



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 698 19 460 T2** 2004.08.26

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 925 580 B1**

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

(21) Deutsches Aktenzeichen: **698 19 460.8**

(86) PCT-Aktenzeichen: **PCT/IB98/00923**

(96) Europäisches Aktenzeichen: **98 923 009.9**

(87) PCT-Veröffentlichungs-Nr.: **WO 99/003097**

(86) PCT-Anmeldetag: **11.06.1998**

(87) Veröffentlichungstag
der PCT-Anmeldung: **21.01.1999**

(97) Erstveröffentlichung durch das EPA: **30.06.1999**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **05.11.2003**

(47) Veröffentlichungstag im Patentblatt: **26.08.2004**

(30) Unionspriorität:
97202166 **11.07.1997** **EP**

(84) Benannte Vertragsstaaten:
DE, FR, GB, IT, SE

(73) Patentinhaber:
**Koninklijke Philips Electronics N.V., Eindhoven,
NL**

(72) Erfinder:
**TAORI, Rakesh, NL-5656 AA Eindhoven, NL;
SLUIJTER, Johannes, Robert, NL-5656 AA
Eindhoven, NL; GERRITS, Johannes, Andreas,
NL-5656 AA Eindhoven, NL**

(74) Vertreter:
Meyer, M., Dipl.-Ing., Pat.-Ass., 52076 Aachen

(54) Bezeichnung: **ÜBERTRAGER MIT VERBESSERTEM SPRACHKODIERER UND DEKODIERER**

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

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

Beschreibung

[0001] Die vorliegende Erfindung bezieht sich auf ein Übertragungssystem mit einem Sender mit einem Sprachcodierer mit Analysenmitteln zum periodischen Ermitteln von Analysenkoeffizienten aus dem Sprachsignal, so dass der Sender Übertragungsmittel aufweist zum Übertragen der genannten Analysenkoeffizienten über ein Übertragungsmedium zu einem Empfänger, so dass der genannte Empfänger einen Sprachdecoder aufweist mit Rekonstruktionsmitteln zum Herleiten eines rekonstruierten Sprachsignals auf Basis der Analysenkoeffizienten.

[0002] Die vorliegende Erfindung bezieht sich ebenfalls auf einen Sender, einen Empfänger, einen Sprachcodierer, einen Sprachdecoder, ein Sprachcodiererverfahren, ein Sprachdecodiererverfahren und auf ein fühlbares Medium mit einem Computerprogramm, das die genannten Methoden implementiert.

[0003] Ein Übertragungssystem der eingangs beschriebenen Art ist aus EP 259 950 bekannt.

[0004] Ein derartiges Übertragungssystem und derartige Sprachcodierer werden in Applikationen verwendet, in denen Sprachsignale über ein Übertragungsmedium mit einer begrenzten Übertragungskapazität übertragen werden sollen oder auf einem Speichermedium mit einer begrenzten Speicherkapazität gespeichert werden sollen. Beispiele derartiger Applikationen sind die Übertragung von Sprachsignalen über das Internet, die Übertragung von Sprachsignalen von einem Schnurlostelephon zu einer Basisstation und umgekehrt und Speicherung von Sprachsignalen auf einem CD-ROM, in einem Festwertspeicher oder auf einer Festplatte.

[0005] Es wurden verschiedene Arbeitsprinzipien von Sprachcodierern versucht zum Erzielen einer angemessenen Sprachqualität zu einer gemäßigten Bitrate. Bei einem dieser Betriebsverfahren wird zwischen stimmhaften und stimmlosen Sprachsignalen ein Unterschied gemacht. Diese zwei Arten von Sprachsignalen werden unter Verwendung verschiedener Codierer codiert, wobei jeder derselben für die Eigenschaften des entsprechenden Typs von Sprachsignalen optimiert wird.

[0006] Ein anderen Betriebstyp ist der sog. CELP-Codierer, wobei ein Sprachsignal mit einem synthetischen Sprachsignal verglichen wird, das dadurch erhalten wird, dass ein Synthesefilter durch ein Anregungssignal angeregt wird, das aus einer Anzahl Anregungssignale hergeleitet wird, die in einem Codebuch gespeichert sind. Damit periodische Signale, wie stimmhafte Sprachsignale behandelt werden können, wird ein sog. adaptives Codebuch verwendet.

[0007] In den beiden Typen Sprachcodierer sollen Analysenparameter bestimmt werden, die Sprachsignale zu beschreiben. Wenn die verfügbare Bitrate für den Sprachcodierer abnimmt, verschlechtert die erhaltene Sprachqualität der rekonstruierten Sprache rapide.

[0008] Es ist nun u. a. eine Aufgabe der vorliegenden Erfindung ein Übertragungssystem für Sprachsignale zu schaffen, wobei die Beeinträchtigung der Sprachqualität bei verringerter Bitrate reduziert wird.

[0009] Deswegen weist das Übertragungssystem nach der vorliegenden Erfindung das Kennzeichen auf, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln, in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment oder umgekehrt, und dass die Rekonstruktionsmittel dazu vorgesehen sind, ein rekonstruiertes Sprachsignal auf Basis der öfter ermittelten Analysenkoeffizienten herzuleiten.

[0010] Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, dass eine wichtige Quelle der Beeinträchtigung der Qualität des Sprachsignals die nicht ausreichende Befolgung von Änderungen in den Analysenparametern bei einem Übergang von stimmhafter Sprache in stimmlose Sprache oder umgekehrt ist. Durch Steigerung der Aktualisierungsrate der Analysenparameter in der Nähe eines derartigen Übergangs wird die Sprachqualität wesentlich verbessert. Weil Übergänge nicht sehr oft auftreten, ist die zusätzliche Bitrate, erforderlich zum Behandeln der öfteren Aktualisierung der Analysenparameter mäßig. Es hat sich herausgestellt, dass es möglich ist, dass die Frequenz der Ermittlung der Analysenkoeffizienten gesteigert wird, bevor der Übergang wirklich auftritt, dass es aber ebenfalls möglich ist, dass die Frequenz der Ermittlung der Analysenkoeffizienten gesteigert wird, nachdem der Übergang stattfindet. Eine Kombination der obenstehenden Art und Weise der Steigerung der Frequenz der Ermittlung der Analysenkoeffizienten ist ebenfalls möglich.

[0011] Eine Ausführungsform der vorliegenden Erfindung weist das Kennzeichen auf, dass der Sprachcodierer einen stimmhaften Sprachcodierer zum Codieren stimmhafter Sprachsegmente aufweist und dass der Sprachcodierer einen stimmlosen Sprachcodierer zum Codieren stimmloser Sprachelemente aufweist.

[0012] Versuche haben gezeigt, dass die Verbesserungen, die dadurch erhalten werden können, dass die Aktualisierungsrate der Analysenparameter in der Nähe eines Übergangs gesteigert wird, besonders vorteilhaft ist für Sprachcodierer, bei denen ein stimmhafter und ein stimmloser Sprachcodierer verwendet wird. Mit derartigen Sprachcodierern ist die mögliche Verbesserung beträchtlich.

[0013] Eine weitere Ausführungsform der vorliegenden Erfindung weist das Kennzeichen auf, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln für zwei Segmente nach dem Übergang.

[0014] Es hat sich herausgestellt, dass dadurch, dass nach dem Übergang die Analysenkoeffizienten für zwei Frames öfter ermittelt werden, dies zu einer wesentlich höheren Sprachqualität führt.

[0015] Noch eine andere Ausführungsform der vorliegenden Erfindung weist das Kennzeichen auf, dass die Analysenmittel dazu vorgesehen sind, die Frequenz der Ermittlung der Analysenkoeffizienten bei einem Übergang zwischen einem stimmhaften und einem stimmlosen Segment und umgekehrt zu verdoppeln.

[0016] Es hat sich herausgestellt, dass eine Verdopplung der Frequenz der Ermittlung der Analysenkoeffizienten ausreicht zum Erhalten einer wesentlich besseren Sprachqualität.

[0017] Ausführungsbeispiele der Erfindung sind in der Zeichnung dargestellt und werden im vorliegenden Fall näher beschrieben. Es zeigen:

[0018] **Fig. 1** ein Übertragungssystem, wobei die vorliegende Erfindung angewandt werden kann,

[0019] **Fig. 2** einen Sprachcodierer **4** nach der vorliegenden Erfindung,

[0020] **Fig. 3** einen stimmhaften Codierer **16** nach der vorliegenden Erfindung,

[0021] **Fig. 4** LPC-Rechenmittel **30** zur Verwendung in dem stimmhaften Sprachcodierer **16** nach **Fig. 3**,

[0022] **Fig. 5** Tonhöhenabstimmittel **32** zur Verwendung in dem Sprachcodierer nach **Fig. 3**,

[0023] **Fig. 6** einen Sprachcodierer **14** für stimmlose Sprache zur Verwendung in dem Sprachcodierer nach **Fig. 2**,

[0024] **Fig. 7** einen Sprachdecoder **14** zur Verwendung in dem System nach **Fig. 1**,

[0025] **Fig. 8** einen stimmhaften Sprachdecoder **94** zur Verwendung in dem Sprachdecoder **14**,

[0026] **Fig. 9** Graphiken von Signalen an einer Anzahl Stellen in dem stimmhaften Sprachdecoder **94**,

[0027] **Fig. 10** einen stimmlosen Sprachdecoder **96** zur Verwendung in dem Sprachdecoder **14**.

[0028] In dem Übertragungssystem nach **Fig. 1** wird ein Sprachsignal einem Eingang eines Senders **2** zugeführt. In dem Sender **2** wird das Sprachsignal in einem Sprachcodierer **4** codiert. Das codierte Sprachsignal an dem Ausgang des Sprachcodierers **4** wird den Übertragungsmitteln **6** zugeführt. Die Übertragungsmittel **6** sind vorgesehen zum Durchführen von Kanalcodierung, zum Verschachteln und zum Modulieren des codierten Sprachsignals.

[0029] Das Ausgangssignal der Übertragungsmittel **6** wird dem Ausgang des Senders zugeführt und wird über ein Übertragungsmittel zu einem Empfänger **5** übertragen. Am Empfänger **5** wird das Ausgangssignal des Kanals Empfangsmitteln **7** zugeführt. Diese Empfangsmittel **7** schaffen HF-Verarbeitung, wie Abstimmung und Demodulation, Entschachtelung (falls anwendbar) und Kanaldecodierung. Das Ausgangssignal der Empfangsmittel **7** wird dem Sprachdecoder **9** zugeführt, der sein Eingangssignal in ein rekonstruiertes Sprachsignal verwandelt.

[0030] Das Eingangssignal $s_s[n]$ des Sprachcodierers **4** nach **Fig. 2** wird von einem DC Kerbfilter **10** gefiltert zum Eliminieren unerwünschter DC-Offsets von dem Eingang. Das genannte DC Kerbfilter hat eine Grenzfrequenz (-3 dB) von 15 Hz. Das Ausgangssignal des DC Kerbfilter **10** wird einem Eingang eines Puffers **11** zugeführt. Der Puffer **11** bietet einem stimmhaften Sprachcodierer **16** nach der vorliegenden Erfindung Blöcke von 400 DC gefilterten Sprachabtastwerten. Der genannte Block von 400 Abtastwerten umfasst 5 Frames von 10 ms Sprache (jeweils 80 Abtastwerte). Er umfasst das aktuelle, zu codierende Frame, zwei vorhergehende und zwei nachfolgende Frames. Der Puffer **11** bietet in jedem Frame-Intervall das jüngste empfangene Frame von 80 Abtastwerten einem Eingang eines 200 Hz Hochpassfilters **12**. Der Ausgang des Hochpassfilters **12** ist mit einem Eingang eines stimmlosen Sprachcodierers **14** verbunden sowie mit einem Eingang eines stimmhaften/stimmlosen Detektors **28**. Das Hochpassfilter **12** schafft Blöcke von 360 Abtastwerten zu dem stimmhaften/stimmlosen Detektor **28** und Blöcke von 160 Abtastwerten (wenn der Sprachcodierer **4** in einer 5,2 kbit/s Mode arbeitet) oder 240 Abtastwerten (wenn der Sprachcodierer **4** in einer 3,2 kbit/s Mode arbeitet) zu dem stimmlosen Sprachcodierer **14**. Die Beziehung zwischen den jeweiligen oben genannten Blöcken von Abtastwerten und dem Ausgang des Puffers **11** ist in der nachfolgenden Tafel angegeben.

Element	5,2 kbit/s		3,2 kbit/s	
	#Abtastwerte	Start	# Abtastwerte	Start
Hochpassfilter 12	80	320	80	320
Stimmhaft/stimmloser Detektor 28	360	0..40	360	0..40
Stimmhafter Sprachcodierer 16	400	0	400	0
Stimmloser Sprachcodierer 14	160	120	240	120
Aktuelles zu codierendes Frame	80	160	80	160

[0031] Der stimmhafte/stimmlose Detektor **28** bestimmt, ob das aktuelle Frame stimmhafte oder stimmlose Sprache aufweist und bietet das Ergebnis als stimmhafter/stimmloser Merker an. Dieser Merker wird einem Multiplexer **22**, dem stimmlosen Sprachcodierer **14** und dem stimmhaften Sprachcodierer **16** zugeführt. Je

nach dem Wert des stimmhaften stimmlosen Merkers wird der stimmhafte Codierer **16** oder der stimmlose Sprachcodierer **14** aktiviert.

[0032] In dem stimmhaften Sprachcodierer **16** wird das Eingangssignal als eine Anzahl harmonisch relatierter sinusförmiger Signale dargestellt. Der Ausgang des stimmhaften Sprachcodierers schafft einen Tonhöhenwert, einen Verstärkungswert und eine Darstellung von 16 Prädiktionsparametern. Der Tonhöhenwert und der Verstärkungswert werden entsprechenden Eingängen eines Multiplexers **22** zugeführt.

[0033] In der 5,2 kbit/s Mode wird die LPC-Berechnung alle 10 ms durchgeführt. In der 3,2 kbit/s Mode wird die LPC-Berechnung alle 20 ms durchgeführt, ausgenommen, wenn es zwischen stimmloser und stimmhafter Sprache oder umgekehrt einen Übergang gibt. Wenn ein derartiger Übergang auftritt, wird in der 3,2 kbit/s Mode die LPC-Berechnung ebenfalls alle 10 ms durchgeführt.

[0034] Die LPC-Koeffizienten an dem Ausgang des stimmhaften Sprachcodierers werden von einem Huffman-Codierer **24** codiert. Die Länge der Huffman-codierten Sequenz wird verglichen mit der Länge der entsprechenden Eingangssequenz, und zwar durch eine Vergleichsschaltung in dem Huffman-Codierer **24**. Wenn die Länge der Huffman-codierten Sequenz länger ist als die Eingangssequenz, wird entschieden, die nicht codierte Sequenz zu übertragen. Sonst wird entschieden, die Huffman-codierte Sequenz zu übertragen. Die genannte Entscheidung wird durch ein "Huffman-Bit" dargestellt, das einem Multiplexer **26** sowie einem Multiplexer **22** zugeführt wird. Der Multiplexer **26** ist vorgesehen um die Huffman-codierte Sequenz oder die Eingangssequenz dem Multiplexer **22** zuzuführen, und zwar in Abhängigkeit von dem Wert des "Huffman-Bits". Die Verwendung des "Huffman-Bits" in Kombination mit dem Multiplexer **26** hat den Vorteil, dass gewährleistet wird, dass die Länge der Darstellung der Prädiktionskoeffizienten einen vorbestimmten Wert nicht übersteigt. Ohne die Verwendung des "Huffman-Bits" und des Multiplexers **26** könnte es passieren, dass die Länge der Huffman-codierten Sequenz die Länge der Eingangssequenz derart übersteigt, dass die codierte Sequenz überhaupt nicht mehr in das Übertragungsframe passt, in dem eine begrenzte Anzahl Bits für die Übertragung der LPC-Koeffizienten reserviert ist.

[0035] In dem stimmlosen Sprachcodierer **14** werden ein Verstärkungswert und 6 Prädiktionskoeffizienten bestimmt um das stimmlose Sprachsignal darzustellen. Die 6 LPC-Koeffizienten werden durch einen Huffman-Codierer **18** codiert, der an dem Ausgang eine Huffman-codierte Sequenz und ein "Huffman-Bit" anbietet. Die Huffman-codierte Sequenz und die Eingangssequenz des Huffman-Codierers **18** werden einem Multiplexer **20** zugeführt, der durch das "Huffman-Bit" gesteuert wird. Die Wirkungsweise der Kombination des Huffman-Codierers **18** und des Multiplexers **20** ist der Gleiche wie die Wirkungsweise des Huffman-Codierers **24** und des Multiplexers **20**.

[0036] Das Ausgangssignal des Multiplexers **20** und das "Huffman-Bit" wird den entsprechenden Eingängen des Multiplexers **22** zugeführt. Der Multiplexer **22** ist dazu vorgesehen, das codierte stimmhafte Sprachsignal oder das codierte stimmlose Sprachsignal zu selektieren, und zwar je nach der Entscheidung des stimmhaften-stimmlosen Detektors **28**. An dem Ausgang des Multiplexers **22** ist das codierte Sprachsignal verfügbar.

[0037] In dem stimmhaften Sprachcodierer **16** nach Fig. 3 sind die Analysenmittel nach der vorliegenden Erfindung durch den "LPC Parameter Computer" **30**, den "Refined Pitch Computer" **32** und den "Pitch Estimator" **38** gebildet. Das Sprachsignal $s[n]$ wird einem Eingang des LPC-Parametercomputers **30** zugeführt. Der LPC-Parametercomputer **30** bestimmt die Prädiktionskoeffizienten $a[i]$, die quantisierten Prädiktionskoeffizienten $aq[i]$, erhalten nach Quantisierung, Codierung und Decodierung $a[i]$ und LPC-Codes $C[i]$, wobei i Werte von 0–15 haben kann.

[0038] Die Tonhöhenbestimmungsmittel nach der vorliegenden Erfindung umfassen Ausgangs-Tonhöhenbestimmungsmittel, in diesem Fall einen Tonhöhenschätzer **38**, und Tonhöhenabstimmungsmittel, in diesem Fall einen "Pitch Range Computer" **34** und einen "Refined Pitch Computer" **32**. Der Tonhöhenschätzer **38** bestimmt einen groben Tonhöhenwert, der in dem Pitch Range Computer **34** verwendet wird zum Bestimmen der Tonhöhenwerte, die in den Tonhöhenabstimmungsmitteln versucht werden sollen, die weiterhin als "Refined Pitch Computer" **32** bezeichnet wird, zur Bestimmung des schlussendlichen Tonhöhenwertes. Der Tonhöhenschätzer **38** schafft eine grobe Tonhöhenperiode, ausgedrückt in einer Anzahl Abtastwerte. Die in dem Refined Pitch Computer **32** zu verwendenden Tonhöhenwerte werden durch den Computer **34** aus der groben Tonhöhenperiode entsprechend der nachfolgenden Tafel ermittelt.

Grobe Tonhöhenperiode p	Frequenz (Hz)	Suchbereich	Schrittgröße	# Kandidaten
$20 \leq p \leq 39$	400...200	$p-3 \dots p+3$	0,25	24
$40 \leq p \leq 79$	200...100	$p-2 \dots p+2$	0,25	16
$80 \leq p \leq 200$	100...40	P	1	1

[0039] In dem Amplitudenspektrum-Computer **36** wird ein gefensteretes Sprachsignal S_{HAM} aus dem Signal $s[i]$ ermittelt, und zwar entsprechend:

$$S_{\text{HAM}}[i - 120] = w_{\text{HAM}}[i] \cdot s[i]$$

(1)

[0040] In (1) ist $w_{\text{HAM}}[i]$ gleich:

$$w_{\text{HAM}} = 0,54 - 0,46 \cos \left\{ \frac{2\pi((i + 0,5) - 120)}{160} \right\}; 120 \leq i < 280 \quad (2)$$

[0041] Das gefensterete Sprachsignal $s_{\text{HAM}}[i]$ ist in die Frequenzdomäne umgeformt, und zwar unter Verwendung eines 512 Punkte FFT. Das Spektrum S_w , das durch die genannte Transformation erhalten worden ist, entspricht der nachstehenden Formel:

$$S_w[k] = \sum_{m=0}^{159} s_{\text{HAM}}[m] \cdot e^{-j2\pi km / 512} \quad (3)$$

[0042] Das in dem Refined Pitch Computer **32** zu verwendende Amplitudenspektrum wird wie folgt berechnet:

$$|S_w[k]| = \sqrt{(\Re\{S_w[k]\})^2 + (\Im\{S_w[k]\})^2} \quad (4)$$

[0043] Der Refined Pitch Computer **32** bestimmt aus den a-Parametern, die von dem LPC Parameter Computer **30** geliefert werden, und aus dem groben Tonhöhenwert einen feinen Tonhöhenwert, was zu einem minimalen Fehlersignal zwischen dem Amplitudenspektrum nach (4) und dem Amplitudenspektrum eines Signals mit einer Anzahl harmonisch relatierter sinusförmiger Signale führt, deren Amplituden durch Abtastung des LPC-Spektrums durch die genannte feine Tonhöhenperiode ermittelt wird.

[0044] In dem Verstärkungscomputer **40** wird die optimale Verstärkung zum Decken des Zielspektrums genau berechnet, und zwar aus dem Spektrum des neu synthetisierten Sprachsignals, unter Anwendung der quantisierten a-Parameter, statt der Anwendung der nicht quantisierten a-Parameter, wie in dem "Refined Pitch Computer" **32**.

[0045] An dem Ausgang des stimmhaften Sprachcodierers **40** sind die **16** LPC-Codes, die feine Tonhöhe und die von dem Verstärkungscomputer **40** berechnete Verstärkung verfügbar. Die Wirkungsweise des LPC-Parametercomputers **30** und des Refined Pitch Computers **32** werden nachstehend eingehend erläutert.

[0046] In dem LPC-Computer **30** nach **Fig. 4** wird ein Fenstervorgang an dem Signal $s[n]$ durchgeführt, und zwar durch einen Fensterprozessor **50**. Nach einem Aspekt der vorliegenden Erfindung ist die Analysenlänge abhängig von dem Wert des stimmhaften/stimmlosen Merkers. In der 5,2 kbit/s Mode wird die LPC-Berechnung alle 10 ms durchgeführt. In der 3,2 kbit/s Mode wird die LPC-Berechnung alle 20 ms durchgeführt, ausgenommen während des Übergangs von dem stimmhaften zu dem stimmlosen oder umgekehrt. Wenn es einen derartigen Übergang gibt, wird die LPC-Berechnung alle 10 ms durchgeführt.

[0047] In der nachfolgenden Tafel sind die Anzahl Abtastwerte angegeben, die an der Ermittlung der Prädiktionkoeffizienten beteiligt sind.

Bitrate und Mode	Analysenlänge N_A und erforderliche Abtastwerte	Aktualisierungsintervall
5,2 kbit/s	160 (120-280)	10 ms
3,2 kbit/s (Übergang)	160 (120-280)	10 ms
3,2 kbit/s (keinen Übergang)	240 (120-360)	20 ms

[0048] Für das Fenster in dem Fall von 5,2 kbit/s und in dem Fall von 3,2 kbit/s, wobei es einen Übergang gibt, kann Folgendes geschrieben werden:

$$w_{\text{HAM}} = 0,54 - 0,46 \cos \left\{ \frac{2\pi((i + 0,5) - 120)}{160} \right\}; 120 \leq i < 280 \quad (5)$$

[0049] Für das gefensterete Sprachsignal wird Folgendes gefunden:

$$S_{\text{HAM}}[i - 120] = w_{\text{HAM}}[i] \cdot s[i]; 120 \leq i < 280$$

(6)

[0050] Wenn es in dem 3,2 kbit/s Fall keinen Übergang gibt, wird in die Mitte des Fensters ein flacher oberer Teil von 80 Abtastwerten eingeführt, wobei das Fenster sich derart erstreckt, dass es 240 Abtastwerte überspannt, ausgehend bei dem Abtastwert 120 und endend vor dem Abtastwert 360. Auf diese Weise wird ein Fenster w'_{HAM} wie folgt erhalten:

$$w'_{\text{HAM}} = \begin{cases} w_{\text{HAM}}[i] ; 120 \leq i < 200 \\ 1 & ; 200 \leq i < 280 \\ w_{\text{HAM}}[i] ; 280 \leq i < 360 \end{cases} \quad (7)$$

für das gefenstertere Sprachsignal kann Folgendes geschrieben werden:

$$S_{\text{HAM}}[i - 120] = w_{\text{HAM}}[i] \cdot s[i]; \quad 120 \leq i < 360$$

(8)

[0051] Der "Autocorrelation Function Computer" **58** bestimmt die Autokorrelationsfunktion R_{ss} des gefenster-ten Sprachsignals. Die Anzahl zu berechnender Korrelationskoeffizienten entspricht der Anzahl Prädiktionskoeffizienten +1. Wenn es ein stimmhaftes Sprachframe gibt, ist die Anzahl zu berechnender Autokorrelationskoeffizienten **17**. Wenn es ein stimmloses Sprachframe gibt, ist die Anzahl zu berechnender Autokorrelationskoeffizienten **7**. Das Vorhandensein eines stimmhaften oder stimmlosen Sprachframes wird dem "Autocorrelation Function Computer" **58** gemeldet, und zwar durch den stimmhaften/stimmlosen Merker.

[0052] Die Autokorrelationskoeffizienten werden mit einem sog. Verzögerungsfenster gefenstert, damit eine bestimmte spektrale Glättung des durch die Autokorrelationskoeffizienten dargestellten Spektrums erhalten wird. Die geglätteten Autokorrelationskoeffizienten $\rho[i]$ werden wie folgt berechnet:

$$\rho[i] = R_{\text{sss}}[i] \cdot \exp\left(\frac{-\pi f_{\mu} i}{8000}\right); \quad 0 \leq i \leq P \quad (9)$$

[0053] In (9) ist f_{μ} die spektrale Glättungskonstante mit einem Wert von 46,4 Hz. Die gefensternten Autokorrelationswerte $\rho[i]$ werden Drain-Elektrode, "Schur recursion" Modul **62** zugeführt, das die Reflexionskoeffizienten $k[1]$ bis $k[P]$ in rekursiver Weise berechnet. Die Schur-Rekursion ist dem Fachmann durchaus bekannt.

[0054] In einem Wandler **66** werden die P Reflexionskoeffizienten $\rho[i]$ in a -Parameter umgewandelt, und zwar zur Verwendung in dem "Refined Pitch Computer" **32** in **Fig. 3**. In einem Quantisierer **64** werden die Reflexionskoeffizienten in "Log Area Ratios" umgewandelt und diese Verhältnisse werden daraufhin einheitlich quantisiert. Die resultierenden LPC-Codes $C[1] \dots C[P]$ werden dem Ausgang des LPC-Parametercomputers zur weiteren Übertragung zugeführt.

[0055] In dem örtlichen Decoder **54** werden die LPC-Codes $C[1] \dots C[P]$ durch einen Rekonstruktor **54** für Reflexionskoeffizienten in rekonstruierte Reflexionskoeffizienten $\hat{k}[i]$ umgewandelt. Daraufhin werden die rekonstruierten Reflexionskoeffizienten $\hat{k}[i]$ in (quantisierte) a -Parameter umgewandelt, und zwar durch einen Reflexionskoeffizienten-zu- a -Parameterwandler **56** umgewandelt.

[0056] Diese örtliche Decodierung wird durchgeführt, damit die gleichen a -Parameter in dem Sprachcodierer **4** und in dem Sprachdecoder **14** verfügbar sind.

[0057] In dem "Refined Pitch Computer" **32** nach **Fig. 5** bestimmt ein "Pitch Frequenz Kandidat Selektor" **70** aus der Anzahl Kandidaten, dem Startwert und der Schrittgröße, wie diese aus dem "Pitch Range Computer" **34** empfangen wurden, die in dem Refined Pitch Computer **32** zu verwendenden Kandidat-Tonhöhenwerte. Für jeden der Kandidaten bestimmt der Pitchfrequenz Kandidat Selektor **70** eine Grundfrequenz $f_{0,i}$.

[0058] Unter Verwendung der Kandidatfrequenz $f_{0,i}$ wird die von den LPC-Koeffizienten beschriebene spektrale Hülle an harmonischen Stellen durch den Spektrumlüllensampler **72** abgetastet. Für $m_{i,k}$, wobei es sich um die Amplitude der k . Harmonischen des i . Kandidaten $f_{0,i}$ handelt, kann Folgendes geschrieben werden:

$$m_{i,k} = \left| \frac{1}{A(z)} \right|_{z=2\pi k \cdot f_{0,i}} \quad (10)$$

[0059] In (10) ist $A(z)$ gleich:

$$A(z) = 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} + \dots + a_p \cdot z^{-P}$$

(11)

[0060] Mit $z = e^{j\Theta_{i,k}} = \cos\Theta_{i,k} + j \cdot \sin\Theta_{i,k}$ und $\Theta_{i,k} = 2\pi k f_{0,i}$ ändert sich (11) in die nachfolgende Formel:

$$A(z)|_{\Theta=\Theta_{i,k}} = 1 + a_1(\cos\Theta_{i,k} + j \cdot \sin\Theta_{i,k}) + \dots + a_p(\cos\Theta_{p,k} + j \cdot \sin\Theta_{p,k})$$

(12)

[0061] Wenn (12) in einen Realteil und einen imaginären Teil aufgeteilt wird, können die Amplituden $m_{i,k}$ wie folgt erhalten werden:

$$m_{i,k} = \frac{1}{\sqrt{R^2(\Theta_{i,k}) + I^2(\Theta_{i,k})}} \quad (13)$$

wobei

$$R(\theta_{i,k}) = 1 + a_1(\cos\theta_{i,k}) + \dots + a_p(\cos\theta_{i,k})$$

(14)

[0062] Und

$$I(\theta_{i,k}) = 1 + a_1(\sin\theta_{i,k}) + \dots + a_p(\sin\theta_{i,k})$$

(15)

[0063] Das Kandidatspektrum $|\hat{S}_{w,i}|$ wird dadurch ermittelt, dass die Spektrallinien $m_{i,k}$ ($1 \leq k \leq L$) mit einer Spektralfensterfunktion W gefaltet werden, welcher der 8192 Punkt FFT der 160 Punkte Hamming Fensters nach (5) und (7) ist, und zwar abhängig von der aktuellen Betriebsart des Codierers. Es ist ersichtlich, dass die 8192 Punkte FFT vorberechnet werden und dass das Ergebnis in ROM gespeichert werden kann. In dem Faltungsprozess wird ein Herunter-Abtastvorgang durchgeführt, weil das Kandidatspektrum mit 256 Punkten des Bezugsspektrums verglichen werden soll, was die Berechnung von mehr als 256 Punkten nutzlos macht. Folglich kann für $|\hat{S}_{w,i}|$ Folgendes geschrieben werden:

$$|\hat{S}_{w,i}[f]| = \sum_{k=1}^L m_{i,k} \cdot W(16 \cdot f - k \cdot f_{0,i}) ; 0 \leq f < 256 \quad (16)$$

[0064] Der Ausdruck (16) gibt nur die allgemeine Form des Amplitudenspektrums für den Tomhöhenkandidaten i , nicht aber die Amplitude. Folglich soll das Spektrum $|\hat{S}_{w,i}|$ um einen Verstärkungsfaktor g_i korrigiert werden, der durch einen MSE-Verstärkungsberechner **78** wie folgt berechnet wird:

$$g_i = \frac{\sum_{j=0}^{256} S_w[j] \cdot \hat{S}_{w,i}[j]}{\sum_{j=0}^{256} (S_w[j])^2} \quad (17)$$

[0065] Es ist ein Verstärker **82** vorgesehen zum Skalieren des Spektrums $|\hat{S}_{w,i}|$ mit dem Verstärkungsfaktor g_i . Ein Subtrahierer **84** berechnet die Differenz zwischen den Koeffizienten des Zielspektrums, wie von dem Amplitudenspektrumcomputer **36** und dem Ausgang des Multiplizierers **82** bestimmt. Daraufhin berechnet ein summierender Quadrierer ein quadriertes Fehlersignal E_i entsprechend:

$$E_i = E(f_{0,i}) = \sum_{j=0}^{255} (|S_w[j]| - g_i \cdot |\hat{S}_{w,i}[j]|)^2 \quad (18)$$

[0066] Die Kandidat-Grundfrequenz $f_{0,i}$, die dazu führt, dass der minimale Wert für die feine Grundfrequenz oder für die feine Tonhöhe selektiert wird. In dem Codierer nach der vorliegenden Erfindung sind insgesamt 368 Tonhöhenperioden möglich, was zum Codieren 9 Bits erfordert. Die Tonhöhe wird alle 10 ms aktualisiert, unabhängig von der Mode des Sprachcodierers. In dem Verstärkungsrechner **40** nach **Fig. 3** wird die zu dem Decoder zu übertragende Verstärkung auf dieselbe Art und Weise berechnet, wie oben anhand der Verstär-

kung g_i beschrieben worden ist, aber nun werden die quantisierten a-Parameter statt der nicht quantisierten a-Parameter verwendet, die bei der Berechnung der Verstärkung g_i benutzt werden. Der zu dem Decoder zu übertragende Verstärkungsfaktor nicht Linear in 6 Bits quantisiert, so dass für geringe Werte von g_i kleine Quantisierungsschritte verwendet werden und für größere Werte von g_i größere Quantisierungsschritte verwendet werden.

[0067] In einem stimmlosen Sprachcodierer **14** nach **Fig. 6** ist die Wirkungsweise des LPC-Parametercomputers **82** ähnlich wie die Wirkungsweise des LPC-Parametercomputers **30** nach **Fig. 4**. Der LPC-Parametercomputer **82** funktioniert mit dem hochpassgefilterten Sprachsignal statt des ursprünglichen Sprachsignals, wie dies durch den LPC-Parametercomputer **30** gemacht wird. Weiterhin ist die Prädiktionsordnung des LPC-Computers **82** 6 statt 16, wie in dem LPC-Parameterpitchcomputer **30**.

[0068] Der Zeitdomäne-Fensterprozessor **84** berechnet ein Hanning gefensteretes Sprachsignal entsprechend:

$$s_w[n] = s[n] \cdot \left(0,5 - 0,5 \cos\left(\frac{2 \cdot \pi(i + 0,5) - 120}{160}\right) \right); 120 \leq i < 280 \quad (19)$$

[0069] In einem RMS-Wertcomputer **86** wird ein Mittelwert g_{uv} der Amplitude eines Sprachsignals berechnet, und zwar entsprechend:

$$g_{uv} = \frac{1}{4} \sqrt{\frac{1}{N} \sum_{i=0}^{159} s_w^2[i]} \quad (20)$$

[0070] Der zu dem Decoder zu übertragende Verstärkungsfaktor g_{uv} ist nicht linear quantisiert in 5 Bits, so dass für geringe Werte von g_{uv} kleine Quantisierungsschritte angewandt werden, und für größere Werte von g_{uv} werden größere Quantisierungsschritte angewandt. Von dem stimmlosen Sprachcodierer **14** werden keine Anregungsparameter bestimmt.

[0071] In dem Sprachdecoder **14** nach **Fig. 7** werden einem Huffmann-Decoder **90** die Huffman-codierten LPC-Codes und ein stimmhafter/stimmloser Merker zugeführt. Der Huffmann-Decoder **90** ist vorgesehen zum Decodieren der Huffman-codierten LPC-Codes entsprechend der von dem Huffman-Codierer **18** verwendeten Huffman-Tafel, wenn der stimmhafte/stimmlose Merker ein stimmloses Signal angibt. Der Huffmann-Decoder **90** ist vorgesehen zum Decodieren der Huffman-codierten LPC-Codes entsprechend der Huffman-Tafel, die von dem Huffman-Codierer **24** angewandt wird, wenn der stimmhafte/stimmlose Merker ein stimmhaftes Signal angibt. In Abhängigkeit von dem Wert des Huffman-Bits werden die empfangenen LPC-Codes von dem Huffmann-Decoder **90** decodiert. Der Verstärkungswert und der feine Tonhöhenwert werden ebenfalls dem Demultiplexer **92** zugeführt.

[0072] Wenn der stimmhafte/stimmlose Merker ein stimmhaftes Sprachframe angibt, werden die feinen Tonhöhen-codes, die Verstärkungscodes und die **16** LPC-Codes einem harmonischen Sprachsynthesizer **94** zugeführt. Wenn der stimmhafte/stimmlose Merker ein stimmloses Sprachframe angibt, werden die Verstärkungscodes und die **6** LPC-Codes einem stimmlosen Sprachsynthesizer **96** zugeführt. Das synthetisierte stimmhafte Sprachsignal $\hat{s}_{v,k}[n]$ an dem Ausgang des harmonischen Sprachsynthesizers **94** und das synthetisierte Sprachsignal $\hat{s}_{uv,k}[n]$ an dem Ausgang des stimmlosen Sprachsynthesizers **96** werden entsprechenden Eingängen eines Multiplexers **98** zugeführt.

[0073] In der stimmhaften Mode führt der Multiplexer **98** das Ausgangssignal $\hat{s}_{v,k}[n]$ des harmonischen Sprachsynthesizers **94** dem Eingang des Überlappungs- und Zufügungs-Synthesizerblocks **100** zu. In der stimmlosen Mode führt der Multiplexer **98** das Ausgangssignal $\hat{s}_{uv,k}[n]$ des stimmlosen Sprachsynthesizers **96** dem Eingang des Überlappungs- und Zufügungs-Synthesizerblocks **100** zu. In diesem Block **100** werden sich teilweise überlappende stimmhafte und stimmlose Sprachsegmente hinzugefügt. Für das Ausgangssignal $\hat{s}[n]$ des Überlappungs-und-Zufügungssyntheseblocks **100** kann Folgendes geschrieben werden:

$$\hat{s}[n] = \begin{cases} \hat{s}_{uv,k-1}[n + N_s / 2] + \hat{s}_{uv,k}[n]; v_{k-1} = 0, v_k = 0 \\ \hat{s}_{uv,k-1}[n + N_s / 2] + \hat{s}_{v,k}[n]; v_{k-1} = 0, v_k = 1 \\ \hat{s}_{v,k-1}[n + N_s / 2] + \hat{s}_{uv,k}[n]; v_{k-1} = 1, v_k = 0 \\ \hat{s}_{v,k-1}[n + N_s / 2] + \hat{s}_{v,k}[n]; v_{k-1} = 1, v_k = 1 \end{cases} \quad (21)$$

[0074] In (21) ist N_s die Länge des Sprachframes, v_k ist der stimmhafte/stimmlose Merker für das vorhergehende Sprachframe und v_k ist der stimmhafte/stimmlose Merker für das aktuelle Sprachframe.

[0075] Das Ausgangssignal $\hat{s}[n]$ des Überlappungs-und-Zufügungsblocks wird einem Nachfilter **102** zugeführt. Das Nachfilter ist vorgesehen zum Verbessern der empfangenen Sprachqualität, indem Störung außer-

halb der formierenden Gebiete unterdrückt wird.

[0076] In dem stimmhaften Sprachdecoder **94** nach **Fig. 8** wird die von dem Demultiplexer **92** empfangene codierte Tonhöhe decodiert und von einem Tonhöhendecoder **104** in eine Tonhöhenperiode umgewandelt. Die von dem Tonhöhendecoder **104** bestimmte Tonhöhenperiode wird einem Eingang eines Phasensynthesizers **106**, einem Eingang einer harmonischen Oszillatorbank **108** und einem ersten Eingang eines LPF Spektrumhüllendenabtasters **110** zugeführt.

[0077] Die von dem Demultiplexer **92** empfangenen LPC-Koeffizienten werden von dem LPC-Decoder **112** decodiert. Die Art und Weise der Decodierung der LPC-Koeffizienten ist abhängig davon, ob das aktuelle Sprachframe stimmhafte oder stimmlose Sprache enthält. Deswegen wird der stimmhafte/stimmlose Merker einem zweiten Eingang des LPC-Decoders **112** zugeführt. Der LPC-Decoder führt die quantisierten a-Parameter einem zweiten Eingang des LPC-Spektrumhüllendenabtasters **110** zu. Die Wirkungsweise des LPC-Spektrumhüllendenabtasters **112** ist durch (13), (14) und (15) beschrieben, weil dieselbe Wirkungsweise in dem "Refined Pitch Computer" **32** durchgeführt wird.

[0078] Der Phasensynthesizer **106** ist vorgesehen zum Berechnen der Phase $\phi_k[i]$ des i. sinusförmigen Signals der L Signale, die das Sprachsignal darstellen. Die Phase $\phi_k[i]$ wird derart gewählt, dass das i. sinusförmige Signal nach wie vor von dem einen Frame zu dem nächsten Frame kontinuierlich ist. Das stimmhafte Sprachsignal wird dadurch synthetisiert, dass sich überlappende Frames, die je 160 gefensterte Abtastwerte enthalten, kombiniert werden. Zwischen zwei benachbarten Frames gibt es eine 50%ige Überlappung, wie aus der Graphik **118** und der Graphik **122** in **Fig. 9** hervorgeht. In den Graphiken **118** und **122** ist das verwendete Fenster gestrichelt dargestellt. Der Phasensynthesizer ist nun vorgesehen um an der Stelle, wo die Überlappung die größte Auswirkung hat, eine kontinuierliche Phase zu schaffen. Mit der hier verwendeten Fensterfunktion ist diese Stelle am Abtastwert **119**. Für die Phase $\phi_k[i]$ des aktuellen Frames kann nun Folgendes geschrieben werden:

$$\phi_k[i] = \phi_{k-1}[i] + i \cdot 2\pi \cdot f_{0,k-1} \frac{3N_s}{4} - i \cdot 2\pi \cdot f_{0,k} \frac{N_s}{4}; 1 \leq i \leq 100 \quad (22)$$

[0079] In dem nun beschriebenen Sprachcodierer ist der Wert von N_s 160. Für das allererste stimmhafte Sprachframe wird der Wert von $\phi_k[i]$ zu einem vorbestimmten Wert initialisiert. Die Phasen $\phi_k[i]$ werden immer aktualisiert, sogar wenn ein stimmloses Sprachframe empfangen wird. In dem genannten Fall wird $F_{0,k}$ auf 50 Hz gesetzt.

[0080] Die harmonische Oszillatorbank **108** erzeugt die vielen harmonisch relatierter Signale $\hat{s}_{v,k}[n]$, die das Sprachsignal darstellen. Diese Berechnung wird unter Anwendung der harmonischen Amplituden $\hat{m}[i]$ durchgeführt, wobei die Frequenz f_0 und die synthetisierten Phasen $\hat{\phi}[i]$ Folgendem entsprechen:

$$\hat{s}'_{v,k}[n] = \sum_{i=1}^L \hat{m}[i] \cos\{(i \cdot 2\pi \cdot f_0) \cdot n + \hat{\phi}[i]\}; 0 \leq n < N_s \quad (23)$$

[0081] Das Signal $\hat{s}_{v,k}[n]$ wird unter Verwendung eines Hanning-Fensters in dem Zeitdomänen-Fensterblock **114** gefenstert. Das gefensterte Signal ist in der Graphik **120** der **Fig. 9** dargestellt. Das Signal $\hat{s}'_{v,k+1}[n]$ wird unter Verwendung eines Hanning-Fensters, das um $N_s/2$ Abtastwerte in der Zeit verschoben ist, gefenstert. Dieses gefensterte Signal ist in der Graphik **124** in **Fig. 9** dargestellt. Die Ausgangssignale des Zeitdomänenfensterblocks **144** werden dadurch erhalten, dass die oben genannten gefensterten Signale addiert werden. Dieses Ausgangssignal ist in der Graphik **126** in **Fig. 9** dargestellt. Ein Verstärkungsdecoder **118** leitet einen Verstärkungswert g_v von dem Eingangssignal her, und das Ausgangssignal des Zeitdomänenfensterblocks **144** wird durch den genannten Verstärkungsfaktor g_v skaliert, und zwar durch den Signalskalierungsblock **116** zum Erhalten des rekonstruierten stimmhaften Sprachsignals $\hat{s}_{v,k}[n]$.

[0082] In dem stimmlosen Sprachsynthesizer **96** werden die LPC-Codes und der stimmhafte/stimmlose Merker einem LPC-Decoder **130** zugeführt. Der LPC-Decoder **130** liefert einem LPC-Synthesefilter **134** eine Vielzahl von 6 a-Parametern. Ein Ausgang eines Gausschen "White-Noise"-Generators **132** ist mit einem Eingang des LPC-Synthesefilters **143** verbunden. Das Ausgangssignal des LPC-Synthesefilters **134** wird durch ein Hanning-Fenster in dem Zeitdomänenfensterblock **140** gefenstert.

[0083] Ein stimmloser Verstärkungsdecoder **136** leitet einen Verstärkungswert \hat{g}_{uv} her, der die gewünschte Energie des vorhandenen stimmlosen Frames darstellt. Aus dieser Verstärkung und der Energie des gefensterten Signals wird ein Skalierungsfaktor \hat{g}'_{uv} für das gefensterte Sprachsignal ermittelt zum Erhalten eines Sprachsignals mit der richtigen Energie. Für diesen Skalierungsfaktor kann Folgendes geschrieben werden:

$$\hat{g}'_{uv} = \sqrt{\frac{\hat{g}_{uv}}{\sum_{n=1}^{N_s-1} (\hat{s}_{uv,k}[n] \cdot w[n])^2}} \quad (24)$$

[0084] Der Signalskalierungsblock **142** ermittelt das Ausgangssignal $\hat{s}_{uv,k}$ dadurch, dass das Ausgangssignal des Zeitdomänenfensterblocks **140** mit dem Skalierungsfaktor \hat{g}'_{uv} multipliziert wird.

[0085] Das oben beschriebene Sprachcodierungssystem kann modifiziert werden zum Liefern einer niedrigeren Bitrate oder einer höheren Sprachqualität. Ein Beispiel eines Sprachcodierungssystems, das eine niedrigere Bitrate erfordert ist ein 2 kbit/s Codierungssystem. Ein derartiges System kann dadurch erhalten werden, dass die Anzahl Prädiktionskoeffizienten, die für stimmhafte Sprache erforderlich sind, von 16 auf 12 reduziert wird und dass differenzielle Codierung der Prädiktionskoeffizienten, der Verstärkung und der feineren Tonhöhe angewandt wird. Differenzielle Codierung bedeutet, dass die zu codierenden Daten nicht einzeln codiert werden, sondern dass nur die Differenz zwischen entsprechenden Daten von aufeinander folgenden Frames übertragen wird. Bei einem Übergang von stimmhafter zu stimmloser Sprache oder umgekehrt, werden in dem ersten neuen Frame alle Koeffizienten einzeln codiert, damit ein Startwert für die Decodierung erhalten wird.

[0086] Es ist ebenfalls möglich, einen Sprachcodierer mit einer gesteigerten Sprachqualität mit einer Bitrate von 6 kbit/s zu erhalten. Die Modifikationen sind hier die Ermittlung der Phase der ersten 8 Harmonischen der Vielzahl von harmonisch relatierten sinusförmigen Signalen. Die Phase $\phi[i]$ wird wie folgt berechnet:

$$\phi[i] = \arctan \frac{I(\Theta_i)}{R(\Theta_i)} \quad (25)$$

[0087] Darin ist $\theta_i = 2\pi f_0 \cdot i \cdot R(\theta_i)$ und $I(\theta_i)$ sind gleich:

$$R(\Theta_i) = \sum_{n=0}^{N-1} s_w[n] \cdot \cos(\Theta_i \cdot n) \quad (26)$$

und

$$I(\Theta_i) = - \sum_{n=0}^{N-1} s_w[n] \cdot \sin(\Theta_i \cdot n) \quad (27)$$

[0088] Die auf diese Art und Weise erhaltenen 8 Phasen $\phi[i]$ werden einheitlich quantisiert zu 6 Bits und in dem Ausgangsbitstrom eingeschlossen.

[0089] Eine weitere Modifikation in dem 6 kbit/s Codierer ist die Übertragung zusätzlicher Verstärkungswerte in der stimmlosen Mode. Normalerweise wird statt einmal je Frame eine Verstärkung alle 2 ms übertragen. In dem ersten Frame unmittelbar hinter einem Übergang werden 10 Verstärkungswerte übertragen, von denen 5 das aktuelle stimmlose Frame darstellen, und von denen 5 das vorhergehende stimmhafte Frame darstellen, das von dem stimmlosen Sprachcodierer verarbeitet wird. Die Verstärkungen werden aus 4 ms überlappenden Fenstern ermittelt.

[0090] Es hat sich herausgestellt, dass die Anzahl LPC Koeffizienten **12** ist und dass wo möglich differenzielle Codierung angewandt wird.

Text in der Zeichnung

Fig. 6

V/Uv Entscheidung
 Hochpass-gefilterte Sprache
82 LPC Parameter Computer
84 Zeitdomänenfenster
86 RMS-Wertcomputer
 LPC-Codes (**6**)
 Verstärkung

Fig. 2

Abgetastete Daten
12 Hochpassfilter
28 stimmhaft/stimmlos Detektor
 V/Uv Entscheidung
 DC Kerbfilter
14 stimmloser Sprachcodierer
 Verstärkung

LPC-Codes
Huffman Codierer
Huffman Bit
Feine Tonhöhe
Verstärkung
LPC-Codes
16 Stimmhafter Sprachcodierer
24 Huffman Coder
Huffman Bit

Fig. 3

V/Uv Entscheidung
30 LPC Parameter Computer
32 Analyse durch Synthese feinere Tonhöhe Computer
Zielspektrum
36 Amplitudenspektrumcomputer
38 Tonhöhenschätzer
LPC Codes
Kandidaten
Schrittgröße
Startwert
34 Tonhöhenbereichscomputer
grobe Tonhöhe
40 Verstärkungscomputer
LPC Codes (**16**)
Feinere Tonhöhe
Verstärkung

Fig. 4

50 Zeitdomänenfenster
58 Autokorrelationsfunktionscomputer
60 Verzögerungsfenster
54 Reflexionskoeffizienten-Rekonstruktor
64 Reflexionskoeffizientenquantifizierer
62 Schur-Rekursion
66 Reflexionskoeffizienten zu a-Parametern
56 Reflexionskoeffizienten zu a-Parametern
a-Parameter (quantisiert)
LPC Codes
a-Parameter

Fig. 5

Kandidaten
Schrittgröße
Startwert
70 Tonhöhenfrequenz Kandidatenselektor
Zielspektrum
72 LPC Spektrumumhüllendenabtaster
76 Größenspektrumrekonstruktor
78 MSE Verstärkungsrechner
74 Minimiere MSE
Quadratfehlercomputer
Feine Tonhöhenperiode

Fig. 7

V/Uv Entscheidung

Huffman Bit
 LPC Codes
 Verstärkung
 Feine Tonhöhe
 V/Uv Entscheidung
90 Huffman-Decoder
 feine Tonhöhe
 Verstärkung
 LPC-Codes (**16**)
 Verstärkung
 LPC-Codes (**6**)
94 Harmonische Sprache Synthesizer
96 Harmonische Sprache Synthesizer
100 Überlauppungs- und Zufügungssynthese
102 Nachfilter
 synthetisierte Sprache

Fig. 8

Verstärkung
 V/Uv Entscheidung
 Feine Tonhöhenperiode
 LPC Codes (**16**)
104 LPC Decoder
106 Phasensynthesizer
108 Harmonische Oszillatorbank
110 LPC Spektralumhüllendenabtaster
112 LPC Decoder
114 Zeitdomänenfenster
116 Signalskalierung
118 Signalskalierung

Fig. 9 Frame

Fig. 10

Verstärkung
 V/Uv Entscheidung
 LPC Codes (**6**)
130 LPC Decoder
132 Gaussches weißes Rauschen Generator
134 LPC Synthesenfilter
136 stimmlose Verstärkungsdecoder
138 Verstärkungsdeckung
140 Zeitdomänenfenster
142 Signalskalierung

Patentansprüche

1. Übertragungssystem mit einem Sender mit einem Sprachcodierer mit Analysenmitteln zum periodischen Ermitteln von Analysenkoeffizienten aus dem Sprachsignal, so dass der Sender Übertragungsmittel aufweist zum Übertragen der genannten Analysenkoeffizienten über ein Übertragungsmedium zu einem Empfänger, so dass der genannte Empfänger einen Sprachdecoder aufweist mit Rekonstruktionsmitteln zum Herleiten eines rekonstruierten Sprachsignals auf Basis der Analysenkoeffizienten, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln, in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment oder umgekehrt, und dass die Rekonstruktionsmittel dazu vorgesehen sind, ein rekonstruiertes Sprachsignal auf Basis der öfter ermittelten Analysenkoeffizienten herzuleiten.

2. Übertragungssystem nach Anspruch 1, dadurch gekennzeichnet, dass der Sprachcodierer einen stimm-

haften Sprachcodierer zum Codieren stimmhafter Sprachsegmente aufweist und dass der Sprachcodierer einen stimmlosen Sprachcodierer zum Codieren stimmloser Sprachelemente aufweist.

3. Übertragungssystem nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln für zwei Segmente nach dem Übergang.

4. Übertragungssystem nach Anspruch 1, 2 oder 3, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, die Frequenz der Ermittlung der Analysenkoeffizienten bei einem Übergang zwischen einem stimmhaften und einem stimmlosen Segment und umgekehrt zu verdoppeln.

5. Übertragungssystem nach Anspruch 4, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, alle 20 ms die Analysenkoeffizienten zu ermitteln, wenn kein Übergang stattfindet, und dass die Analysenmittel dazu vorgesehen sind, alle 10 ms die Analysenkoeffizienten zu ermitteln, wenn ein Übergang stattfindet.

6. Sender mit einem Sprachcodierer mit Analysenmitteln zum periodischen Ermitteln von Analysenkoeffizienten aus dem Sprachsignal, so dass der Sender Übertragungsmittel aufweist zum Übertragen der genannten Analysenkoeffizienten, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment und umgekehrt.

7. Empfänger zum Empfangen eines codierten Sprachsignals mit einer Anzahl Analysenkoeffizienten, so dass der genannte Empfänger einen Sprachdecoder aufweist mit Rekonstruktionsmitteln zum Herleiten eines rekonstruierten Sprachsignals auf Basis von Analysenkoeffizienten, extrahiert aus dem empfangenen Signal, dadurch gekennzeichnet, dass das codierte Sprachsignal die Analysenkoeffizienten öfter trägt in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsignal und einem stimmlosen Sprachsignal oder umgekehrt, und dass die Rekonstruktionsmittel dazu vorgesehen sind, ein rekonstruiertes Sprachsignal herzuleiten, und zwar auf Basis der öfter verfügbaren Analysenkoeffizienten.

8. Sprachcodieranordnung mit Analysenmitteln zum periodischen Ermitteln von Analysenkoeffizienten aus dem Sprachsignal, dadurch gekennzeichnet, dass die Analysenmittel dazu vorgesehen sind, die Analysenkoeffizienten öfter zu ermitteln in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment und umgekehrt.

9. Sprachdecoderanordnung zum Decodieren eines codierten Sprachsignals mit einer Anzahl Analysenkoeffizienten, so dass die genannte Sprachdecoderanordnung Rekonstruktionsmittel aufweist zum Herleiten eines rekonstruierten Sprachsignals auf Basis von Analysenkoeffizienten, extrahiert aus dem empfangenen Signal, dadurch gekennzeichnet, dass das codierte Sprachsignal die Analysenkoeffizienten öfter trägt in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment und umgekehrt, und dass die Rekonstruktionsmittel dazu vorgesehen sind, ein rekonstruiertes Sprachsignal herzuleiten, und zwar auf Basis der öfter verfügbaren Analysenkoeffizienten.

10. Sprachcodierverfahren, wobei dieses Verfahren die nachfolgenden Verfahrensschritte umfasst: das periodische Ermitteln von Analysenkoeffizienten aus dem Sprachsignal, dadurch gekennzeichnet, dass das Verfahren weiterhin das Ermitteln von Analysenkoeffizienten umfasst, und zwar öfter in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachelement oder umgekehrt.

11. Sprachdecodierverfahren zum Decodieren eines codierten Sprachsignals mit einer Anzahl Analysenkoeffizienten, so dass das genannte Verfahren das Herleiten eines rekonstruierten Sprachsignals umfasst, und zwar auf Basis von Analysenkoeffizienten, extrahiert aus dem empfangenen Signal, dadurch gekennzeichnet, dass das codierte Sprachsignal die Analysenkoeffizienten öfter in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment und umgekehrt trägt, und dass Herleitung des rekonstruierten Sprachsignals auf Basis öfter verfügbarer Analysenkoeffizienten durchgeführt wird.

12. Codiertes Sprachsignal mit einer Anzahl Analysenkoeffizienten, periodisch in das codierte Sprachsignal eingeführt, dadurch gekennzeichnet, dass das codierte Sprachsignal die Analysenkoeffizienten öfter in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment oder umgekehrt trägt.

13. Fühlbares Medium mit einem Computerprogramm zum Durchführen eines Sprachcodierverfahrens mit

periodischer Ermittlung von Analysenkoeffizienten aus dem Sprachsignal, dadurch gekennzeichnet, dass das Verfahren das öftere Ermitteln der Analysenkoeffizienten in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegmentes und eines stimmlosen Sprachsegmentes oder umgekehrt umfasst.

14. Fühlbares Medium mit einem Computerprogramm zum Durchführen eines Sprachdecodierverfahrens zum Decodieren eines Sprachsignals mit einer Anzahl Analysenkoeffizienten, so dass das genannte Verfahren das Herleiten eines rekonstruierten Sprachsignals auf Basis von Analysenkoeffizienten, extrahiert aus dem empfangenen Signal umfasst, dadurch gekennzeichnet, dass das codierte Sprachsignal die Analysenkoeffizienten öfter trägt in der Nähe eines Übergangs zwischen einem stimmhaften Sprachsegment und einem stimmlosen Sprachsegment oder umgekehrt, und dass Herleitung des rekonstruierten Sprachsignals auf Basis der öfter verfügbarer Analysenkoeffizienten durchgeführt wird.

Es folgen 9 Blatt Zeichnungen

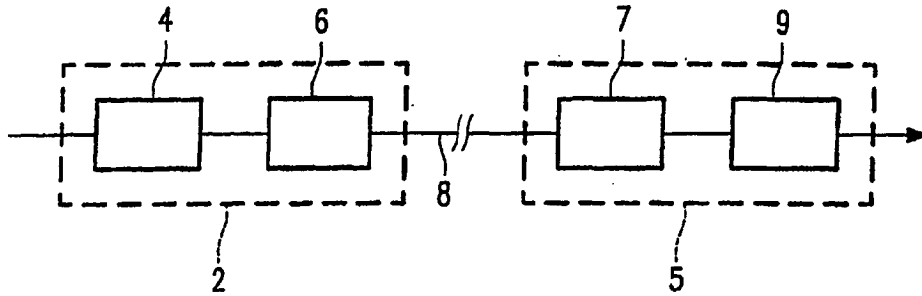


FIG. 1

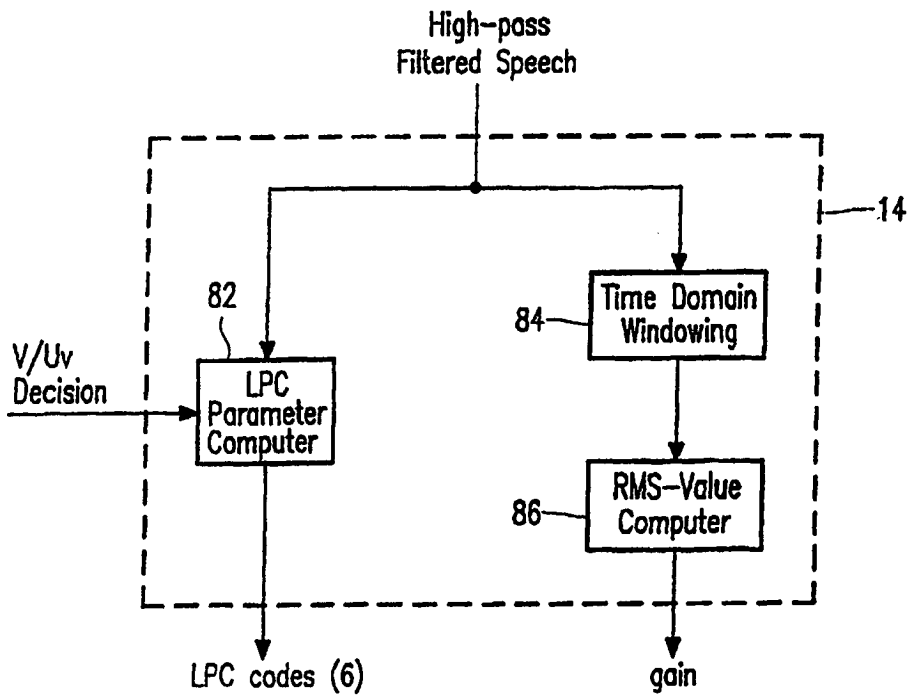


FIG. 6

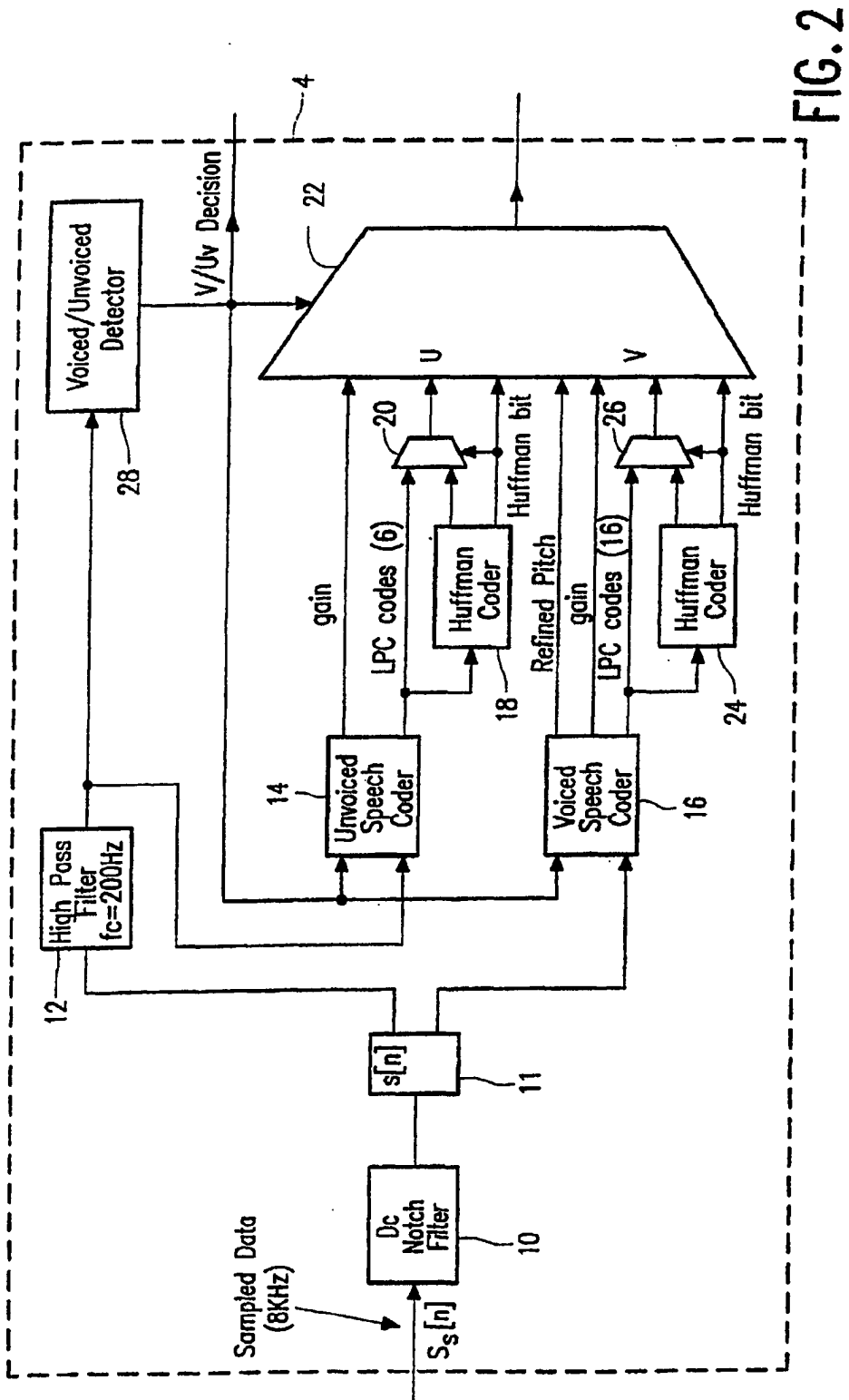


FIG. 2

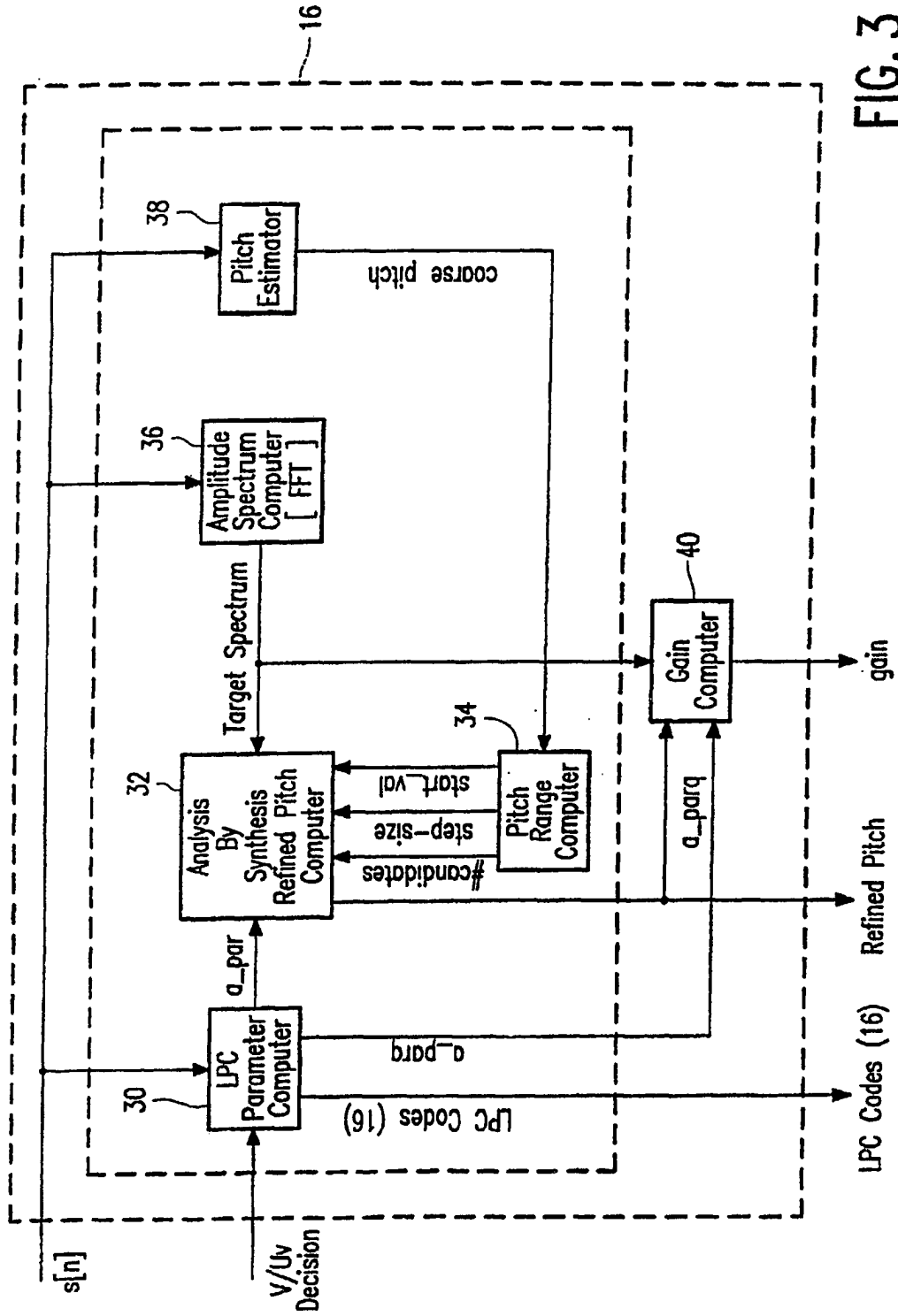


FIG. 3

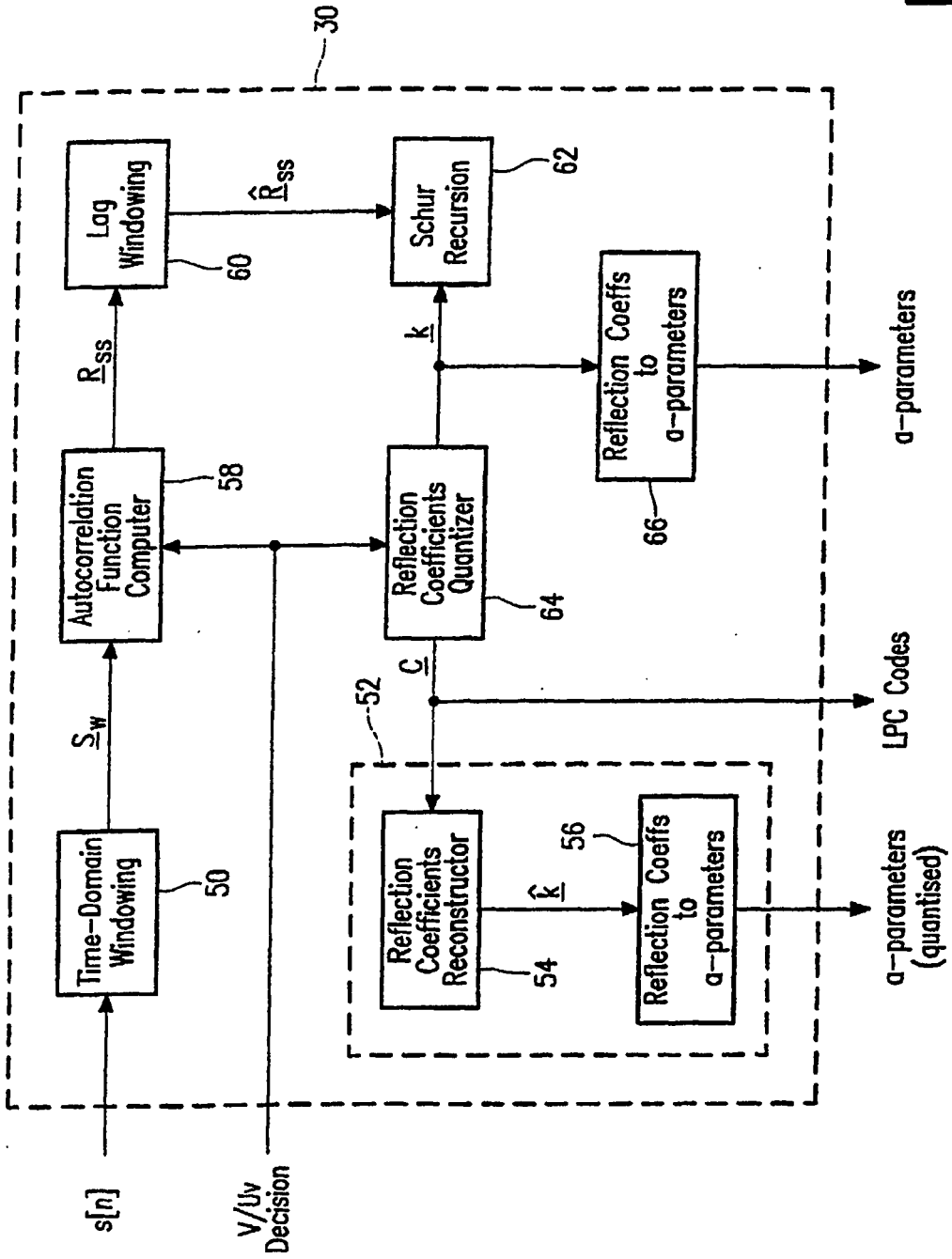


FIG. 4

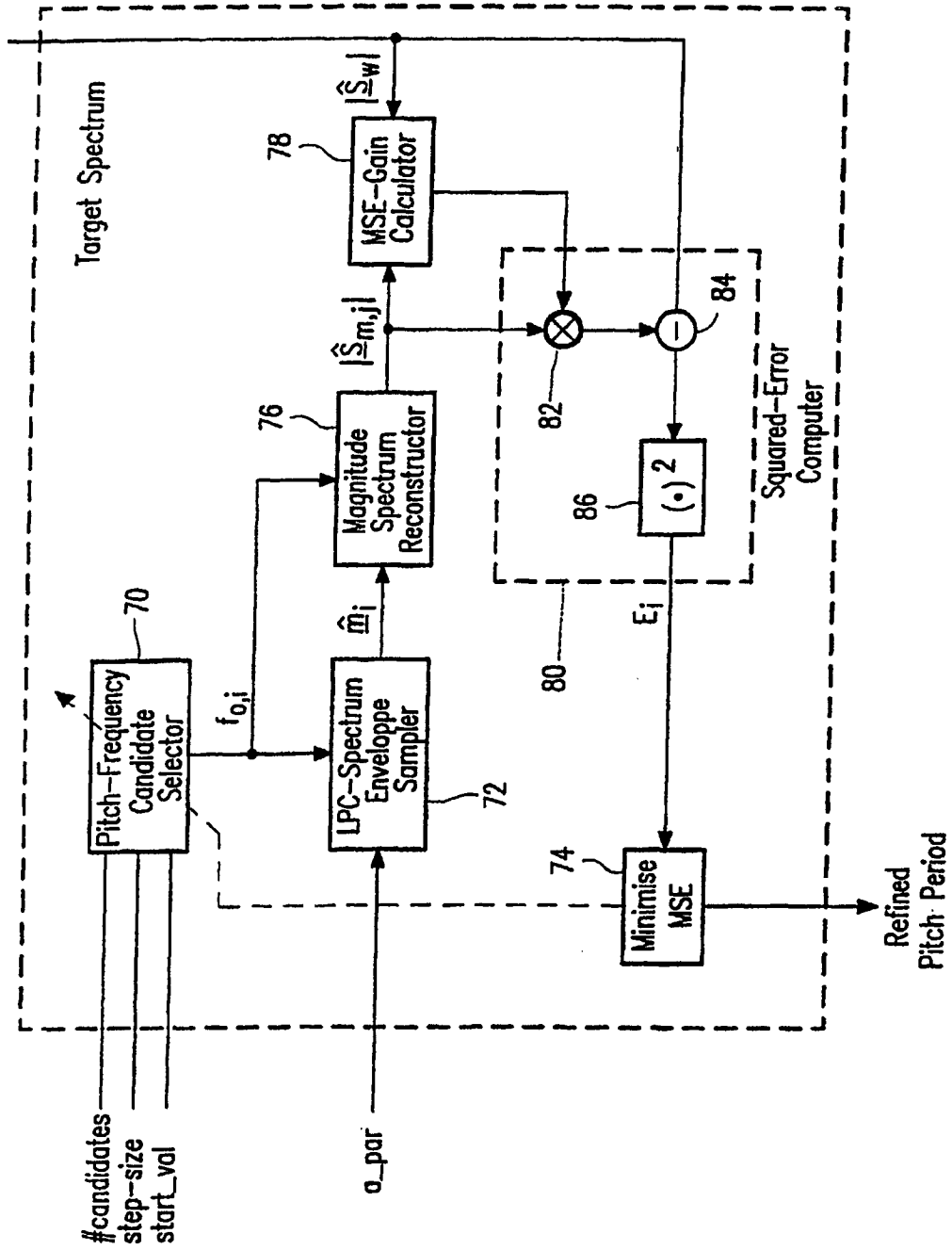


FIG. 5

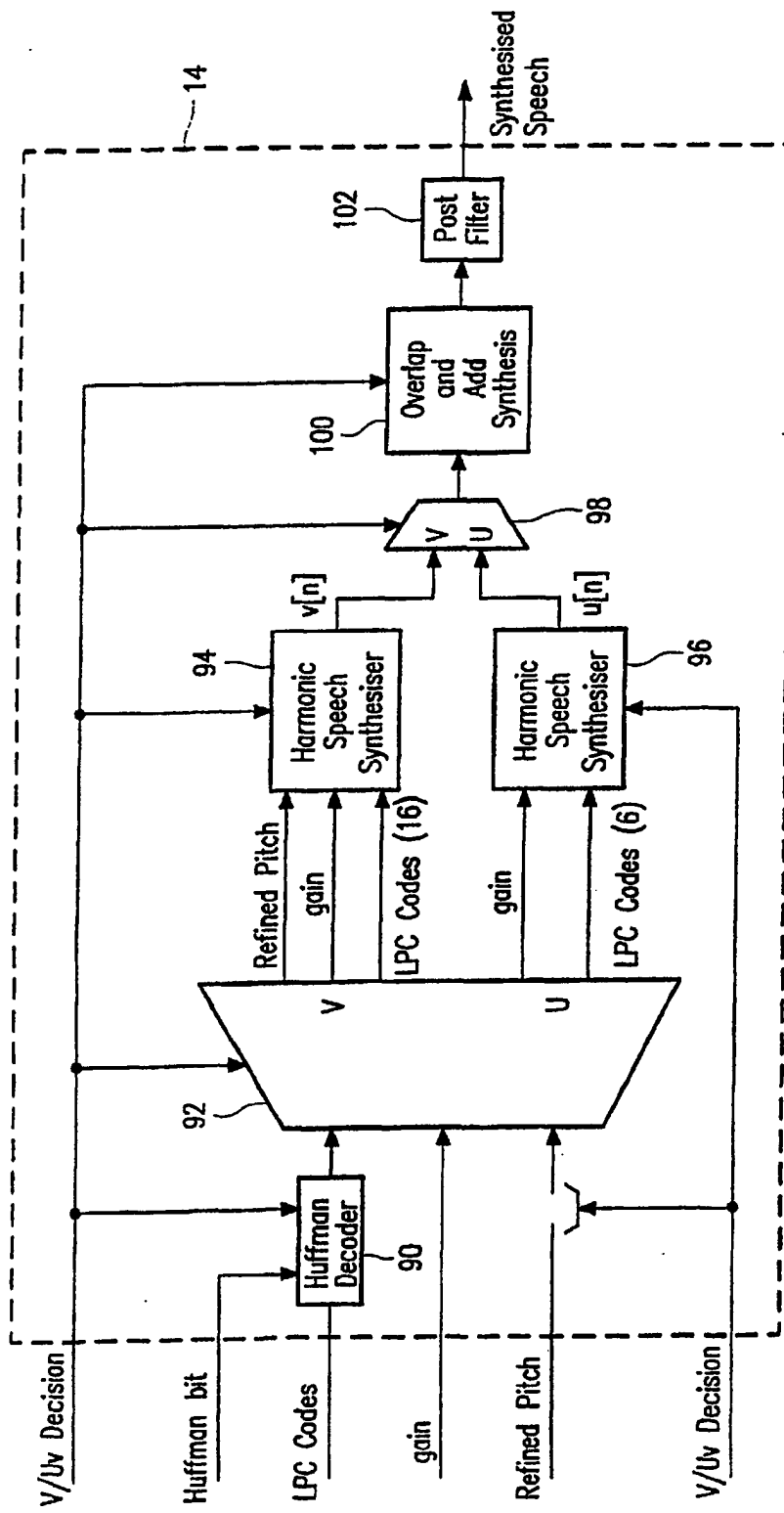


FIG. 7

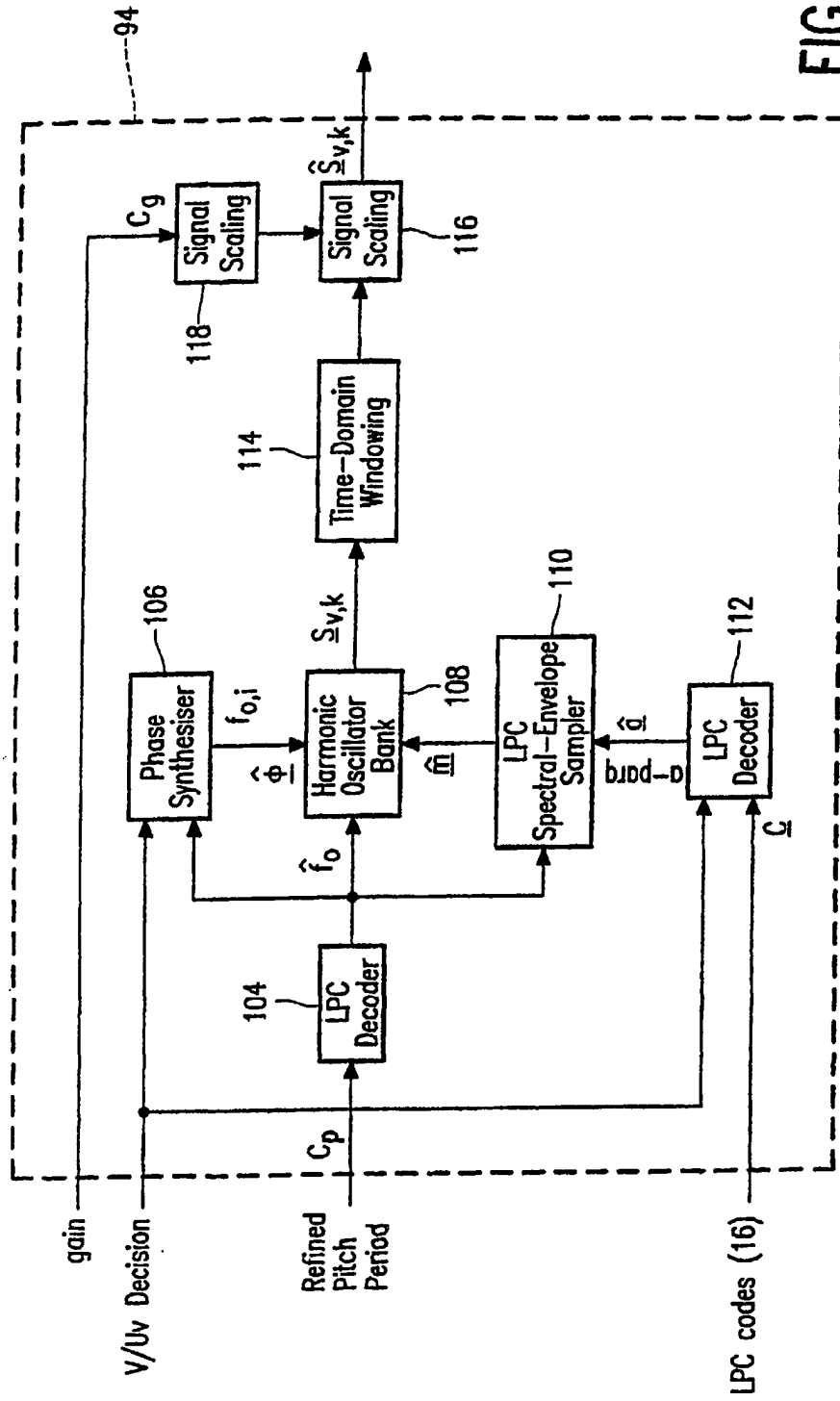


FIG. 8

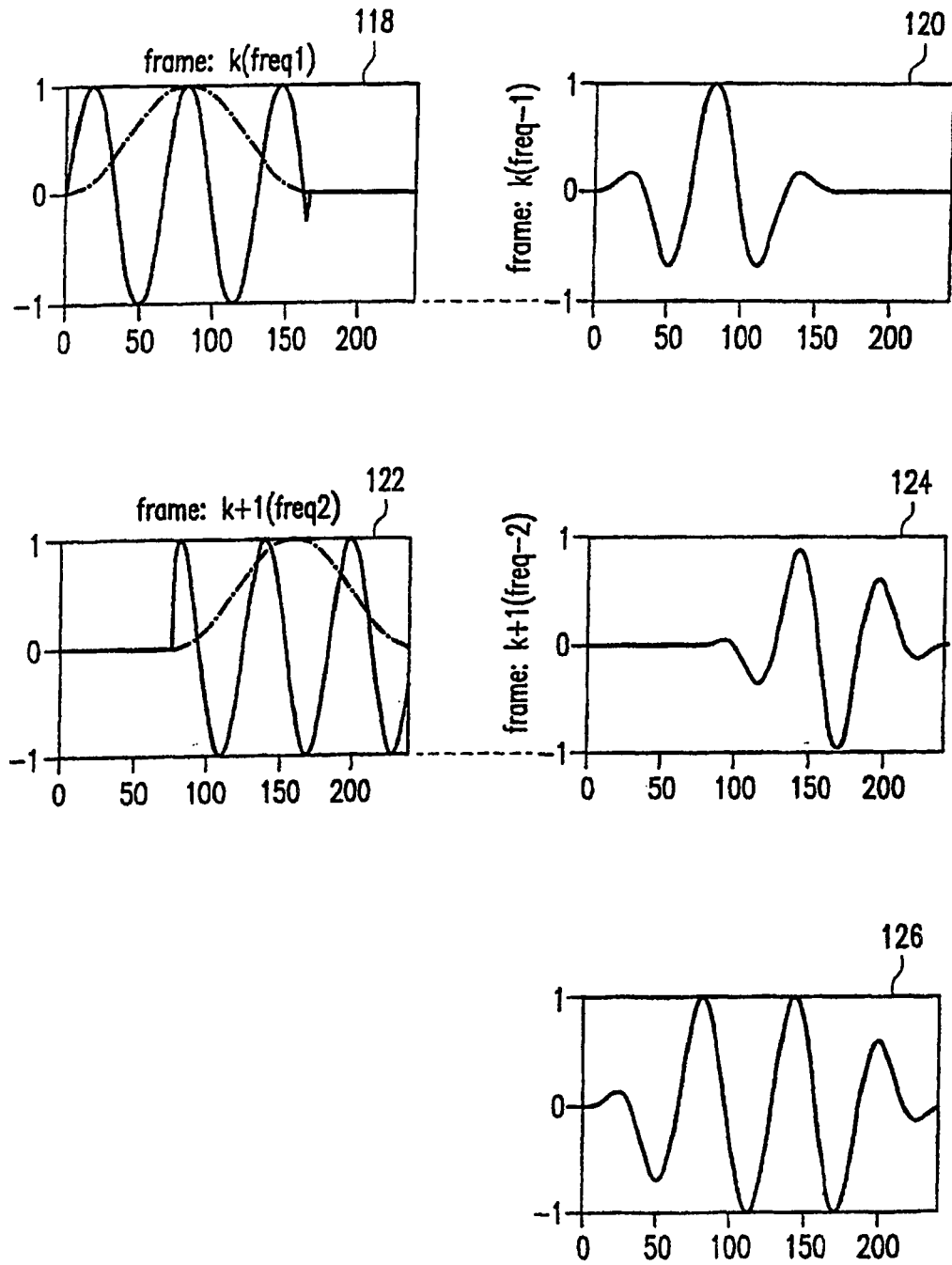


FIG. 9

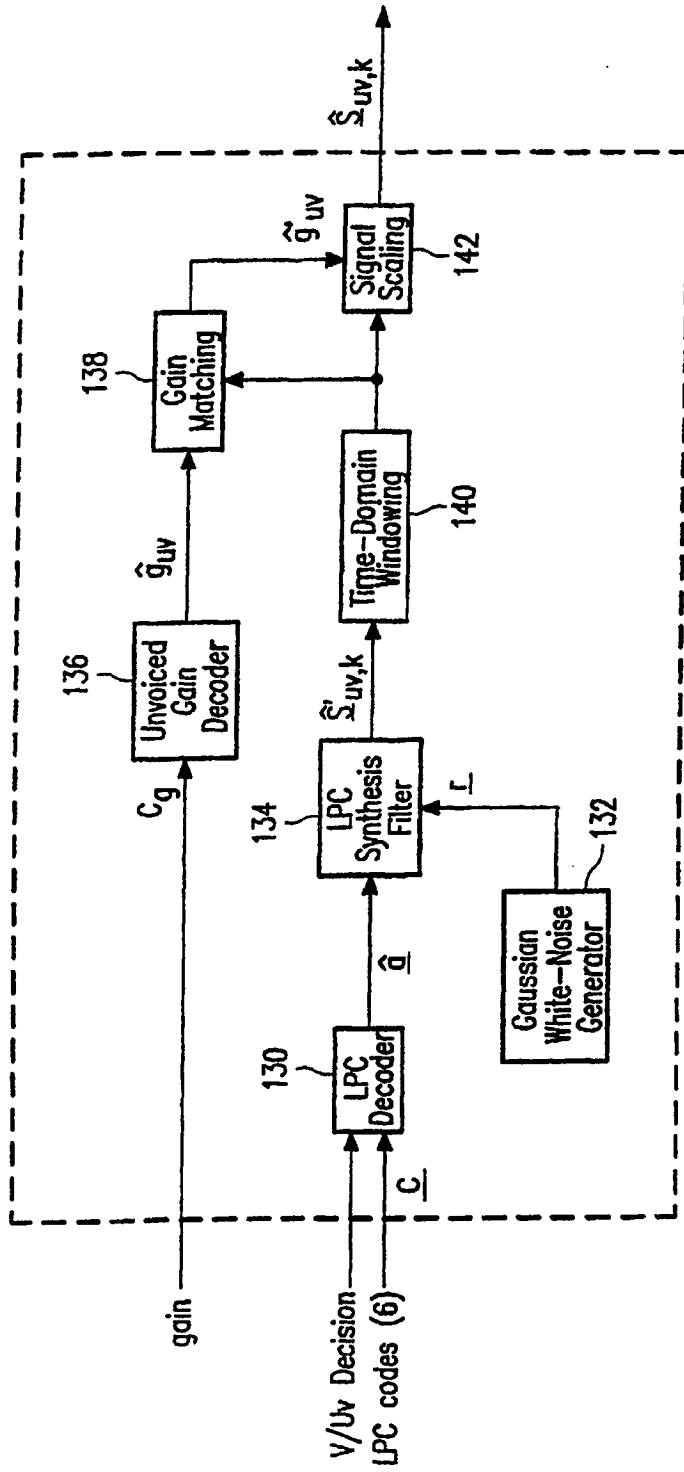


FIG. 10