



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 699 10 239 T2 2004.06.24**

(12)

## Übersetzung der europäischen Patentschrift

(97) **EP 1 125 276 B1**

(51) Int Cl.<sup>7</sup>: **G10L 11/04**

(21) Deutsches Aktenzeichen: **699 10 239.1**

(86) PCT-Aktenzeichen: **PCT/CA99/01008**

(96) Europäisches Aktenzeichen: **99 952 199.0**

(87) PCT-Veröffentlichungs-Nr.: **WO 00/25298**

(86) PCT-Anmeldetag: **27.10.1999**

(87) Veröffentlichungstag  
der PCT-Anmeldung: **04.05.2000**

(97) Erstveröffentlichung durch das EPA: **22.08.2001**

(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **06.08.2003**

(47) Veröffentlichungstag im Patentblatt: **24.06.2004**

(30) Unionspriorität:  
**2252170 27.10.1998 CA**

(84) Benannte Vertragsstaaten:  
**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,  
LI, LU, MC, NL, PT, SE**

(73) Patentinhaber:  
**Voiceage Corp., Ville Mont-Royal, Quebec, CA**

(72) Erfinder:  
**BESSETTE, Bruno, Rock Forest, CA; SALAMI,  
Redwan, Sherbrooke, CA; LEFEBVRE, Roch,  
Canton de Magog, CA**

(74) Vertreter:  
**Betten & Resch, 80333 München**

(54) Bezeichnung: **VERFAHREN UND VORRICHTUNG ZUR ADAPTIVEN BANDBREITENABHÄNGIGEN GRUNDFRE-  
QUENZSUCHE FÜR DIE KODIERUNG BREITBANDIGER SIGNALE**

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****HINTERGRUND DER ERFINDUNG**

## 1. Gebiet der Erfindung:

[0001] Die vorliegende Erfindung bezieht sich auf eine effiziente Technik zum digitalen Codieren eines Breitbandsignals, insbesondere aber nicht ausschließlich eines Sprachsignals, im Hinblick auf das Übertragen oder das Speichern und das Synthetisieren dieses Breitband-Tonsignals. Spezifischer beschäftigt sich die Erfindung mit einer verbesserten Tonhöhenuchvorrichtung und einem verbesserten Tonhöhenuchverfahren.

## 2. Kurzbeschreibung des Standes der Technik:

[0002] Der Bedarf an effizienten digitalen Breitband-Sprach/Audio-Codierungstechniken mit einem guten Kompromiss bezüglich subjektiver Qualität/Bitrate nimmt für zahlreiche Anwendungen zu, wie z. B. sowohl Audio/Video-Telekonferenzen, Multimedia und drahtlose Anwendungen als auch Internet- und Paketnetz-Anwendungen. Bis vor kurzem wurden in Sprachcodierungsanwendungen hauptsächlich Telephonbandbreiten, die im Bereich von 200–3400 Hz gefiltert wurden, verwendet. Es gibt jedoch einen zunehmenden Bedarf an Breitband-sprachanwendungen, um die Verständlichkeit und Natürlichkeit der Sprachsignale zu vergrößern. Es wurde festgestellt, dass eine Bandbreite im Bereich von 50–7000 Hz für die Lieferung einer persönlichen Sprachqualität ausreichend ist. Für Audiosignale ergibt dieser Bereich eine annehmbare Audioqualität, aber immer noch niedriger als die CD-Qualität, die im Bereich 20–20000 Hz arbeitet.

[0003] Ein Sprachcodierer setzt ein Sprachsignal in einen digitalen Bitstrom um, der über einen Kommunikationskanal übertragen wird (oder in einem Speichermedium gespeichert wird). Das Sprachsignal wird digitalisiert (abgetastet und normalerweise mit 16 Bits pro Abtastwert quantisiert), wobei der Sprachcodierer die Aufgabe hat, diese digitalen Abtastwerte mit einer kleineren Anzahl von Bits darzustellen, während er eine gute subjektive Sprachqualität aufrechterhält. Der Sprachdecodierer oder Synthetisieren wirkt auf den übertragenen oder gespeicherten Bitstrom und setzt ihn wieder in ein Tonsignal um.

[0004] Eine der besten Techniken des Standes der Technik, die einen guten Qualität/Bitrate-Kompromiss erreichen kann, ist die so genannte codeerregte lineare Vorhersagetechnik (CELP-Technik). Als ein Beispiel offenbart EP-A-0 421 444 einen CELP-gestützten Codieren. Entsprechend dieser Technik wird das abgetastete Sprachsignal in aufeinander folgenden Blöcken aus L Abtastwerten verarbeitet, die normalerweise als Rahmen bezeichnet werden, wobei L irgendeine vorgegebene Zahl ist (die 10–30 ms der Sprache entspricht). In der CELP wird ein lineares Vorhersagesynthesefilter (LP-Synthesefilter) für jeden Rahmen berechnet und übertragen. Der Rahmen aus L Abtastwerten wird dann in kleinere Blöcke unterteilt, die als Unterrahmen mit der Größe von N Abtastwerten bezeichnet werden, wobei  $L = kN$  gilt, wobei k die Anzahl der Unterrahmen in einem Rahmen ist (N entspricht normalerweise 4–10 ms der Sprache). Ein Erregungssignal wird in jedem Unterrahmen bestimmt, das normalerweise aus zwei Komponenten besteht: eine aus der früheren Erregung (die außerdem als Tonhöhenbeitrag oder adaptives Codebuch oder Tonhöhen-Codebuch bezeichnet wird) und die andere von einem innovativen Codebuch (das außerdem als festes Codebuch bezeichnet wird). Dieses Erregungssignal wird übertragen und am Decodieren als die Eingabe des LP-Synthesefilters verwendet, um die synthetisierte Sprache zu erhalten.

[0005] Ein innovatives Codebuch im CELP-Kontext ist eine indexierte Menge aus Sequenzen, die N Abtastwerte lang sind, die als N-dimensionale Codevektoren bezeichnet werden. Jede Codebuch-Sequenz wird durch eine ganze Zahl k indexiert, die von 1 bis M variiert, wobei M die Größe des Codebuchs darstellt, die oft als eine Anzahl von Bits b dargestellt wird, wobei  $M = 2^b$  gilt.

[0006] Um die Sprache entsprechend der CELP-Technik zu synthetisieren, wird jeder Block aus N Abtastwerten synthetisiert, indem ein geeigneter Codevektor aus dem Codebuch durch zeitveränderliche Filter gefiltert wird, die die spektralen Eigenschaften des Sprachsignals modellieren. Auf der Codierseite wird die Syntheseausgabe für alle Codevektoren aus dem Codebuch oder eine Teilmenge der Codevektoren aus dem Codebuch berechnet (Codebuch-Suche). Der gehaltene Codevektor ist der eine, der die Syntheseausgabe erzeugt, die am genauesten am ursprünglichen Sprachsignal entsprechend einem wahrnehmungsmäßig gewichteten Verzerrungsmaß liegt. Diese wahrnehmungsmäßige Gewichtung wird unter Verwendung eines so genannten Wahrnehmungsgewichtungsfilters ausgeführt, das normalerweise aus dem LP-Synthesefilter abgeleitet wird.

[0007] Das CELP-Modell ist beim Codieren von Telephonband-Tonsignalen sehr erfolgreich gewesen, wobei mehrere CELP-gestützte Standards in einem weiten Bereich von Anwendungen, insbesondere in digitalen Zellenanwendungen, vorhanden sind. Im Telephonband wird das Tonsignal auf 200–3400 Hz bandbegrenzt und mit 8000 Abtastwerten/s abgetastet. In Breitband-Sprach/Audio-Anwendungen wird das Tonsignal auf 50–7000 Hz bandbegrenzt und mit 16000 Abtastwerten/s abgetastet.

[0008] Es ergeben sich einige Schwierigkeiten, wenn das für das Telephonband optimierte CELP-Modell auf

Breitbandssignale angewendet wird, wobei zusätzliche Merkmale zu dem Modell hinzugefügt werden müssen, um Breitbandssignale in hoher Qualität zu erhalten. Breitbandssignale zeigen im Vergleich zu Telefonbandsignalen einen viel breiteren Dynamikbereich, der zu Genauigkeitsproblemen führt, wenn eine Festkomma-Implementierung des Algorithmus erforderlich ist (die in drahtlosen Anwendungen wesentlich ist). Ferner verbraucht das CELP-Modell oft die meisten seiner Codierungsbits im Niederfrequenzbereich, der normalerweise höhere Energieinhalte besitzt, dies führt zu einem Tiefpass-Ausgangssignal. Um dieses Problem zu überwinden, ist das Wahrnehmungsgewichtungsfilter zu modifizieren, damit es sich für Breitbandssignale eignet, wobei Vorverzerrungstechniken, die die Hochfrequenzbereiche verstärken, wichtig werden, um den Dynamikbereich zu verringern, dies ergibt eine einfachere Festkomma-Implementierung, und um eine bessere Codierung der Inhalte mit höheren Frequenzen des Signals zu sichern. Ferner erstrecken sich die Tonhöheninhalte im Spektrum stimmhafter Segmente in den Breitbandsignalen nicht über den ganzen Bereich des Spektrums, wobei der Betrag der Stimmhaftigkeit im Vergleich zum Schmalbandsignalen mehr Abweichung zeigt. Deshalb sind im Fall von Breitbandsignalen die vorhandenen Tonhöhenstruktturen nicht sehr effizient. Folglich ist es wichtig, die Tonhöhenanalyse mit geschlossener Schleife zu verbessern, um die Abweichungen im Stimmpegel besser unterzubringen.

#### DIE AUFGABEN DER ERFINDUNG

[0009] Es ist deshalb eine Aufgabe der vorliegenden Erfindung, ein Verfahren und eine Vorrichtung zum effizienten Codieren von Breitbandtonsignalen (7000 Hz) unter Verwendung von CELP-Codierungstechniken zu schaffen, die eine verbesserte Tonhöheninhalte verwenden, um rekonstruierte Tonsignale in hoher Qualität zu erhalten.

#### ZUSAMMENFASSUNG DER ERFINDUNG

[0010] Spezifischer wird gemäß der vorliegenden Erfindung, wie sie in den Ansprüchen 1–63 beansprucht ist, ein Verfahren zum Auswählen einer optimalen Menge von Tonhöhencodebuch-Parametern, die einem Signalpfad, der den niedrigsten berechneten Tonhöhenvorhersagefehler besitzt, zugeordnet sind, aus wenigstens zwei Signalpfaden geschaffen. Der Tonhöhenvorhersagefehler wird in Reaktion auf einen Tonhöhencodevektor von einer Tonhöhencodebuch-Suchvorrichtung berechnet. In wenigstens einem der zwei Signalpfade wird der Tonhöhenvorhersagefehler gefiltert, bevor der Tonhöhencodevektor für die Berechnung des Tonhöhenvorhersagefehlers des einen Pfades geliefert wird. Schließlich werden die in den wenigstens zwei Signalpfaden berechneten Tonhöhenvorhersagefehler verglichen, der Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler wird gewählt und die Menge der Codebuchparameter, die dem gewählten Signalpfad zugeordnet ist, wird ausgewählt.

[0011] Die Tonhöhenanalysevorrichtung der Erfindung zum Erzeugen einer optimalen Menge von Tonhöhencodebuch-Parametern umfasst:

- a) wenigstens zwei Signalpfade, die jeweiligen Mengen von Tonhöhencodebuch-Parametern zugeordnet sind, wobei:
  - i) jeder Signalpfad eine Tonhöhenvorhersagefehler-Berechnungsvorrichtung zum Berechnen eines Tonhöhenvorhersagefehlers eines Tonhöhencodevektors von einer Tonhöhencodebuch-Suchvorrichtung umfasst; und
  - ii) wenigstens einer der beiden Pfade ein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung des Pfades umfasst; und
- b) eine Auswahleinrichtung zum Vergleichen der in den Signalpfaden berechneten Tonhöhenvorhersagefehler, um den Signalpfad auszuwählen, der den niedrigsten berechneten Tonhöhenvorhersagefehler hat, und um die Menge von Tonhöhencodebuch-Parametern, die dem ausgewählten Signalpfad zugeordnet sind, zu wählen.

[0012] Das neue Verfahren und die neue Vorrichtung, die die effiziente Modellierung der harmonischen Struktur des Sprachspektrums ausführen, verwenden mehrere Formen von Tiefpassfiltern, die auf die frühere Erregung angewendet werden, wobei diejenige, die die höhere Vorhersageverstärkung liefert, ausgewählt wird. Wenn die Unterabtast-Tonhöhenauflösung verwendet wird, können die Tiefpassfilter in die Interpolationsfilter aufgenommen sein, die verwendet werden, um die höhere Tonhöhenauflösung zu erhalten.

[0013] In einer bevorzugten Ausführungsform der Erfindung umfasst jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung der obenbeschriebenen Tonhöhenanalysevorrichtung:

- a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit einem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;
- b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Re-

- aktion auf den gefalteten Tönhöhencodevektor und einen Tönhöhen-Suchzielvektor;
- c) einen Verstärker zum Multiplizieren des gefalteten Tönhöhencodevektors mit der Tönhöhenverstärkung, um dadurch einen verstärkten gefalteten Tönhöhencodevektor zu erzeugen; und
- d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tönhöhencodevektors mit dem Tönhöhen-Suchzielvektor, um dadurch den Tönhöhenvorhersagefehler zu erzeugen.

[0014] In einer weiteren bevorzugten Ausführungsform der Erfindung umfasst die Tönhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tönhöhenverstärkung  $b^{(i)}$  unter Verwendung der folgenden Beziehung:

$$b^{(i)} = x^t y^{(i)} / \|y^{(i)}\|^2$$

wobei  $0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tönhöhen-Suchzielvektor ist und  $y^{(i)}$  der gefaltete Tönhöhencodevektor ist.

[0015] Die vorliegende Erfindung bezieht sich ferner auf einen Codieren, der die obenbeschriebene Tönhöhenanalysevorrichtung besitzt, um ein Breitband-Eingangssignal zu codieren, wobei der Codieren umfasst:

- a) eine Berechnungseinrichtung für ein lineares Vorhersagesynthesefilter, das auf das Breitbandsignal anspricht, um Koeffizienten für ein lineares Vorhersagesynthesefilter zu erzeugen;
- b) ein Wahrnehmungsgewichtungsfiler, das auf das Breitbandsignal und auf die Koeffizienten des linearen Vorhersagesynthesefilters anspricht, um ein durch Wahrnehmung gewichtetes Signal zu erzeugen;
- c) einen Impulsantwortgenerator, der auf die Koeffizienten für das lineare Vorhersagesynthesefilter anspricht, um ein gewichtetes Synthesefilter-Impulsantwortsignal zu erzeugen;
- d) eine Tönhöhen-Sucheinheit zum Erzeugen von Tönhöhencodebuch-Parametern, wobei die Tönhöhen-Sucheinheit umfasst:
  - i) die Tönhöhencodebuch-Suchvorrichtung, die auf das durch Wahrnehmung gewichtete Signal und auf die Koeffizienten für das lineare Vorhersagesynthesefilter anspricht, um den Tönhöhencodevektor und einen innovativen Suchzielvektor zu erzeugen; und
  - ii) die Tönhöhenanalysevorrichtung, die auf den Tönhöhencodevektor anspricht, um aus den Mengen von Tönhöhencodebuch-Parametern diejenige Menge von Tönhöhencodebuch-Parametern auszuwählen, die dem Pfad mit dem niedrigsten berechneten Tönhöhenvorhersagefehler zugeordnet ist;
- d) eine Vorrichtung zum Suchen eines innovativen Codebuchs, die auf das gewichtete Synthesefilter-Impulsantwortsignal und auf den innovativen Suchzielvektor anspricht, um innovative Codebuchparameter zu erzeugen, und
- e) eine Signalformungsvorrichtung zum Erzeugen eines codierten Breitbandsignals, das die Menge von Tönhöhencodebuch-Parametern, die dem Pfad mit dem niedrigsten Tönhöhenvorhersagefehler zugeordnet sind, die innovativen Codebuchparameter und die Koeffizienten für das lineare Vorhersagesynthesefilter umfasst.

[0016] Die vorliegende Erfindung bezieht sich noch weiter auf ein Zellenkommunikationssystem, eine mobile Zellen-Sender/Empfänger-Einheit, ein Zellennetzelement und ein bidirektionales drahtloses Kommunikationssystem, das den obenbeschriebenen Decodieren umfasst.

[0017] Die Aufgaben, Vorteile und anderen Merkmale der vorliegenden Erfindung werden durch das Lesen der folgenden nicht einschränkenden Beschreibung einer ihrer bevorzugten Ausführungsformen offensichtlicher, die lediglich beispielhaft unter Bezugnahme auf die beigefügte Zeichnung gegeben wird.

#### KURZBESCHREIBUNG DER ZEICHNUNG

[0018] In der beigefügten Zeichnung ist:

[0019] **Fig. 1** ein schematischer Blockschaltplan einer bevorzugten Ausführungsform der Breitband-Codierungsvorrichtung;

[0020] **Fig. 2** ein schematischer Blockschaltplan einer bevorzugten Ausführungsform der Breitband-Decodierungsvorrichtung;

[0021] **Fig. 3** ein schematischer Blockschaltplan einer bevorzugten Ausführungsform der Tönhöhenanalysevorrichtung; und

[0022] **Fig. 4** ein vereinfachter schematischer Blockschaltplan eines Zellenkommunikationssystems, in dem die Breitband-Codierungsvorrichtung nach **Fig. 1** und die Breitband-Decodierungsvorrichtung nach **Fig. 2** verwendet werden können.

## AUSFÜHRLICHE BESCHREIBUNG DER BEVORZUGTEN AUSFÜHRUNGSFORM

[0023] Wie den Durchschnittsfachleuten auf dem Gebiet wohl bekannt ist, schafft ein Zellenkommunikationssystem, wie z. B. **401** (siehe **Fig. 4**), einen Telekommunikationsdienst über einen großen geographischen Bereich, indem dieser große geographische Bereich in eine Anzahl C von kleineren Zellen unterteilt wird. Diese C kleineren Zellen können durch entsprechende Zellenbasisstationen **402<sub>1</sub>**, **402<sub>2</sub>**, ..., **402<sub>C</sub>** bedient werden, um jede Zelle mit den Funkzeichengabe-, Audio- und Datenkanälen zu versehen.

[0024] Die Funkzeichengabe-Kanäle werden verwendet, um mobile Funktelefone (mobile Sender/Empfänger-Einheiten), wie z. B. **403**, innerhalb der Grenzen des Versorgungsbereichs (der Zelle) der Zellenbasisstationen **402** zu rufen und um Anrufe zu anderen Funktelefonen **403**, die sich entweder innerhalb oder außerhalb der Zelle der Basisstation befinden, oder zu einem anderen Netz, wie z. B. dem öffentlichen Fernsprechnetz (PSTN) **404**, einzuleiten.

[0025] Sobald ein Funktelefon **403** einen Anruf erfolgreich eingeleitet oder empfangen hat, wird ein Audio- oder Datenkanal zwischen diesen Funktelefon **403** und der Zellenbasisstationen **402**, die der Zelle entspricht, in der sich das Funktelefon **403** befindet, hergestellt, wobei die Kommunikation zwischen der Basisstation **402** und dem Funktelefon **403** über diesen Audio- oder Datenkanal geleitet wird. Das Funktelefon **403** kann außerdem Steuer- oder Synchronisierungsinformationen über einen Zeichengabekanal empfangen, während ein Anruf im Gange ist.

[0026] Falls ein Funktelefon **403** eine Zelle verlässt und in eine weitere benachbarte Zelle eintritt, während ein Anruf im Gange ist, reicht das Funktelefon **403** den Anruf an einen verfügbaren Audio- oder Datenkanal der neuen Zellenbasisstationen **402** weiter. Falls ein Funktelefon **403** eine Zelle verlässt und in eine weitere benachbarte Zelle eintritt, während kein Anruf im Gange ist, sendet das Funktelefon **403** eine Steuernachricht über den Zeichengabekanal, um sich in der Basisstation **402** der neuen Zelle einzubuchen. In dieser Weise ist eine Mobilkommunikation über einen weiten geographischen Bereich möglich.

[0027] Das Zellenkommunikationssystem **401** umfasst ferner ein Steuerendgerät **405**, um die Kommunikation zwischen den Zellenbasisstationen **402** und dem PSTN **404** zu steuern, z. B. während einer Kommunikation zwischen einem Funktelefon **403** und dem PSTN **404** oder zwischen einem Funktelefon **403**, das sich in einer ersten Zelle befindet, und einem Funktelefon **403**, das sich in einer zweiten Zelle befindet.

[0028] Selbstverständlich ist ein bidirektionales drahtloses Funkkommunikations-Untersystem erforderlich, um einen Audio- oder Datenkanal zwischen einer Basisstation **402** einer Zelle und einem Funktelefon **403**, das sich in dieser Zelle befindet, herzustellen. Wie in **Fig. 4** in sehr vereinfachter Form veranschaulicht ist, umfasst ein derartiges bidirektionales drahtloses Funkkommunikations-Untersystem im Funktelefon **403** typischerweise:

- einen Sender **406**, der enthält:
- einen Codieren **407**, der das Sprachsignal codiert; und
- eine Sendeschaltung **408**, die das codierte Sprachsignal vom Codieren **407** durch eine Antenne, wie z. B. **409**, sendet; und
- einen Empfänger **410**, der enthält:
- eine Empfangsschaltung **411**, die ein gesendetes codiertes Sprachsignal normalerweise durch die gleiche Antenne **409** empfängt; und
- einen Decodieren **412**, der das empfangene codierte Sprachsignal von der Empfangsschaltung **411** decodiert.

[0029] Das Funktelefon umfasst ferner andere herkömmliche Funktelefonschaltungen **413**, mit denen der Codieren **407** und der Decodieren **412** verbunden sind und die die Signale von diesen verarbeiten, wobei diese Schaltungen **413** den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind, wobei sie demzufolge in der vorliegenden Beschreibung nicht weiter beschrieben sind.

[0030] Außerdem umfasst ein derartiges bidirektionales drahtloses Funkkommunikations-Untersystem in der Basisstation **402** typischerweise:

- einen Sender **414**, der enthält:
- einen Codieren **415**, der das Sprachsignal codiert; und
- eine Sendeschaltung **416**, die das codierte Sprachsignal vom Codieren **415** durch eine Antenne, wie z. B. **417**, sendet; und
- einen Empfänger **418**, der enthält:
- eine Empfangsschaltung **419**, die ein gesendetes codiertes Sprachsignal durch die gleiche Antenne **417** oder durch eine weitere (nicht gezeigte) Antenne empfängt; und
- einen Decodieren **420**, der das empfangene codierte Sprachsignal von der Empfangsschaltung **419** decodiert.

[0031] Die Basisstation **402** umfasst ferner typischerweise eine Basisstations-Stuereinrichtung **421** zusam-

men mit ihrer zugeordneten Datenbank **422**, die die Kommunikation zwischen dem Steuerendgerät **405** und dem Sender **414** und dem Empfänger **418** steuert.

[0032] Wie den Durchschnittsfachleuten auf dem Gebiet wohl bekannt ist, ist die Sprachcodierung erforderlich, um die Bandbreite zu reduzieren, die notwendig ist, um das Tonsignal, z. B. das Sprachsignal, wie z. B. die Sprache, über das bidirektionale drahtlose Funkkommunikations-Untersystem, d. h. zwischen einem Funktelefon **403** und einer Basisstation **402**, zu übertragen.

[0033] Die LP-Sprachcodierer (wie z. B. **415** und **407**) arbeiten typischerweise mit 13 kbit/s und darunter, wobei z. B. codeerregte lineare Vorhersagecodierer (CELP-Codierer) typischerweise ein LP-Synthesefilter verwenden, um die kurzfristige spektrale Enveloppe des Sprachsignals zu modellieren. Die LP-Informationen werden typischerweise alle 10 oder 20 ms zum Decodieren (wie z. B. **420** und **412**) gesendet, wobei sie auf der Decodierseite extrahiert werden.

[0034] Die in der vorliegenden Beschreibung offenbarten neuartigen Techniken können für verschiedene LP-gestützte Codierungssysteme gelten. In der bevorzugten Ausführungsform wird jedoch ein CELP-Codierungssystem für den Zweck verwendet, eine nichteinschränkende Veranschaulichung dieser Techniken darzustellen. In der gleichen Weise können derartige Techniken sowohl mit Tonsignalen, die anders als Stimme und Sprache sind, als auch mit anderen Typen von Breitbandsignalen verwendet werden.

[0035] **Fig. 1** zeigt einen allgemeinen Blockschaltplan einer CELP-Sprachcodierungsvorrichtung **100**, die modifiziert worden ist, um sie besser an Breitbandsignale anzupassen.

[0036] Das abgetastete Eingangssprachsignal **114** wird in aufeinander folgende Blöcke aus L Abtastwerten, die als "Rahmen" bezeichnet werden, unterteilt. In jedem Rahmen werden verschiedene Parameter, die das Sprachsignal in dem Rahmen darstellen, berechnet, codiert und übertragen. Die LP-Parameter, die das LP-Synthesefilter darstellen, werden normalerweise einmal für jeden Rahmen berechnet. Der Rahmen wird weiterhin kleinere Blöcke aus N Abtastwerten unterteilt (Blöcke der Länge N), in denen die Erregungsparameter (den Tonhöhe und die Innovation) bestimmt werden. In der CELP-Literatur werden diese Blöcke der Länge N als "Unterrahmen" bezeichnet, wobei die Signale aus N Abtastwerten in den Unterrahmen als N-dimensionale Vektoren bezeichnet werden. In dieser bevorzugten Ausführungsform entspricht die Länge N 5 ms, während die Länge L 20 ms entspricht, was bedeutet, dass ein Rahmen vier Unterrahmen enthält (N = 80 bei der Abtastrate von 16 kHz und 64 nach der Unterabtastung auf 12,8 kHz). In der Codierungsprozedur treten mehrere N-dimensionale Vektoren auf. Sowohl eine Liste der Vektoren, die in den **Fig. 1** und **2** auftreten, als auch eine Liste der übertragenen Parameter sind im Folgenden angegeben:

#### Liste der wichtigsten N-dimensionalen Vektoren

s Breitbandsignal-Eingangssprachvektor (nach der Unterabtastung, Vorverarbeitung und Vorverzerrung);  
 $s_w$  gewichteter Sprachvektor;  
 $s_0$  Nulleingangsverhalten des gewichteten Synthesefilters;  
 $s_p$  unterabgetastetes vorverarbeitetes Signal; überabgetastetes synthetisiertes Sprachsignal;  
 $s'$  Synthesesignal vor der Rückentzerrung;  
 $s_d$  rückentzerrtes Synthesesignal;  
 $s_h$  Synthesesignal nach der Rückentzerrung und der Nachverarbeitung;  
 $x$  Zielvektor für die Tonhöhensuche;  
 $x'$  Zielvektor für die Innovationssuche;  
 $h$  gewichtete Synthesefilter-Impulsantwort;  
 $v_T$  adaptiver Codebuchvektor (Tonhöhen-Codebuchvektor) bei der Verzögerung T  
 $y_T$  gefilterter Tonhöhen-Codebuchvektor ( $v_T$ , gefaltet mit  $h$ );  
 $c_k$  innovativer Codevektor beim Index k (k-ter Eintrag vom Innovations-Codebuch);  
 $c_f$  verbesserter skaliertes Innovations-Codevektor;  
 $u$  Erregungssignal (skaliertes Innovations- und Tonhöhen-Codevektoren);  
 $u'$  verbesserte Erregung;  
 $z$  Bandpass-Rauschsequenz;  
 $w'$  weiße Rauschsequenz; und  
 $w$  skalierte Rauschsequenz.

#### Liste der übertragenen Parameter

STP kurzfristige Vorhersageparameter (definieren  $A(z)$ );  
 T Tonhöhenachteiligung (oder Tonhöhen-Codebuchindex);  
 b Tonhöhenverstärkung (oder Tonhöhen-Codebuchverstärkung);  
 j Index des mit dem Tonhöhen-Codevektor verwendeten Tiefpassfilters;  
 k Codevektor-Index (Eintrag im Innovations-Codebuch); und

g Innovations-Codebuchverstärkung.

[0037] In dieser bevorzugten Ausführungsform werden die STP-Parameter einmal pro Rahmen übertragen, während der Rest der Parameter vielmals pro Rahmen (jeden Unterrahmen) übertragen wird.

#### DIE CODIERERSEITE

[0038] Das abgetastete Sprachsignal wird durch die Codierungsvorrichtung **100** nach **Fig. 1**, die in elf von **101** bis **111** nummerierte Module aufgespalten ist, blockweise codiert.

[0039] Die Eingangssprache wird in den oben erwähnten Blöcken aus  $L$  Abtastwerten, die als Rahmen bezeichnet werden, verarbeitet.

[0040] In **Fig. 1** wird das abgetastete Eingangssprachsignal **114** in einem Unterabtastmodul **101** unterabtastet. Das Signal wird z. B. von 16 kHz auf 12,8 kHz unter Verwendung von den Durchschnittsfachleuten auf dem Gebiet wohl bekannten Techniken unterabtastet. Die Unterabtastung herab zu einer anderen Frequenz ist selbstverständlich vorstellbar. Die Unterabtastung vergrößert den Codierungswirkungsgrad, weil eine kleinere Frequenzbandbreite codiert wird. Dies reduziert außerdem die algorithmische Komplexität, weil die Anzahl der Abtastwerte in einem Rahmen verringert wird. Die Verwendung der Unterabtastung wird signifikant, wenn die Bitrate unter 16 kbit/s reduziert wird, obwohl die Unterabtastung nicht wesentlich über 16 kbit/s liegt.

[0041] Nach der Unterabtastung wird der Rahmen aus 320 Abtastwerten von 20 ms auf einen Rahmen aus 256 Abtastwerten reduziert (Unterabtastverhältnis von 4/5).

[0042] Der Eingangsrahmen wird dann zum optionalen Vorverarbeitungsblock **102** geliefert. Der Vorverarbeitungsblock **102** kann aus einem Hochpassfilter mit einer Grenzfrequenz von 50 Hz bestehen. Das Hochpassfilter **102** beseitigt unerwünschte Tonkomponenten unter 50 Hz.

[0043] Das unterabtastete vorverarbeitete Signal wird durch  $s_p(n)$ ,  $n = 0, 1, 2, \dots, L - 1$ , bezeichnet, wobei  $L$  die Länge des Rahmens ist (256 bei einer Abtastfrequenz von 12,8 kHz). In einer bevorzugten Ausführungsform des Vorverzerrungsfilters **103** wird das Signal  $s_p(n)$  unter Verwendung eines Filters vorverzerrt, das die folgende Übertragungsfunktion besitzt:

$$P(z) = 1 - \mu z^{-1},$$

wobei  $\mu$  ein Vorverzerrungsfaktor mit einem Wert ist, der sich zwischen 0 und 1 befindet (ein typischer Wert ist  $\mu = 0,7$ ). Es könnte außerdem ein Filter höherer Ordnung verwendet werden. Es sollte darauf hingewiesen werden, dass das Hochpassfilter **102** und das Vorverzerrungsfilters **103** vertauscht werden können, um effizientere Festkomma-Implementierungen zu erhalten.

[0044] Die Funktion des Vorverzerrungsfilters **103** ist, die Hochfrequenzinhalte des Eingangssignals zu verbessern. Es reduziert außerdem den Dynamikbereich des Eingangssprachsignals, das es für die Festkomma-Implementierung geeigneter wiedergibt. Ohne die Vorverzerrung ist die Festkomma-LP-Analyse unter Verwendung einfachgenauer Arithmetik schwierig zu implementieren.

[0045] Die Vorverzerrung spielt außerdem eine wichtige Rolle beim Erreichen einer passenden wahrnehmungsmäßigen Gesamtgewichtung des Quantisierungsfehlers, was zu einer verbesserten Tonqualität beiträgt. Dies ist im Folgenden ausführlicher erklärt.

[0046] Das Ausgangssignal des Vorverzerrungsfilters **103** wird als  $s(n)$  bezeichnet. Dieses Signal wird für die Ausführung der LP-Analyse im Rechnermodul **104** verwendet. Die LP-Analyse ist eine den Durchschnittsfachleuten auf dem Gebiet wohl bekannte Technik. In dieser bevorzugten Ausführungsform wird der Autokorrelationszugang verwendet. Im Autokorrelationszugang wird das Signal  $s(n)$  zuerst unter Verwendung eines Hamming-Fensters (das normalerweise eine Länge in der Größenordnung von 30–40 ms besitzt) mit Fenstern versehen. Die Autokorrelationen werden aus dem mit Fenstern versehenen Signal berechnet, wobei die Levinson-Durbin-Rekursion verwendet wird, um die LP-Filterkoeffizienten  $a_i$  zu berechnen, wobei  $i = 1, \dots, p$  gilt, und wobei  $p$  die LP-Ordnung ist, die in der Breitbandcodierung typischerweise 16 beträgt. Die Parameter  $a_i$  sind die Koeffizienten der Übertragungsfunktion des LP-Filters, die durch die folgende Beziehung gegeben ist:

$$A(z) = 1 + \sum_{i=1}^p a_i z^{-i}.$$

[0047] Die LP-Analyse wird im Rechnermodul **104** ausgeführt, das außerdem die Quantisierung und die Interpolation der LP-Filterkoeffizienten ausführt. Die LP-Filterkoeffizienten werden zuerst in einen weiteren äquivalenten Bereich transformiert, der für die Zwecke der Quantisierung und Interpolation geeigneter ist. Der Linienspektralpaar-Bereich (LSP-Bereich) und der Immittanzspektralpaar-Bereich (ISP-Bereich) sind zwei Bereiche, in denen die Quantisierung und die Interpolation effizient ausgeführt werden können. Die 16 LP-Filterkoeffizienten  $a_i$  können in der Größenordnung von 30 bis 50 Bits unter Verwendung der Spalt- oder Mehrstu-

fen-Quantisierung oder einer Kombination daraus quantisiert werden. Der Zweck der Interpolation ist, die Aktualisierung der LP-Filterkoeffizienten für jeden Unterrahmen zu ermöglichen, während sie jedem Rahmen einmal übertragen werden, dies verbessert die Leistung des Codierers, ohne die Bitrate zu vergrößern. Es wird angenommen, dass die Quantisierung und die Interpolation der LP-Filterkoeffizienten den Durchschnittsfachleuten auf dem Gebiet anderweitig wohl bekannt sind, wobei sie demzufolge in der vorliegenden Beschreibung nicht weiter beschrieben sind.

[0048] Die folgenden Abschnitte beschreiben den Rest der auf einer Unterrahmen-Grundlage ausgeführten Codierungsoperationen. In der folgenden Beschreibung bezeichnet das Filter  $A(z)$  das nicht quantisierte interpolierte LP-Filter des Unterrahmens, während das Filter  $\hat{A}(z)$  das quantisierte interpolierte LP-Filter des Unterrahmens bezeichnet.

Die wahrnehmungsmäßige Gewichtung:

[0049] In den Analyse-während-der-Synthese-Codierern werden die optimalen Tonhöhen- und Innovationsparameter gesucht, indem der mittlere quadratische Fehler zwischen der Eingangssprache und der synthetisierten Sprache in einem wahrnehmungsmäßig gewichteten Bereich minimiert wird. Dies ist zum Minimieren des Fehlers zwischen der gewichteten Eingangssprache und der gewichteten Synthesprache äquivalent.

[0050] Das gewichtete Signal  $s_w(n)$  wird in einem Wahrnehmungsgewichtungsfiler **105** berechnet. Traditionell wird das gewichtete Signal  $s_w(n)$  durch ein Gewichtungsfiler berechnet, das eine Übertragungsfunktion  $W(z)$  in der Form besitzt:

$$W(z) = A(z/\gamma_1)/A(z/\gamma_2),$$

wobei  $0 < \gamma_2 < \gamma_1 \leq 1$  gilt.

[0051] Wie den Durchschnittsfachleuten auf dem Gebiet wohl bekannt ist, zeigt in den Analyse-während-der-Synthese-Codierern (AbS-Codierern) die Analyse, dass der Quantisierungsfehler durch eine Übertragungsfunktion  $W^{-1}(z)$  gewichtet ist, die das Inverse der Übertragungsfunktion des Wahrnehmungsgewichtungsfilters **105** ist. Dieses Ergebnis ist von B. S. Atal und M. R. Schnöder in "Predictive coding of speech and subjective error criteria", IEEE Transaction ASSP, Bd. 27, Nr. 3, S. 247–254, Juni 1979, gut beschrieben. Die Übertragungsfunktion  $W^{-1}(z)$  zeigt einiges der Formantenstruktur des Eingangssprachsignals. Folglich wird die Maskierungseigenschaft des menschlichen Ohrs ausgenutzt, indem der Quatisierungsfehler so geformt wird, dass er mehr Energie in den Formantenbereichen besitzt, in denen er durch die in diesen Bereichen vorhandene starke Signalenergie maskiert wird. Der Betrag der Gewichtung wird durch die Faktoren  $\gamma_2$  und  $\gamma_1$  gesteuert.

[0052] Das obige herkömmliche Wahrnehmungsgewichtungsfiler **105** arbeitet gut mit Telefonbandsignalen. Es ist jedoch festgestellt worden, dass dieses herkömmliche Wahrnehmungsgewichtungsfiler **105** für die effiziente wahrnehmungsmäßige Gewichtung von Breitbandsignalen nicht geeignet ist. Es ist außerdem festgestellt worden, dass das herkömmliche Wahrnehmungsgewichtungsfiler **105** inhärente Einschränkungen beim gleichzeitigen Modellieren der Formantenstruktur und der erforderlichen spektralen Neigung besitzt. Die spektrale Neigung ist in Breitbandsignalen, zurückzuführen auf den breiten Dynamikbereich zwischen den niedrigen und den hohen Frequenzen, ausgeprägter. Der Stand der Technik hat vorgeschlagen, ein Neigungsfiler in  $W(z)$  hinzuzufügen, um die Neigung und die Formantengewichtung des Breitband-Eingangssignals separat zu steuern.

[0053] Eine neuartige Lösung für dieses Problem ist, gemäß der vorliegenden Erfindung ein Vorverzerrungsfiler **103** am Eingang einzufügen, basierend auf der vorverzerrten Sprache  $s(n)$  das LP-Filter  $A(z)$  zu berechnen und ein modifiziertes Filter  $W(z)$  durch die Fixierung seines Nenners zu verwenden.

[0054] Die LP-Analyse wird im Modul **104** am vorverzerrten Signal  $s(n)$  ausgeführt, um das LP-Filter  $A(z)$  zu erhalten. Außerdem wird ein neues Wahrnehmungsgewichtungsfiler **105** mit festem Nenner verwendet. Ein Beispiel der Übertragungsfunktion für das Wahrnehmungsgewichtungsfiler **104** ist durch die folgende Beziehung gegeben:

$$W(z) = A(z/\gamma_1)/(1 - \gamma_2 z^{-1}),$$

wobei  $0 < \gamma_2 < \gamma_1 \leq 1$  gilt.

[0055] Im Nenner kann eine höhere Ordnung verwendet werden. Diese Struktur entkoppelt im Wesentlichen die Formantengewichtung von der Neigung.

[0056] Es wird angemerkt, dass, weil  $A(z)$  basierend auf dem vorverzerrten Sprachsignal  $s(n)$  berechnet wird, die Neigung des Filters  $1/A(z/\gamma_1)$  im Vergleich zu dem Fall, in dem  $A(z)$  basierend auf der ursprünglichen Sprache berechnet wird, weniger ausgeprägt ist. Weil die Rückentzerrung auf der Decodierseite unter Verwendung eines Filters ausgeführt wird, das die Übertragungsfunktion



$$P^{-1}(z) = 1(1 - \mu z^{-1})$$

besitzt, wird das Spektrum des Quantisierungsfehlers durch ein Filter geformt, das eine Übertragungsfunktion  $W = (z)P^{-1}(z)$  besitzt. Wenn  $\gamma_2$  gleich  $\mu$  gesetzt wird, was typischerweise der Fall ist, wird das Spektrum des Quantisierungsfehlers durch ein Filter geformt, dessen Übertragungsfunktion  $1/A(z/\gamma_1)$  ist, wobei  $A(z)$  basierend auf dem vorverzerrten Sprachsignal berechnet wird. Das subjektive Hören zeigte, dass diese Struktur, um die Fehlerformung einer Kombination der Vorverzerrung und der modifizierten Gewichtungsfaltung auszuführen, außer den Vorteilen der Leichtigkeit der algorithmischen Festkomma-Implementierung für die Codierung von Breitbandsignalen sehr effizient ist.

#### Die Tonhöhenanalyse:

[0057] Um die Tonhöhenanalyse zu vereinfachen, wird zuerst die Tonhöhenabweichung  $T_{OL}$  mit offener Schleife im Tonhöhenforschmodul **106** mit offener Schleife unter Verwendung des gewichteten Sprachsignals  $s_w(n)$  geschätzt. Dann wird die Tonhöhenanalyse mit geschlossener Schleife, die im Tonhöhenforschmodul **107** mit geschlossener Schleife auf einer Unterrahmen-Grundlage ausgeführt wird, um die Tonhöhenabweichung  $T_{OL}$  mit offener Schleife eingeschränkt, was die Suchkomplexität der LTP-Parameter  $T$  und  $b$  (Tonhöhenabweichung und Tonhöhenverstärkung) signifikant reduziert. Die Tonhöhenanalyse mit offener Schleife wird im Modul **106** normalerweise einmal jede 10 ms (zwei Unterrahmen) unter Verwendung von Techniken ausgeführt, die den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind.

[0058] Der Zielvektor  $x$  für die LTP-Analyse (langfristige Vorhersage-Analyse) wird zuerst berechnet. Dies wird normalerweise ausgeführt, indem das Nulleingangsverhalten  $s_0$  des gewichteten Synthesefilters  $W(z)/\hat{A}(z)$  vom gewichteten Sprachsignal  $s_w(n)$  subtrahiert wird. Dieses Nulleingangsverhalten  $s_0$  wird durch eine Nulleingangsverhalten-Berechnungseinrichtung **108** berechnet. Spezifischer wird der Zielvektor  $x$  unter Verwendung der folgenden Beziehung berechnet:

$$x = s_w - s_0,$$

wobei  $x$  der  $N$ -dimensionale Zielvektor ist,  $s_w$  der gewichtete Sprachvektor im Unterrahmen ist und  $s_0$  das Nulleingangsverhalten des Filters  $W(z)/\hat{A}(z)$  ist, das die Ausgabe des kombinierten Filters  $W(z)/\hat{A}(z)$ , zurückzuführen auf seine Anfangszustände, ist. Die Nulleingangsverhalten-Berechnungseinrichtung **108** berechnet in Reaktion auf das quantisierte interpolierte LP-Filter  $\hat{A}(z)$  aus der LP-Analyse-, Quantisierungs- und Interpolations-Berechnungseinrichtung **104** und die Anfangszustände des gewichteten Synthesefilters  $W(z)/\hat{A}(z)$ , die im Speichermodul **111** gespeichert sind, das Nulleingangsverhalten  $s_0$  (den Teil des Verhaltens, der auf die Anfangszustände zurückzuführen ist, wie sie durch das Einstellen der Eingänge gleich null bestimmt sind) des Filters  $W(z)/\hat{A}(z)$ . Diese Operation ist den Durchschnittsfachleuten auf dem Gebiet wohl bekannt, wobei sie demzufolge nicht weiter beschrieben ist.

[0059] Selbstverständlich können alternative aber mathematisch äquivalente Zugänge verwendet werden, um den Zielvektor  $x$  zu berechnen.

[0060] Ein  $N$ -dimensionaler Impulsantwort-Vektor  $h$  des gewichteten Synthesefilters  $W(z)/\hat{A}(z)$  wird im Impulsantwort-Generator **109** unter Verwendung der LP-Filterkoeffizienten  $A(z)$  und  $\hat{A}(z)$  vom Modul **104** berechnet. Abermals ist diese Operation den Durchschnittsfachleuten auf dem Gebiet wohl bekannt, wobei sie demzufolge in der vorliegenden Beschreibung nicht weiter beschrieben ist.

[0061] Die Tonhöhenparameter mit geschlossener Schleife (oder die Tonhöhencodebuch-Parameter)  $b$ ,  $T$  und  $j$  werden im Tonhöhenforschmodul **107** mit geschlossener Schleife berechnet, das den Zielvektor  $x$ , den Impulsantwort-Vektor  $h$  und die Tonhöhenabweichung  $T_{OL}$  mit offener Schleife als Eingaben verwendet. Traditionell ist die Tonhöhenvorhersage durch ein Tonhöhenfilter dargestellt worden, das die folgende Übertragungsfunktion besitzt:

$$1/(1 - bz^{-T}),$$

wobei  $b$  die Tonhöhenverstärkung ist, während  $T$  die Tonhöhenverzögerung oder -abweichung ist. In diesem Fall ist der Tonhöhenbeitrag zum Erregungssignal  $u(n)$  durch  $bu(n - T)$  gegeben, wobei die Gesamterregung durch

$$u(n) = bu(n - T) + gc_k(n)$$

gegeben ist, wobei  $g$  die innovative Codebuch-Verstärkung ist, während  $c_k(n)$  der innovative Codevektor beim Index  $k$  ist.

[0062] Diese Darstellung besitzt Einschränkungen, falls die Tonhöhenabweichung  $T$  kürzer als die Unterrahmen-Länge  $N$  ist. In einer anderen Darstellung kann der Tonhöhenbeitrag als ein Tonhöhen-Codebuch gese-

hen werden, das das frühere Erregungssignal enthält. Im Allgemeinen ist jeder Vektor im Tonhöhen-Codebuch eine um eins verschobene Version des vorausgehenden Vektors (Verwerfen eines Abtastwertes und Hinzufügen eines neuen Abtastwertes). Für Tonhöhen-nacheilungen  $T > N$  ist das Tonhöhen-Codebuch zur Filterstruktur  $1/(1 - bz^{-T})$  äquivalent, wobei ein Tonhöhen-Codebuchvektor  $v_T(n)$  bei der Tonhöhen-nacheilung  $T$  durch

$$v_T(n) = u(n - T), n = 0, \dots, N - 1,$$

gegeben ist. Für Tonhöhen-nacheilungen  $T$ , die kürzer als  $N$  sind, wird ein Vektor  $v_T(n)$  aufgebaut, indem die verfügbaren Abtastwerte von der früheren Erregung wiederholt werden, bis der Vektor vollständig ist (dies ist zur Filterstruktur nicht äquivalent).

[0063] In neueren Codierern wird eine höhere Tonhöhenauflösung verwendet, die die Qualität der stimmhaften Tonsegmente signifikant verbessert. Dies wird erreicht, indem das frühere Erregungssignal unter Verwendung von mehrphasigen Interpolationsfiltern überabgetastet wird. In diesem Fall entspricht der Vektor  $v_T(n)$  normalerweise einer interpolierten Version der früheren Erregung, wobei die Tonhöhen-nacheilung  $T$  eine nicht ganzzahlige Verzögerung ist (z. B. 50,25).

[0064] Die Tonhöhen-suche umfasst das Feststellen der besten Tonhöhen-nacheilung  $T$  und der besten Tonhöhen-verstärkung  $b$ , die den mittleren quadratischen gewichteten Fehler  $E$  zwischen dem Zielvektor  $x$  und der skalierten gefilterten früheren Erregung minimieren. Der Fehler  $E$  wird als:

$$E = \|x - by_T\|^2$$

ausgedrückt, wobei  $y_T$  der gefilterte Tonhöhen-Codebuchvektor bei der Tonhöhen-nacheilung  $T$  ist:

$$y_T(n) = v_T(n) * h(n) = \sum_{i=0}^n v_T(i)h(n-i), n = 0, \dots, N-1.$$

[0065] Es kann gezeigt werden, dass der Fehler  $E$  minimiert wird, indem das Suchkriterium maximiert wird

$$C = \frac{x^t y_T}{\sqrt{y_T^t y_T}},$$

wobei  $t$  die Vektor-Transponierung bezeichnet.

[0066] In der bevorzugten Ausführungsform der vorliegenden Erfindung wird eine 1/3-Unterabtast-Tonhöhenauflösung verwendet, wobei die Tonhöhen-suche (Tonhöhen-Codebuchsuche) aus drei Stufen aufgebaut ist.

[0067] In der ersten Stufe wird die Tonhöhen-nacheilung  $T_{OL}$  mit offener Schleife in einem Tonhöhen-suchmodul **106** mit offener Schleife in Reaktion auf das gewichtete Sprachsignal  $s_w(n)$  geschätzt. Wie in der vorausgehenden Beschreibung angegeben worden ist, wird diese Tonhöhenanalyse mit offener Schleife normalerweise einmal jede 10 ms (zwei Unterrahmen) unter Verwendung von Techniken ausgeführt, die den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind.

[0068] In der zweiten Stufe wird das Suchkriterium  $C$  im Tonhöhen-suchmodul **107** mit geschlossener Schleife für ganzzahlige Tonhöhen-nacheilungen um die geschätzte Tonhöhen-nacheilung  $T_{OL}$  mit offener Schleife (normalerweise  $\pm 5$ ) gesucht, was die Suchprozedur signifikant vereinfacht. Es wird eine einfache Prozedur verwendet, um den gefilterten Codevektor  $y_T$  ohne die Notwendigkeit, die Faltung für jede Tonhöhen-nacheilung zu berechnen, zu aktualisieren.

[0069] Sobald eine optimale ganzzahlige Tonhöhen-nacheilung in der zweiten Stufe gefunden worden ist, prüft eine dritte Stufe der Suche (das Modul **107**) die Bruchteile um diese optimale ganzzahlige Tonhöhen-nacheilung.

[0070] Wende der Tonhöhen-Prädiktor durch ein Filter der Form  $1/(1 - bz^{-T})$  dargestellt wird, was eine gültige Annahme für Tonhöhen-nacheilungen  $T > N$  ist, zeigt das Spektrum des Tonhöhenfilters eine harmonische Struktur über den ganzen Frequenzbereich mit einer harmonischen Frequenz, die mit  $1/T$  in Beziehung steht. Im Fall von Breitbandsignalen ist diese Struktur nicht sehr effizient, weil die harmonische Struktur in Breitbandsignalen nicht das ganze ausgedehnte Spektrum abdeckt. Die harmonische Struktur ist nur bis zu einer bestimmten Frequenz vorhanden, abhängig vom Sprachsegment. Folglich muss, um eine effiziente Darstellung des Tonhöhenbeitrags in den stimmhaften Segmenten von Breitband-Sprache zu erreichen, das Tonhöhen-Vorhersagefilter die Flexibilität besitzen, den Betrag der Periodizität über das Breitbandspektrum zu variieren.

[0071] Ein neues Verfahren, das die effiziente Modellierung der harmonische Struktur des Sprachspektrums von Breitbandsignalen ausführt, ist in der vorliegenden Beschreibung offenbart, wodurch mehrere Formen von Tiefpassfiltern auf die frühere Erregung angewendet werden, wobei das Tiefpassfilter mit der höheren Vorher-

sageverstärkung ausgewählt wird.

[0072] Wenn die Unterabtast-Tonhöhenauflösung verwendet wird, können die Tiefpassfilter in die Interpolationsfilter aufgenommen werden, die verwendet werden, um die höhere Tonhöhenauflösung zu erhalten. In diesem Fall wird die dritte Stufe der Tonhöhenauflösung, in der die Bruchteile um die gewählte ganzzahlige Tonhöhenannäherung geprüft werden, für einige Interpolationsfilter wiederholt, die verschiedene Tiefpasskennlinien besitzen, wobei der Bruchteil und der Filterindex, die das Suchkriterium C maximieren, ausgewählt werden.

[0073] Ein einfacherer Zugang besteht darin, die Suche in den drei obenbeschriebenen Stufen abzuschließen, um die optimale gebrochene Tonhöhenannäherung unter Verwendung nur eines Interpolationsfilters mit einem bestimmten Frequenzgang zu bestimmen, und die optimale Tiefpass-Filterform am Ende auszuwählen, indem die verschiedenen vorgegebenen Tiefpassfilter auf den gewählten Tonhöhen-Codebuchvektor  $v_T$  angewendet werden, und das Tiefpassfilter auszuwählen, das den Tonhöhen-Vorhersagefehler minimiert. Dieser Zugang ist im Folgenden ausführlich erörtert.

[0074] **Fig. 3** veranschaulicht einen schematischen Blockschaltplan einer bevorzugten Ausführungsform des vorgeschlagenen Zugangs.

[0075] Im Speichermodul **303** ist das frühere Erregungssignal  $u(n)$ ,  $n < 0$ , gespeichert. Das Tonhöhen-Codebuch-Suchmodul **301** führt in Reaktion auf den Zielvektor  $x$ , die Tonhöhenannäherung  $T_{OL}$  mit offener Schleife und das frühere Erregungssignal  $u(n)$ ,  $n < 0$ , vom Speichermodul **303** eine Tonhöhen-Codebuchsuche (Tonhöhen-Codebuchsuche) aus, die das oben definierte Suchkriterium C minimiert. Aus dem Ergebnis der im Modul **301** ausgeführten Suche erzeugt das Modul **302** den optimalen Tonhöhen-Codebuchvektor  $v_T$ . Es wird angemerkt, dass, weil eine Unterabtast-Tonhöhenauflösung verwendet wird (gebrochene Tonhöhe), das frühere Erregungssignal  $u(n)$ ,  $n < 0$ , interpoliert wird, wobei der Tonhöhen-Codebuchvektor  $v_T$  dem interpolierten früheren Erregungssignal entspricht. In dieser bevorzugten Ausführungsform besitzt das Interpolationsfilter (im Modul **301**, aber nicht gezeigt) eine Tiefpassfilter-Kennlinie, die die Frequenzinhalte über 7000 Hz beseitigt.

[0076] In einer bevorzugten Ausführungsform werden K Filterkennlinien verwendet; diese Filterkennlinien könnten Tiefpassfilter- oder Bandpassfilter-Kennlinien sein. Sobald der optimale Codevektor  $v_T$  bestimmt und durch den Tonhöhen-Codevektorgenerator **302** geliefert worden ist, werden K gefilterte Versionen von  $v_T$  jeweils unter Verwendung K verschiedener Frequenzformungsfilter, wie z. B. **305<sup>(j)</sup>**, berechnet, wobei  $j = 1, 2, \dots, K$  gilt. Diese gefilterten Versionen werden als  $v_f^{(j)}$  bezeichnet, wobei  $j = 1, 2, \dots, K$  gilt. Die verschiedenen Vektoren  $v_f^{(j)}$  werden in entsprechenden Modulen **304<sup>(j)</sup>**, wobei  $j = 1, 2, \dots, K$  gilt, mit der Impulsantwort  $h$  gefaltet, um die Vektoren  $y^{(j)}$  zu erhalten, wobei  $j = 1, 2, \dots, K$  gilt. Um den mittleren quadratischen Tonhöhen-Vorhersagefehler für jeden Vektor  $y^{(j)}$  zu berechnen, wird der Wert  $y^{(j)}$  mittels eines entsprechenden Verstärkers **307<sup>(j)</sup>** mit der Verstärkung  $b$  multipliziert, wobei der Wert  $by^{(j)}$  vom Zielvektor  $x$  mittels eines entsprechenden Subtrahierers **308<sup>(j)</sup>** subtrahiert wird. Die Wähleinrichtung **309** wählt das Frequenzformungsfilter **305<sup>(j)</sup>**, das den mittleren quadratischen Tonhöhen-Vorhersagefehler minimiert,

$$e^{(j)} = \|x - b^{(j)}y^{(j)}\|^2,$$

$$j = 1, 2, \dots, K.$$

[0077] Um den mittleren quadratischen Tonhöhen-Vorhersagefehler  $e^{(j)}$  für jeden Wert von  $y^{(j)}$  zu berechnen, wird der Wert  $y^{(j)}$  mittels eines entsprechenden Verstärkers **307<sup>(j)</sup>** mit der Verstärkung  $b$  multipliziert, wobei der Wert  $b^{(j)}y^{(j)}$  mittels der Subtrahieren **308<sup>(j)</sup>** vom Zielvektor  $x$  subtrahiert wird. Jede Verstärkung  $b^{(j)}$  wird in einer entsprechenden Verstärkungsberechnungseinrichtung **306<sup>(j)</sup>** in Verbindung mit dem Frequenzformungsfilter beim Index  $j$  unter Verwendung der folgenden Beziehung berechnet:

$$b(j) = x^T y^{(j)} / \|y^{(j)}\|^2.$$

[0078] In der Wähleinrichtung **309** werden die Parameter  $b$ ,  $T$  und  $j$  basierend auf  $v_T$  oder  $v_f^{(j)}$  gewählt, was den mittleren quadratischen Vorhersagefehler  $e$  minimiert.

[0079] In **Fig. 1** wird der Tonhöhen-Codebuchindex  $T$  codiert und zum Multiplexer **112** gesendet. Die Tonhöhenverstärkung  $b$  wird quantisiert und zum Multiplexer **112** gesendet. Bei diesem neuen Zugang sind zusätzliche Informationen notwendig, um den Index  $j$  des gewählten Frequenzformungsfilters im Multiplexer **112** zu codieren. Wenn z. B. drei Filter verwendet werden ( $j = 0, 1, 2, 3$ ), dann sind zwei Bits notwendig, um diese Informationen darzustellen. Die Filterindexinformationen  $j$  können außerdem gemeinsam mit der Tonhöhenverstärkung  $b$  codiert werden.

Die innovative Codebuch-Suche:

[0080] Sobald die Tonhöhen- oder LTP-Parameter (die langfristigen Vorhersageparameter)  $b$ ,  $T$  und  $j$  bestimmt worden sind, ist der nächste Schritt, mittels des Suchmoduls **110** nach **Fig. 1** nach der optimalen innovativen Erregung zu suchen. Zuerst wird der Zielvektor  $x$  aktualisiert, indem der LTP-Beitrag subtrahiert wird:

$$x' = x - by_T,$$

wobei  $b$  die Tonhöhenverstärkung ist, während  $y_T$  der gefilterte Tonhöhen-Codebuchvektor ist (die frühere Erregung mit der Verzögerung  $T$ , gefiltert mit dem ausgewählten Tiefpassfilter und mit der Impulsantwort  $h$  gefaltet, wie unter Bezugnahme auf **Fig. 3** beschrieben worden ist).

[0081] Die Suchprozedur in der CELP wird ausgeführt, indem der optimale Erregungs-Codevektor  $c_k$  und die optimale Verstärkung  $g$ , die den mittleren quadratischen Fehler zwischen dem Zielvektor und dem skalierten gefilterten Codevektor minimieren, festgestellt werden,

$$E = \|x' - gHc_k\|^2,$$

wobei  $H$  eine aus dem Impulsantwort-Vektor  $h$  abgeleitete untere Dreiecks-Faltungsmatrix ist.

[0082] In der bevorzugten Ausführungsform der vorliegenden Erfindung wird die innovative Codebuch-Suche im Modul **110** mittels eines algebraischen Codebuchs ausgeführt, wie es in den US-Patenten Nr. 5.444.816 (Adoul u. a.), erteilt am 22. August 1995; 5.699.482, erteilt am 17. Dezember 1997 an Adoul u. a.; 5.754.976, erteilt am 19. Mai 1998 an Adoul u. a.; und 5.701.392 (Adoul u. a.), datiert vom 23. Dezember 1997, beschrieben ist.

[0083] Sobald der optimale Erregungs-Codevektor  $c_k$  und seine Verstärkung  $g$  durch das Modul **110** gewählt worden sind, werden der Codebuch-Index  $k$  und die Verstärkung  $g$  codiert und zum Multiplexer **112** gesendet.

[0084] In **Fig. 1** werden die Parameter  $b$ ,  $T$ ,  $j$ ,  $\hat{A}^{(z)}$ ,  $k$  und  $g$  durch den Multiplexer **112** multiplexiert, bevor sie durch einen Kommunikationskanal übertragen werden.

#### Die Speicheraktualisierung:

[0085] Im Speichermodul **111** (**Fig. 1**) werden die Zustände des gewichteten Synthesefilters  $W(z)/\hat{A}(z)$  durch Filterung des Erregungssignals  $u = gc_k + bv_T$  durch das gewichtete Synthesefilter aktualisiert. Nach dieser Filterung werden die Zustände des Filters gespeichert und im nächsten Unterrahmen als Anfangszustände für die Berechnung des Nulleingangsverhaltens im Rechnermodul **108** verwendet.

[0086] Wie im Fall des Zielvektors  $x$  können alternative aber mathematische äquivalente Zugänge, die den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind, verwendet werden, um die Filterzustände zu aktualisieren.

#### DIE DECODIERERSEITE

[0087] Die Sprachdecodierungsvorrichtung **200** nach **Fig. 2** veranschaulicht die verschiedenen Schritte, die zwischen dem digitalen Eingang **222** (dem Eingangstrom zum Demultiplexer **217**) und der abgetasteten Ausgangssprache **223** (die Ausgabe des Addierers **221**) ausgeführt werden.

[0088] Der Demultiplexer **217** extrahiert die Synthesemodellparameter aus den vom digitalen Eingangskanal empfangenen binären Informationen. Aus jedem empfangenen binären Rahmen sind die extrahierten Parameter:

- die kurzfristigen Vorhersageparameter (STP-Parameter)  $\hat{A}^{(z)}$  (einmal pro Rahmen);
- die langfristigen Vorhersageparameter (LTP-Parameter)  $T$ ,  $b$  und  $j$  (für jeden Unterrahmen); und
- der Innovations-Codebuchindex  $k$  und die Verstärkung  $g$  (für jeden Unterrahmen).

[0089] Das aktuelle Sprachsignal wird basierend auf diesen Parametern synthetisiert, wie im Folgenden erklärt ist.

[0090] Das innovative Codebuch **218** erzeugt in Reaktion auf den Index  $k$  den Innovations-Codevektor  $c_k$ , der durch einen Verstärker **224** durch den decodierten Verstärkungsfaktor  $g$  skaliert wird. In der bevorzugten Ausführungsform wird ein innovatives Codebuch **218**, wie es in den oben erwähnten US-Patenten Nr. 5.444.816; 5.699.482; 5.754.976; und 5.701.392 beschrieben ist, verwendet, um den innovativen Codevektor  $c_k$  darzustellen.

[0091] Der erzeugte skalierte Codevektor  $gc_k$  am Ausgang des Verstärkers **224** wird durch ein Innovationsfilter **205** verarbeitet.

#### Die Periodizitätsverbesserung:

[0092] Der erzeugte skalierte Codevektor am Ausgang des Verstärkers **224** wird durch eine frequenzabhängige Tonhöhen-Verbesserungseinrichtung **205** verarbeitet.

[0093] Die Verbesserung der Periodizität des Erregungssignals  $u$  verbessert die Qualität im Fall stimmhafter Segmente. Dies wurde in der Vergangenheit durch die Filterung des Innovationsvektors vom innovativen Co-

debuch (festen Codebuch) **218** durch ein Filter in der Form  $1/(1 - \epsilon bz^{-1})$  ausgeführt, wobei  $\epsilon$  ein Faktor unter 0,5 ist, der den Betrag der eingeführten Periodizität steuert. Dieser Zugang ist in dem Fall von Breitbandsignalen weniger effizient, weil er die Periodizität über das ganze Spektrum einführt. Ein neuer alternativer Zugang, der Teil der vorliegenden Erfindung ist, ist offenbart, wodurch die Periodizitätsverbesserung durch Filterung des innovativen Codevektors  $c_k$  vom innovativen (festen) Codebuch durch ein Innovationsfilter **205** ( $F(z)$ ) ausgeführt wird, dessen Frequenzgang die höheren Frequenzen mehr als die niedrigeren Frequenzen hervorhebt. Die Koeffizienten von  $F(z)$  stehen mit dem Betrag der Periodizität im Erregungssignal  $u$  in Beziehung.

[0094] Es sind viele den Fachleuten auf dem Gebiet bekannte Verfahren verfügbar, um gültige Periodizitätskoeffizienten zu erhalten. Der Wert der Verstärkung  $b$  schafft z. B. eine Anzeige der Periodizität. Das heißt, falls die Verstärkung  $b$  nah bei 1 liegt, ist die Periodizität des Erregungssignals  $u$  hoch, während, falls die Verstärkung  $b$  kleiner als 0,5 ist, die Periodizität niedrig ist.

[0095] Einer weitere effiziente Art, die Koeffizienten des Filters  $F(z)$  abzuleiten, die in einer bevorzugten Ausführungsform verwendet wird, ist, sie mit dem Betrag des Tonhöhenbeitrags im Gesamterregungssignal  $u$  in Beziehung zu setzen. Dies führt zu einem Frequenzgang, der von der Unterrahmen-Periodizität abhängig ist, wobei für höhere Tonhöhenverstärkungen höhere Frequenzen stärker hervorgehoben werden (stärkere Gesamtneigung). Das Innovationsfilter **205** besitzt die Wirkung der Verringerung der Energie des innovativen Codevektors  $c_k$  bei niedrigen Frequenzen, wenn das Erregungssignal  $u$  periodischer ist, was die Periodizität des Erregungssignals  $u$  bei niedrigeren Frequenzen mehr als bei höheren Frequenzen verbessert. Die vorgeschlagenen Formen des Innovationsfilters **205** sind

$$(1) F(z) = 1 - \sigma z^{-1}$$

oder

$$(2) F(z) = -\alpha z + 1 - \alpha z^{-1},$$

wobei  $\sigma$  oder  $\alpha$  Periodizitätsfaktoren sind, die vom Periodizitätsniveau des Erregungssignals  $u$  abgeleitet sind.

[0096] Die zweite Form von  $F(z)$  mit drei Termen wird in einer bevorzugten Ausführungsform verwendet. Der Periodizitätsfaktor  $\alpha$  wird im Stimmfaktorgenerator **204** berechnet. Es können mehrere Verfahren verwendet werden, um den Periodizitätsfaktor  $\alpha$  basierend auf der Periodizität des Erregungssignals  $u$  abzuleiten. Zwei Verfahren sind im Folgenden dargestellt.

Das Verfahren 1:

[0097] Das Verhältnis des Tonhöhenbeitrags zum Gesamterregungssignal  $u$  wird zuerst im Stimmfaktorgenerator **204** durch

$$R_p = \frac{b^2 \mathbf{v}_T^t \mathbf{v}_T}{\mathbf{u}^t \mathbf{u}} = \frac{b^2 \sum_{n=0}^{N-1} v_T^2(n)}{\sum_{n=0}^{N-1} u^2(n)}$$

berechnet, wobei  $\mathbf{v}_T$  der Tonhöhen-Codebuchvektor ist,  $b$  die Tonhöhenverstärkung ist und  $u$  das Erregungssignal  $u$  ist, das am Ausgang des Addierers **219** durch

$$u = g c_k + b v_T$$

gegeben ist.

[0098] Es wird angemerkt, dass der Term  $b v_T$  seine Quelle im Tonhöhen-Codebuch (Tonhöhen-Codebuch) **201** in Reaktion auf die Tonhöhenneigung  $T$  und den früheren Wert von  $u$ , der im Speicher **203** gespeichert ist, besitzt. Der Tonhöhen-Codevektor  $\mathbf{v}_T$  aus den Tonhöhen-Codebuch **201** wird dann durch ein Tiefpassfilter **202** verarbeitet, dessen Grenzfrequenz mittels des Index  $j$  vom Demultiplexer **217** eingestellt wird. Der resultierende Codevektor  $\mathbf{v}_T$  wird dann durch einen Verstärker **226** mit der Verstärkung  $b$  vom Demultiplexer **217** multipliziert, um das Signal  $b v_T$  zu erhalten.

[0099] Der Faktor  $\alpha$  wird durch den Stimmfaktorgenerator **204** durch

$$\alpha = q R_p, \text{ beschränkt durch } \alpha < p$$

berechnet, wobei  $q$  ein Faktor ist, der den Betrag der Verbesserung steuert ( $q$  ist in dieser bevorzugten Ausführungsform auf 0,25 gesetzt).

## Das Verfahren 2:

[0100] Ein weiteres in einer bevorzugten Ausführungsform der Erfindung verwendetes Verfahren für die Berechnung des Periodizitätsfaktors  $\alpha$  ist im Folgenden erörtert.

[0101] Zuerst wird im Stimmfaktorgenerator **204** ein Stimmfaktor  $r_v$  durch

$$r_v = (E_v - E_c)/(E_v + E_c)$$

berechnet, wobei  $E_v$  die Energie des skalierten Tonhöhen-Codevektors  $bv_T$  ist und  $E_c$  die Energie des skalierten innovativen Codevektors  $gc_k$  ist. Das heißt,

$$E_v = b^2 v_T^t v_T = b^2 \sum_{n=0}^{N-1} v_T^2(n)$$

und

$$E_c = g^2 c_k^t c_k = g^2 \sum_{n=0}^{N-1} c_k^2(n).$$

[0102] Es wird angemerkt, dass der Wert von  $r_v$  zwischen  $-1$  und  $1$  liegt ( $1$  entspricht rein stimmhaften Signalen, während  $-1$  rein stimmlosen Signalen entspricht).

[0103] In dieser bevorzugten Ausführungsform wird der Faktor  $\alpha$  dann im Stimmfaktorgenerator **204** durch

$$\alpha = 0,125(1 + r_v)$$

berechnet, der einem Wert von  $0$  für rein stimmlose Signale und  $0,25$  für rein stimmhafte Signale entspricht.

[0104] In der ersten Form von  $F(z)$  mit zwei Termen kann der Periodizitätsfaktor  $\sigma$  unter Verwendung von  $\sigma = 2\alpha$  in den obigen Verfahren 1 und 2 approximiert werden. In einem derartigen Fall wird der Periodizitätsfaktor  $\sigma$  im obigen Verfahren 1 wie folgt berechnet:

$$\sigma = 2qR_p, \text{ beschränkt durch } \sigma < 2q.$$

[0105] Im Verfahren 2 wird der Periodizitätsfaktor  $\sigma$  wie folgt berechnet:

$$\sigma = 0,25(1 + r_v).$$

[0106] Das verbesserte Signal  $c_f$  wird deshalb durch Filterung des skalierten innovativen Codevektors  $gc_k$  durch das Innovationsfilter **205** ( $F(z)$ ) berechnet.

[0107] Das verbesserte Erregungssignal  $u'$  wird durch den Addieren **220** als:

$$u' = c_f + bv_T$$

berechnet.

[0108] Es wird angemerkt, dass dieser Prozess nicht im Codierer **100** ausgeführt wird. Folglich ist es wesentlich, die Inhalte des Tonhöhen-Codebuchs **201** unter Verwendung des Erregungssignals  $u$  ohne Verbesserung zu aktualisieren, um die Synchronisation zwischen dem Codierer **100** und den Decodieren **200** zu erhalten. Deshalb wird das Erregungssignal  $u$  verwendet, um den Speicher **203** des Tonhöhen-Codebuchs **201** zu aktualisieren, während das verbesserte Erregungssignal  $u'$  am Eingang des LP-Synthesefilters **206** verwendet wird.

## Die Synthese und die Rückentzerrung

[0109] Das synthetisierte Signal  $s'$  wird durch Filterung des verbesserten Erregungssignals  $u'$  durch das LP-Synthesefilter **206** berechnet, das die Form  $1/\hat{A}(z)$  besitzt, wobei  $A(z)$  das interpolierte LP-Filter im aktuellen Unterrahmen ist. Wie in **Fig. 2** ersichtlich ist, werden die quantisierten LP-Koeffizienten  $\hat{A}(z)$  auf der Leitung 225 vom Demultiplexer **217** zum LP-Synthesefilter **206** geliefert, um die Parameter des LP-Synthesefilters **206** dementsprechend einzustellen. Das Rückentzerrungsfilter **207** ist das Inverse des Vorverzerrungsfilters **103** nach **Fig. 1**. Die Übertragungsfunktion des Rückentzerrungsfilters **207** ist durch

$$D(z) = 1/(1 - \mu z^{-1})$$

gegeben, wobei  $\mu$  ein Vorverzerrungsfaktor mit einem Wert ist, der sich zwischen 0 und 1 befindet (ein typischer Wert ist  $\mu = 0,7$ ). Es könnte außerdem ein Filter höherer Ordnung verwendet werden.

[0110] Der Vektor  $s'$  wird durch das Rückentzerrungsfilter  $D(z)$  (das Modul **207**) gefiltert, um den Vektor  $s_h$  zu erhalten, der durch das Hochpassfilter **208** geleitet wird, um die unerwünschten Frequenzen unter 50 Hz zu entfernen und ferner  $s_h$  zu erhalten.

### Die Überabtastung und die Hochfrequenz-Regenerierung

[0111] Das Überabtaastmodul **209** führt den inversen Prozess des Unterabtaastmoduls **101** nach **Fig. 1** aus. In dieser bevorzugten Ausführungsform setzt die Überabtastung von der 12,8-kHz-Abtastrate zur ursprünglichen 16-kHz-Abtastrate unter Verwendung von Techniken um, die den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind. Das überabgetastete Synthesesignal wird als  $\hat{s}$  bezeichnet. Das Signal  $\hat{s}$  wird außerdem als das synthetisierte Breitband-Zwischensignal bezeichnet.

[0112] Das überabgetastete Synthesesignal  $\hat{s}$  enthält nicht die höheren Frequenzkomponenten, die durch den Unterabtaastprozess (das Modul **101** nach **Fig. 1**) im Codieren **100** verloren wurden. Dies ergibt eine Tiefpass-Wahrnehmung des synthetisierten Sprachsignals. Um das volle Band des ursprünglichen Signals wiederherzustellen, wird eine Hochfrequenz-Erzeugungsprozedur offenbart. Diese Prozedur wird in den Modulen **210** bis **216** und dem Addieren **221** ausgeführt, wobei sie die Eingabe vom Stimmfaktorgenerator **204** (**Fig. 2**) erfordert.

[0113] In diesem neuen Zugang werden die Hochfrequenzinhalte erzeugt, indem der obere Teil des Spektrums mit einem weißen Rauschen gefüllt wird, das im Erregungsbereich passend skaliert und dann in den Sprachbereich umgesetzt wird, vorzugsweise, indem es mit dem gleichen LP-Synthesefilter geformt wird, das für das Synthetisieren des unterabgetasteten Signals  $\hat{s}$  verwendet wird.

[0114] Die Hochfrequenz-Erzeugungsprozedur gemäß der vorliegenden Erfindung im Folgenden beschrieben.

[0115] Der Zufallsrauschgenerator **213** erzeugt unter Verwendung von Techniken, die den Durchschnittsfachleuten auf dem Gebiet wohl bekannt sind, eine weiße Rauschsequenz  $w'$  mit einem ebenen Spektrum über die ganze Frequenzbandbreite. Die erzeugte Sequenz besitzt die Länge  $M$ , die die Unterrahmen-Länge im ursprünglichen Bereich ist. Es wird angemerkt, dass  $N$  die Unterrahmen-Länge im unterabgetasteten Bereich ist. In dieser bevorzugten Ausführungsform gilt  $N = 64$  und  $N' = 80$ , was 5 ms entspricht.

[0116] Die weiße Rauschsequenz wird im Verstärkungseinstellmodul **214** passend skaliert. Die Verstärkungseinstellung umfasst die folgenden Schritte. Zuerst wird die Energie der erzeugten weißen Rauschsequenz  $w'$  gleich Energie des verbesserten Regungssignals  $u'$  gesetzt, die durch ein Energieberechnungsmodul **210** berechnet wird, wobei die resultierende skalierte Rauschsequenz durch

$$w(n) = w'(n) \sqrt{\frac{\sum_{n=0}^{N-1} u'^2(n)}{\sum_{n=0}^{N'-1} w'^2(n)}},$$

$n = 0, \dots, M - 1$ , gegeben ist.

[0117] Der zweite Schritt bei der Verstärkungsskalierung ist, die Hochfrequenzinhalte des synthetisierten Signals am Ausgang des Stimmfaktorgenerators **204** zu berücksichtigen, um die Energie des erzeugten Rauschens im Fall von stimmhaften Segmenten (wo im Vergleich zu stimmlosen Segmenten wenige Energie bei hohen Frequenzen vorhanden ist) zu reduzieren. In dieser bevorzugten Ausführungsform ist das Messen der Hochfrequenzinhalte durch das Messen der Neigung des Synthesesignals durch eine Spektralneigungs-Berechnungseinrichtung **212** und die Reduzierung der Energie dementsprechend implementiert. Andere Messungen, wie z. B. Nulldurchgangsmessungen, können ebenso verwendet werden. Wenn die Neigung sehr stark ist, was stimmhaften Segmenten entspricht, wird die Rauschenergie weiter reduziert. Der Neigungsfaktor wird im Modul **212** als der erste Korrelationskoeffizient des Synthesesignals  $s_h$  berechnet, wobei er durch

$$\text{Neigung} = \frac{\sum_{n=1}^{N-1} s_h(n)s_h(n-1)}{\sum_{n=0}^{N-1} s_h^2(n)},$$

abhängig von Neigung  $\geq 0$  und Neigung  $\geq r_v$ , gegeben ist, wobei der Stimmfaktor  $r_v$  durch

$$r_v = (E_v - E_c)/(E_v + E_c)$$

gegeben ist, wobei  $E_v$  die Energie des skalierten Tonhöhen-Codevektors  $bv_T$  ist und  $E_c$  die Energie des skalierten innovativen Codevektors  $gc_k$  ist, wie vorausgehend beschrieben worden ist. Der Stimmfaktor  $r_v$  ist am häufigsten kleiner als die Neigung, aber diese Bedingung wurde als eine Vorsichtsmaßnahme gegen Hochfrequenzöne eingeführt, bei denen der Neigungswert negativ und der Wert von  $r_v$  hoch ist. Deshalb reduziert diese Bedingung die Rauschenergie für derartige Klangsignale.

[0118] Der Neigungswert ist im Fall eines ebenen Spektrums 0 und im Fall stark stimmhafter Signale 1, während er im Fall stimmloser Signale negativ ist, in denen bei hohen Frequenzen mehr Energie vorhanden ist.

[0119] Es können verschiedene Verfahren verwendet werden, um den Skalierungsfaktor  $g_f$  aus der Menge der Hochfrequenzinhalte abzuleiten. In dieser Erfindung werden zwei Verfahren basierend auf der obenbeschriebenen Neigung des Signals angegeben.

#### Das Verfahren 1:

[0120] Der Skalierungsfaktor  $g_t$  wird aus der Neigung durch

$$g_t = 1 - \text{Neigung}, \text{ beschränkt durch } 0,2 \leq g_t \leq 1,0$$

abgeleitet. Für stark stimmhafte Signale, bei denen sich die Neigung 1 nähert, ist  $g_t$  0,2, während für stark stimmlose Signale  $g_t$  1,0 wird.

#### Das Verfahren 2:

[0121] Der Neigungsfaktor  $g_t$  wird zuerst eingeschränkt, damit er größer oder gleich null ist, wobei dann der Skalierungsfaktoren aus der Neigung durch

$$g_t = 10^{-0,6\text{Neigung}}$$

abgeleitet wird.

[0122] Die im Verstärkungseinstellmodul **214** erzeugte skalierte Rauschsequenz  $w_g$  ist deshalb durch

$$w_g = g_t w$$

gegeben.

[0123] Wenn sich die Neigung nahe bei null befindet, ist der Skalierungsfaktor  $g_t$  nahe bei 1, was nicht zu einer Energiereduzierung führt. Wenn der Neigungswert 1 ist, führt der Skalierungsfaktor  $g_t$  zu einer Reduzierung von 12 dB in der Energie des erzeugten Rauschens.

[0124] Sobald das Rauschen ( $w_g$ ) passend skaliert ist, wird es unter Verwendung der spektralen Formungseinrichtung **215** in den Sprachbereich gebracht. In der bevorzugten Ausführungsform wird dies durch Filterung des Rauschens  $w_g$  durch eine bandbreitenerweiterte Version des gleichen LP-Synthesefilters ausgeführt, das im unterabgetastetem Bereich verwendet wird ( $1/\hat{A}(z/0,8)$ ). Die entsprechenden bandbreitenerweiterten LP-Filterkoeffizienten werden in der spektralen Formungseinrichtung **215** berechnet.

[0125] Die gefilterte skalierte Rauschsequenz  $w_f$  wird dann in den erforderlichen Frequenzbereich bandpassgefiltert, um unter Verwendung des Bandpassfilters **216** wiederhergestellt zu werden. In der bevorzugten Ausführungsform schränkt das Bandpassfilter **216** die Rauschsequenz auf den Frequenzbereich 5,6–7,2 kHz ein. Die resultierende bandpassgefilterte Rauschsequenz  $r$  wird im Addieren **221** zum überabgetasteten synthetisierten Sprachsignal  $\hat{s}$  addiert, um das endgültige rekonstruierte Tonsignal  $s_{\text{out}}$  am Ausgang **223** zu erhalten.

[0126] Obwohl die vorliegende Erfindung vorausgehend mittels einer ihrer bevorzugten Ausführungsformen beschrieben worden ist, kann diese Ausführungsform nach Wunsch innerhalb des Umfangs der beigefügten Ansprüche modifiziert werden. Auch wenn die bevorzugte Ausführungsform die Verwendung von Breitband-Sprachsignalen erörtert, wird es für die Fachleute auf dem Gebiet offensichtlich sein, dass der Erfindungsgegenstand außerdem auf andere Ausführungsformen unter Verwendung von Breitbandsignalen im Allgemeinen gerichtet ist, und dass er nicht notwendigerweise auf Sprachanwendungen eingeschränkt ist.

### Patentansprüche

1. Tonhöhenanalysevorrichtung zum Erzeugen einer optimalen Menge von Tonhöhencodebuch-Parametern in Reaktion auf ein Breitbandsignal, die umfaßt:

a) wenigstens zwei Signalpfade, die jeweiligen Mengen von Tonhöhencodebuch-Parametern zugeordnet sind, wobei:

i) jeder Signalpfad eine Tonhöhenvorhersagefehler-Berechnungsvorrichtung (**307**, **308**) zum Berechnen eines



Tonhöhenvorhersagefehlers eines Tonhöhencodevektors von einer Tonhöhencodebuch-Suchvorrichtung (301) umfaßt; und

ii) wenigstens einer der beiden Pfade ein Filter (305) zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung des einen Pfades umfaßt; und

b) eine Auswahleinrichtung (309) zum Vergleichen der in den wenigstens zwei Signalpfaden berechneten Tonhöhenvorhersagefehler, um den Signalpfad auszuwählen, der den niedrigsten berechneten Tonhöhenvorhersagefehler hat, und um die Menge von Tonhöhencodebuch-Parametern, die dem ausgewählten Signalpfad zugeordnet sind, zu wählen.

2. Tonhöhenanalysevorrichtung nach Anspruch 1, bei der einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

3. Tonhöhenanalysevorrichtung nach Anspruch 1, bei der die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

4. Tonhöhenanalysevorrichtung nach Anspruch 3, bei der die Filter der mehreren Pfade aus der Gruppe ausgewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht, und bei der die Filter unterschiedliche Frequenzgänge besitzen.

5. Tonhöhenanalysevorrichtung nach Anspruch 1, bei der jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt:

a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit einem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;

b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und einen Tonhöhen-Suchzielvektor;

c) einen Verstärker zum Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und

d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

6. Tonhöhenanalysevorrichtung nach Anspruch 5, bei der die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

7. Tonhöhenanalysevorrichtung nach Anspruch 1, bei der die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei der die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler umfaßt.

8. Tonhöhenanalysevorrichtung nach Anspruch 5, bei der:

a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;

b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und

c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

9. Tonhöhenanalysevorrichtung nach Anspruch 1, bei der das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter verwendet wird, um eine Unterab-tast-Version des Tonhöhencodevektors zu erzeugen.

10. Tonhöhenanalyseverfahren zum Erzeugen einer optimalen Menge von Tonhöhencodebuch-Parame-

tern in Reaktion auf ein Breitbandssignal, das umfaßt:

- a) in wenigstens zwei Signalpfaden, die jeweiligen Mengen von Tonhöhencodebuch-Parametern zugeordnet sind, Berechnen eines Tonhöhenvorhersagefehlers eines Tonhöhencodevektors von einer Tonhöhencodebuch-Suchvorrichtung für jeden Signalpfad;
- b) in wenigstens einem der zwei Signalpfade Filtern des Tonhöhencodevektors, bevor der Tonhöhencodevektor geliefert wird, um den Tonhöhenvorhersagefehler des einen Pfades zu berechnen; und
- c) Vergleichen der Tonhöhenvorhersagefehler, die in den wenigstens zwei Signalpfaden berechnet werden, Wählen des Signalpfades mit dem niedrigsten berechneten Tonhöhenvorhersagefehler und Auswählen der Menge von Tonhöhencodebuch-Parametern, die dem gewählten Signalpfad zugeordnet sind.

11. Tonhöhenanalyseverfahren nach Anspruch 10, bei dem in einem der wenigstens zwei Pfade keine Filterung des Tonhöhencodevektors ausgeführt wird, bevor der Tonhöhencodevektor zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung geliefert wird.

12. Tonhöhenanalyseverfahren nach Anspruch 10, bei dem die Signalpfade mehrere Signalpfade umfassen und bei dem das Filtern des Tonhöhencodevektors in jedem der mehreren Signalpfade ausgeführt wird, bevor der Tonhöhencodevektor zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades geliefert wird.

13. Tonhöhenanalyseverfahren nach Anspruch 12, das ferner das Auswählen der Filter der mehreren Pfade aus der Gruppe, die aus Tiefpaß- und Bandpaßfiltern besteht, umfaßt und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

14. Tonhöhenanalyseverfahren nach Anspruch 10, bei dem das Berechnen eines Tonhöhenvorhersagefehlers in jedem Signalpfad umfaßt:

- a) Falten des Tonhöhencodevektors mit einem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher Berechnen eines gefalteten Tonhöhencodevektors;
- b) Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und einen Tonhöhen-Suchzielvektor;
- c) Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und
- d) Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

15. Tonhöhenanalyseverfahren nach Anspruch 14, bei dem die Tonhöhenverstärkungs-Berechnung das Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

16. Tonhöhenanalyseverfahren nach Anspruch 10, bei dem das Berechnen des Tonhöhenvorhersagefehlers in jedem Signalpfad das Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem das Vergleichen des Tonhöhenvorhersagefehlers das Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und das Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehlerumfaßt.

17. Tonhöhenanalyseverfahren nach Anspruch 14, bei dem:

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;
- b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und
- c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

18. Tonhöhenanalyseverfahren nach Anspruch 10, bei dem das Filtern des Tonhöhencodevektors in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter verwendet wird, um eine Unterabtast-Version des Tonhöhencodevektors zu erzeugen.

19. Codieren, der eine Tonhöhenanalysevorrichtung nach Anspruch 1 besitzt, um ein Breitband-Eingangs-

signal zu codieren, wobei der Codieren umfaßt:

- a) eine Berechnungseinrichtung für ein lineares Vorhersagesynthesefilter, das auf das Breitbandsignal anspricht, um Koeffizienten für ein lineares Vorhersagesynthesefilter zu erzeugen;
- b) ein Wahrnehmungsgewichtungsfiler, das auf das Breitbandsignal und auf die Koeffizienten des linearen Vorhersagesynthesefilters anspricht, um ein durch Wahrnehmung gewichtetes Signal zu erzeugen;
- c) einen Impulsantwortgenerator, der auf die Koeffizienten für das lineare Vorhersagesynthesefilter anspricht, um ein gewichtetes Synthesefilter-Impulsantwortsignal zu erzeugen;
- d) eine Tonhöhen-Sucheinheit zum Erzeugen von Tonhöhencodebuch-Parametern, wobei die Tonhöhen-Sucheinheit umfaßt:
  - i) die Tonhöhencodebuch-Suchvorrichtung, die auf das durch Wahrnehmung gewichtete Signal und auf die Koeffizienten für das lineare Vorhersagesynthesefilter anspricht, um den Tonhöhencodevektor und einen innovativen Suchzielvektor zu erzeugen; und
  - ii) die Tonhöhenanalysevorrichtung, die auf den Tonhöhencodevektor anspricht, und aus den Mengen von Tonhöhencodebuch-Parametern diejenige Menge von Tonhöhencodebuch-Parametern auswählt, die dem Pfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler zugeordnet ist;
- d) eine Vorrichtung zum Suchen eines innovativen Codebuchs, die auf das gewichtete Synthesefilter-Impulsantwortsignal und auf den innovativen Suchzielvektor anspricht, um innovative Codebuchparameter zu erzeugen, und
- e) eine Signalformungsvorrichtung zum Erzeugen eines codierten Breitbandsignals, das die Menge von Tonhöhencodebuch-Parametern, die dem Pfad mit dem niedrigsten Tonhöhenvorhersagefehler zugeordnet sind, die innovativen Codebuchparameter und die Koeffizienten für das lineare Vorhersagesynthesefilter umfaßt.

20. Codieren nach Anspruch 19, bei dem einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

21. Codieren nach Anspruch 19, bei dem die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

22. Codieren nach Anspruch 21, bei dem die Filter der mehreren Pfade aus der Gruppe gewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

23. Codieren nach Anspruch 19, bei dem jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt:

- a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit dem gewichteten Synthesefilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;
- b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und den Tonhöhen-Suchzielvektor;
- c) einen Verstärker zum Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und
- d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

24. Codieren nach Anspruch 23, bei dem die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

25. Codieren nach Anspruch 19, bei dem die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler umfaßt.

26. Codieren nach Anspruch 23, bei dem:

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;
- b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und
- c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

27. Codieren nach Anspruch 19, bei dem das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter zum Erzeugen einer Unterabtast-Version des Tonhöhencodevektors verwendet wird.

28. Zellenkommunikationssystem zum Bedienen eines großen geographischen Gebiets, das in mehrere Zellen unterteilt ist, wobei das System umfaßt:

- a) mobile Sender/Empfänger-Einheiten;
- b) Zellenbasisstationen, die sich jeweils in den Zellen befinden;
- c) ein Steuerendgerät zum Steuern der Kommunikation zwischen den Zellenbasisstationen;
- d) ein bidirektionales drahtloses Kommunikationsundersystem zwischen jeder Mobileinheit, die sich in einer Zelle befindet, und der Zellenbasisstation der einen Zelle, wobei das bidirektionale drahtlose Kommunikationsundersystem sowohl in der Mobileinheit als auch in der Zellenbasisstation umfaßt:
- i) einen Sender, der einen Codierer zum Codieren eines Breitbandsignals nach Anspruch 19 und eine Sende-schaltung zum Senden des codierten Breitbandsignals enthält; und
- ii) einen Empfänger, der eine Empfängerschaltung zum Empfangen eines gesendeten codierten Breitbandsignals und einen Decodieren zum Decodieren des empfangenen codierten Breitbandsignals enthält.

29. Zellenkommunikationssystem nach Anspruch 28, bei dem einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

30. Zellenkommunikationssystem nach Anspruch 28, bei dem die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

31. Zellenkommunikationssystem nach Anspruch 30, bei der die Filter der mehreren Pfade aus der Gruppe gewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht, und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

32. Zellenkommunikationssystem nach Anspruch 28, bei dem die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehlerumfaßt.

33. Zellenkommunikationssystem nach Anspruch 32, bei dem die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

34. Zellenkommunikationssystem nach Anspruch 28, bei dem die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehlerumfaßt.

35. Zellenkommunikationssystem nach Anspruch 32, bei dem:

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;

- b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und  
 c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

36. Zellenkommunikationssystem nach Anspruch 28, bei dem das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter zum Erzeugen einer Unterab-tast-Version des Tonhöhencodevektors verwendet wird.

37. Mobile Zellen-Sender/Empfänger-Einheit, die umfaßt:

- a) einen Sender, der einen Codieren zum Codieren eines Breitbandsignals und eine Sendeschaltung zum Senden des codierten Breitbandsignals enthält; und  
 b) einen Empfänger, der eine Empfängerschaltung zum Empfangen eines gesendeten codierten Breitbandsignals nach Anspruch 19 und einen Decodieren zum Decodieren des empfangenen decodierten Breitbandsignals enthält.

38. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 37, bei der einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

39. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 37, bei der die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

40. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 39, bei der die Filter der mehreren Pfade aus der Gruppe gewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

41. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 37, bei der jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt: jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt:

- a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit dem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;  
 b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und den Tonhöhen-Suchzielvektor;  
 c) einen Verstärker zum Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und  
 d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

42. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 41, bei der die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

43. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 37, bei der die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler umfaßt.

44. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 41, bei der

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;  
 b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und  
 c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

45. Mobile Zellen-Sender/Empfänger-Einheit nach Anspruch 37, bei der das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter zum Erzeugen einer Unterabtast-Version des Tonhöhencodevektors verwendet wird.

46. Zellennetzelement, das umfaßt:

- a) einen Sender, der einen Codieren zum Codieren eines Breitbandsignals nach Anspruch 19 und eine Sendeschaltung zum Senden des codierten Breitbandsignals enthält; und
- b) einen Empfänger, der eine Empfängerschaltung zum Empfangen eines gesendeten codierten Breitbandsignals und einen Decodieren zum Decodieren des empfangenen decodierten Breitbandsignals enthält.

47. Zellennetzelement nach Anspruch 46, bei dem einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

48. Zellennetzelement nach Anspruch 46, bei dem die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

49. Zellennetzelement nach Anspruch 48, bei dem die Filter der mehreren Pfade aus der Gruppe gewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

50. Zellennetzelement nach Anspruch 46, bei dem jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt:

- a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit dem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;
- b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und den Tonhöhen-Suchzielvektor;
- c) einen Verstärker zum Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und
- d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

51. Zellennetzelement nach Anspruch 50, bei dem die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhencodevektor ist.

52. Zellennetzelement nach Anspruch 46, bei dem die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler umfaßt.

53. Zellennetzelement nach Anspruch 50, bei dem:

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;
- b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und
- c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

54. Zellennetzelement nach Anspruch 46, bei dem das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter zum Erzeugen einer Unterabtast-Version des Tonhöhencodevektors verwendet wird.

55. Bidirektionales drahtloses Kommunikationsuntersystem mit einem Zellennetzsystem für

die Bedienung eines großen geographischen Gebiets, das in mehrere Zellen unterteilt ist, wobei das Untersystem umfaßt: mobile Sender/Empfänger-Einheiten; Zellenbasisstationen, die sich in den jeweiligen Zellen befinden; und ein Steuerendgerät zum Steuern der Kommunikation zwischen den Zellenbasisstationen; wobei sich das bidirektionale drahtlose Kommunikationsundersystem zwischen jeder Mobileinheit, die sich in einer Zelle befindet, und der Zellenbasisstation der einen Zelle befindet, wobei das bidirektionale drahtlose Kommunikationsundersystem sowohl in der Mobileinheit als auch in der Zellenbasisstation umfaßt:

- a) einen Sender, der einen Codierer zum Codieren eines Breitbandsignals nach Anspruch 19 und eine Sende-schaltung zum Senden des codierten Breitbandsignals enthält; und
- b) einen Empfänger, der eine Empfängerschaltung zum Empfangen eines gesendeten codierten Breitbandsignals und einen Decodierer zum Decodieren des empfangen codierten Breitbandsignals enthält.

56. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 55, bei dem einer der wenigstens zwei Pfade kein Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt.

57. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 55, bei dem die Signalpfade mehrere Signalpfade umfassen, wovon jeder mit einem Filter zum Filtern des Tonhöhencodevektors vor der Lieferung des Tonhöhencodevektors zu der Tonhöhenvorhersagefehler-Berechnungsvorrichtung desselben Pfades versehen ist.

58. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 57, bei dem die Filter der mehreren Pfade aus der Gruppe gewählt sind, die aus Tiefpaß- und Bandpaßfiltern besteht und bei dem die Filter unterschiedliche Frequenzgänge besitzen.

59. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 55, bei dem jede Tonhöhenvorhersagefehler-Berechnungsvorrichtung umfaßt:

- a) eine Faltungseinheit zum Falten des Tonhöhencodevektors mit dem gewichteten Synthetisierungsfilter-Impulsantwortsignal und daher zum Berechnen eines gefalteten Tonhöhencodevektors;
- b) eine Tonhöhenverstärkungs-Berechnungseinrichtung zum Berechnen einer Tonhöhenverstärkung in Reaktion auf den gefalteten Tonhöhencodevektor und den Tonhöhen-Suchzielvektor;
- c) einen Verstärker zum Multiplizieren des gefalteten Tonhöhencodevektors mit der Tonhöhenverstärkung, um dadurch einen verstärkten gefalteten Tonhöhencodevektor zu erzeugen; und
- d) eine Kombinationsschaltung zum Kombinieren des verstärkten gefalteten Tonhöhencodevektors mit dem Tonhöhen-Suchzielvektor, um dadurch den Tonhöhenvorhersagefehler zu erzeugen.

60. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 59, bei dem die Tonhöhenverstärkungs-Berechnungseinrichtung eine Einrichtung zum Berechnen der Tonhöhenverstärkung  $b^{(j)}$  unter Verwendung der folgenden Beziehung umfaßt:

$$b^{(j)} = x^t y^{(j)} / \|y^{(j)}\|^2$$

wobei  $j = 0, 1, 2, \dots, K$  und wobei  $K$  einer Anzahl von Signalpfaden entspricht, und wobei  $x$  der Tonhöhen-Suchzielvektor ist und  $y^{(j)}$  der gefaltete Tonhöhen-Codevektor ist.

61. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 55, bei dem die Tonhöhenvorhersagefehler-Berechnungsvorrichtung jedes Signalpfades eine Einrichtung zum Berechnen einer Energie des entsprechenden Tonhöhenvorhersagefehlers umfaßt und bei dem die Auswahleinrichtung eine Einrichtung zum Vergleichen der Energien der Tonhöhenvorhersagefehler der verschiedenen Signalpfade und zum Wählen des Signalpfades mit der niedrigsten berechneten Energie des Tonhöhenvorhersagefehlers als den Signalpfad mit dem niedrigsten berechneten Tonhöhenvorhersagefehler umfaßt.

62. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 59, bei dem:

- a) jedes der Filter der mehreren Signalpfade durch einen Filterindex identifiziert wird;
- b) der Tonhöhencodevektor durch einen Tonhöhencodebuch-Index identifiziert wird; und
- c) die Tonhöhencodebuch-Parameter den Filterindex, den Tonhöhencodebuch-Index und die Tonhöhenverstärkung umfassen.

63. Bidirektionales drahtloses Kommunikationsundersystem nach Anspruch 55, bei dem das Filter in ein Interpolationsfilter der Tonhöhencodebuch-Suchvorrichtung integriert ist, wobei das Interpolationsfilter zum Er-

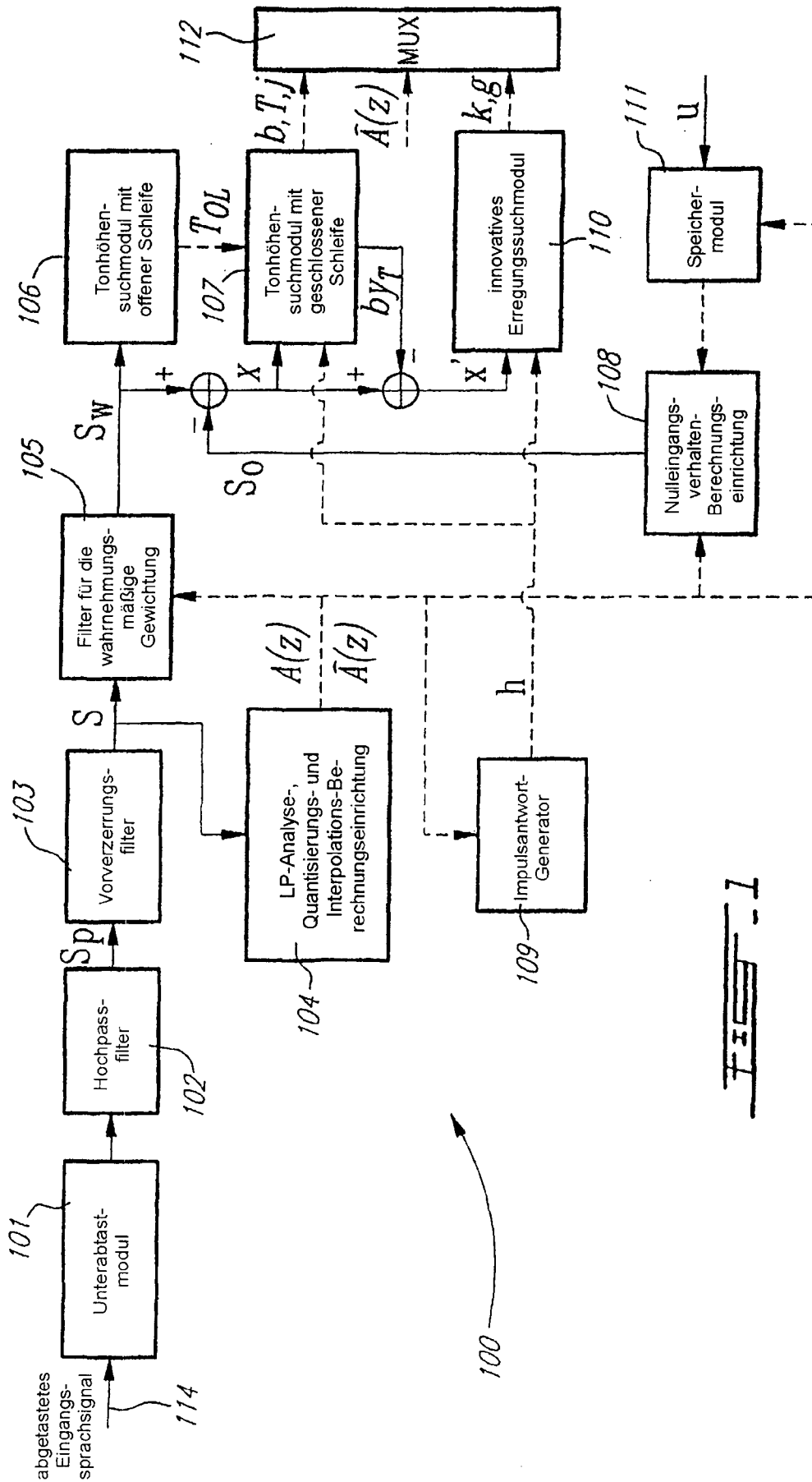
DE 699 10 239 T2 2004.06.24

zeugen einer Unterabtast-Version des Tonhöhencodevektors verwendet wird.

Es folgen 4 Blatt Zeichnungen

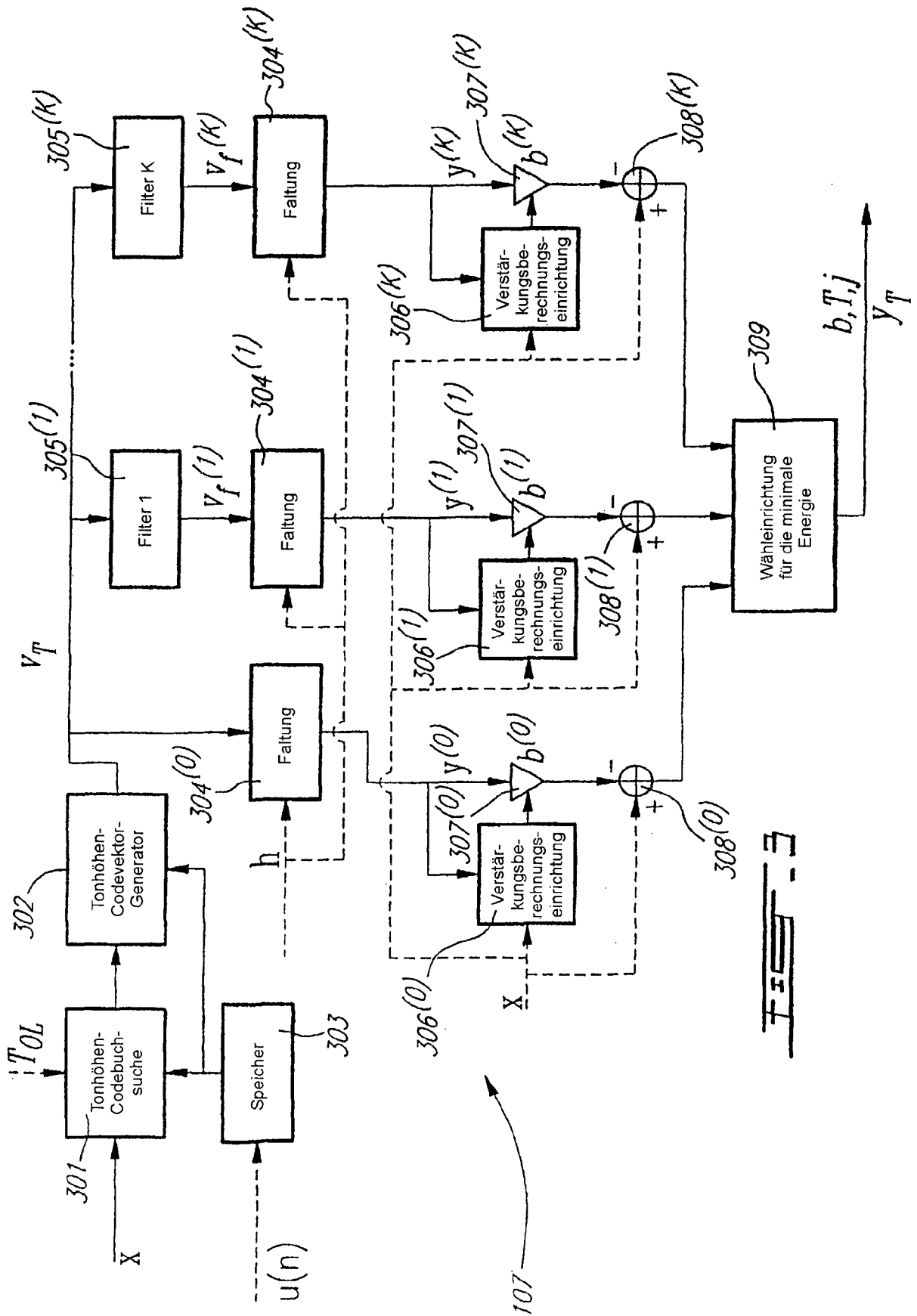


Anhängende Zeichnungen



**F I E S S**





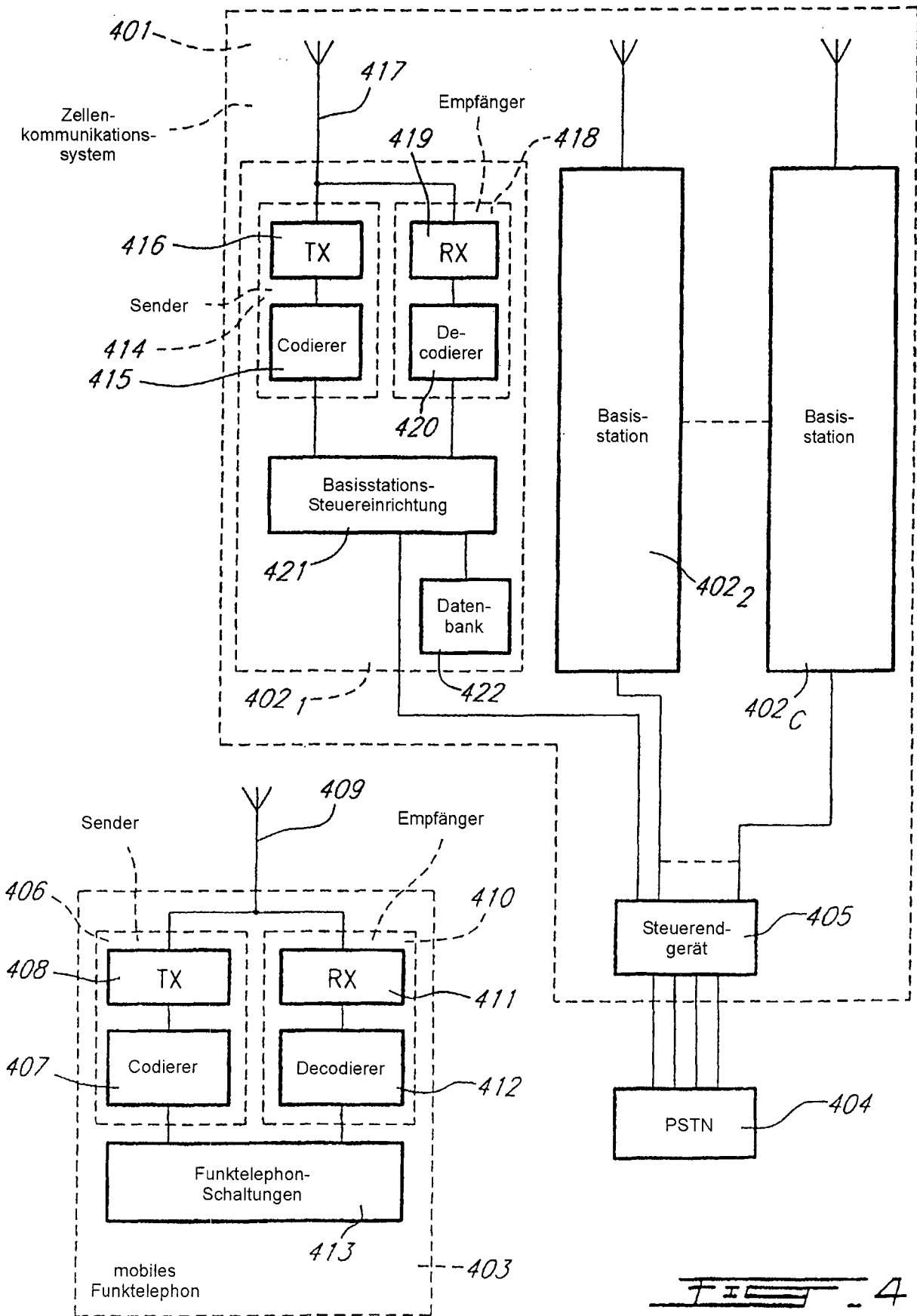


FIG. 4