel enthält des weiteren eine Datenverbesserungseinheit **36** zum Verbessern der Eigenschaftsparameter und der Codierverzerrungen, die im Codierverzerrungsspeicherabschnitt **35** gespeichert sind, eine statistische Verarbeitungseinheit **37** zur statistischen Verarbeitung der Eigenschaftsparameter und der Codierverzerrungen, die in der Codierverzerrungsspeichereinheit **35** gespeichert sind, und Aufbereiten der Codiermodulauswahlregel, und eine Auswahlregelspeichereinheit **38** zum zeitweiligen Speichern der Codiermodulauswahlregel.

[0064] In der Datenverbesserungseinheit **36** werden die Eigenschaftsparameter und die Codierverzerrungen, die in der Codierverzerrungsspeichereinheit **35** gespeichert sind, korrigiert von einer Bedienperson, falls dies erforderlich ist. Darüber hinaus kann der Wille einer Bedienperson beeinflusst werden bezüglich der Eigenschaftsparameter und der Codierverzerrungen. Beispielsweise in Fällen, bei denen die Anzahl von Informationsbits in einem Codiermodul A größer ist als in einem Codiermodul B und ein Störabstand, der eine Codierverzerrung für den Codiermodul A aufzeigt, geringfügig höher als derjenige für das Codiermodul B ist, das Codiermodul A als geeignetes Codiermodul ausgewählt, weil eine Tonqualität, wiedergegeben gemäß dem Codiermodul A, geringfügig besser ist als diejenige, die gemäß dem Modul B wiedergegeben wird. Wenn jedoch eine Bedienperson beurteilt, daß der Abfall an Informationsbits wichtiger ist als eine geringfügig verbesserte Tonqualität, wird das geeignete Codiermodul neu geschrieben in das Codiermodul B von der Bedienperson, weil die Anzahl von Informationsbits im Codiermodul B gering ist. Wenn folglich das passende Codiermodul neu geschrieben ist in das Codiermodul B bei jeder der großen Anzahl von Analysierperioden, kann ein Auswahlverhältnis vom Codiermodul B menschlich erhöht werden, und eine Vielzahl von Synthesesprachen können vorzugsweise mit einer geringeren durchschnittlichen Bitrate gewonnen werden.

[0065] In der statistischen Verarbeitungseinheit **37** wird die Codiermodulauswahlregel statistisch aufbereitet unter Verwendung eines neuronalen Netzes oder gemäß einer Vektorquantisierung, einer Lernvektorquantisierung oder einer Bayesian-Beurteilung. In diesem Ausführungsbeispiel wird die Aufbereitung der Codiermodulauswahlregel ausgeführt unter Verwendung des einfachsten neuronalen Netzes, in dem Perceptronen in drei Schichten angeordnet sind.

[0066] Wenn eine Vielzahl von eingegebenen Datenstücken I_i zur Eingabe von Neuronen eingerichtet sind in einer Eingabeschicht, wie in **Fig. 8** gezeigt, werden die eingegebenen Daten I_i im allgemeinen mit einer Vielzahl von Wichtungsfaktoren W_{ji} gewichtet und gemäß einer Sigmoid-Funktion im neuronalen Netz verarbeitet, und eine Vielzahl von Ausgabedatenstücken O_j , die die erforderliche Information zur Beurteilung aufzeigen, werden abgegeben von Ausgabeneuronen einer Ausgabeschicht. Die Verarbeitung im neuronalen Netz wird mit folgender Formel angegeben.

$$O_j = f\left(\sum_i W_{ji} * I_i\right)$$

$F(s) = 1/(1 + \exp(-x))$: Sigmoid-Funktion

[0067] Der Bereich von i reicht hier von 1 bis i_{\max} (i_{\max} ist die Anzahl von Eingangsneuronen), und der Bereich von j reicht von 1 bis j_{\max} (j_{\max} ist die Anzahl von Ausgangsneuronen).

[0068] Ein neuronales Netz, in dem eine Vielzahl von Ausgangsdatenstücken O_j Informationen aufzeigen, die in diesem Ausführungsbeispiel erforderlich sind zum Bestimmen des passenden Codiermoduls, werden abgegeben von einer Vielzahl von Ausgangsneuronen einer Ausgabeschicht, wenn eine Vielzahl von Eigenschaftsparametern für eine Analysierperiode als Eingangsdaten I_i eingegeben wird und aufgebaut ist in der statistischen Verarbeitungseinheit **37**, und das neuronale Netz aufgebaut ist in der Modulauswahleinheit **18**.

[0069] In einer statistischen Verarbeitung, in der das neuronale Netz verwendet wird, werden die Wichtungsfaktoren W_{ji} bestimmt durch Eingabe einer großen Anzahl von Eigenschaftsparametersätzen (sogenanntes Lernen). In diesem Ausführungsbeispiel wird ein Lernen gemäß einem Fehler invers ausbreitendem Verfahren

ausgeführt. Im einzelnen werden die Wichtungsfaktoren W_{ji} , die Anfangswerte haben, die bestimmt sind von einem Zufallszahlengenerator, in das neuronale Netz gegeben. Eine Vielzahl von Eigenschaftsparametern für eine Analysierperiode, gespeichert in der Codierverzerrungsspeichereinheit **35**, werden danach in die Eingabeneuronen der Eingabeschicht gegeben, und eine Vielzahl von Lernsignalen werden an die Vielzahl der Ausgabeneuronen der Ausgabeschicht gegeben, um ein Ausgabeneuron zu stimulieren, gemäß einem geeigneten Codiermodul für die Analysierperiode. Wenn beispielsweise fünf Codiermodule vorhanden sind, dann werden fünf Ausgabeneuronen in der Abgabeschicht aufbereitet ($j = 1$ bis 5), ein Lernsignal mit einem hohen Wert von "1" wird an ein Ausgabeneuron entsprechend dem geeigneten Codiermodul gegeben, und vier Lernsignale jeweils mit einem niedrigen Wert von "0" werden an die anderen vier Ausgabeneuronen gegeben. Folglich werden die Wichtungsfaktoren W_{ji} geändert in neue Wichtungsfaktoren

$$W'_{ji} = W_{ji} + \Delta W_{ji}$$

$$\Delta W_{ji} = \alpha * W_{ji} - \frac{\mu * \partial \left(\sum_k (T_k - O_k)^2 \right)}{\partial W_{ji}}$$

[0070] Hier bedeutet α ein Glättungsfaktor (Konstante), μ bedeutet einen Beschleunigungsfaktor (Konstante), ∂ bedeutet Partialdifferential, T_k bedeutet einen Wert eines Lernsignals, das an ein k -tes Abgabeneuron gegeben wurde, und O_k bedeutet einen Ausgangswert aus dem k -ten Ausgabeneuron.

[0071] Danach werden die Wichtungsfaktoren W'_{ji} als Wichtungsfaktoren W_{ji} behandelt, und die Änderung der Wichtungsfaktoren W_{ji} wird wiederholt durch Eingeben eines anderen Satzes von Eigenschaftsparametern für eine Vielzahl von Analysierperioden. Werte der Wichtungsfaktoren W_{ji} werden folglich konvergiert, und eine Codiermodulauswahlregel, die sich zusammensetzt aus den Wichtungsfaktoren W_{ji} , die konvergiert sind, können heuristisch aufbereitet werden. Die Codiermodulauswahlregel wird zeitweilig gespeichert in der Auswahlregelspeichereinheit **38** und wird gesendet an die Auswahlregelspeichereinheit **16**.

[0072] Im obigen Aufbau der Vorrichtung **11** zur Sprachcodierung ist eine Operation der Vorrichtung **11** beschrieben.

[0073] Wird eine Stimme oder Sprache empfangen in der Eingabesprachempfangseinheit **12**, dann wird die Stimme oder die Sprache umgesetzt in eine Vielzahl digitaler Sprachsignale, und jedes der digitalen Sprachsignale wird verarbeitet in der Sprachanalyseinheit **22**, um eine Vielzahl von Eigenschaftsparametern zu errechnen, die die Spracheigenschaften für jedes der digitalen Analysiersignale aufzeigen. Danach werden die Eigenschaftsparameter für jedes der digitalen Analysiersignale errechnet und an die Eingabeneuronen der Eingabeschicht des neuronalen Netzes gegeben, welches aufgebaut ist in der statistischen Verarbeitungseinheit **37**, und in der Modulauswahleinheit **18** eingerichtet ist. Auch eine Codiermodulauswahlregel, zusammengesetzt aus den Wichtungsfaktoren W_{ji} , wird aus der Auswahlregelspeichereinheit **16** in das neuronale Netz der Modulauswahleinheit **18** übertragen. Danach wird das neuronale Netz mit den Eigenschaftsparametern und den Wichtungsfaktoren W_{ji} betrieben, die die Vorrichtung **31** gewonnen hat, nur eines der Ausgabeneuronen in der Abgabeschicht des neuronalen Netzes wird stimuliert. Da die Anzahl der stimulierten Ausgabeneuronen der Anzahl der passenden Codiermodule entspricht, wird das geeignete Codiermodul ausgewählt aus den Codiermodulen, die in der Codiermodulspeichereinheit **17** gespeichert sind. Mit anderen Worten, weil ein Satz von Eigenschaftsparametern für das digitale Analysiersignal aufzeigt, daß das digitale Analysiersignal einem speziellen Sprachabschnitt entspricht, wie einem Ruheabschnitt der Stimme, wird ein Verschlusslaut eines stimmlosen Konsonanten der Sprache, ein Reibelaut des stimmlosen Konsonanten der Sprache, ein stimmhafter Konsonant der Stimme, ein stationärer Vokalabschnitt eines Vokals der Stimme oder ein Vokalübergangsabschnitt des Vokals der Stimme, ein Codiermodul gemäß dem speziellen Sprachabschnitt als passendes Codiermodul ausgewählt.

[0074] Danach wird jedes der digitalen Analysiersignale im passenden Codiermodul codiert, um ein codiertes Signal zu gewinnen. In diesem Falle enthält beispielsweise jedes der Codiermodule das adaptive Codierbuch **104**, das probabilistische Codebuch **105**, den Addierer **106**, die LPC-Synthetisierereinheit **107**, die Vergleichseinheit **108** und die Parametercodiereinheit **109**. Danach wird eine Serie von codierten Signalen für jedes der digitalen Sprachsignale in der Signalvereinheitlichungseinheit **20** erzeugt und abgegeben an die Sendeleitung **21**.

[0075] Als nächstes anhand der Fig. 9 und 10 beschrieben ist ein Schätzergebnis für die Auswahl des geeigneten Codiermoduls auf der Grundlage des neuronalen Netzes.

[0076] Bei dieser Schätzung, wie sie in **Fig. 9** gezeigt ist, sind fünf Codiermodule, betrieben gemäß dem CELP-Codierverfahren, in der Codiereinheit **19** vorgesehen. Eine Abtastrate für die eingegebene Sprache beträgt 8 kHz, und eine Analysierperiode beträgt 21 Millisekunden. Das Codiermodul M1 ist geeignet für einen Vokalübergangsabschnitt und einen stimmhaften Konsonanten, und hat eine Bitrate von 8 kbps, das Codiermodul M2 ist geeignet für einen Verschlusslaut eines stimmlosen Konsonanten, einen Kopfabschnitt eines Wortes und einen stimmhaften Klang bei einem rückwärtigen Abschnitt eines Wortes und hat eine Bitrate von 8 kbps, das Codiermodul M3 ist geeignet für einen stationären Vokalabschnitt und hat eine Bitrate von 4 kbps, das Codiermodul M4 ist geeignet für einen Reibelaut eines stimmlosen Konsonanten und hat eine Bitrate von 4 kbps, und das Codiermodul M5 ist geeignet für einen Ruheabschnitt und hat eine Bitrate von 2 kbps.

[0077] Zwanzig Arten von Eigenschaftsparametern P1 bis P21 werden auch in der Sprachanalyseinheit **22** für jedes digitale Analysiersignal errechnet und gesendet an die Codiermodulwahlseinheit **18**. Die Eigenschaftsparameter P1 bis P21 sind die folgenden:

- P1: Stimmleistung eines aktuellen digitalen Analysiersignals bei einer aktuellen Analysierperiode
- P2: Verhältnis einer Stimmleistung des aktuellen digitalen Analysiersignals an eine Stimmleistung eines vorherigen digitalen Analysiersignals bei einer vorherigen Analysierperiode, die vor der aktuellen Analysierperiode liegt
- P3: Verhältnis einer Stimmleistung des aktuellen digitalen Analysiersignals beim Kopfabschnitt der aktuellen Analysierperiode zu derjenigen beim rückwärtigen Abschnitt der aktuellen Analysierperiode
- P4: Primärautokorrelationskoeffizient für das aktuelle digitale Analysiersignal
- P5: Sekundärautokorrelationskoeffizient für das aktuelle digitale Analysiersignal
- P6: Maximalwert eines langfristigen Prädiktionskoeffizienten für das aktuelle digitale Analysiersignal
- P7: Verhältnis einer langfristigen Prädiktionsverschiebung für das aktuelle digitale Analysesignal zu demjenigen des vorherigen digitalen Analysesignals.
- P8: eine Stimmleistung des vorherigen digitalen Analysesignals
- P9: ein Verhältnis der Stimmleistung des vorherigen digitalen Analysesignals am Kopfabschnitt der vorherigen Analyseperiode zu derjenigen eines hinteren Abschnitts der vorherigen Analyseperiode
- P10: ein primärer Autokorrelationskoeffizient für das vorherige digitale Analysesignal
- P11: ein sekundärer Autokorrelationskoeffizient für das vorherige digitale Analysesignal
- P12: ein Maximalwert des langfristigen Prädiktionskoeffizienten für das vorherige digitale Analysesignal
- P13: Verhältnis einer langfristigen Vorhersageverzögerung für das vorherige digitale Analysesignal zu demjenigen eines zweiten vorherigen digitalen Analysesignals vor dem vorigen digitalen Analysesignal
- P14: Störabstand für das vorherige digitale Analysesignal
- P15: Stimmleistung des zweiten vorherigen digitalen Analysesignals
- P16: Verhältnis einer Stimmleistung des zweiten vorherigen digitalen Analysesignals am Kopfabschnitt einer zweiten vorherigen Analyseperiode zu derjenigen eines hinteren Abschnitts von der zweiten vorherigen Analyseperiode
- P17: primärer Autokorrelationskoeffizient für das zweite vorherige digitale Analysesignal
- P18: sekundärer Autokorrelationskoeffizient für das zweite vorherige digitale Analysesignal
- P19: Maximalwert eines langfristigen Prädiktionskoeffizienten für das zweite vorherige digitale Analysesignal
- P20: Störabstand für das zweite vorherige digitale Analysesignal

[0078] Ein neuronales Netz, das in der Codiermodulwahlseinheit **18** vorgesehen ist, ist aufgebaut aus zwanzig Eingabeneuronen in einer Eingabeschicht, fünfzehn Zwischenneuronen in einer Zwischenschicht und fünf Ausgabeneuronen in einer Ausgabeschicht. In diesem Falle werden zum Auswählen eines geeigneten Codiermoduls die Multiplikation und Addition 395 Mal erforderlich, das Errechnen der Sigmoid-Funktion wird 15 Mal erforderlich und das Teilen in eine Vielzahl von Fällen ist 4 Mal erforderlich. Die Größe einer Nurlesespeicherzone, die erforderlich ist zum Speichern der Wichtungsfaktoren W_{ji} beträgt 395 Wörter.

[0079] Um eine Codiermodulwahlregel in der Vorrichtung **341** aufzubereiten, wird auch eine Vielzahl von Abtastkurzsätzen durch **24** Leute an Stücken von Lerndaten gelesen zur Eingabe einer Vielzahl digitaler Abtastsprachsignale in die Sprachaufzeichnungseinheit **32**. In diesem Falle wird der Beschleunigungsfaktor geändert in einen Bereich von 0,01 auf 0,002, und der Glättungsfaktor wird geändert in einem Bereich von 0,5 auf 0,2, so daß die Wichtungsfaktoren W_{ji} schnell konvergieren. In Fällen, bei denen eine Erkennungssprache von der Vorrichtung **11** zur Sprachcodierung zu erkennen ist und diese in die Vorrichtung **11** als Stücke der Erkennungsdaten eingegeben wird, die den Lerndaten äquivalent sind, ist ein Erkennungsergebnis der Sprache in **Fig. 10** gezeigt. In diesem Falle wird ein Erkennungskoeffizient, der bedeutet, daß ein Sprachabschnitt korrekt erkannt ist, aufgezeigt durch eine Wahrscheinlichkeit (%), daß das geeignete Codiermodul korrekt in der Codiermodulwahlseinheit **18** ausgewählt ist. Beispielsweise beträgt in Fällen, bei denen das geeignete Codier-

modul das Codiermodus M1 ist, der Erkennungskoeffizient 88,6 %.

[0080] Auch wird eine Vielzahl kurzer Sätze, die sich von Abtastkurzsätzen unterscheiden, von 8 Leuten gelesen, um einen Originalsprachklang für jeden der Kurzsätze zu erzeugen. Jeder Originalsprachklang wird in der Vorrichtung **11** zur Sprachcodierung codiert, um eine Serie codierter Signale für jeden der kurzen Sätze zu erzeugen. In diesem Falle werden die Wichtungsfaktoren W_{ji} , gewonnen aus den Abtastkurzsätzen, in der Speichereinheit **18** für Codiermodulauswahlregeln gespeichert und in das neuronale Netz der Codiermodulauswahleinheit **18** gegeben. Danach werden die codierten Signale in einer Codiervorrichtung (nicht dargestellt) decodiert, um einen synthetisierten Klang für jeden der kurzen Sätze wiederzugeben. Jeder der synthetisierten Klänge wird subjektiv von 19 Leuten (Männer und Frauen) geschätzt. Das heißt, jeder Mann oder jede Frau gibt einen von fünf Schätzgraden für jeden der synthetisierten Klänge an, und ein mittlerer Meinungswert wird gewonnen für die synthetisierten Klänge. In der selben Weise werden ein mittlerer Meinungswert für die Originalsprachklänge und ein mittlerer Meinungswert für die herkömmlich synthetisierten Klänge, gewonnen in einer herkömmlichen Sprachcodiervorrichtung und betrieben gemäß dem QCELP-Codierverfahren, gewonnen. Obwohl der mittlere Meinungswert für die Originalsprachklänge 3,61 beträgt, wird als Schätzergebnis der mittlere Meinungswert für die herkömmlich synthetisierten Klänge weitestgehend abgesenkt auf 3,08. Im Gegensatz dazu wird der mittlere Meinungswert für die synthetisierten Klänge leicht abgesenkt auf 3,38. Folglich können in der Vorrichtung **11** zur Sprachcodierung synthetisierte Klänge mit hervorragender Tonqualität gewonnen werden.

[0081] Da eine Abtastsprache in der Sprachanalyseinheit **22** analysiert wird, um eine Vielzahl von Eigenschaftsparametern zu gewinnen, die die Spracheigenschaften der Abtastsprache für jede Analysierperiode aufzeigen, und weil eine Codierverzerrung (oder ein Störabstand) gewonnen wird aus den Eigenschaftsparametern in jedem der Codiermodule, kann folglich ein geeignetes Codiermodul, bei dem die Codierverzerrung minimiert ist oder der Störabstand maximiert ist, leicht für jede Analysierperiode spezifiziert werden.

[0082] Weil die Eigenschaftsparameter und die Codierverzerrung statistisch in der Statistikverarbeitungseinheit **37** verarbeitet werden, kann auch eine Codiermodulauswahlregel leicht erzielt werden. Weil die Eigenschaftsparameter für eine Vielzahl von Eingabeneuronen eines neuronalen Netzes als ein Satz von Eingabedaten vergeben werden und eine Vielzahl von Lernsignalen in eine Vielzahl von Ausgabeneuronen des neuronalen Netzes gegeben werden, um beispielsweise einen der Ausgabeneuronen gemäß dem geeigneten Codiermodul zu stimulieren, kann eine Vielzahl von Wichtungsfaktoren W_{ji} , die eine Funktion des neuronalen Netzes spezifizieren, leicht konvergiert und erzielt werden als eine Codiermodulauswahlregel.

[0083] Da auch eine Unterscheidungssprache in der Sprachanalyseinheit **22** analysiert wird, um eine Vielzahl von Eigenschaftsparametern zu erzielen, die Spracheigenschaften der Unterscheidungssprache für jede Analysierperiode aufzeigen, und weil ein geeignetes Codiermodul aus einer Vielzahl von Codiermodulen in der Codiermodulauswahleinheit **18** ausgewählt wird durch Eingabe der Eigenschaftsparameter in die Eingabeneuronen des neuronalen Netzes, in dem die Wichtungsfaktoren W_{ji} , die die Codiermodulauswahlregel bedeuten, eingesetzt und spezifiziert werden für ein stimuliertes Ausgabeneuron gemäß dem geeigneten Codiermodul, kann die Auswahl des geeigneten Codiermoduls leicht mit geringem Rechenumfang erfolgen.

[0084] Da auch die Eigenschaftsparameter der Codierverzerrung von der Bedienperson in der Datenverbesserungseinheit **36** vor dem Ausführen der statistischen Verarbeitung verbessert werden können in der statistischen Verarbeitungseinheit **37**, kann die der Bedienperson wiedergegebene Codiermodulauswahlregel gewonnen werden. Eine Sprache kann in effizienter Weise in der Vorrichtung **11** zur Sprachcodierung codiert werden, wenn die Bedienperson dies wünscht.

[0085] Als nächstes beschrieben ist ein zweites Ausführungsbeispiel, das nicht in den Schutzbereich der vorliegenden Erfindung fällt.

[0086] Fig. 11 ist ein Blockdiagramm einer Vorrichtung zur Sprachcodierung gemäß einem zweiten Ausführungsbeispiel, das nicht in den Umfang der vorliegenden Erfindung fällt.

[0087] Wie in Fig. 11 gezeigt, ist eine Vorrichtung **41** zur Sprachcodierung ausgestattet mit: der Eingabesprachempfangseinheit **102** zum Aufnehmen einer Eingangssprache oder einer Stimme und zum Erzeugen von Sprachdatenstücken; der LPC-Analyseinheit **103** zum Ausführen einer Autokorrelationsanalyse und einer linearen Prädiktionskoeffizientenanalyse, um einen linearen Prädiktionskoeffizienten für alle Sprachdaten zu gewinnen; einem adaptiven Codebuch **42** zum Speichern einer Vielzahl vergangener Syntheseklangquellen (oder ver-

gangener Synthesecodevektoren);
 einem probabilistischen Codebuch **43** zum Speicher einer Vielzahl fester Klangquellen (oder fester Codevektoren);
 einer Klangquellenaddiereinheit **44** zum Herausnehmen der letzten Syntheseklangquellen aus dem adaptiven Codebuch **42** als Kandidat für eine geeignete Syntheseklangquelle der Vergangenheit, Errechnen einer optimalen Verstärkung für den Kandidaten, Auslesen alle festen Klangquellen in einer Normalrichtung vom probabilistischen Codebuch **43** als Kandidat für eine geeignete feste Klangquelle, Errechnen einer optimalen Verstärkung für den Kandidaten, Auslesen alle festen Klangquellen in umgekehrter Richtung gegenüber der Normalrichtung aus dem probabilistischen Codebuch **43**, um eine umgekehrte feste Klangquelle als anderen Kandidaten für die geeignete feste Klangquelle zu gewinnen, Errechnen der optimalen Verstärkung für den Kandidaten, Hinzufügen aller vergangenen Syntheseklangquellen und aller festen Klangquellen, die mit den optimalen Verstärkungen leistungsjustiert sind, zum Erzeugen einer ersten Syntheseklangquelle, und Hinzufügen einer jeden vergangenen Syntheseklangquelle und jeder festen Umkehrklangquelle, die leistungsjustiert sind mit den optimalen Verstärkungen, zum Erzeugen einer zweiten Syntheseklangquelle;
 einer LPC-Synthetisierereinheit **45** zum Filtern aller ersten Syntheseklangquellen und aller zweiten Syntheseklangquellen, gewonnen in der Klangquellenaddiereinheit **44**, mit dem wiedergegebenen linearen Prädiktionskoeffizienten, gewonnen in der LPC-Analysiereinheit **103**, um eine erste synthetische Sprache und eine zweite synthetische Sprache zu erzeugen;
 einer Synthesep Sprachauswahleinheit **46** zum Errechnen eines Abstands zwischen dem zugehörigen Sprachdatenstück, aufgenommen in der Eingabespracheempfangseinheit **103**, und jeder der ersten und zweiten Synthesep Sprachen für alle Sprachdaten, Auswählen einer speziellen Synthesep Sprache, die dem speziellen Abstand entspricht, der einen Minimalwert unter jenen der Abstände aufweist unter der ersten und zweiten Synthesep Sprache für alle Sprachdaten, Auswählen einer speziellen vergangenen Syntheseklangquelle gemäß der speziellen Synthesep Sprache unter den vergangenen Syntheseklangquellen als geeignete vergangene Syntheseklangquelle und Auswählen einer speziellen festen Klangquelle gemäß der speziellen Synthesep Sprache aus den festen Klangquellen als geeignete feste Klangquelle;
 einer Parametercodiereinheit **47** zum Codieren der geeigneten Verstärkungen der geeigneten vergangenen Syntheseklangquellen und der geeigneten festen Klangquelle zum Erzielen beider Verstärkungscodes, Erzeugen einer speziellen Syntheseklangquelle aus den Verstärkungscodes und den geeigneten vergangenen Synthesen und festen Klangquellen für alle Sprachdaten und Senden der speziellen Syntheseklangquelle an das adaptive Codebuch **42**, um die spezielle vergangene Klangquelle zu ersetzen durch die spezielle Syntheseklangquelle;
 einer Synthesep Sprachvereinheitlichungseinheit **48** zum Vereinheitlichen der speziellen Synthesep Sprachen, ausgewählt in der Synthesep Sprachauswahleinheit **46**, in einer Serie von Synthesep Sprachen gemäß einer codierten Sprache, und Senden der Serie von Synthesep Sprachen auf eine Sendeleitung **49** zur Wiedergabe der Sprache in einer Wiedergabevorrichtung (nicht dargestellt).

[0088] In der LPC-Synthetisierereinheit **45** wird das akustische Gefühl für die Sprachdaten gewichtet mit dem linearen Prädiktionskoeffizienten, einem Frequenzanhebungsfilterkoeffizienten und einem langfristigen Prädiktionskoeffizienten, gewonnen durch Ausführen einer langfristigen Prädiktionsanalyse für die Sprachdaten. Die vergangene und die feste Klangquelle werden herausgenommen aus den Unterrahmen, gewonnen durch Einteilen eines jeden Analyseblocks im adaptiven Codebuch **42** und im probabilistischen Codebuch **43**.

[0089] In der obigen Konfiguration ist der Betrieb der Sprachcodiervorrichtung **41** beschrieben worden.

[0090] Wenn eine Sprache von der Empfangseinheit **102** aufgenommen wird, erfolgt das Erzeugen eines linearen Wiedergabeprediktionskoeffizienten in der LPC-Analysiereinheit **103** und wird gesandt zur LPC-Synthetisierereinheit **45**, und ein Sprachdatenstück wird an die Synthesep Sprachauswahleinheit **46** gesandt. Jede der Syntheseklangquellen wird aus dem adaptiven Codebuch **42** herausgenommen zur Klangquellenaddiereinheit **44** als Kandidat für eine geeignete Vergangenheitssyntheseklangquelle.

[0091] Wie in **Fig. 12** gezeigt, wird jede der festen Klangquellen, gespeichert im probabilistischen Codebuch **43**, in Normalrichtung ausgelesen und zur Einheit **44** als Kandidat für eine geeignete feste Klangquelle gesandt, wobei alle festen Klangquellen in der Einheit **43** gespeichert sind und in umgekehrter Richtung ausgelesen werden, um eine feste Umkehrklangquelle zu bilden, und jede der festen Umkehrklangquellen wird zur Einheit **44** als weiterer Kandidat für die geeignete feste Klangquelle gesandt. Das heißt, wenn ein Abtastzeiger zum Zeigen einer festen Klangquelle in Normalrichtung verschoben wird, erfolgt das Senden der festen Klangquelle zur Einheit **44**. Auch wenn der Abtastzeiger in Umkehrrichtung verschoben wird, erfolgt das Senden der festen Umkehrklangquelle zur Einheit **44**. Die beiden Klangquellen werden erzeugt aus einer festen Klangquelle als zwei Kandidaten für die passende feste Klangquelle.

[0092] Die Syntheseklangquelle wird danach erzeugt aus allen Kandidaten für die geeignete vergangene Klangquelle und aus allen Kandidaten für die geeignete feste Klangquelle in der Einheit **44**, eine Synthesep Sprache wird erzeugt aus jeder der Syntheseklangquellen gemäß dem wiedergegebenen linearen Prädiktionskoeffizienten in der LPC-Synthetisiereinheit **44** für alle Sprachdaten. Danach wird ein Abstand zwischen einem zugehörigen Sprachdatenstück und jeder ersten und zweiten Sprache errechnet und eine spezielle Synthesep Sprache, geeignet für die zugehörigen Sprachdaten, wird aus den Synthesep Sprachen in der Synthesep Sprachauswahleinheit **46** für alle Sprachdaten ausgewählt.

[0093] Danach werden die speziellen Synthesep Sprachen für die Sprachdaten vereinheitlicht zu einer Serie spezieller Synthesep Sprachen, die eine codierte Sprache in der Synthesep Sprachvereinheitlichungseinheit **48** aufzeigen, und die Serie spezieller Synthesep Sprachen wird zur Sendeleitung **49** gegeben, um die Sprache wiederzugeben.

[0094] Auch die geeignete vergangene Syntheseklangquelle und die geeignete feste Klangquelle der speziellen Synthesep Sprache werden ausgewählt aus den vergangenen Syntheseklangquellen und festen Klangquellen in der Einheit **46** für alle Sprachdaten, zwei Verstärkungscodes werden erzeugt aus den geeigneten Verstärkungen für die geeignete vergangene Syntheseklangquelle und die geeignete feste Klangquelle, und eine spezielle Syntheseklangquelle wird erzeugt aus den Verstärkungscodes und den geeigneten vergangenen Synthesen und festen Klangquellen in der Parametercodiereinheit **47** für alle Sprachdaten. Die geeignete vergangene Syntheseklangquelle, die im adaptiven Codierbuch **42** gespeichert ist, wird ersetzt durch die spezielle Syntheseklangquelle.

[0095] Da zwei Klangquellen aus einer festen Klangquelle erzeugt werden, kann die Anzahl fester Klangquellen, die als Kandidaten für die geeignete feste Klangquelle dienen, folglich im wesentlichen auf die doppelte Anzahl von festen Klangquellen erhöht werden, die im probabilistischen Codebuch **43** gespeichert sind. Die Größe des probabilistischen Codebuchs **43** kann folglich auf die Hälfte derjenigen vom probabilistischen Codebuch **105** reduziert werden unter der Bedingung, daß eine wiedergegebene Sprache nicht verschlechtert wird.

[0096] Da zwei Klangquellen aus einer festen Klangquelle erzeugt werden durch Verschieben des Abtastzeigers in zwei Richtungen, die entgegengesetzt sind, kann auch die Größe des probabilistischen Codebuchs **43** auf die Hälfte reduziert werden, ohne den Rechenumfang zu erhöhen.

[0097] Als nächstes beschrieben ist eine Vorrichtung zur Verbesserung der Klangquelle, in der alle festen Klangquellen in eine ideale feste Klangquelle verbessert werden.

[0098] In Fällen, bei denen eine feste Klangquelle symmetrisch ist, verschlechtert sich die Qualität der wiedergegebenen Sprache unter Verwendung einer festen Umkehrklangquelle, die der symmetrischen festen Klangquelle entgegengesetzt ist. Auch im Falle, bei dem die umgekehrte feste Klangquelle gemäß einer anderen festen Klangquelle gleicht, verschlechtert sich die Qualität der wiedergegebenen Sprache unter Verwendung der festen Umkehrklangquelle. In Fällen, bei denen jede feste Klangquelle nicht symmetrisch ist und jede feste Umkehrklangquelle nicht der festen Klangquelle gleicht, kann folglich die Qualität der wiedergegebenen Sprache verbessert werden.

[0099] **Fig. 13** ist ein Blockdiagramm einer Klangquellenverbesserungsvorrichtung nach dem zweiten Ausführungsbeispiel.

[0100] Wie in **Fig. 13** gezeigt, enthält eine Klangquellenverbesserungsvorrichtung **51** eine Leistungsrecheneinheit **52** zum Errechnen einer Vorderstimmeleistung in einem Vorderabschnitt einer jeden festen Klangquelle und zum Errechnen einer hinteren Stimmeleistung an einem hinteren Abschnitt einer jeden festen Klangquelle, eine Klangquellenumkehrereinheit **53** zum Neueinrichten jeder festen Klangquelle, wobei die hintere Stimmeleistung höher als die vordere Stimmeleistung ist, in einem Unterrahmen in Umkehr zum Erzeugen einer idealen Klangquelle, in der die vordere Stimmeleistung gleich oder höher ist als hintere Stimmeleistung, und eine Klangquellenspeichereinheit **54** zum Speichern der idealen Klangquellen, erzeugt in der Klangquellenumkehrereinheit **53**, und Speichern einer Vielzahl fester Klangquellen, die von der Einheit **53** nicht neu umgekehrt eingerichtet sind, als ideale Klangquellen.

[0101] In der obigen Konfiguration arbeiten die festen Klangquellen als Codevektoren. Jede feste Klangquelle, in der die hintere Stimmeleistung höher ist als die vordere Stimmeleistung, wird rückwärts angeordnet, und jede andere feste Klangquelle, in der die vordere Stimmeleistung gleich oder höher als hintere Stimmeleistung

ist, wird nicht neu geordnet. Die vordere Stimmleistung wird folglich gleich oder höher als die hintere Stimmleistung in jeder festen Klangquelle. Das heißt, es wird eine Schwerpunktgruppe fester Klangquellen gebildet. Die festen Klangquellen, die im probabilistischen Codebuch **105** gespeichert sind, werden ersetzt durch die idealen Klangquellen, die in der Klangquellenspeichereinheit **54** gespeichert sind.

[0102] Eine Neuordnungstechnik für eine Population von Vektoren gemäß einem LBG-Algorithmus ist in der Literatur offenbart in "IEEE Transactions on Communications", Band COM-28, Nr. 1, Januar 1980.

[0103] Da die vordere Stimmleistung gleich oder höher als die hintere Stimmleistung in jeder festen Klangquelle wird, sind alle festen Klangquellen nicht symmetrisch, und alle festen Umkehrklangquellen gleichen nicht irgendeiner festen Tonquelle.

[0104] Die Qualität wiedergegebener Sprache kann folglich in zuverlässiger Weise verbessert werden.

[0105] Eine andere Klangquellenverbesserungsvorrichtung, in der alle festen Klangquellen verbessert sind zu idealen festen Klangquellen ist nachstehend gemäß einem Lernverfahren beschrieben.

[0106] Fig. 14 ist ein Blockdiagramm einer anderen Klangquellenverbesserungsvorrichtung nach dem zweiten Ausführungsbeispiel.

[0107] Wie in Fig. 14 gezeigt, enthält eine Klangquellenverbesserungsvorrichtung **55** eine geeignete Klangquellenerzeugungseinheit **56** zum Anweisen der Klangquellenaddiereinheit **44**, der LPC-Synthetisiereinheit **45** und der Synthesep Sprachauswahleinheit **46**, die geeignete vergangene Klangquelle auszusuchen, die aus dem adaptiven Codebuch **42** gelesen wird, und eine passende feste Klangquelle, die aus dem probabilistischen Codebuch **43** gelesen wird, für jedes eingegebene Sprachdatenstück, eine ideale Klangquellenerzeugungseinheit **57** zum Erzeugen einer idealen Klangquelle aus allen eingegebenen Sprachdaten durch Inversfilterung aller eingegebenen Sprachdaten mit einem umgekehrten linearen Prädiktionskoeffizienten, der dem linearen Prädiktionskoeffizienten umgekehrt ist, erzeugt in der LPC-Analysiereinheit **103**, einen Subtrahierer **58**, der alle geeigneten vergangenen Klangquellen des adaptiven Codebuchs **42** von allen idealen Klangquellen abzieht, um eine ideale probabilistische Codebuchklangquelle für alle eingegebenen Sprachdaten zu erzeugen, und eine Lerneinheit **59** zum Erzeugen einer idealen Klangquelle diese probabilistischen Codebuchs **43** für alle eingegebenen Sprachdaten aus allen idealen probabilistischen Codebuchklangquellen durch Konvergieren der idealen Klangquelle gemäß einem Lernverfahren.

[0108] Das Lernverfahren wird gemäß nachstehender Gleichung ausgeführt:

$$C_{ji}(n) = C_{ji}(n-1) * (1-\mu) + D_i/g_i * \mu$$

[0109] Hier bedeutet $C_{ji}(n)$ einen Vektor der idealen Klangquelle, gewonnen in einer n-ten Umsetzstufe, i bedeutet die Anzahl von Sprachdatenstücken, eingegeben als Abtastdatenstücke ($i = 1$ bis L), L bedeutet die Länge des Vektors der idealen Klangquelle, j bedeutet die Anzahl passender fester Klangquellen, ausgewählt für die Sprachdaten, D_i bedeutet die ideale probabilistischen Codebuchklangquellen, g_i bedeutet die geeignete Verstärkung für die passenden festen Klangquellen und μ bedeutet einen Lernfaktor ($0 < \mu < 1$).

[0110] In der obigen Konfiguration konvergiert jede der idealen Klangquellen des probabilistischen Codebuchs **43** in der Lerneinheit **59** durch allmähliches Absenken des Lernfaktors μ bis auf Null. Danach werden die festen Klangquellen, die im probabilistischen Codebuch **105** gespeichert sind, ersetzt durch ideale Klangquellen.

[0111] Da die idealen Klangquellen erzeugt werden, während die idealen Klangquellen angepaßt werden an die eingegebenen Sprachdaten und die festen Klangquellen, gespeichert im probabilistischen Codebuch **105**, ersetzt werden durch die idealen Klangquellen, kann die Qualität der wiedergegebenen Sprache folglich in zuverlässiger Weise verbessert werden.

[0112] Als nächstes beschrieben ist eine lineare Prädiktionskoeffizienten-Analysiervorrichtung, die verwendet wird, um einen linearen Prädiktionskoeffizienten für alle Sprachsignale in einer Sprachcodiervorrichtung zu gewinnen, gemäß eines dritten Ausführungsbeispiels, das nicht in den Umfang der vorliegenden Erfindung fällt.

[0113] Fig. 15 ist ein Blockdiagramm einer linearen Prädiktionskoeffizienten-Analysiervorrichtung nach einem dritten Ausführungsbeispiel, das nicht in den Umfang der vorliegenden fällt.

[0114] Wie in **Fig. 15** gezeigt, ist eine lineare Prädiktionskoeffizienten-Analysiervorrichtung **61** ausgestattet mit.

einer Eingabesprachempfangseinheit **62** zum Empfangen von Sprache und zum Umsetzen der Sprache in eine Vielzahl von Sprachsignalen, die in Analysierperioden eingeteilt werden;
 einer Partialanalysierperiodenausleseseinheit **63** zum Auslesen einer Vielzahl von Partialanalysierperioden mit verschiedenen Zeitlängen, von denen alle Analysierperioden unter der Bedingung stehen, daß es zulässig ist, daß die Partialanalysierperioden sich untereinander überlappen, einer Vielzahl von Partialsprachsignalen gemäß den Partialanalysierperioden, die von allen Sprachsignalen ausgelesen werden;
 einer Fensterkoeffizientenmultipliziereinheit **64** zum Multiplizieren aller Partialsprachsignale mit einem Fensterkoeffizienten, um ein fensterverarbeitetes Partialsprachsignal für alle Partialanalysierperioden zu erzeugen;
 einer Autokorrelationsanalysiereinheit **65** zum Analysieren einer Autokorrelation für alle fensterverarbeiteten Partialsprachsignale, um eine Autokorrelationsfunktion für alle Partialanalysierperioden zu erzeugen;
 einer Autokorrelationsfunktionssynthetisiereinheit **66** zum Wichten einer jeden Autokorrelationsfunktion mit einem Wichtungsfaktor zum Erzeugen einer gewichteten Autokorrelationsfunktion für alle Partialanalysierperioden und zum Addieren der gewichteten Autokorrelationsfunktionen untereinander zum Erzeugen einer synthetisierten Autokorrelationsfunktion für alle Analysierperioden; und mit
 einer linearen Prädiktionskoeffizienten-Analysiereinheit **67** zum Ausführen einer linearen Prädiktionsanalyse für die synthetisierte Autokorrelationsfunktion gemäß einem Autokorrelationsverfahren für alle Sprachsignale.

[0115] Anwendbar ist, daß sich die Partialanalysierperioden, ausgelesen von der Partialanalysierperiodenausleseseinheit **63**, untereinander überlappen. Ein Koeffizient in einem Hamming-Fenster, einem Hanning-Fenster, einem Blackman-Harris-Fenster oder dergleichen wird verwendet als Fensterkoeffizient. Die lineare Prädiktionsanalyse, die ausgeführt wird von der linearen Prädiktionskoeffizienten-Analysiereinheit **67**, ist offenbart in verschiedenen Dokumenten der Sprachinformationsverarbeitung, wie beispielsweise in "The Autokorrelation Method" von L. R. Labiner und R. W. Schafer, "Digital Processing of Speech Signals", Seiten 401 bis 403.

[0116] Wenn in der obigen Konfiguration eine Sprache in die Empfangseinheit **62** gelangt, wird eine Vielzahl von Sprachsignalen mit Analysierperioden zeitseriell erzeugt. Danach wird eine Vielzahl Partialanalysierperioden mit unterschiedlichen Zeitlängen aus allen Analysierperioden in der Ausleseseinheit **63** ausgelesen, und eine Vielzahl von Partialsprachsignalen $PX_{i,k}$ gemäß den Partialanalysierperioden werden für alle Sprachsignale X_i ausgelesen. In diesem Falle, wie er in **Fig. 16** gezeigt ist, ist es zulässig, daß sich die Partialanalysierperioden untereinander überlappen. Danach werden alle Partialsprachsignale $PX_{i,k}$ multipliziert mit einem Fensterkoeffizienten $H_{i,k}$, um ein fensterverarbeitetes Partialsprachsignal $PY_{i,k}$ für alle Partialanalysierperioden in der Multipliziereinheit **64** zu erzeugen:

$$PY_{i,k} = H_{i,k} * PX_{i,k}$$

[0117] Hier bedeutet k die Anzahl von Partialsprachsignalen eines aktuellen Sprachsignals in einer aktuellen Analysierperiode ($k = 1$ bis K_{max}), und i bedeutet die Anzahl aktueller Sprachsignale unter den Sprachsignalen ($i = 1$ bis I_{max}).

[0118] Danach wird eine Autokorrelation aller fensterverarbeiteten Partialsprachsignale $PY_{i,k}$ in der Analysiereinheit **65** analysiert, um eine Autokorrelationsfunktion $PV_{i,j}$ für alle Partialanalysierperioden zu erzeugen:

$$PV_{i,j} = \sum_{k=j+1}^{K_{max}} \{PY_{i,k} * PY_{i,(k-j)}\}$$

[0119] Hier bedeutet j die Anzahl von Autokorrelationsfunktionen in der aktuellen Analysierperiode ($j = 1$ bis J_{max}).

[0120] Danach wird in der Synthetisiereinheit **66** jede der Autokorrelationsfunktionen $PV_{i,j}$ gewichtet mit einem Wichtungsfaktor W_j , um eine gewichtete Autokorrelationsfunktion für alle Partialanalysierperioden zu erzeugen, und um die gewichteten Autokorrelationsfunktionen untereinander zu addieren zum Erzeugen einer synthetisierten Autokorrelationsfunktion U_i für alle Analysierperioden:

$$U_i = \sum_j^{J_{max}} (W_j * PV_{i,j})$$

[0121] Danach wird eine lineare Prädiktionsanalyse für die synthetisierte Autokorrelationsfunktion U_i ausge-

führt gemäß einem Autokorrelationsverfahren für alle Sprachsignale in der Analysiereinheit **67**.

[0122] Da irgendein nachfolgendes Sprachsignal folglich nicht mit einem Fensterkoeffizienten für das aktuelle Sprachsignal multipliziert ist, tritt keinerlei Codierprozeßverzögerung bei allen Codes der Sprachsignale auf.

[0123] Da alle Sprachsignale mit einer Vielzahl von Fensterkoeffizienten multipliziert sind, kann auch ein Informationsstück über das gesamte Sprachsignal in zuverlässiger Weise auf einem Informationsstück für ein zugehöriges fensterverarbeitetes Sprachsignal ohne irgendeine Codierverarbeitungsverzögerung wiedergegeben werden.

[0124] Als nächstes wird ein Bestimmungsverfahren von Werten der Wichtungsfaktoren beschrieben.

[0125] Werte der Wichtungsfaktoren für die Autokorrelationsfunktionen in einer Analysierperiode hängen ab von den Eigenschaften eines Sprachsignals in der Analysierperiode. In Fällen, bei denen es beispielsweise erforderlich ist, den hinteren Abschnitt des Sprachsignals in einer Analysierperiode anzuheben, wird eine Partialanalysierperiode, die am hinteren Abschnitt der Analysierperiode plaziert ist, in der Ausleseeinheit **63** ausgelesen, um eine kurze Zeitlänge aufzuweisen, eine Autokorrelationsfunktion für die Partialanalysierperiode wird erzeugt in der Analysiereinheit **65**, und die Autokorrelation wird gewichtet mit einem Wichtungsfaktor, der einen hohen Wert hat, in der Synthesiereinheit **66**.

[0126] Da ein Sprachsignal folglich örtlich angehoben werden kann in jeder Analysierperiode unter Verwendung eines Wichtungsfaktors mit einem hohen Wert, wird ein linearer Prädiktionskoeffizient gewonnen, der ein Informationsstück wiedergibt, in dem das Sprachsignal örtlich angehoben ist.

[0127] In Fällen, bei denen ein linearer Prädiktionskoeffizient, der ein gesamtes Sprachsignal repräsentiert, für alle Analysierperioden erforderlich ist, werden Werte einer Vielzahl von Wichtungsfaktoren für eine Vielzahl von Partialsprachsignalen, die alle die Analysierperiode haben, eingesetzt in die Synthesiereinheit **66**, um einheitlich das Sprachsignal bei jeder Analysierperiode zu wichten. Genauer gesagt, eine Summe von Fensterkoeffizienten für die Partialsprachsignale eines jeden Sprachsignals werden auf einen Wert gebracht, der nahe an Eins liegt. Das heißt, die Werte von Wichtungsfaktoren W_j werden bestimmt zum Minimieren eines Fehlers E , und zwar folgendermaßen:

$$E = \sum_i^{\text{imax}} \left\{ 1 - \sum_j^{\text{jmax}} (W_j * H_{i,j}) \right\}^2$$

[0128] Um die Werte der Wichtungsfaktoren nach der obigen Gleichung zu bestimmen, wird eine Vielzahl von Simultangleichungen, gewonnen durch Partialdifferenzieren des Fehlers E in Hinsicht auf jeden der Wichtungsfaktoren, gelöst. Die Simultangleichungen, die einem der k -Werte entsprechen, werden folgendermaßen ausgedrückt:

$$\sum_i^{\text{imax}} H_{i,k} = \sum_j^{\text{jmax}} W_j \sum_i^{\text{imax}} (H_{i,j} * H_{i,k})$$

[0129] Hier bedeutet k die Anzahl von Partialsprachsignalen eines aktuellen Sprachsignals in einer aktuellen Analysierperiode ($k = 1$ bis K_{max}).

[0130] Als nächstes beschrieben ist ein Schätzergebnis von Codier- und Decodieroperationen, die für Sprachsignale ausgeführt werden, die die Empfangseinheit **62** erhält.

[0131] Schätzbedingungen sind folgende. Eine Abtastrate für die Sprache in der Empfangseinheit **62** beträgt 8kHz. Das heißt, die Sprachsignale werden mit der Abtastrate erzeugt. Alle Analysierperioden werden eingeteilt in 256 Abtastperioden. Ein erstes Hamming-Fenster mit einer ersten Zeitlänge, die den 32 Abtastperioden äquivalent ist, wird für ein Partialsprachsignal mit einer Partialanalysierperiode eingesetzt, die sich im Bereich von einer ersten Abtastperiode bis zu einer 32-ten Abtastperiode befindet, und mit einer anderen Partialsprachsignale mit einer Partialanalysierperiode im Bereich von einer 225-ten Abtastperiode bis zu einer 256-ten Abtastperiode. Ein zweites Hamming-Fenster mit einer zweiten Zeitlänge, die 64 Abtastperioden äquivalent ist, wird eingesetzt für ein Sprachsignal im Bereich einer ersten Abtastperiode bis zu einer 64-ten Abtastperiode, und ein anderes Partialsprachsignal im Bereich von einer 193-ten Abtastperiode bis zu einer 256-ten Abtastperiode. Ein drittes Hamming-Fenster mit einer dritten Zeitlänge, die 128 Abtastperioden äquivalent ist, wird eingesetzt für ein Partialsprachsignal im Bereich von einer ersten Abtastperiode bis zu einer 128-ten Abtast-

periode, und ein anderes Partialsprachsignal im Bereich von einer 129-ten Abtastperiode bis zu einer 256-ten Abtastperiode. Ein viertes Hamming-Fenster mit einer vierten Zeitlänge, die 256 Abtastperioden äquivalent ist, wird eingesetzt für ein Partialsprachsignal im Bereich einer ersten Abtastperiode bis zu einer 256-ten Abtastperiode. Eine Summe von Fensterkoeffizienten für die Abtastperioden jeden Sprachsignals wird auf einen Wert nahe Eins gebracht.

[0132] Ein linearer Prädiktionskoeffizient wird gewonnen für jedes der Sprachsignale, das die Empfangseinheit **92** erzeugt, gemäß den obigen Schätzbedingungen in der Vorrichtung **61**, codierte Sprachsignale werden gewonnen unter Verwendung der linearen Prädiktionskoeffizienten, die gemäß dem dritten Ausführungsbeispiel gewonnen werden, und die codierten Sprachsignale werden in einer Decodiervorrichtung decodiert, um Wiedergabesprache zu gewinnen.

[0133] Obwohl lokal in wiedergegebener Sprache Allophonie auftritt in Fällen, bei denen ein linearer Prädiktionskoeffizient für alle Sprachsignale in einer Vorrichtung gewonnen wird, wird die lokale Allophonie weitestgehend in Fällen reduziert, bei denen ein linearer Prädiktionskoeffizient für alle Sprachsignale gemäß den obigen Schätzbedingungen in der Analysiervorrichtung **61** für lineare Prädiktionskoeffizienten gewonnen wird. Das heißt, die Qualität der wiedergegebenen Sprache wird im dritten Ausführungsbeispiel weitestgehend verbessert.

[0134] Als nächstes beschrieben ist eine Vorrichtung zur Geräuschreduzierung, bei der ein Geräusch in einer Analysierperiode für eine Sprache enthalten ist, die in einer Vorrichtung zur Sprachkodierung kodiert wird, oder enthalten ist in einer anderen Analysierperiode für eine Nichtsprache und reduziert wird gemäß einem vierten Ausführungsbeispiel, das nicht in den Umfang der vorliegenden Erfindung fällt.

[0135] **Fig. 17** ist ein Blockdiagramm der Vorrichtung zur Geräuschreduzierung nach einem vierten Ausführungsbeispiel, das nicht in den Umfang der vorliegenden Erfindung fällt.

[0136] Wie in **Fig. 17** gezeigt, ist die Vorrichtung **71** zur Geräuschreduzierung ausgestattet mit:
 einem Analog-Digital-Umsetzer **72** (A/D-Umsetzer) zum Umsetzen einer Vielzahl von Rahmen analoger eingegebener Sprachsignale S_{in} in eine Vielzahl von Rahmen digitaler Eingangssignale, eine Sprache und ein Geräusch ist in diesem Rahmen enthalten;
 einer Eingangssignalwellenformeinheit **73** zum Einstellen von Wellenformen der Rahmen digitaler Eingangssignale, gewonnen im Umsetzer **72** in einem Speicher, einer nach dem anderen;
 einer linearen Prädiktionskoeffizientenanalysiereinheit **74** zum Ausführen einer linearen Prädiktionskoeffizientenanalyse für jede der Wellenformen der Rahmen digitaler eingegebener Signale, eingesetzt in die Eingangssignalwellenformeinheit **73**, um einen linearen Prädiktionskoeffizienten für jeden Rahmen digitaler Eingangssignale zu erhalten;
 einer Fourier-Transformationseinheit **75** zum Transformieren einer jeden der Wellenformen der Rahmen digitaler Eingangssignale, eingesetzt in die Eingangssignalwellenformeinheit **73** und zum Erzielen eines Eingangsspektrums und eines Phasenspektrums gemäß einem jeden Rahmen der digitalen Eingangssignale für alle Frequenzwerte;
 einer Rauschannahmeeinheit **76** zum Annehmen eines ersten Geräuschspektrums und eines zweiten Geräuschspektrums für alle Frequenzwerte gemäß einem aktuell eingegebenen Spektrum und mehreren in der Vergangenheit eingegebenen Spektren, gewonnen aus den aktuellen und den vergangenen Bildern der digitalen Eingangssignale in die Fourier-Transformationseinheit **75**;
 einer Spektrumspeichereinheit **77** zum Speichern des ersten und zweiten Geräuschspektrums, das die Annahmeeinheit **76** für jeden Frequenzwert angenommen hat;
 einer Stabilisierungskoeffizientenspeichereinheit **78** zum Speichern eines Stabilisierungskoeffizienten gemäß einem jeden Rahmen digitaler Eingangssignale, wobei der Stabilisierungskoeffizient eine höhere Stabilität vom ersten Geräuschspektrum aufzeigt, das die Geräuschannahmeeinheit **76** als einen Wert angenommen hat, wenn ein Wert des Stabilisierungskoeffizienten klein wird, und der Stabilisierungskoeffizient zwangsweise auf einen höheren Wert von der Geräuschannahmeeinheit **76** in Fällen eingestellt wird, bei denen die Geräuschleistung, gewonnen durch Addieren des ersten Geräuschspektrums für alle Frequenzwerte gemeinsam aufzeigt, daß keinerlei Geräusch im laufenden und in den vergangenen Rahmen existiert;
 einer Geräuschreduzierkoeffizientenjustiereinheit **79** zum Justieren eines Geräuschreduzierkoeffizienten, der verwendet wird zum Steuern eines Verminderungsgrades für ein Geräusch bei jedem Rahmen digitaler Eingangssignale gemäß dem Stabilisierungskoeffizienten, der in der Stabilisierungskoeffizientenspeichereinheit **78** gespeichert ist und allmählich den Stabilisierungskoeffizienten vermindert, der in der Speichereinheit **78** bei jedem digitalen Eingangssignal gespeichert wird, den der A/D-Umsetzer **72** gewinnt;
 einer Geräuschreduziereinheit **80** zum Reduzieren des Geräuschs, das in jedem Rahmen digitaler Eingangs-

signale gemäß dem Geräuschreduzierkoeffizienten enthalten ist, der in den die Geräuschreduzierkoeffizientenjustiereinheit **79** für jeden Frequenzwert eingestellt hat durch Subtrahieren des Geräuschspektrums, das in der Geräuschspektrumspeichereinheit **77** gespeichert ist, von jedem eingegebenen Spektrum, das die Fourier-Transformationseinheit **75** gewonnen hat, und zum Ausheben eines geräuschreduzierten Eingangsspektrums gemäß einem jeden Rahmen für alle Frequenzwerte;

einer Spektrumkompensationseinheit **81** zum Kompensieren eines jeden der geräuschverminderten Eingangsspektren aus der Reduziereinheit **80** für eine exzessive Subtraktion des Geräuschspektrums und zum Abgeben eines Ausgangsspektrums gemäß einem jeden Rahmen für alle Frequenzwerte;

einer Spektrumspeichereinheit **82** für vorherige Eingangssignale zum Speichern eines jeden Ausgangsspektrums, das die Spektrumkompensationseinheit **81** als vorheriges eingegebene Spektrum für alle Frequenzwerte abgibt, wobei die Kompensation in der Kompensationseinheit **81** unter Verwendung des zuvor eingegebenen Spektrums für jeden Frequenzwert erfolgt;

einer inversen Fourier-Transformationseinheit **83** zum Ausführen einer inversen Fourier-Transformation für jedes der Ausgangsspektren, gesendet von der Spektrumkompensationseinheit **81** gemäß dem Phasenspektrum, das gewonnen wird für jeden Rahmen digitaler Eingangssignale in die Fourier-Transformationseinheit **75** und zum Erzielen einer Gruppe von Ausgangssignalen erster Ordnung für jeden Rahmen digitaler Eingangssignale;

einer Spektrumanhebeeinheit **84** zum Verstärken eines Spektrums einer jeden Gruppe von Ausgangssignalen erster Ordnung, gewonnen in der inversen Fourier-Transformationseinheit **83** durch Ausführen einer stark betonenden Filterverarbeitung, einer hochfrequenzbetonenden Filterverarbeitung und einem Leistungsbetonungsfilter, das für jede Gruppe von Ausgangssignalen erster Ordnung gemäß dem linearen Prädiktionskoeffizienten verarbeitet, gewonnen für jeden Rahmen der digitalen Eingangssignale in der linearen Prädiktionskoeffizienten-Analysiereinheit **74** um zum Erzielen einer Gruppe von Ausgangssignalen zweiter Ordnung für jeden Rahmen digitaler Eingangssignale;

einer Speichereinheit **85** für zuvor gelesene Wellenformen, die einen Kopfabschnitt des aktuellen Rahmens von Ausgangssignalen zweiter Ordnung speichert, die von der Spektrumbetonungseinheit **84** betont wurden; und mit

einer Wellenformanpaßeinheit **86** zum Aufnehmen eines aktuellen Rahmens von Ausgangssignalen zweiter Ordnung und eines Kopfabschnitts eines nachfolgenden Rahmens von Ausgangssignalen zweiter Ordnung, die die Spektrumbetonungseinheit **84** betont hat, Deformieren einer Wellenform des aktuellen Rahmens, empfangen gemäß dem Kopfabschnitt vom aktuellen Rahmen, der in der Speichereinheit **85** für zuvor gelesene Wellenformen gespeichert ist, um die Wellenform des aktuellen Rahmens an diejenige eines vorangehenden Rahmens anzupassen, der dem aktuellen Rahmen vorangeht, Abgeben eines angepaßten Rahmens von Ausgangssignalen zweiter Ordnung als aktuellen Rahmen von Ausgangssignalen So für jeden Rahmen digitaler Eingangssignale und Senden des Kopfabschnitts vom nachfolgenden Rahmen an die Speichereinheit **85**, um den Kopfabschnitt des aktuellen Rahmens mit dem Kopfabschnitt des nachfolgenden Rahmens zu ersetzen.

[0137] In der obigen Konfiguration ist ein Betrieb der Vorrichtung **71** zur Geräuschverminderung kurz beschrieben worden.

[0138] Eine Vielzahl von Rahmen analoger Eingangssprachsignale \sin werden umgesetzt in eine Vielzahl von Rahmen digitaler Eingangssignale im A/D-Umsetzer **72** in aufeinanderfolgender Reihe, und Wellenformen der Rahmen digitaler Eingangssignale werden in einen Speicher der Einstelleinheit **73** gesetzt. Ein Stabilisierungskoeffizient, der die Stabilität einer Geräuschannahme aufzeigt, wird auf einen Anfangswert gesetzt und in der Speichereinheit **78** für Stabilisierungskoeffizienten gespeichert. In der Justiereinheit **79** wird ein Geräuschreduzierkoeffizient justiert gemäß dem Stabilisierungskoeffizienten, und der Stabilisierungskoeffizient wird allmählich vermindert, immer wenn ein Rahmen digitaler Eingangssignale im A/D-Umsetzer **72** gewonnen ist.

[0139] Ein linearer Prädiktionskoeffizient wird in der Analysiereinheit **74** für jeden Rahmen digitaler Eingangssignale erzeugt, die in die Einstelleinheit **73** gesetzt wurden, und eine Fourier-Transformation erfolgt für jeden Rahmen digitaler Eingangssignale in der Transformationseinheit **75**, um ein Eingangsspektrum und ein Phasenspektrum gemäß einem jeden Rahmen digitaler Eingangssignale für alle Frequenzwerte zu erzeugen. Danach wird ein Geräuschspektrum, das in einem aktuellen oder in mehreren vergangenen Rahmen des Eingangsspektrums enthalten ist, in der Annahmeeinheit **76** angenommen und in der Geräuschspektrumspeichereinheit **77** für jeden Frequenzwert gespeichert. In Fällen, bei denen die Geräuschannahmeeinheit **76** beurteilt, daß keinerlei Geräusch in aktuellen oder den vergangenen Rahmen gemäß dem Geräuschspektrum für alle Frequenzwerte vorhanden ist, wird der Stabilisierungskoeffizient, gespeichert in der Stabilisierungskoeffizientenspeichereinheit **78**, zwangsweise auf einen hohen Wert eingestellt, weil das Geräuschspektrum sehr instabil angenommen wird. Jedes der in der Speichereinheit **77** gespeicherte Geräuschspektrum wird danach von

dem aktuellen Eingangsspektrum subtrahiert bei dem laufenden Rahmen digitaler Eingangssignale gemäß dem Geräuschreduzierkoeffizienten, der in der Justiereinheit **79** gespeichert ist, und eine Vielzahl von geräuschreduzierten Spektren bei dem aktuellen Rahmen für alle Frequenzwerte wird abgegeben. Das heißt, ein in jedem Rahmen digitaler Eingangssignale enthaltenes Geräusch wird reduziert gemäß dem Geräuschreduzierkoeffizienten. In diesem Falle gibt es die Wahrscheinlichkeit, daß die Intensität eines geräuschreduzierten Eingangsspektrums auf einen negativen Wert vermindert wird. In Fällen, bei denen die Intensität vom geräuschreduzierten Eingangsspektrum auf einen negativen Wert vermindert wird, erfolgt das Kompensieren des geräuschreduzierten Eingangsspektrums in der Kompensationseinheit **81** gemäß dem zuvor eingegebenen Spektrum, das die Speichereinheit **82** speichert. Danach wird eine inverse Fourier-Transformation für alle eingegebenen Spektren ausgeführt, gewonnen in der Kompensationseinheit **81** gemäß einem jeden Phasenspektrum. Folglich wird eine Vielzahl von Gruppen von Ausgangssignalen erster Ordnung nacheinander gewonnen. Jede Gruppe von Ausgangssignalen erster Ordnung ist zusammengesetzt aus einem Rahmen von Ausgangssignalen erster Ordnung und einem Kopfabschnitt eines nachfolgenden Rahmens von Ausgangssignalen erster Ordnung. Danach wird ein Spektrum einer jeden Gruppe von Ausgangssignalen erster Ordnung in der Betonungseinheit **84** gemäß einem linearen Prädiktionskoeffizienten betont, der für jeden Rahmen digitaler Eingangssignale in der Analysiereinheit **74** für lineare Prädiktionskoeffizienten gewonnen wird. Danach wird eine Wellenform eines aktuellen Rahmens von Ausgangssignalen zweiter Ordnung in der Betonungseinheit **84** gewonnen und so angepaßt, daß ein vorangehender Rahmen von Ausgangssignalen zweiter Ordnung in der Anpaßeinheit **86** erfolgt, und der aktuelle Rahmen der Ausgangssignale zweiter Ordnung wird als aktueller Rahmen von den Ausgangssignalen So abgegeben.

[0140] Als nächstes beschrieben ist ein Algorithmus in der Vorrichtung **71** in Einzelheiten. Namen verschiedener fester Parameter und Einstellbeispiele sind in der Tabelle 1 aufgelistet.

Tabelle 1

Feste Parameter	Einstellbeispiele
Rahmenlänge eines aktuellen Rahmens von Ausgangssignalen	160 Punkte, die 20 Millisekunden für den aktuellen Rahmen äquivalent sind, abgetastet mit 8 kHz
Länge des Kopfabschnitt vom nachfolgenden Rahmen, gelesen mit dem aktuellen Rahmen	60 Punkte äquivalent zu 7,5 Millisekunden für den nachfolgenden Rahmen, abgetastet mit 8 kHz

Tabelle 1 (Forts.)

Feste Parameter	Einstellbeispiele
Bestimmter Geräuschreduzierkoeffizient	20,0
Prädiktionsreihenfolge des linearen Prädiktionskoeffizienten	10
Bezugsfortsetzungszahl von Geräuschspektren	100
Bestimmte minimale Klangleistung	10,0
Bewegungsdurchschnitt-betonungskoeffizient (MA-Koeffizient)	0,6
Autoregressiver Betonungskoeffizient (AR-Koeffizient)	0,7
Hochfrequenzbetonungskoeffizient	0,3
Klangleistungsbetonungskoeffizient	1,2
Bezugsgeräuschleistung	2.700,0

[0141] Nachdem die festen Parameter eingesetzt sind, werden die statischen RAM-Zonen der Speichereinheit **78** für Stabilisierungskoeffizienten, die Geräuschspektrumspeichereinheit **77**, die Speichereinheit **82** für vorheriges Spektrum und die Speichereinheit **85** für vorher gelesene Wellenform gelöscht und auf die Anfangswerte gesetzt. Dies geschieht, weil der Stabilisierungskoeffizient die Stabilität für die Annahme eines Geräuschs aufzeigt, das im Eingangssignal enthalten, ein Stabilisierungskoeffizient vom Wert 2,0 wird anfänglich in die Stabilisierungskoeffizientenspeichereinheit **78** eingesetzt.

[0142] Je kleiner der Wert des Stabilisierungskoeffizienten ist, um so stabiler ist die Geräuschannahme in der Geräuschannahmeeinheit **76**. Die Speichereinheit **77** für das Geräuschspektrum hat einen Bereich zum Speichern eines ersten Geräuschspektrums als Kandidat erster Stelle für jeden Frequenzwert, ein zweites Geräuschspektrum als Kandidaten zweiter Stelle für jeden Frequenzwert, wobei die Anzahl von Rahmen (erste Fortsetzungszahl genannt), in der das erste Geräuschspektrum beibehalten wird auf einem Minimalwert unter Werten von Eingangsspektren in den Rahmen und die Anzahl der Rahmen (zweite Fortsetzungszahl genannt), in der das zweite Geräuschspektrum beibehalten wird auf einem zweiten Minimalwert unter Werten der Eingangsspektren in den Rahmen. Hinreichend hohe Werte werden folglich anfänglich in der Einheit **77** als Werte erster und zweiter Rauschspektren gespeichert, und die Bezugsfortsetzungszahl für die erste und zweite Fortsetzungszahl. Die Speichereinheit **82** für vorherige Spektren hat einen Bereich zum Speichern eines Prädiktionsingangsspektrums für einen vorangehenden Rahmen und eine vorangehende Klangleistung für den vorangehenden Rahmen. Folglich wird eine bestimmte minimale Klangleistung anfänglich in der Speichereinheit für jeden der Frequenzwerte gespeichert, und ein Nullwert wird anfänglich in der Einheit **82** als vorangehende Klangleistung gespeichert. Die Speichereinheit **85** für zuvor gelesene Wellenform hat einen Bereich zum Speichern eines Kopfabchnitts eines nachfolgenden Rahmens, der mit dem aktuellen Rahmen gelesen wurde (oder einen Kopfabchnitt eines aktuellen Rahmens, der mit einem vorangehenden Rahmen gelesen wurde)

zum Anpassen des aktuellen Rahmens geräuschreduzierter Eingangssignale zum vorangehenden Rahmen geräuschreduzierter Eingangssignale. Folglich wird anfänglich ein Null-Wert in der Einheit **85** gespeichert.

[0143] Als nächstes beschrieben ist ein Geräuschreduzieralgorithmus in der Vorrichtung **71**.

[0144] Ein aktueller Rahmen analoger Eingangssignale und ein Abschnitt eines nachfolgenden Rahmens analoger Eingangssignale werden umgesetzt in einen aktuellen Rahmen digitaler Eingangssignale (160 Punkte) in einen Abschnitt nachfolgenden Rahmens digitaler Eingangssignale (60 Punkte) in A/D-Umsetzer **72**. Danach wird ein Geräuschreduzierkoeffizient Q in der Justiereinheit **79** gemäß einem Stabilisierungskoeffizienten C errechnet, der in der Einheit **78** gespeichert ist, und der bestimmte Geräuschreduzierkoeffizient P , wie er nachstehend angegeben ist.

$$C' = C * 0,8$$

$$Q = P * (1,0 - C)$$

$$Q = 1,0$$

$$Q = 0,0$$

(im Falle von $C < 1,0$)

(im Falle von $1,0 < C < 2,0$)

(im Falle von $2,0 < C$)

[0145] Ein erneuerter Stabilisierungskoeffizient C' wird in der Speichereinheit **78** als Stabilisierungskoeffizient C gespeichert. Das heißt, der Stabilisierungskoeffizient C wird allmählich vermindert, und zwar bei einem Rahmen digitaler Eingangssignale, der in A/D-Umsetzer **72** gewonnen wird, weil die Stabilität einer Rauschannahme erhöht ist, die in der Geräuschannahmeeinheit **76** ausgeführt wird.

[0146] In der Einstelleinheit **73** werden die digitalen Eingangssignale, gewonnen im A/D-Umsetzer **72**, in einen rückwärtigen Abschnitt eines ersten Speichers geschrieben, der eine Datenpunktlänge von 2^n hat (n ist eine natürliche Zahl), und ein Kopfabschnitt vom ersten Speicher wird ausgefüllt mit einer Serie von digitalen "0"-Werten. Das heißt, digitale "0"-Werte werden in den Kopfabschnitt geschrieben, dessen Bereich von 0 bis 35 Positionen reicht, und die digitalen Eingangssignale werden in den rückwärtigen Abschnitt geschrieben, der den Bereich von 36 – 255 Positionen abdeckt, um eine erste Datenanordnung im ersten Speicher zu bilden. Diese erste Datenanordnung wird verwendet als Realkomponente bei einer schnellen Fourier-Transformation achter Ordnung, die die Transformationseinheit **75** ausführt. Eine zweite Datenanordnung, die zusammengesetzt aus digitalen "0"-Werten, wird ebenfalls anfänglich in einem zweiten Speicher ausgeführt, der dieselbe Datenpunktlänge von 2^n hat. Die zweite Datenanordnung wird verwendet als Imaginärkomponente bei der schnellen Fourier-Transformation achter Ordnung.

[0147] In der Analysiereinheit **74** für lineare Prädiktionskoeffizienten wird ein Humming-Fenster auf die erste Datenanordnung gegeben. Das heißt, die digitalen Eingangssignale werden im ersten Speicher der Einheit **73** multipliziert mit einem Fensterkoeffizienten, um eine Vielzahl digitaler fensterverarbeiteter Signale zu erzeugen, eine Autokorrelationsanalyse erfolgt für die digitalen fensterverarbeiteten Signale, um eine Vielzahl von Autokorrelationskoeffizienten zu erzeugen, eine lineare Prädiktionskoeffizientenanalyse auf der Grundlage eines Autokorrelationsverfahren wird für die Autokorrelationskoeffizienten ausgeführt, und ein lineare Prädiktionskoeffizient α_i wird gewonnen. Danach werden ein MA Koeffizient $\alpha(MA)_i$ und ein AR -Koeffizient $\alpha(AR)_i$ in einem stark betonenden Filter, verwendet in der Betonungseinheit **84**, gemäß dem linearen Prädiktionskoeffizienten α_i errechnet, ein MA-Betonungskoeffizient β bestimmt und ein AR-Betonungskoeffizient γ vorbestimmt.

$$\alpha(MA)_i = \alpha_i * \beta^i$$

$$\alpha(AR)_i = \alpha_i * \gamma^i$$

[0148] Hier bedeutet i einen Grad.

[0149] In der Fourier-Transformationseinheit **75** wird eine direkte Fourier-Transformation für die erste Datenanordnung ausgeführt, gebildet als Realkomponente in der zweiten Datenanordnung, gebildet als Imaginärkomponente in der Einstelleinheit **73** zum Erzielen einer Realkomponente und einer Imaginärkomponente eines komplexen Spektrums, das ausgedrückt ist durch einen komplexen Vektor für jeden Frequenzwert, und ein erster Absolutwert der Realkomponente und ein zweiter Absolutwert der Imaginärkomponente werden miteinander addiert. Weil ein Prozeß, der eine Quadratwurzel eines quadrierten absoluten Wertes ist, gewonnen durch Addieren der quadrierten Realkomponente und der quadrierten Imaginärkomponente miteinander, folglich errechnet, um ein Amplitudenspektrum mit einem Wert der Quadratwurzel für jeden Frequenzwert zu bekommen, wird folglich nicht ausgeführt, ein Pseudoamplitudenspektrum (nachstehend als Eingangsspektrum bezeichnet) mit einem Wert einer Summe des ersten Absolutwerts und des zweiten Absolutwerts wird für

jeden Frequenzwert gewonnen. Auch eine Summe der Eingangsspektren für alle Frequenzwerte wird als Eingangsklangleistung für jeden Rahmen gewonnen.

[0150] Weil die Errechnung zum Erzielen der Quadratwurzel eines quadrierten Absolutwertes folglich fortgelassen werden kann, kann der Errechnungsumfang, der erforderlich ist, das Eingangsspektrum zu bekommen, beträchtlich vermindert werden.

[0151] Als nächstes beschrieben ist die Verarbeitung in der Geräuschannahmeeinheit **76**.

[0152] Da das Geräusch ständig in jedem Rahmen analoger Eingangssprachsignale S_{in} als Hintergrund vorhanden ist, hat das Geräusch immer eine minimale Intensität unter jenen der analogen Eingangssprachsignale S_{in} . In diesem Ausführungsbeispiel wird folglich das erste und das zweite Geräuschspektrum, das das Geräusch repräsentiert, folgendermaßen bestimmt. Ein erstes Geräuschspektrum, das in der Geräuschspektrumspeichereinheit **77** als Kandidat erster Stelle für jeden Frequenzwert gespeichert ist, wird festgelegt als Eingangsspektrum mit einem Minimalwert unter den Werten eingegebener Spektren im laufenden Rahmen und einem oder mehreren vergangenen Rahmen der Zahl, die der ersten Fortsetzungszahl gleicht. Das heißt, der Minimalwert $N_{i,t}$ des ersten Geräuschspektrums wird folgendermaßen formuliert:

$$N_{i,t} = \min \{S_{i,p}\} \quad (p = t \text{ bis } t-T)$$

[0153] Hier bedeutet i die Anzahl von Frequenzwerten, t bedeutet ein aktueller Rahmen, T bedeutet die erste Fortsetzungszahl der Rahmen aus dem aktuellen Rahmen zum nächstletzten Rahmen, und $S_{i,p}$ bedeutet ein Eingabespektrum eines Rahmens p für einen i -ten Frequenzwert.

[0154] Auch ein zweites Geräuschspektrum, das in der Geräuschspektrumspeichereinheit **77** als Kandidat an zweiter Stelle für jeden Frequenzwert gespeichert ist, wird festgelegt als Eingangsspektrum mit einem zweiten Minimalwert unter Werten eingegebener Spektren im aktuellen Rahmen und einem oder mehreren vergangenen Rahmen der Zahl, die der zweiten Fortsetzungszahl gleicht.

[0155] Da die Bedingungen des Geräuschs stetig in analogen Eingabesprachsignalen S_{in} vorhanden sind, und sich ständig während einer langen Periode ändern, wird das erste oder zweite Geräuschspektrum geändert, wenn die erste oder zweite Fortsetzungszahl größer als die Bezugsfortsetzungszahl wird.

[0156] In Einzelheiten: die erste Fortsetzungszahl für das erste Geräuschspektrum, das als Kandidat erster Stelle gespeichert ist, und die zweite Fortsetzungszahl für das zweite Geräuschspektrum, das als Kandidat zweiter Stelle in der Speichereinheit **77** gespeichert ist, wird jeweils inkrementiert in der Annahmeeinheit **76** bei jeder ausgeführten Fourier-Transformation für jeden Rahmen digitaler Eingangssignale in der Transformationsseinheit **75**. Im Falle, daß die erste Fortsetzungszahl für eines der Frequenzwerte größer ist als die Bezugsfortsetzungszahl ist, werden danach das zweite Geräuschspektrum und die zweite Fortsetzungszahl als Kandidaten zweiter Stelle in ein erstes Geräuschspektrum verbracht, und eine erste Fortsetzungszahl als Kandidat an erster Stelle in die Annahmeeinheit **76**, ein drittes Geräuschspektrum als Kandidat dritter Stelle wird in ein zweites Geräuschspektrum als Kandidat zweiter Stelle gegeben, und eine Fortsetzungszahl als Kandidat dritter Stelle wird auf null gesetzt. In diesem Falle wird ein Geräuschspektrum mit einem Wert verwendet, der geringfügig höher ist als derjenige des zweiten Geräuschspektrums, verwendet als drittes Geräuschspektrum für den Kandidaten an dritter Stelle. Beispielsweise wird ein Geräuschspektrum mit einem Wert des 1,2-fachen über demjenigen des zweiten Geräuschspektrums automatisch auf das dritte Geräuschspektrum als Kandidat dritter Stelle gesetzt.

[0157] In Fällen, bei denen die zweite Fortsetzungszahl für einen der Frequenzwerte größer ist als die Bezugsfortsetzungszahl, wird ein drittes Geräuschspektrum für einen Kandidaten dritter Stelle in ein zweites Geräuschspektrum als Kandidat zweiter Stelle gesetzt, und eine dritte Fortsetzungszahl als Kandidat dritter Stelle auf null gesetzt.

[0158] In Fällen, in denen im Gegensatz dazu weder das erste noch die zweite Fortsetzungszahl für einen der Frequenzwerte nicht größer als die Bezugsfortsetzungszahl ist, wird das erste Geräuschspektrum verglichen mit einem Eingangsspektrum für den aktuellen Rahmen in der Geräuschannahmeeinheit **76** für jeden Frequenzwert. In Fällen, bei denen ein Wert des Eingabespektrums kleiner ist als derjenige des ersten Geräuschspektrums, werden das erste Geräuschspektrum und die erste Fortsetzungszahl als Kandidat erster Stelle eingesetzt in ein zweites Geräuschspektrum und eine zweite Fortsetzungszahl als Kandidat zweiter Stelle, das Eingabespektrum wird an ein erstes Geräuschspektrum als Kandidat erster Stelle eingegeben, und eine erste

Fortsetzungszahl als Kandidat erster Stelle wird auf null gesetzt. In Fällen, bei denen ein Wert eingegebenen Spektrums nicht kleiner als derjenige des ersten Spektrums ist, wird das zweite Geräuschspektrum verglichen mit dem eingegebenen Spektrum in die Geräuschannahmeeinheit **76** für jeden Frequenzwert. In Fällen, bei denen ein Wert des eingegebenen Spektrums kleiner als derjenige des zweiten Geräuschspektrums ist, wird das Eingabespektrum auf ein zweites Geräuschspektrum als Kandidat zweiter Stelle eingesetzt, und eine Fortsetzungszahl als Kandidat zweiter Stelle wird auf null gesetzt.

[0159] Danach werden das erste und das zweite Geräuschspektrum und die erste und die zweite Fortsetzungszahl in die Annahmeeinheit **276** verbracht und in der Speichereinheit **77** gespeichert.

[0160] In diesem Ausführungsbeispiel wird ein Geräuschspektrum $N_{j,t}$ bei einem Rahmen t für einen j -ten Frequenzwert in der Annahmeeinheit **77** aus vier eingegebenen Spektren angenommen, entsprechend den vier Frequenzwerten, die einander im selben Rahmen t benachbart sind, um eine Speicherkapazität eines RAM zu vermindern, die erforderlich ist zum Speichern des Geräuschspektrums in der Speichereinheit **77**. Beispielsweise wird ein Fourier-Transformierer mit 256 Punkten verwendet, und die Geräuschspektren $N_{j,t}$ werden folgendermaßen angegeben:

$$N_{j,t} = \min\{\min(S_{4j,p}, S_{4j+1,p}, S_{4j+2,p}, S_{4j+3,p})\} \quad (p = t \text{ to } t-T)$$

[0161] Hier hat j den Bereich von 0 bis 31, und i von $S_{i,p}$ hat den Bereich von 0 bis 127.

[0162] Da in diesem Falle ein Pseudoamplitudenspektrum ein Eingabespektrum aufrief, das symmetrisch zur Frequenzachse ist, wenn die Geräuschspektren $N_{i,t}$ für alle Frequenzwerte angenommen werden, weil Spektren und Fortsetzungszahlen für 128 Frequenzwerte in der Speichereinheit **77** gespeichert werden, eine Speicherkapazität des RAM aufsummiert wird auf 512 Wörter $\{128 \text{ (Frequenzwerte)} * 2 \text{ (Spektrum- und Fortsetzungsnummer)} * 2 \text{ (Kandidaten erster Stelle und zweiter Stelle)}\}$. Wenn im Gegensatz dazu die Geräuschspektren $N_{j,t}$ aus vier Eingabespektren angenommen werden, vermindert sich die Speicherkapazität vom RAM auf 128 Wörter $\{32 \text{ (Frequenzwerte)} * 2 \text{ (Spektrum- und Fortsetzungsnummer)} * 2 \text{ (Kandidaten erster und zweiter Stelle)}\}$

[0163] Obwohl in diesem Falle die Frequenzauflösung und die Geräuschspektren vermindert sind, ist sichergestellt, daß eine Geräuschverminderungsfunktion in der Vorrichtung **71** nicht substantiell verschlechtert wird, obwohl die Rauschspektren $N_{j,t}$ in der Annahmeeinheit **76** angenommen werden.

[0164] Weil jedes der Geräuschspektren $N_{j,i}$ nicht von einem eingegebenen Spektrum für einen Fortsetzungszahl angenommen wird, wenn ein stationärer Klang, wie eine Sinuswelle, ein Vokal oder dergleichen für eine längere Zeit fortgesetzt wird, gibt es auch den Effekt, daß ein eingegebenes Spektrum daran gehindert wird, als Geräuschspektrum angenommen zu werden.

[0165] Nachdem das erste und das zweite Geräuschspektrum angenommen ist, wird eine Summe des ersten Geräuschspektrums für alle Frequenzwerte errechnet als Geräuschleistung in der Annahmeeinheit **76**. Wenn die Geräuschleistung geringer als die Bezugsleistung ist, wird in der Annahmeeinheit beurteilt, daß irgendwelche Geräusche nicht vorhanden sind im aktuellen Rahmen und den vergangenen Rahmen gemäß dem ersten Geräuschspektrum, und ein in der Stabilisierungskoeffizientenspeichereinheit **78** gespeicherter Stabilisierungskoeffizient wird zwangsweise auf einen höheren Wert von 3,0 gesetzt. Danach wird der auf 3,0 gesetzte Stabilisierungskoeffizient zur Geräuschverminderungskoeffizienteneinstelleinheit **79** gesandt, um die Justiereinheit **79** darüber zu informieren, daß das Geräusch im aktuellen Rahmen des eingegebenen Signals und der vergangenen Rahmen eingegebener Sprachsignale S_{in} nicht vorhanden ist. Wenn irgendein Geräusch nicht vorhanden ist, kann folglich der Geräuschverminderungskoeffizient auf einen niedrigeren Wert in der Justiereinheit **79** eingestellt werden. Das heißt, eine Spektralverzerrung in einer Sprache, verursacht durch die Geräuschverminderung, kann vermindert werden.

[0166] Als nächstes beschrieben ist die Arbeitsweise, die die Geräuschverminderungseinheit **80** ausführt.

[0167] Nachdem ein aktueller Rahmen digitaler Eingabesignale transformiert ist in ein aktuelles Eingabespektrum in der Transformiereinheit **75**, wird das aktuelle Eingabespektrum übertragen zur Geräuschannahmeeinheit **76** und zur Geräuschverminderungseinheit **80**. In der Geräuschannahmeeinheit **76** wird eine Vielzahl von ersten Geräuschspektren $N_{i,t}$ beim aktuellen Rahmen für alle Frequenzwerte angenommen und gespeichert als Kandidaten an erster Stelle in der Geräuschspektrumspeichereinheit **77**. Danach werden erste Geräuschspektren übertragen zur Geräuschverminderungseinheit **80**. In der Geräuschverminderungseinheit

80 werden die ersten Geräuschspektren jeweils multipliziert mit dem Geräuschverminderungskoeffizienten, der in der Justiereinheit **79** justiert wurde, um ein multipliziertes Geräuschspektrum für jeden Frequenzwert zu erzeugen, und das multiplizierte Geräuschspektrum wird vom aktuellen Eingabespektrum subtrahiert, um ein geräuschreduziertes Eingabespektrum für jeden Frequenzwert zu erhalten. In Fällen, bei denen die ersten Geräuschspektren $N_{j,t}$ im laufenden Rahmen für alle nummernreduzierte Frequenzwerte in der Geräuschannahmeeinheit **76** angenommen und als Kandidaten erster Stelle in der Geräuschspektrumspeichereinheit **77** gespeichert werden, erfolgt im Gegensatz dazu das Multiplizieren der ersten Geräuschspektren $N_{j,t}$ jeweils mit dem Geräuschverminderungskoeffizienten, um eine Vielzahl multiplizierter Geräuschspektren zu erzeugen, und jedes der multiplizierten Geräuschspektren wird von dem laufenden Eingabespektrum entsprechend demselben Frequenzwert subtrahiert, um eine Vielzahl von geräuschverminderten Eingabespektren für die Anzahl reduzierter Frequenzwerte zu erzeugen.

[0168] Als nächstes beschrieben ist die Arbeitsweise, die von der Spektrumkompensationseinheit **81** ausgeführt wird.

[0169] Eine Vielzahl von Eingabespektren, gewonnen in der Transformationseinheit **75** für alle Frequenzwerte, erfahren eine Speicherung als Vielzahl vorhergehender Eingabespektren in der Speichereinheit **82** für vorherige Spektren im voraus. Nachdem ein aktuelles Eingabespektrum entsprechend einem aktuellen Rahmen digitaler Eingangssignale errechnet ist in der Fourier-Transformationseinheit **75** für jeden der Frequenzwerte, wird ein in der Speichereinheit **78** gespeicherter Stabilisierungskoeffizient als Bezug genommen von der Spektrumkompensationseinheit **81**. Ist der Stabilisierungskoeffizient höher als 1,4, dann beurteilt die Kompensationseinheit **81**, daß eine Geräuschannahme für einen vorangehenden Rahmen in der Geräuschannahmeeinheit ausgeführt wird und instabil ist, und die vorherigen Eingabespektren, die in der Speichereinheit **82** im voraus gespeichert wurden, sind ebenfalls instabil. Die speziellen vorherigen Eingabespektren, die in der Speichereinheit **82** gespeichert sind, werden folglich ersetzt durch eine Vielzahl aktueller Eingabespektren für einen aktuellen Rahmen, und die aktuellen Eingabespektren, die in der Speichereinheit **82** gespeichert sind, werden häufig von vorhergehenden Eingabespektren für alle Frequenzwerte aufgerufen. In Fällen, bei denen der Stabilisierungskoeffizient im Gegensatz dazu nicht höher als 1,4 ist, werden die vorherigen Eingabespektren beibehalten, die in der Speichereinheit **82** gespeichert sind.

[0170] Wenn danach eine Vielzahl aktueller geräuschverminderter Eingangsspektren entsprechend dem aktuellen Rahmen digitaler Eingabesignale aller Frequenzwerte gesendet werden von der geräuschvermindernden Einheit **80** an die Kompensationseinheit **81**, erfolgt das Überprüfen von Werten aktueller geräuschverminderter Eingangsspektren. In Fällen, bei denen ein Wert eines speziellen aktuellen geräuschverminderten Eingangsspektrums für einen speziellen Frequenzwert negativ ist, wird ein Eingangsspektrum mit einem niedrigeren Wert zwischen einem aktuellen Eingabespektrum für einen speziellen Frequenzwert von der Sendeeinheit **75** gesendet, und ein vorheriges Eingabespektrum für den speziellen Frequenzwert, gespeichert in der Speichereinheit **82**, wird ausgewählt als Ausgabespektrum für den speziellen Frequenzwert. In Fällen, in denen im Gegensatz dazu der Wert des speziellen aktuellen geräuschverminderten Eingangsspektrums für den speziellen Frequenzwert nicht negativ ist, wird ein spezielles aktuelles geräuschvermindertes Eingangsspektrum für den speziellen Frequenzwert verwendet als Ausgabespektrum für den speziellen Frequenzwert. Jedes der ausgegebenen Spektren für alle Frequenzwerte werden danach in der Speichereinheit **82** gespeichert, um das vorherige Eingabespektrum zu ersetzen, das in der Speichereinheit **82** gespeichert ist, mit dem Ausgabespektrum. Die Ausgabespektren, die in der Einheit **82** gespeichert sind, werden eine Vielzahl vorheriger Eingabespektren aufrufen. Die Ausgabespektren werden auch an eine inverse Fourier-Transformationseinheit **83** gesandt.

[0171] In diesem Ausführungsbeispiel ist die Speichereinheit **82** für vorhergehendes Spektrum erforderlich, um die vorhergehenden Eingabespektren zu speichern. In Fällen, bei denen multiplizierte Eingabespektren gewonnen werden durch Multiplizieren der aktuellen Eingabespektren mit einem Koeffizienten, werden jedoch behandelt wie bei den vorherigen Eingabespektren, und die Speichereinheit **82** ist nicht erforderlich, eine Speicherkapazität, die in der Vorrichtung **71** erforderlich ist, kann verkleinert werden. In Fällen, bei denen ein spezielles aktuelles Eingabespektrum für einen speziellen Frequenzwert als spezielles Ausgabespektrum für den speziellen Frequenzwert eingesetzt ist, wenn ein spezielles aktuelles geräuschvermindertes Eingangsspektrum für den speziellen Frequenzwert negativ ist, dann ist die Speichereinheit **82** nicht erforderlich.

[0172] Auch in Fällen, bei denen irgendeine Sprache nicht in einem aktuellen Rahmen enthalten ist, und nur ein Geräusch im aktuellen Rahmen enthalten ist, wird eine Klangleistung des aktuellen Rahmens plötzlich vermindert, wenn das Geräusch in der Geräuschverminderungseinheit **80** vermindert wurde, und Allophonie tritt im aktuellen Rahmen auf, wegen der plötzlichen Änderung der Klangleistung. Um das Auftreten von Allophonie

zu verhindern, wird die Klangleistung bei einer Serie von Rahmen, die den aktuellen Rahmen enthalten, geglättet unter Verwendung einer Klangleistung der zuvor eingegebenen Spektren, die die Speichereinheit **82** speichert. In Einzelheiten: eine Summe von Werten der Ausgangsspektren entsprechend dem aktuellen Rahmen wird errechnet als Ausgabeklangleistung. Auf den in der Speichereinheit **78** gespeicherten Stabilisierungskoeffizienten wird danach Bezug genommen. In Fällen, bei denen der Stabilisierungskoeffizient über 1,4 liegt, wird die abgegebene Klangleistung, die errechnet ist, in der Speichereinheit **82** als vorherige Klangleistung gespeichert. In Fällen, bei denen im Gegensatz dazu der Stabilisierungskoeffizient nicht höher als 1,4 ist, wird die Klangleistung, gewonnen durch Addierwerte der vorherigen Eingabespektren, gespeichert in der Einheit **82**, gemeinsam in der Speichereinheit gespeichert als vorherige Klangleistung.

[0173] In Fällen, bei denen die Ausgangsklangleistung danach geringer als das 1,8-fache der vorherigen Klangleistung ist, gespeichert in der Speichereinheit **82**, und in Fällen, bei denen die vorherige Klangleistung geringer ist als eine multiplizierte Klangleistung, gewonnen durch Multiplizieren der Geräuschleistung, die die Annahmeeinheit **76** errechnet hat, um einen Wert gleich dem 1,4-fachen des Geräuschverminderungskoeffizienten, wird die Glättung der Ausgangsklangleistung ausgeführt durch Ändern der vorherigen Klangleistung D_p , die in der Speichereinheit **82** gespeichert ist, auf eine neue vorherige Klangleistung D_q , und zwar folgendermaßen:

$$D_q = D_p * 0,8 + A_p * 0,2$$

[0174] Hier bedeutet A_p die Ausgangsklangleistung.

[0175] Danach wird jedes der Ausgabespektren für alle Frequenzwerte des laufenden Rahmens multipliziert durch einen Teilwert D_q/D_p , und eine Vielzahl multiplizierter Ausgangsspektren für alle Frequenzwerte bei dem aktuellen Rahmen werden gewonnen. Danach wird die neue vorherige Klangleistung D_q ersetzt durch die vorherige Klangleistung D_p , die in der Speichereinheit **82** gespeichert ist, und die multiplizierten Ausgangsspektren werden ersetzt durch die vorherigen Eingabespektren, die die Speichereinheit **82** speichert. Auch werden die multiplizierten Ausgangsspektren behandelt als Vielzahl von Ausgangsspektren und verarbeitet in der inversen Fourier-Transformationseinheit **83**.

[0176] In der inverse Fourier-Transformationseinheit **83** wird eine Vielzahl von komplexen Spektren aus den Phasenspektren erzeugt, die die Fourier-Transformationseinheit **75** erzielt, und den Ausgabespektren, die die Spektrumkompensationseinheit **81** erzielt, und eine inverse schnelle Fourier-Transformation erfolgt für die komplexen Spektren, um eine Gruppe von Ausgangssignalen erster Ordnung zu erhalten.

[0177] Als nächstes beschrieben ist die Arbeitsweise, die die Spektrumbetonungseinheit **84** für die Gruppe von Ausgangssignalen erster Ordnung ausführt.

[0178] Die Arbeitsweise in einem stark betonenden Filter erfolgt für die Gruppe von Ausgangssignalen erster Ordnung, gewonnen in der Transformationseinheit **83**, mit den MA-Koeffizienten $\alpha(MA)_i$ und den AR-Koeffizienten $\alpha(AR)_i$, gewonnen in der Analysiereinheit **74**. Eine Übertragungsfunktion $Ft1$ des hochbetonenden Filters wird folgendermaßen dargestellt:

$$Ft1 = \{1 + \alpha(MA)_1 Z^{-1} + \alpha(MA)_2 Z^{-2} + \dots + \alpha(MA)_j Z^{-j}\} / \{1 + \alpha(AR)_1 Z^{-1} + \alpha(AR)_2 Z^{-2} + \dots + \alpha(AR)_j Z^{-j}\}$$

[0179] Um darüber hinaus Hochfrequenzkomponenten der Gruppe von Ausgangssignalen erster Ordnung zu betonen, wird eine Operation in einem Hochfrequenzbetonungsfilter ausgeführt für die Gruppe von Ausgangssignalen erster Ordnung mit einem Hochfrequenzkomponenten betonenden Koeffizienten δ , der vorbestimmt ist. Eine Übertragungsfunktion $Ft2$ des hochfrequenzbetonenden Filters wird folgendermaßen ausgedrückt:

$$Ft2 = 1 - \delta z^{-1}$$

[0180] Um darüber hinaus eine Klangleistung der Gruppe von Ausgangssignalen erster Ordnung zu betonen, wird die Gruppe von Ausgangssignalen erster Ordnung, in der die hochfrequenten Komponenten vom hochbetonenden Filter multipliziert mit einem Leistungsbetonungskoeffizienten, der zum Bestimmen einer Gruppe von Ausgangssignalen zweiter Ordnung bestimmt ist.

[0181] In der Wellenformanpaßeinheit **86** wird die Gruppe von Ausgangssignalen zweiter Ordnung, gewonnen in der Betonungseinheit **84**, und ein Kopfabschnitt des aktuellen Rahmens von Ausgangssignalen, gespeichert in der Speichereinheit **85** für zuvor ausgelesene Wellenformen einander überlagert. In diesem Falle ist

die Gruppe von Ausgangssignalen zweiter Ordnung zusammengesetzt aus einem aktuellen Rahmen von Ausgangssignalen zweiter Ordnung und einem Kopfabschnitt eines nachfolgenden Rahmens von Ausgangssignalen zweiter Ordnung, die simultan betont werden mit dem aktuellen Rahmen der Ausgangssignale zweiter Ordnung, und der Kopfabschnitt des aktuellen Rahmens von ausgegebenen Signalen, gespeichert in der Speichereinheit **85** für vorausgelesene Wellenformen, wird erzeugt aus einem Kopfabschnitt eines aktuellen Rahmens der Ausgangssignale zweiter Ordnung, simultan betont mit einem vorausgehenden Rahmen von Ausgangssignalen zweiter Ordnung. Wenn die Gruppe von Ausgangssignalen zweiter Ordnung und der Kopfabschnitt des aktuellen Rahmens der Ausgangssignale einander überlagert werden, wird die Gruppe zweiter Ordnung von Ausgangssignalen transformiert in einen aktuellen Rahmen der Ausgangssignale letztlicher Ordnung und einem Kopfabschnitt eines nachfolgenden Rahmens von Ausgangssignalen letztlicher Ordnung, und ein aktueller Rahmen der Ausgangssignale letztlicher Ordnung paßt zu einem vorhergehenden Rahmen der Ausgangssignale letztlicher Ordnung, die zuvor transformiert worden sind. Danach wird der aktuelle Rahmen letztlicher Ordnung zu Ausgangssignalen abgegeben als aktueller Rahmen von Ausgangssignalen, und der Kopfabschnitt des nachfolgenden Rahmens von Ausgangssignalen zweiter Ordnung wird abgegeben als Kopfabschnitt eines nachfolgenden Rahmens von Ausgangssignalen. Auch der Kopfabschnitt des aktuellen Rahmens von Ausgangssignalen, gespeichert in der Speichereinheit **85**, wird ersetzt mit dem Kopfabschnitt des nachfolgenden Rahmens der Ausgangssignale letztlicher Ordnung. Der Kopfabschnitt des nachfolgenden Rahmens von Ausgangssignalen letztlicher Ordnung, gespeichert in der Einheit **85**, wird ein Kopfabschnitt des nachfolgenden Rahmens von Ausgangssignalen.

[0182] Die Anpaßoperation wird gemäß folgender Gleichungen ausgeführt:

$$O_j = (j \cdot D_j + (L-j) \cdot Z_j) / L \quad (j = 0 \text{ bis } L-1)$$

$$O_j = D_j \quad (j = L \text{ bis } L+M-1)$$

[0183] Hier bedeutet O_j Ausgangssignale des aktuellen Rahmens und des Kopfabschnitts des nachfolgenden Rahmens, D_j bedeutet die Ausgangssignale zweiter Ordnung im aktuellen Rahmen und den Kopfabschnitt im nachfolgenden Rahmen, Z_j bedeutet den Kopfabschnitt des aktuellen Rahmens von Ausgangssignalen, die in der Speichereinheit **85** für vorhergelesene Wellenformen gespeichert ist, L bedeutet die Anzahl von Signalen, die es im Kopfabschnitt des aktuellen (oder nachfolgenden) Rahmens gibt, und M bedeutet die Anzahl von Signalen, die sich im aktuellen Rahmen befinden. Auch der Kopfabschnitt des aktuellen Rahmens von Ausgangssignalen Z_j , gespeichert in der Speichereinheit **85** für vorausgelesene Wellenformen, wird ersetzt mit dem Kopfabschnitt des nachfolgenden Rahmens von Ausgangssignalen O_{M+j} , simultan übertragen mit dem aktuellen Rahmen von Ausgangssignalen in der Anpaßeinheit **86**, folgendermaßen:

$$Z_j = O_{M+j} \quad (j = 0 \text{ bis } L-1)$$

[0184] Durch die Ausgangssignale O_j ($j = 0$ bis $L+M-1$) werden in diesem Falle im aktuellen Rahmen der Kopfabschnitt des nachfolgenden Rahmens abgegeben, die Ausgangssignale O_j ($j = 0$ bis $M-1$) im aktuellen Rahmen werden verwendet für eine folgende Vorrichtung (nicht dargestellt). Die Ausgangssignale O_j ($j = M$ bis $L+M-1$) im Kopfabschnitt des nachfolgenden Rahmens werden im Gegensatz dazu aufgegeben, weil andere Eingangssignale O_j ($j = 0$ bis $M-1$) im nachfolgenden Rahmen in einem folgenden Schritt abgegeben werden. Da die Kontinuität der Ausgangssignale O_j ($j = 0$ bis $L+M-1$) im Kopfabschnitt des nachfolgenden Rahmens und die Ausgangssignale O_j ($j = 0$ bis $M-1$) im aktuellen Rahmen jedoch beibehalten werden, können die Ausgangssignale O_j ($j = M$ bis $L+M-1$) für eine Frequenzanalyse verwendet werden, wie für eine lineare Prädiktionanalyse oder wie für eine Filteranalyse.

[0185] Als nächstes beschrieben ist eine Schätzung der Betriebsfähigkeit der Vorrichtung **71** zur Geräuschverminderung.

[0186] Um die Leistungsfähigkeit der Vorrichtung zur Geräuschverminderung **71** abzuschätzen, wird eine Vielzahl von Sprachsignalen S_{in} , die eine Sprache und ein Geräusch enthalten, in die Vorrichtung **71** zur Geräuschverminderung eingegeben, um die Geräusche zu vermindern. Danach wird eine Vielzahl von Ausgangssignalen S_{out} , die die Vorrichtung **71** erzeugt hat, eingegeben in eine Vorrichtung zur Sprachcodierung, in der ein CELP-Codierverfahren ausgeführt wird, und es wird eine Vielzahl codierter Signale gewonnen. Die codierten Signale Danach werden die Signale zur Wiedergabe der Sprache decodiert. Wie zum Stand der Technik anhand **Fig. 4** beschrieben, besteht der Nachteil bei der herkömmlichen Vorrichtung zur Sprachcodierung oder bei einem herkömmlichen Sprachcodec darin, daß in der Sprache vorhandene Geräusche als Allophonie wiedergegeben werden, wenn das Geräusch codiert und decodiert wird, und es ist schwierig, eine synthetische Sprache zu erkennen. Eine große Anzahl von Männern und Frauen hören folglich die synthetische Sprache, und eine subjektive Schätzung wird durchgeführt.

[0187] Eine geräuschverminderte Synthesprache wird in einem Geräuschverminderungssystem **91** erzeugt, wie es in **Fig. 18** dargestellt ist.

[0188] Eine Vielzahl von Sprachsignalen S_{in} , die Geräusche enthalten, werden in die Vorrichtung **71** zur Rauschverminderung eingegeben, und das Geräusch wird vermindert. Danach wird eine Vielzahl von Ausgangssignalen S_o , die die Vorrichtung **71** erzeugt, eingegeben in eine herkömmliche Vorrichtung **92** zur Sprachcodierung, um eine Vielzahl codierter Sprachsignale S_c zu erzeugen. Danach werden die codierten Sprachsignale in einer Vorrichtung **93** zur Sprachdecodierung decodiert, um die Sprache als geräuschverminderte Synthesprache S_s wiederzugeben. Um eine Bezugssynthesprache zu erzeugen, in der das Geräusch nicht vermindert ist, werden die Sprachsignale S_{in} auch direkt in die herkömmliche Vorrichtung **92** zur Sprachcodierung eingegeben, und eine Vielzahl von codierten Bezugssprachsignalen S_{rc} werden gewonnen. Danach werden die codierten Bezugssprachsignale S_{rc} in die Vorrichtung **93** zur Sprachcodierung eingegeben, und eine Bezugssynthesprache S_{rs} wird gewonnen.

[0189] Speicherkapazität eines statischen RAM in der Vorrichtung **71** hat etwa 300 Wörter, die Speicherkapazität eines dynamischen RAHM in der Vorrichtung **71** etwa 1,4 k Wörter, und ein Verarbeitungsumfang in der Vorrichtung **71** beträgt etwa 3 MOPS. Ein japanischer Kurzsatz wird gelesen von 2 Männern und 2 Frauen während 2 oder 3 Sekunden, um eine Vielzahl geräuschfreier Signale zu erzeugen. Auch werden die Sprachsignale, die Geräusche enthalten erzeugt durch Hinzufügen eines Straßengeräuschs den geräuschfreien Sprachsignalen mit einem Störabstand von 20 dB. Die geräuschfreie Synthesprache wird gewonnen durch Eingabe der geräuschfreien Sprachsignale in das System **91**.

[0190] Die geräuschfreie Synthesprache, die Bezugssynthesprache, in der das Geräusch nicht vermindert ist, und die geräuschverminderte Synthesprache werden von Zuhörern abgeschätzt, die sich zusammensetzen aus 16 Männern und Frauen. Alle Hörer geben eine bis fünf Arten von Schätzwerten für jede geräuschfreie Synthesprache ab, für die Bezugssynthesprache und für die geräuschverminderte Synthesprache. In Fällen, bei denen der Hörer fühlt, daß eine Synthesprache sehr schlecht ist, gibt der Hörer einen Schätzwert von 1 ab. In Fällen, bei denen der Hörer fühlt, daß die Synthesprache schlecht ist, gibt der Hörer einen Schätzwert von 2 ab. In Fällen, bei denen der Hörer fühlt, daß die Synthesprache moderat ist, gibt der Hörer einen Schätzwert von 3 ab. In Fällen, bei denen der Hörer fühlt, daß die Sprache gut ist, gibt der Hörer einen Schätzwert von 4 ab. In Fällen, in denen ein Hörer fühlt, daß eine Synthesprache sehr gut ist, gibt der Hörer einen Schätzwert von 5 ab. Danach wird ein Mittelmeinungswert errechnet durch Durchschnittsbildung der Schätzwerte, die die Hörer abgegeben haben für eine jede Synthesprache. Der mittlere Meinungswert ist in Tabelle 2 dargestellt.

Tabelle 2

stimmlose Synthesprache	Bezugs- Synthesprache	geräuschverminderte Synthesprache
3,50	3,13	3,50

[0191] Wie in Tabelle 2 gezeigt, ist der mittlere Meinungswert für die geräuschfreie Synthesprache sehr hoch. Ein Geräusch, das in der Synthesprache enthalten ist, verstärkt sich, ein Gefühl der Synthesprache wird schlecht, weil Allophonie verursacht wird, weil durch das Geräusch Allophonie erzeugt wird. In Fällen, bei denen das Geräusch in den Sprachsignalen enthalten ist und nicht vermindert ist von der Geräuschverminderungsvorrichtung **71**, wird der mittlere Meinungswert beträchtlich herabgesetzt auf 3,13. Das heißt, eine Wiedergabequalität für die Sprache ist beträchtlich verschlechtert. In Fällen, bei denen im Gegensatz dazu das Geräusch, das in den Sprachsignalen enthalten ist, vermindert wurde in der Geräuschverminderungsvorrichtung **71**, wird der mittlere Meinungswert nicht herabgesetzt. Das heißt, eine Wiedergabequalität für die Sprache wird nicht verschlechtert.

[0192] Weil das Eingabespektrum mit einem Minimalwert unter Werten von Eingabespektren bei einem aktuellen Rahmen und weiter vergangenen Rahmen festgelegt ist als Rauschspektrum im aktuellen Rahmen für jeden Frequenzwert und das Geräuschspektrum angenommen wird gemäß der obigen Definition in der Annahmemeinheit **76**, ist folglich ungeachtet der Tatsache, ob Geräusch im aktuellen Rahmen von Sprachsignalen enthalten ist, ein Geräuschspektrum in verlässlicher Weise in der Annahmemeinheit **76** angenommen. Das heißt,

obwohl es schwierig ist, einen Rahmen der eingegebenen Sprachsignale zu spezifizieren, in denen Sprache vorhanden ist, dann ein Geräusch, das in der Sprache vorhanden ist oder mit einem sprachfreien Signal vorhanden ist, in zuverlässiger Weise festgestellt werden, und das Geräusch läßt sich vermindern auf einen moderaten Grad in der Geräuschverminderungseinheit **80** unter Verwendung des Stabilisierungskoeffizienten und des Geräuschverminderungskoeffizienten.

[0193] Merkmale eingegebener Spektrumschleife können betont werden in der Betonungseinheit **84** unter Verwendung auch des linearen Prädiktionskoeffizienten. Obwohl eine Geräuschstärke hoch ist, kann folglich die Verschlechterung der wiedergegebenen Sprachqualität vermieden werden.

[0194] Da die für das Erzielen der Amplitudenspektren erforderliche Rechenleistung nicht ausgeführt wird und die Pseudoamplitudenspektren errechnet werden in der Fourier-Transformationseinheit **75**, kann das Errechnen der gewonnenen Quadratwurzel eines quadrierten Absolutwerts entfallen, und ein Rechenumfang, der erforderlich ist zum Erzielen des Eingabespektrums, kann beträchtlich in der Fourier-Transformationseinheit **75** vermindert werden.

[0195] Weil ein Geräuschspektrum für einen j -ten Frequenzwert aus den vier Eingabespektren für den j -ten, $(j+1)$ -ten, $(j+2)$ -ten und den $(j+3)$ -ten Frequenzwert in der Annahmeeinheit **77** angenommen wird, kann das Geräuschspektrum für vier Frequenzwerte erzielt werden. Die Speicherkapazität eines RAM, die erforderlich ist zum Speichern der Geräuschspektren in der Geräuschspektrumspeichereinheit **77**, kann folglich vermindert werden. Da auch jedes der Geräuschspektren nicht aus dem Eingabespektrum für einen Frequenzwert angenommen wird, läßt es sich vermeiden, daß sich ein stationärer Klang für lange Zeit fortsetzt und fälschlicherweise als Geräusch angenommen wird.

[0196] Weil ein Geräuschspektrum, das in der Speichereinheit **77** gespeichert ist, als erststelliger Kandidat aufgegeben wird, wenn eine Speicherperiode des ersten Geräuschspektrums einen Überschuß einer Bezugsfortsetzungsperiode hat und ein zweites Geräuschspektrum, das in der Speichereinheit **77** gespeichert ist als zweitplazierter Kandidat, als erstplazierter Kandidat eingesetzt wird, obwohl ein Geräusch sich allmählich ändert, kann die Geräuschannahme, ausgeführt für die Änderung des Geräuschs, erfolgen.

[0197] Weil der Stabilisierungskoeffizient C eingestellt ist auf einen niedrigen Wert, wenn ein Geräusch in stabiler Weise vorhanden ist in den eingegebenen Sprachsignalen S_{in} , wird auch der Geräuschverminderungskoeffizient Q auf einen hohen Wert gesetzt. Obwohl die Stärke eines Geräuschs, das in der Geräuschannahmeeinheit **76** angenommen wird, niedriger ist als die aktuelle Geräuschintensität, kann folglich die Geräuschintensität als erhöht angenommen werden durch Multiplizieren des Geräuschspektrums mit dem Geräuschverminderungskoeffizienten. Das in den eingegebenen Sprachsignalen S_{in} vorhandene Geräusch kann folglich reduziert werden auf einen moderaten Grad in der Geräuschverminderungseinheit **80**.

[0198] Da auch der Stabilisierungskoeffizient C auf einen hohen Wert gesetzt ist, wenn ein Geräusch nicht in den eingegebenen Sprachsignalen S_{in} vorhanden ist, wird der Geräuschverminderungskoeffizient Q auf Null gesetzt. Weil ein Wert des Geräuschspektrums folglich nicht von der Geräuschverminderungseinheit **80** vermindert ist, kann eine Spektralverzerrung in einem Sprachsignal auf Grund einer Geräuschverminderung vermieden werden.

[0199] In Fällen, bei denen ein Wert des geräuschverminderten Eingangsspektrums in einem aktuellen Rahmen für einen speziellen Frequenzwert negativ ist, wird auch das geräuschverminderte Eingangsspektrum ersetzt durch ein Eingangsspektrum mit einem niedrigeren Wert zwischen einem aktuellen Eingangsspektrum und einem vorangehenden Eingangsspektrum für den speziellen Frequenzwert, und das geräuschverminderte Eingangsspektrum wird kompensiert für eine überschießende Verminderung des Geräuschs in der Spektralkompensationseinheit **81**. Ein Allophoniegefühl für den speziellen Frequenzwert kann folglich vermindert werden, und die Verschlechterung einer wiedergegebenen Sprachqualität wird vermieden.

[0200] Weil die Ausgabesignale erster Ordnung, die die inverse Fourier-Transformationseinheit **83** erzeugt hat, auch in einem stark betonenden Filter betont werden, kann ein hochfrequenzbetonendes Filter und ein leistungsbetonendes Filter eine Spektrumverzerrung auf Grund der Geräuschverminderung korrigieren, eine wiedergegebene Tonqualität wird verbessert. Obwohl der Grad der Geräuschverminderung hoch ist, weil die Geräuschintensität hoch ist, kann die Verschlechterung der wiedergegebenen Tonqualität vermieden werden.

[0201] Weil jeder Rahmen abgegebener Signale an die benachbarten Rahmen der Ausgangssignale in der Wellenformanpaßeinheit **86** angepaßt ist, kann die wiedergegebene Tonqualität verbessert werden.

[0202] Nachdem die Prinzipien der vorliegenden Erfindung in bevorzugten Ausführungsbeispiele dieser veranschaulicht und beschrieben worden sind, sollte es von jedem Fachmann anerkannt werden, daß die Erfindung in der Anordnung und in Einzelheiten abgewandelt werden kann, ohne daß vom erfinderischen Prinzip abgewichen wird. Wir beanspruchen alle Abwandlungen, die in den Umfang der anliegenden Patentansprüche fallen.

[0203] Eine Sprachanalysiereinheit analysiert Abtastsprache zur Erzielung von Abtasteigenschaftsparametern, und aus den Abtasteigenschaftsparametern eines jeden der Vielzahl von Codiermodulen wird eine Codierverzerrung errechnet. Die Abtasteigenschaftsparameter und die Codierverzerrungen werden statistisch von einer Statistikverarbeitungseinheit verarbeitet, um eine Codiermodulauswahlregel zu gewinnen. Wenn danach eine Sprache von der Sprachanalysiereinheit zur Gewinnung der charakteristischen Parameter analysiert ist, wird ein geeignetes Codiermodul von einer Codiermodulauswahleinheit aus den Codiermodulen gemäß der Codiermodulauswahlregel unter der Bedingung ausgewählt, daß ein Codierverzerrung für die Eigenschaftsparametern im geeigneten Codiermodul minimiert ist. Danach werden die Eigenschaftsparameter der Sprache im passenden Codiermodul codiert, und eine codierte Sprache wird gewonnen. Ist die Codiersprache decodiert, wird eine Wiedergabesprache gewonnen. Weil ein geeignetes Codiermodul aus einer Vielzahl von Codiermodulen gemäß der Codiermodulauswahlregel leicht auswählbar ist, kann das Auftreten von Allophonie in der wiedergegebenen Sprache bei geringem Rechenaufwand vermieden werden.

Patentansprüche

1. Vorrichtung zur Sprachcodierung, mit:

einem Codiermodulspeicher (17) zum Speichern einer Vielzahl von Codiermodulen;
 einem Sprachanalysiermittel (22) zum Analysieren eines Sprachsignals, um eine Vielzahl charakteristischer Parameter zu erzielen, die Spracheigenschaften des Sprachsignals aufzeigen;
 einem Auswahlregelspeichermittel (16), in dem eine im voraus durch statistische Verarbeitung erzielte Codiermodulauswahlregel der Beziehung zwischen den Eigenschaftsparametern und einem bestgeeigneten Codiermodul gespeichert ist;
 einem Codiermodulauswahlmittel (18) zur Auswahl eines der im Codiermodulspeichermittel (17) gespeicherten Codiermodule als geeigneter Codiermodul, in dem eine vom Sprachanalysiermittel erzielte Codierverzerrung für die Eigenschaftsparameter (22) gemäß der Codiermodulauswahlregel minimiert ist; und mit
 einem Codiermittel (19) zum Codieren des durch das Sprachanalysiermittel (22) im vom Codiermodulauswahlmittel (18) ausgewählten geeigneten Codiermodul erzielten Sprachsignals.

2. Vorrichtung zur Sprachcodierung nach Anspruch 1, die die Codiermodulauswahlregel im voraus durch statistische Verarbeitung der Beziehung zwischen Eigenschaftsparametern und einem bestgeeigneten Codiermodul im voraus gewinnt, erzielt als Ergebnis des Codierens eines jeden der Anzahl eingegebenen Sprachklänge unter Verwendung aller im Codiermodulspeichermittel (17) gespeicherten Codiermodule, wobei die Eigenschaftsparameter Spracheigenschaften aufzeigen, bei denen eine Klanganalyse in Verbindung mit der Anzahl eingegebener Sprachklänge die Eigenschaftsparameter erzielt und ein Statistikverarbeitungsmittel (37) die statistische Verarbeitung durchführt.

3. Vorrichtung nach Anspruch 2, bei der das statistische Verarbeitungsmittel (37) ein neuronales Netzwerk besitzt, das ausgestattet ist mit:
 einer Vielzahl von eingegebenen Neuronen, die in einer Eingangsschicht zum Aufnehmen der Eigenschaftsparameter vorgesehen sind;
 einer Vielzahl von Zwischenneuronen, die in einer Zwischenschicht zur Wichtung der von den Eingangsneuronen mit einer Vielzahl von Wichtungsfaktoren empfangenen Eigenschaftsparameter vorgesehen sind; und mit
 einer Vielzahl von Ausgangsneuronen, die in einer Ausgabeschicht zum Empfangen einer Vielzahl von Lehrsignalen vorgesehen sind, um Werte der Wichtungsfaktoren unter der Bedingung anzunähern, daß ein Ausgangsneuron gemäß einem Codiermodul, in dem eine Codierverzerrung mit einem aus den Eigenschaftsparametern in den Codiermodulen durch das Sprachanalysiermittel (22) errechneten Minimalwert unter Werten von Codierverzerrungen aus den Parametern durch ein erregtes Sprachanalysiermittel (22) errechnet wird, wobei die angenäherten Wichtungsfaktoren als Codiermodulauswahlregel behandelt werden und das Codiermodulauswahlmittel (18) das neuronale Netzwerk hat, in dem die angenäherten Faktoren eingestellt sind zum Erregen eines der Ausgangsneuronen gemäß dem geeigneten Codiermodul in Fällen, bei denen die Eingabe der Eigenschaftsparameter in die Eingabeneuronen erfolgt.

4. Vorrichtung nach Anspruch 2 oder 3, die des weiteren ausgestattet ist mit

einem Datenverbesserungsmittel (**36**) zum Verbessern der Eigenschaftsparameter der aus den Eigenschaftsparametern in den Codiermodulen errechneten Codiervverzerrungen, wobei das statistische Verarbeitungsmittel (**37**) die Eigenschaftsparameter der angenäherten Codiervverzerrungen statistisch verarbeitet.

5. Vorrichtung nach einem der vorstehenden Ansprüche, in der jeder der im Codiermodulspeichermittel (**17**) gespeicherten Codiermodule für einen Ruheabschnitt der Sprache, eine Affrizierung eines sprachlosen Konsonanten der Sprache, einen Reibelaut des unbetonten Konsonanten der Sprache, einen betonten Konsonanten der Sprache, einen vokalanhaltenden Abschnitt eines Vokals der Sprache oder für einen Vokalsübergangsabschnitt eines Vokals der Sprache geeignet ist.

Es folgen 9 Blatt Zeichnungen

FIG. 1
(STAND DER TECHNIK)

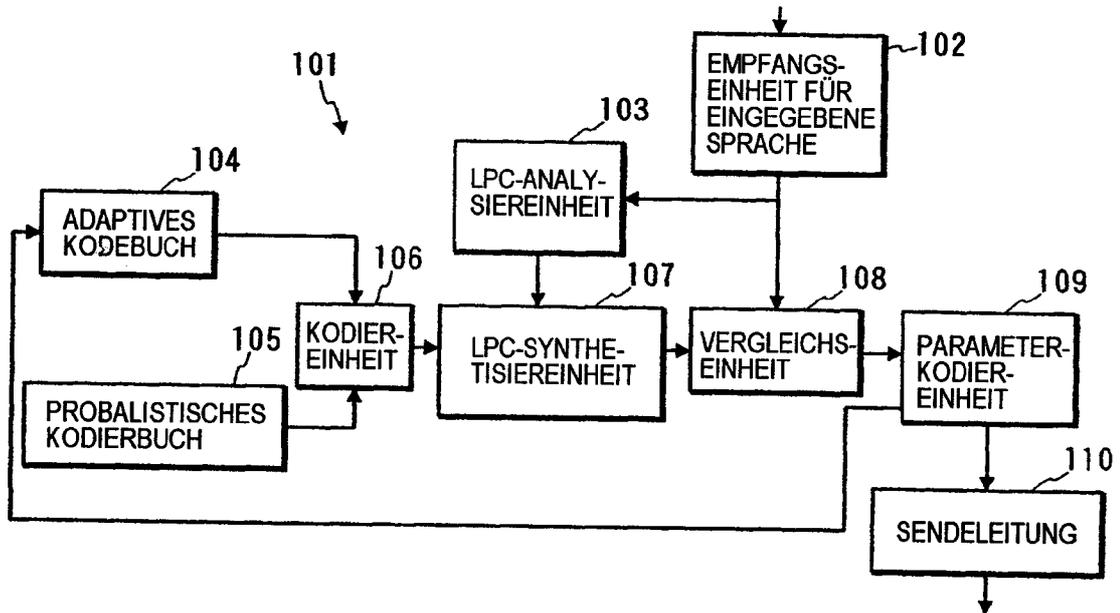


FIG. 2
(STAND DER TECHNIK)

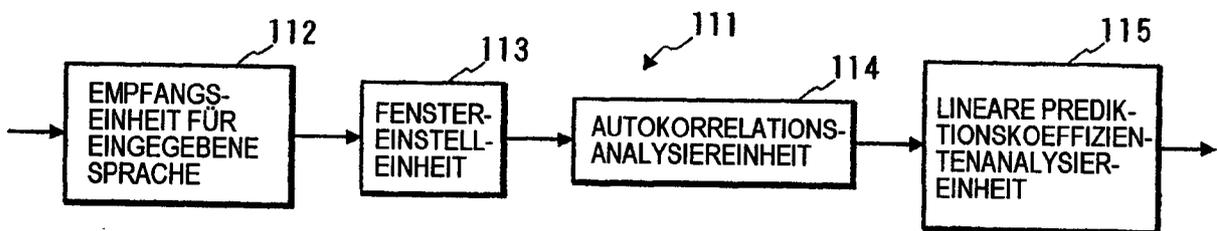


FIG. 3
(STAND DER TECHNIK)

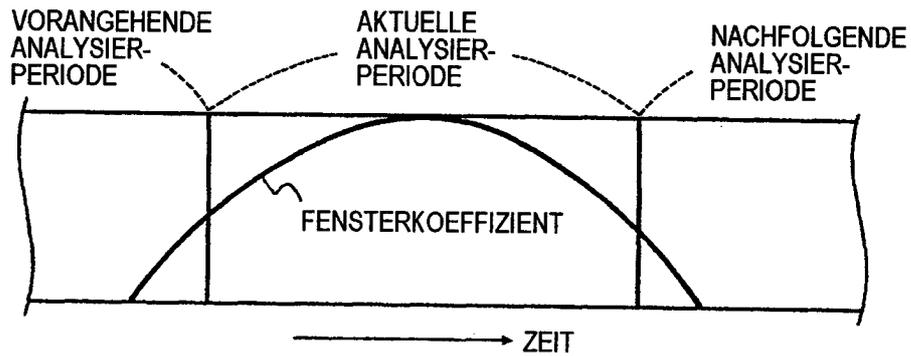


FIG. 4
(STAND DER TECHNIK)

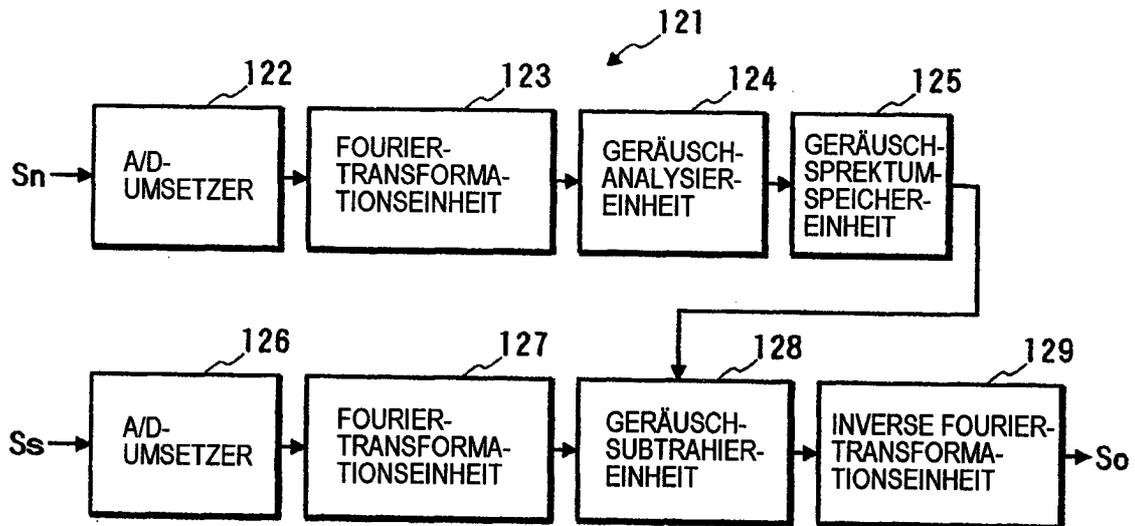


FIG. 5

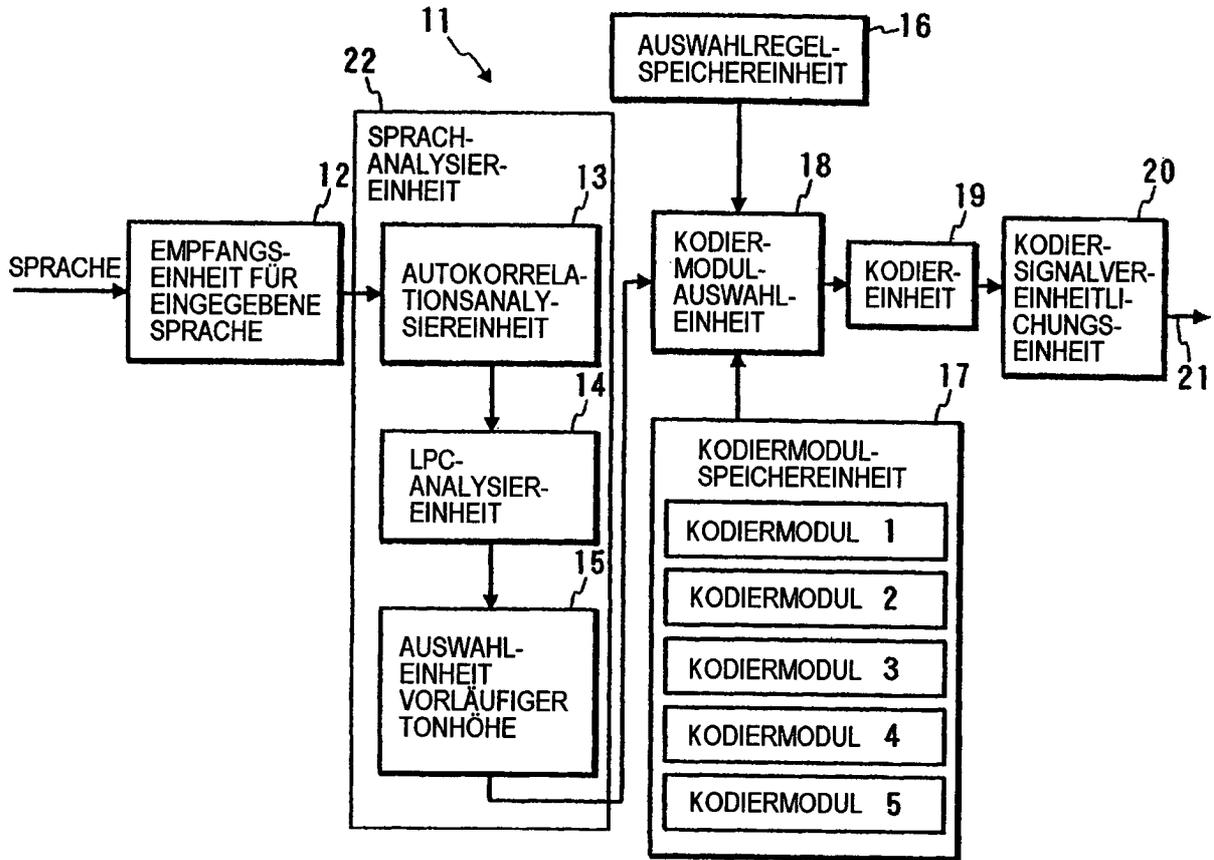


FIG. 6

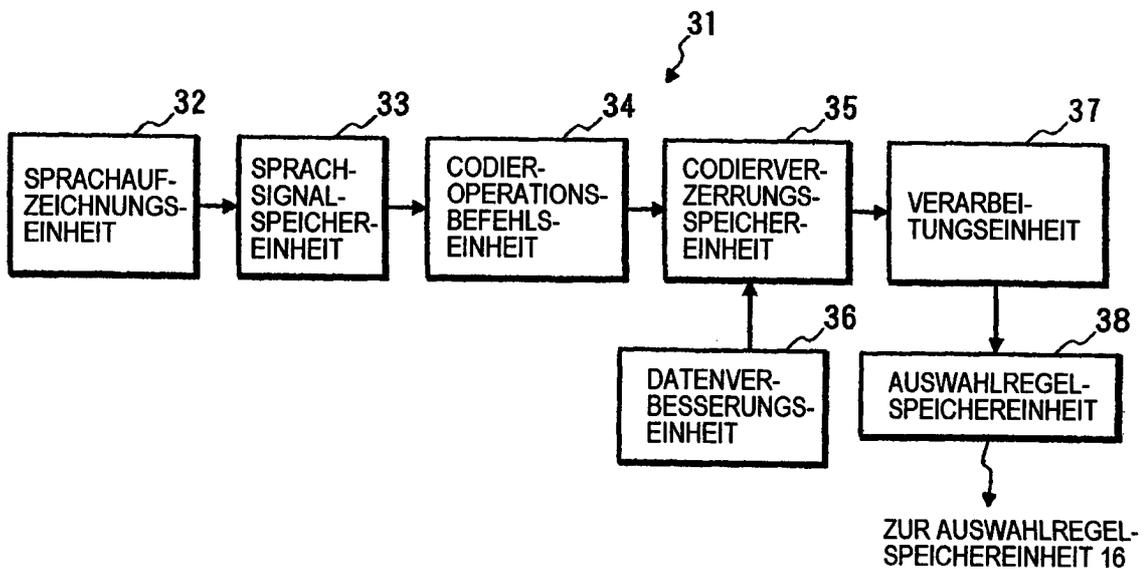


FIG. 7

NR. DER ANALYSIER-PERIODEN	STIMM-LEISTUNG	AUTO-KORRELATIONS-KOEFFIZIENTEN		STIMM-LEISTUNGS-VERHÄLT-NIS	LANG-FRISTIGE PRÄDIKTIONS-KOEFFIZIENTEN	IMP-PULS-WERT	NORMIERUNGS-LEISTUNG	KANDIDAT FÜR TON-HÖHE	AUFEICHNUNGSVERZERRUNG AUFGEZEIGT VOM STÖRABSTAND					NR. PASSEN-DEN KODIER-MODULS	
		PRIMÄR	SEKUN-DÄR						MODUL 1	MODUL 2	MODUL 3	MODUL 4	MODUL 5		
0	2.173	0.713	0.479	0.141	0.148	0.011	0.841	23	0.000	0.000	0.000	0.000	0.000	0.000	5
1	0.061	0.270	0.051	0.174	0.185	0.193	0.688	27	0.000	0.000	0.000	0.000	0.000	0.000	5
2	0.029	-0.074	-0.058	0.130	0.180	0.028	0.571	39	0.000	0.000	0.000	0.000	0.000	0.000	5
3	0.027	-0.003	-0.139	0.152	0.156	0.041	0.463	41	0.000	0.000	0.000	0.000	0.000	0.000	5
4	0.031	-0.247	-0.351	0.150	0.148	0.012	0.749	23	0.000	0.000	0.000	0.000	0.000	0.000	5
5	0.016	-0.241	0.000	0.160	0.070	0.000	0.890	38	0.000	0.000	0.000	0.000	0.000	0.000	5
6	0.029	0.117	-0.119	0.177	0.180	0.042	0.702	62	0.000	0.000	0.000	0.000	0.000	0.000	5
7	0.014	0.003	-0.077	0.247	0.117	0.104	0.770	59	0.000	0.000	0.000	0.000	0.000	0.000	5
8	49016.524	0.478	-0.426	0.151	0.188	0.111	0.788	16	3.495	4.520	0.526	1.721	0.883	0.000	2
9	67035.218	0.436	-0.409	0.165	0.336	0.065	0.167	18	5.250	5.534	0.673	1.995	1.260	0.000	2
10	162991.576	0.451	-0.383	0.256	0.305	0.014	0.289	24	4.881	5.259	1.359	2.521	0.826	0.000	2
11	893765.121	0.792	0.360	0.258	0.312	0.017	0.364	43	5.346	5.090	1.149	2.241	0.491	0.000	1
12	8988145.136	0.771	0.285	0.572	0.312	0.322	0.513	40	9.872	7.311	5.222	2.689	1.064	0.000	1
13	18178059.652	0.758	0.274	0.719	0.266	0.688	0.485	38	9.998	6.700	7.734	3.707	1.270	0.000	1
14	3185444.959	0.785	0.417	0.430	0.242	0.649	0.608	38	7.263	6.091	3.679	4.610	1.146	0.000	1
15	8174221.206	0.805	0.509	0.678	0.281	0.351	0.471	35	10.474	7.131	8.680	3.620	0.817	0.000	1
16	13366907.599	0.811	0.538	0.830	0.188	0.926	0.687	35	13.547	8.983	9.925	4.070	1.805	0.000	1
17	12520034.594	0.831	0.571	0.679	0.164	0.629	0.771	36	14.910	9.173	13.446	4.179	1.483	0.000	1
18	11098012.995	0.893	0.665	0.476	0.203	0.513	0.643	38	13.117	10.677	10.229	4.265	1.506	0.000	1
19	5502417.846	0.913	0.710	0.246	0.297	0.017	0.382	84	12.619	10.637	7.996	4.008	2.152	0.000	1
20	3131338.762	0.892	0.642	0.433	0.219	0.302	0.653	46	12.566	8.789	9.577	5.135	1.717	0.000	1
21	6968244.433	0.873	0.572	0.383	0.266	0.425	0.652	49	10.147	8.251	6.362	3.786	1.702	0.000	1
22	8643054.702	0.853	0.509	0.468	0.234	0.517	0.520	52	9.557	8.661	7.044	3.511	0.891	0.000	1

FIG. 8

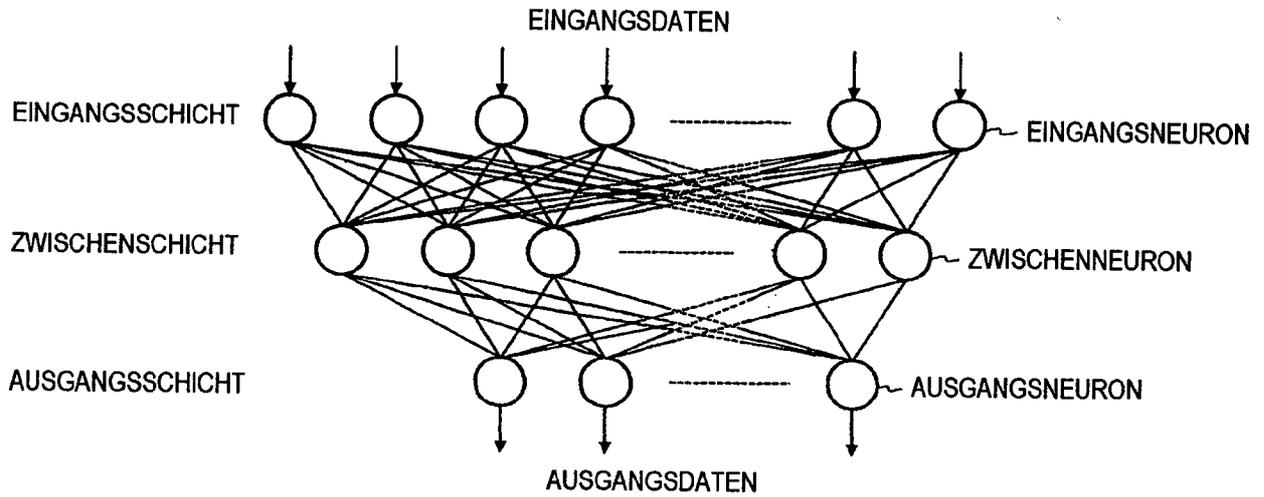


FIG. 9

KODIERMODUL	M1	M2	M3	M4	M5
BIT/RATE	8 kbps		4 kbps		2 kbps
LPC	33		12	7	
STIMMLEISTUNG	7		7		7
UNTERRAHMEN	29 × 4	30 × 4	30 × 2	16 × 4	13 × 2
ADAPTIVES KODIERBUCH	8		9		
PROBABILISTISCHES KODIERBUCH	14	24	14	12	9
VERSTÄRKUNG	7	6	7	4	4

FIG. 10

AUSZUWÄHLENDES KODIERMODUL	AKTUELL AUSGEWÄHLTES KODIERMODUL					ERKENNUNGS- KOEFFIZIENT (%)
	KODIER- MODUL M1	KODIER- MODUL M2	KODIER- MODUL M3	KODIER- MODUL M4	KODIER- MODUL M5	
KODIERMODUL M1	2337	143	157	0	0	88.6
KODIERMODUL M2	160	1436	1	90	2	85.0
KODIERMODUL M3	179	0	285	1	0	61.2
KODIERMODUL M4	0	35	0	464	5	92.0
KODIERMODUL M5	0	2	0	20	1959	98.8

EINHEIT: ANZAHL ABTASTWERTE

FIG. 11

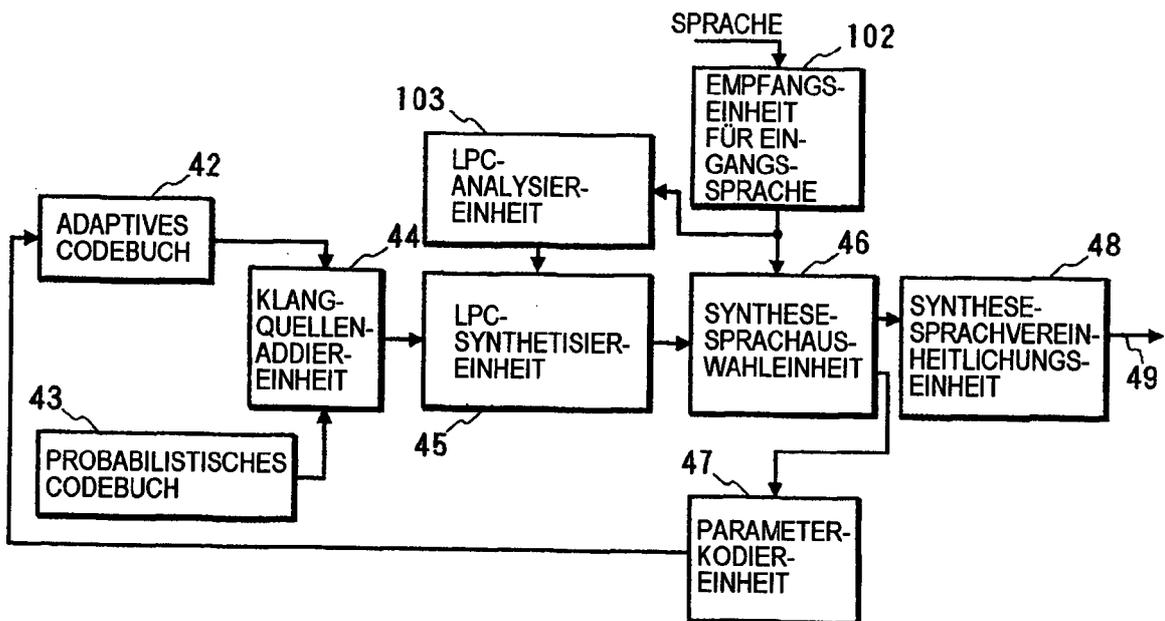


FIG. 12

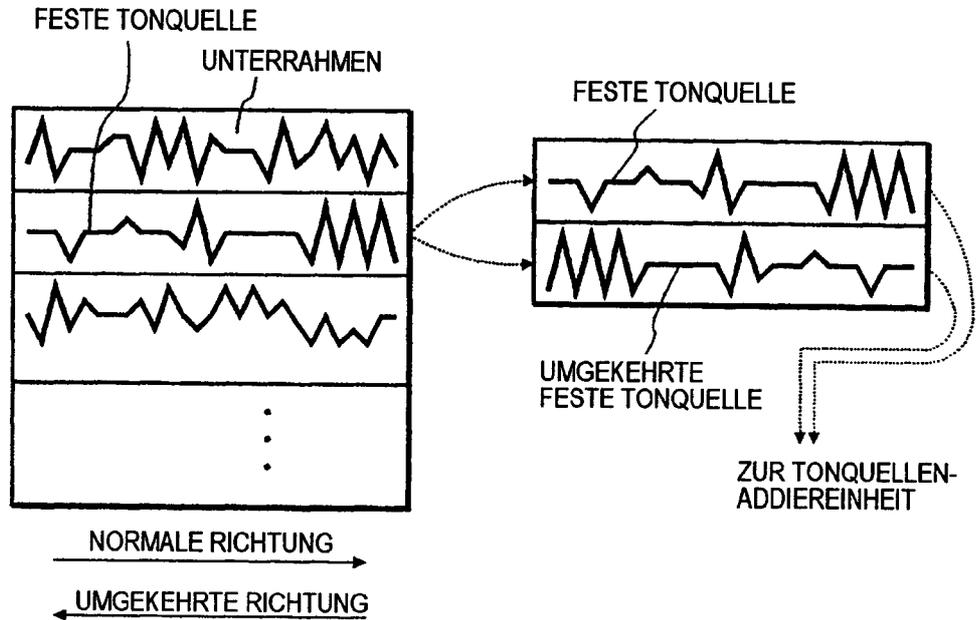


FIG. 13

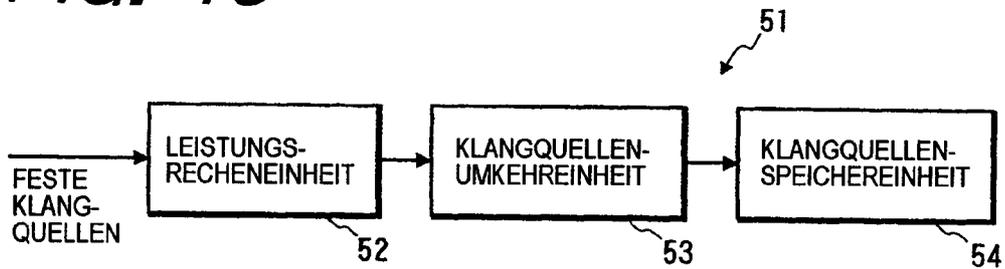


FIG. 14

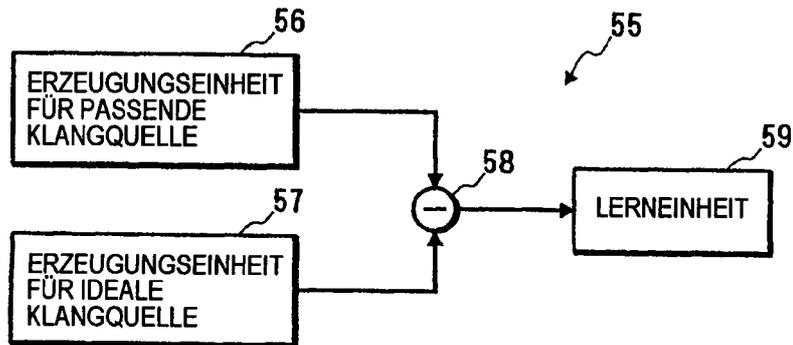


FIG. 15

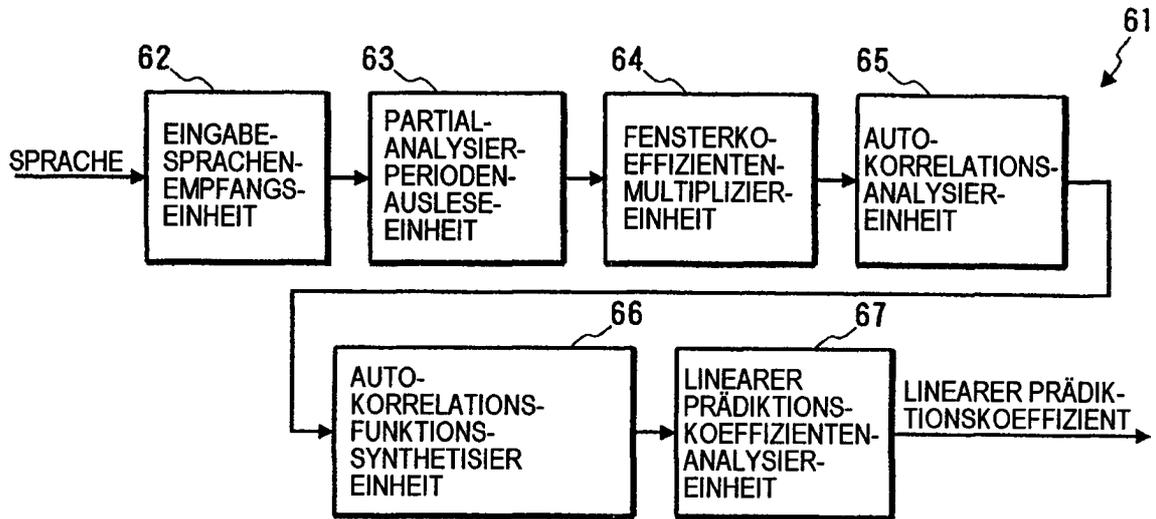


FIG. 16

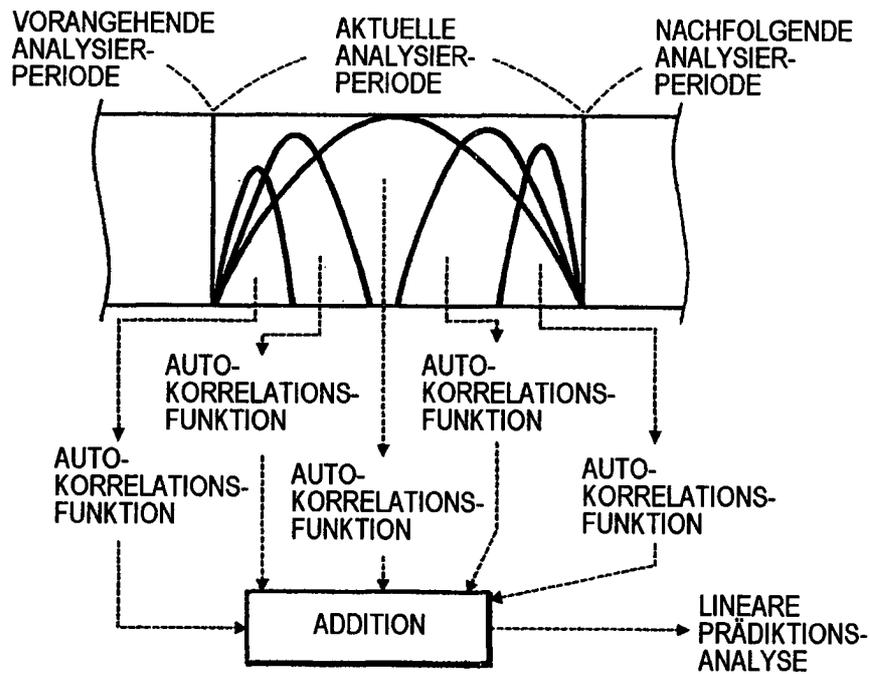


FIG. 17

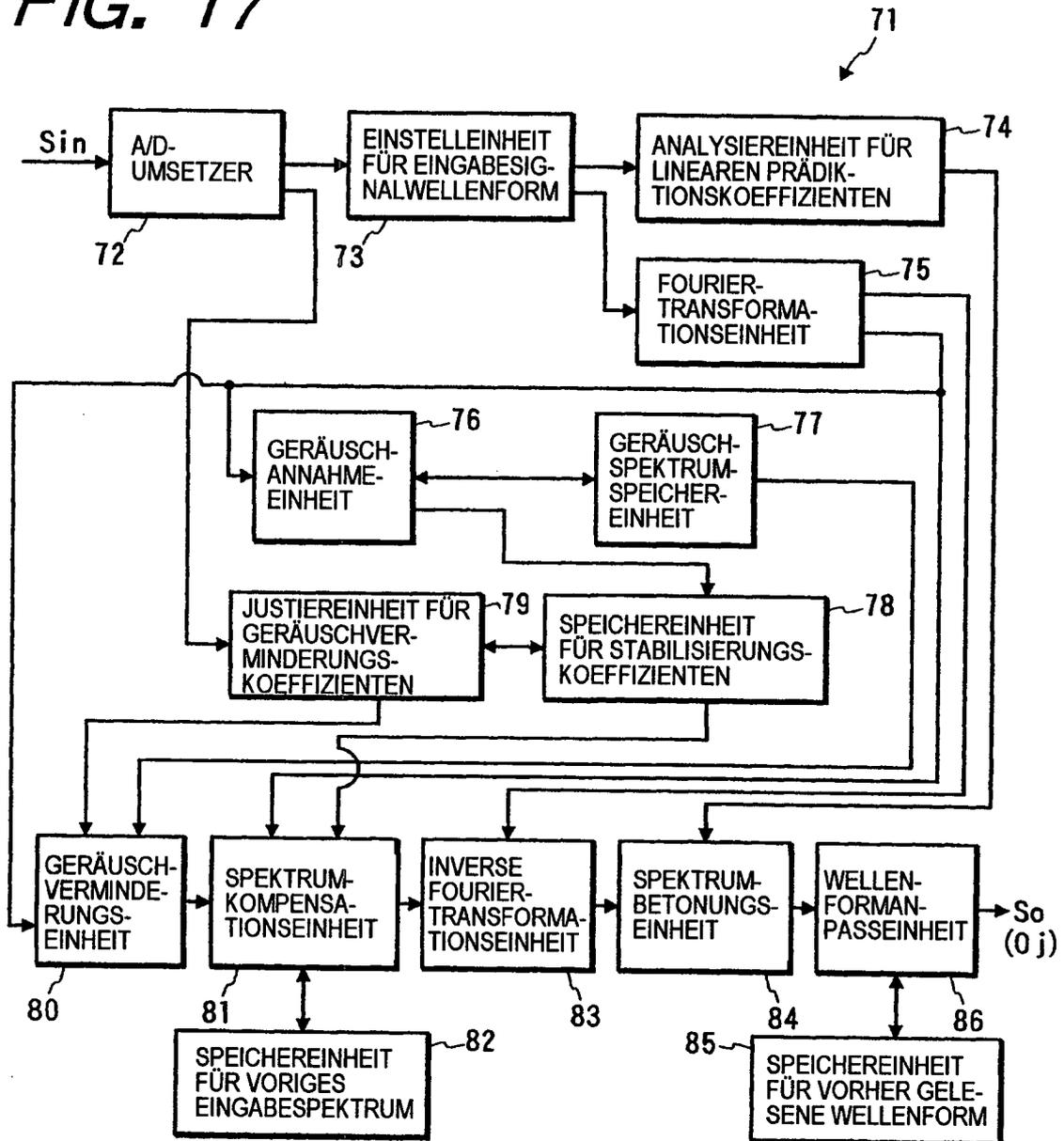


FIG. 18

