



(11) **EP 1 619 665 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
08.09.2010 Bulletin 2010/36

(51) Int Cl.:
G10L 19/06^(2006.01)

(21) Application number: **05015989.6**

(22) Date of filing: **22.07.2005**

(54) **Voice coding apparatus and method using PLP in mobile communications terminal**

Verfahren, Vorrichtung zur Sprachkodierung in einem mobilen Kommunikationsendgerät mittels PLP

Méthode et appareil de codage de la parole dans un terminal mobile de communication utilisant la PLP

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI
SK TR**

(30) Priority: **23.07.2004 KR 2004057739**

(43) Date of publication of application:
25.01.2006 Bulletin 2006/04

(73) Proprietor: **LG Electronics, Inc.
Seoul 150-721 (KR)**

(72) Inventor: **Kim, Chan-Woo
Ilsan-Gu,
Goyang,
Gyeonggi-Do (KR)**

(74) Representative: **Katérlé, Axel et al
Wuesthoff & Wuesthoff
Patent- und Rechtsanwälte
Schweigerstraße 2
81541 München (DE)**

(56) References cited:
EP-A- 1 199 812 US-A1- 2004 128 130

- **GUNAWAN W ET AL: "PLP coefficients can be quantized at 400 bps" 2001 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). SALT LAKE CITY, UT, MAY 7 - 11, 2001, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 6, 7 May 2001 (2001-05-07), pages 77-80, XP010803089 ISBN: 0-7803-7041-4**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 619 665 B1

Description**BACKGROUND OF THE INVENTION**

1. Field of the Invention

[0001] The present invention relates to a coding of a mobile communications terminal, and particularly; to a voice coding apparatus and method using a Perceptual Linear Prediction (PLP).

2. Background of the Related Art

[0002] As mobile communication techniques are developed, mobile communications terminals have provided data communications using numbers, characters, symbols, and the like, and multimedia communications including various image signals as well as voice communications. A plurality of terminal users receive radio channels allocated thereto from a system and transmit and receive required data using radio resources. However, the radio channels have limited bandwidths in order for the plurality of users to use the radio channels at the same time, and accordingly a data bit rate of each user is deservedly limited.

[0003] Therefore, a coding technique has been proposed for transmitting a greater amount of data using above limited data bit rate. Various methods exist as the related art voice coding technique, each of which has several advantages at a certain bit rate.

[0004] For instance, a speech coding using a generic audio coding, a Pulse Code Modulation (PCM), and an Adaptive Delta Pulse Code Modulation (ADPCM) are effectively used at a high-bit rate over 16Kbps, and a Code Excited Linear Prediction (CELP) and other various variations are effectively used at a medium-bit rate at a range of 2.4Kbps to 16Kbps. In particular, a coding method using LD-CELP, CS-ACELP, VSELP and MELP and a wideband speech coding can be used at the medium-bit rate. Also, a Linear Predictive Coding (LPC), Residual Excited Linear Predictive (RELP), formants vocoder and Cepstral vocoder have many advantages at a low-bit rate at a range of 75bps to 2.4Kbps see for example "PL8 Coefficients can be quantized at 400 BPS" from Gunavan et al. in ICASSP 2001, for the parallel use of CS Acelp for speech coding and PLP for speech recognition.

[0005] Thus, in the related art and the present invention as claimed in claims 1 and 5, a method for improving the LPC among coding methods used at the low-bit rate will now be explained.

[0006] Fig. 1 illustrates a structure of the related art LPC encoder.

[0007] As illustrated in the drawing, the related art LPC encoder includes: a correlator 10 for calculating an autocorrelation value $r_x[n]$ of an input signal $x[n]$; an LP coefficient calculator 11 for calculating an LP coefficient a_L and a gain G by processing the autocorrelation value $r_x[n]$; a V/UV determining unit 12 for determining whether the input signal $x[n]$ is a voiced V signal or a unvoiced UV signal; a pitch calculator 13 for calculating a pitch P of the corresponding signal when the input signal $x[n]$ is the voice V signal; a parameter coding unit 14 for outputting a bit stream by coding the LP coefficient a_n , the gain G and the pitch P received from the LP coefficient calculator 11 and the pitch calculator 13 according to a V/UV indication bit outputted from the V/UV determining unit 12.

[0008] An operation of the related art LPC encoder having such construction will now be explained.

[0009] First, the correlator 10 autocorrelates an input signal $x[n]$. The LP coefficient calculator 11 processes an autocorrelation value $r_x[n]$ calculated by the correlator 10 so as to calculate an LP coefficient a_n and a gain G . At this time, the V/UV determining unit 12 determines whether the input signal $x[n]$ is a voiced V signal or a unvoiced UV signal to output a V/UV indication bit, and then outputs only the voiced V signal. The pitch calculator 13 calculates a pitch P of the voiced V signal which is outputted from the V/UV determining unit 12.

[0010] Accordingly, when the V/UV indication bit indicates the voiced V signal, the parameter coding unit 14 outputs a bit stream by coding (encoding by a low-bit rate) the LP coefficient a_n , the gain G , and the pitch P received from the LP coefficient calculator 11 and the pitch calculator 13. Afterwards, a controller (not shown) processes the bit stream to thusly output it to a radio (wireless) unit (not shown). The radio unit converts the signal outputted from the control unit into a radio (wireless) signal and transmits the converted radio signal.

[0011] Thus, in the related art, a mobile communications terminal performs the LPC coding to transmit an audio signal by a low-bit rate. However, in the related art LPC coding, a linear predication coefficient is generally used, which does not consider human auditory sensing features. Therefore, for the related art LPC coding operated using the low-bit rate, a compression efficiency is not very high (i.e., 1200Kbps to 2400Kbps) and good sound quality can not be obtained.

SUMMARY OF THE INVENTION

[0012] Therefore, an object of the present invention is to provide a voice coding apparatus and method of a mobile communications terminal capable of improving compression efficiency and sound quality by performing an LPC coding

using a PLP coefficient.

[0013] To achieve these and other advantages and in accordance with the purpose of the present invention, as embodied and broadly described herein, there is provided a Linear Predictive Coding (LPC) encoder of a mobile communications terminal comprising: a Perceptual Linear Prediction (PLP) coefficient calculator for calculating a PLP coefficient and a gain by processing an input signal; a V/UV determining unit for determining whether the input signal is a voiced signal or a unvoiced signal, and thusly outputting the determination signal and the voiced signal when the input signal is the voiced signal; a pitch calculator for calculating a pitch of the input signal outputted from the V/UV determining unit; and a parameter coding unit for performing a low-bit rate coding using the PLP coefficient, the gain, and the pitch on the basis of the determination signal.

[0014] To achieve these and other advantages and in accordance with the purpose of the present invention, as embodied and broadly described herein, there is provided a low-bit rate voice coding method of a mobile communications terminal comprising: calculating a Perceptual Linear Prediction (PLP) coefficient and a gain by processing an input signal; determining whether the input signal is a voiced signal and a unvoiced signal, and thereby outputting a determination bit value and the voiced signal when the input signal is determined as the voiced signal; calculating a pitch of the input signal outputted from a V/UV determining unit; and performing a low-bit rate coding using the PLP coefficient, the gain and the pitch on the basis of the determination bit value.

[0015] Preferably, the voiced signal is a speech signal.

[0016] Preferably, the PLP coefficient has about a 7th degree for a 8 kHz sampling rate.

[0017] The foregoing and other objects, features, aspects and advantages of the present invention will become more apparent from the following detailed description of the present invention when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute a part of this specification, illustrate embodiments of the invention and together with the description serve to explain the principles of the invention.

[0019] In the drawings:

- Fig. 1 illustrates a structure of a related art LPC encoder using an LP coefficient;
- Fig. 2 illustrates an LPC encoder using a PLP coefficient according to the present invention; and
- Fig. 3 illustrates sequential steps, in detail, of calculating a PLP coefficient in Fig. 2.

DETAILED DESCRIPTION OF THE INVENTION

[0020] Reference will now be made in detail to the preferred embodiments of the present invention, examples of which are illustrated in the accompanying drawings.

[0021] The present invention provides a low-bit rate voice coding using a Perceptual Linear Prediction (PLP) capable of performing a coding of a degree (an order) lower than that of a Linear Predictive Coding (LPC) in order to perform a voice coding having high compressibility.

[0022] First, a difference between the PLP and the LP will now be explained.

[0023] The LP is classically well-known, so that a detailed derived formula therefor will not be described. The LP basically refers to obtaining a LP coefficient a_k so that a Mean Squared Error (MSE), namely, a value of $e[n]$ can be a minimum value according to Formula (1) as follows.

$$e[n] = x[n] - \hat{x}[n] = \sum_{k=0}^{N_{pred}} a_k x[n-k] \text{ -----Formula (1)}$$

[0024] The obtained LP coefficient a_k has about 8th to 12th degrees (orders) for a 8 kHz sampling rate. Therefore, the obtained LP coefficient a_k is used for various coding methods (e.g., LPC, CELP, MELP, RELP, etc) using a Linear Prediction (LP), which is disclosed in more detail in Speech coding and synthesis, Amsterdam, the Netherlands: Elsevier, 1995.

[0025] The PLP was introduced on a paper of Hermansky in 1990 for the first time. The PLP uses human auditory sensing features similar to the existing Mel-Frequency Cepstral Coefficient (MFCC). Therefore, the present invention performs a low-bit rate voice coding using the PLP coefficient in stead of using the LP coefficient upon performing the LPC for a low-bit rate.

[0026] That is, the present invention obtains spectrum using the PLP coefficient. The PLP coefficient reflects a human auditory effect. Accordingly, in aspect of the MSE, a greater error may occur in the spectrum using the PLP coefficient than using the LP. However, the spectrum using the PLP coefficient may have a less error when considering the auditory effect. Also, for coefficient transmissions, in case of LPC, for a typical 8kHz sampling rate, transmissions of about a 10th degree (order) are used, but for PLP, transmissions of about a 7th degree (order) are used, thus the bit rate can be lowered.

[0027] Fig. 2 illustrates a construction of an LPC encoder using the PLP coefficient according to the present invention.

[0028] Referring to the Fig. 2, an LPC encoder using the PLP coefficient is constructed as same as the related art LPC encoder shown in Fig. 1, except of which the correlator 10 is not included and a PLP coefficient calculator 20 replaces the LP coefficient calculator 11.

[0029] The PLP coefficient calculator 20 processes a speech signal $S[n]$ to calculate a PLP coefficient a_p and a gain G in which the auditory effect is considered.

[0030] An operation of the LPC encoder using the PLP coefficient having such construction according to the present invention will now be explained with reference to the accompanying drawing.

[0031] First, the PLP coefficient calculator 20 receives the speech signal $S[n]$, so as to calculate the PLP coefficient a_p and the gain G by sequentially performing operations shown in Fig. 3.

[0032] That is, the PLP coefficient calculator 20 performs a fast Fourier transform (FFT) of the input signal, namely, the speech signal $S[n]$. A critical-bank integration and resampling processing is performed for the Fourier-transformed speech signal to thusly remove noise components from the speech signal $S[n]$ by a frequency unit.

[0033] Once removing the noise components, the PLP coefficient calculator 20 performs equalizing and loudness processing of the Fourier-transformed speech signal into sound components having magnitudes appropriate for human auditory sensing, and then the speech signal is matched with an output power to allow listening by humans.

[0034] When the power matching is completed, the PLP coefficient calculator 20 performs an inverse discrete Fourier transform of the corresponding speech signal to thereafter obtain a set of Linear equations from the corresponding speech signal. Therefore, the PLP coefficient calculator 20 performs a Cepstral Recursion processing for the set of Linear equations, and thus outputs Cepstral Coefficients of a PLP model, namely, the PLP coefficients a_p . In other words, the PLP coefficient calculator 20 outputs to the parameter coding unit 23 a low degree (order) of the PLP coefficients a_p and a gain G reflecting the human auditory sensing features as parameter values.

[0035] At this time, the V/UV determining unit 21 outputs a V/UV Indication bit and transfers the speech signal $S[n]$ to the pitch calculator 22. The pitch calculator 22 calculates a pitch P of the speech signal $S[n]$.

[0036] Accordingly, the parameter coding unit 23 outputs a bit stream by coding (encoding by a low-bit rate) the V/UV Indication bit value, the PLP coefficient a_p , the gain G and the pitch P received from the PLP coefficient calculator 20 and the pitch calculator 22. Preferably, a degree of the transmitted PLP coefficient a_p is about a 7th degree for a 8 kHz sampling rate. Afterwards, a controller (not shown) processes the bit stream and then outputs the processed bit stream to a radio (wireless) unit (not shown). The radio unit converts the signal outputted from the controller into a radio signal (wireless signal) and transmits it.

[0037] As described above, in the present invention, the LPC is performed by using the PLP coefficient, and thus a compressibility can be improved and voice-grade signal can be transmitted by a more efficient low-bit rate.

[0038] In addition, in the present invention, a higher compressibility can be realized and a quality of signal with high sound quality can be expected by using the PLP coefficient as a parameter rather than using the existing LP coefficient.

[0039] Therefore, the voice coding apparatus and method according to the present invention can be used for coding and decoding voice using a low-bit rate, or be used for a device which takes up a small area and performs a voice synthesis using PLP parameters.

[0040] Furthermore, the voice coding apparatus and method according to the present invention can be used for a speech coding for an application as much as a voice itself is not very important but enough to hear. Also, an effective voice conversation can be performed on the Internet which stores data by a high compressibility or requires a low-bit rate in an embedded system with a limited memory.

[0041] As the present invention may be embodied in several forms without departing from the spirit or essential characteristics thereof, it should also be understood that the above-described embodiments are not limited by any of the details of the foregoing description, unless otherwise specified, but rather should be construed broadly within its spirit and scope as defined in the appended claims, and therefore all changes and modifications that fall within the metes and bounds of the claims, or equivalence of such metes and bounds are therefore intended to be embraced by the appended claims.

Claims

1. A voice coding apparatus in a mobile communications terminal comprising:

a Perceptual Linear Prediction (PLP) coefficient calculator (20) for calculating a PLP coefficient and a gain by processing an input signal;
 a V/UV determining unit (21) for determining whether the input signal is a voiced signal or an unvoiced signal, and thus outputting a determination result and the voiced signal when the input signal is the voiced signal;
 a pitch calculator (22) for calculating a pitch of the input signal outputted from the V/UV determining unit (21); and
 a parameter coding unit (23) for performing a low-bit rate coding using the PLP coefficient, the gain, and the pitch on the basis of the determination result.

2. The apparatus of claim 1, wherein the voiced signal is a speech signal.

3. The apparatus of claim 1, wherein the determination result denotes a bit value for whether the input signal is the voiced signal or the unvoiced signal.

4. The apparatus of claim 1, wherein a degree of the PLP coefficient is about a 7th degree for a 8 kHz sampling rate.

5. A voice coding method of a mobile communications terminal comprising:

calculating a Perceptual Linear Prediction (PLP) coefficient and a gain by processing an input signal;
 determining whether the input signal is a voiced signal or an unvoiced signal,
 and thereby outputting a determination signal and the voiced signal when the input signal is determined as the voiced signal;
 calculating a pitch of the Input signal outputted from a V/UV determining unit; and
 performing a low-bit rate coding using the PLP coefficient, the gain and the pitch on the basis of the determination signal.

6. The method of claim 5, wherein the voiced signal is a speech signal.

7. The method of claim 5, wherein the step of calculating the PLP coefficient and the gain comprises:

performing a fast Fourier transform (FFT) for the input signal;
 performing a critical-bank integration and resampling of the Fourier transformed speech signal to thus remove noise components by a frequency unit;
 performing equalizing and loudness processing of the Fourier-transformed speech signal into sound components having magnitudes appropriate for human auditory sensing, and then matching the speech signal with an appropriate output power;
 performing an inverse discrete Fourier transform of the speech signal matched with the output power, and thereby obtaining a set of linear equations; and
 performing a cepstral recursion processing for the set of linear equations, and thereby obtaining a PLP coefficient and a gain.

8. The method of claim 5, wherein a degree of the PLP coefficient is about a 7th degree for a 8 kHz sampling rate.

Patentansprüche

1. Sprachkodierungsvorrichtung in einem mobilen Kommunikationsendgerät, die umfasst:

- eine PLP-, d. h. Perceptual Linear Prediction, Koeffizientenrecheneinheit (20) zum Berechnen eines PLP-Koeffizienten und einer Verstärkung durch Verarbeiten eines Eingangssignals,
- eine Stimmhaft/Stimmlos-Bestimmungseinheit (21) zum Bestimmen, ob das Eingangssignal ein stimmhaftes Signal oder ein stimmloses Signal ist, und damit zum Ausgeben eines Bestimmungsergebnisses und des stimmhaften Signals, wenn das Eingangssignal ein stimmhaftes Signal ist,
- eine Tonhöherecheneinheit (22) zum Berechnen einer Tonhöhe des von der Stimmhaft/Stimmlos-Bestimmungseinheit (21) ausgegebenen Eingangssignals, und
- eine Parameterkodiereinheit (23) zum Durchführen einer Kodierung mit niedriger Bitrate unter Verwendung des PLP-Koeffizienten, der Verstärkung und der Tonhöhe auf Basis des Bestimmungsergebnisses.

2. Vorrichtung nach Anspruch 1,

bei der das stimmhafte Signal ein Sprachsignal ist.

3. Vorrichtung nach Anspruch 1,
bei der das Bestimmungsergebnis einen Bitwert bezeichnet, der Auskunft darüber gibt, ob das Eingangssignal ein
stimmhaftes Signal oder ein stimmloses Signal ist.

4. Vorrichtung nach Anspruch 1,
bei der die Ordnung des PLP-Koeffizienten bei einer Abtastrate von 8 kHz ungefähr eine siebte Ordnung ist.

5. Sprachkodierungsverfahren eines mobilen Kommunikationsendgeräts, das umfasst:

- Berechnen eines PLP, d. h. Perceptual Linear Prediction, Koeffizienten und einer Verstärkung durch Verarbeiten eines Eingangssignals,
- Bestimmen, ob das Eingangssignal ein stimmhaftes Signal oder ein stimmloses Signal ist, und Ausgeben des Bestimmungssignals und des stimmhaften Signals, wenn bestimmt wird, dass das Eingangssignal ein stimmhaftes Signal ist,
- Berechnen einer Tonhöhe des von einer Stimmhaft/Stimmlos-Bestimmungseinheit ausgegebenen Eingangssignals, und
- Durchführen einer Kodierung mit niedriger Bitrate unter Verwendung des PLP-Koeffizienten, der Verstärkung und der Tonhöhe auf Basis des Bestimmungssignals.

6. Verfahren nach Anspruch 5,
bei dem das stimmhafte Signal ein Sprachsignal ist.

7. Verfahren nach Anspruch 5,
bei dem der Schritt des Berechnens des PLP-Koeffizienten und der Verstärkung umfasst:

- Durchführen einer schnellen Fourier-Transformation (FFT) für das Eingangssignal,
- Durchführen einer Critical-Bank-Integration und erneuten Abtastung des Fourier-transformierten Sprachsignals, um dadurch Rauschkomponenten mittels einer Frequenzeinheit zu entfernen,
- Durchführen einer Entzerrungs- und Lautstärkeverarbeitung des Fourier-transformierten Sprachsignals, um Tonkomponenten mit Stärken zu erhalten, die für die menschliche akustische Wahrnehmung geeignet sind, und anschließendes Anpassen des Sprachsignals an eine geeignete Ausgangsleistung,
- Durchführen einer inversen diskreten Fourier-Transformation des an die Ausgangsleistung angepassten Sprachsignals, wodurch ein Satz linearer Gleichungen erhalten wird, und
- Durchführen einer cepstralen Rekursionsverarbeitung für den Satz linearer Gleichungen, wodurch ein PLP-Koeffizient und eine Verstärkung erhalten werden.

8. Verfahren nach Anspruch 5,
bei dem die Ordnung des PLP-Koeffizienten bei einer Abtastrate von 8 kHz ungefähr eine siebte Ordnung ist.

Revendications

1. Appareil de codage de la voix dans un terminal de communications mobile comprenant :

- un calculateur de coefficient de Prédiction Linéaire Perceptuelle (PLP) (20) pour calculer un coefficient de PLP et un gain par traitement d'un signal d'entrée ;
- une unité de détermination V/UV (21) pour déterminer si le signal d'entrée est un signal vocal ou un signal non vocal, et ainsi délivrer en sortie un résultat de détermination et le signal vocal lorsque le signal d'entrée est le signal vocal ;
- un calculateur de hauteur (22) pour calculer une hauteur du signal d'entrée délivré en sortie de l'unité de détermination V/UV (21) ; et
- une unité de codage de paramètres (23) pour effectuer un codage à faible débit binaire en utilisant le coefficient de PLP, le gain et la hauteur sur la base du résultat de détermination.

2. Appareil selon la revendication 1, dans lequel le signal vocal est un signal de parole.

EP 1 619 665 B1

3. Appareil selon la revendication 1, dans lequel le résultat de détermination dénote une valeur de bits pour savoir si le signal d'entrée est le signal vocal ou le signal non vocal.

4. Appareil selon la revendication 1, dans lequel un degré du coefficient de PLP est environ un 7^{ème} degré pour un taux d'échantillonnage de 8 kHz.

5. Procédé de codage de la voix d'un terminal de communications mobile comprenant :

le calcul d'un coefficient de Prédiction Linéaire Perceptuelle (PLP) et d'un gain par traitement d'un signal d'entrée ;

la détermination de savoir si le signal d'entrée est un signal vocal ou un signal non vocal, et ainsi la délivrance en sortie d'un signal de détermination et du signal vocal lorsque le signal d'entrée est déterminé comme le signal vocal ;

le calcul d'une hauteur du signal d'entrée délivré en sortie d'une unité de détermination V/UV ; et

l'exécution d'un codage à faible débit binaire en utilisant le coefficient de PLP, le gain et la hauteur sur la base du signal de détermination.

6. Procédé selon la revendication 5, dans lequel le signal vocal est un signal de parole.

7. Procédé selon la revendication 5, dans lequel l'étape de calcul du coefficient de PLP et du gain comprend :

l'exécution d'une transformée de Fourier rapide (FFT) pour le signal d'entrée ;

l'exécution d'une intégration de bande critique et d'un rééchantillonnage du signal de parole à transformée de Fourier pour ainsi supprimer des composantes de bruit par une unité de fréquence ;

l'exécution d'un traitement d'égalisation et de correction physiologique du signal de parole à transformée de Fourier en composantes de son ayant des grandeurs appropriées à une détection auditive humaine, et ensuite la mise en correspondance du signal de parole avec une puissance de sortie appropriée ;

l'exécution d'une transformée de Fourier discrète inverse du signal de parole mis en correspondance avec la puissance de sortie, et ainsi l'obtention d'un ensemble d'équations linéaires ; et

l'exécution d'un traitement par récursion cepstrale pour l'ensemble d'équations linéaires, et ainsi l'obtention d'un coefficient de PLP et d'un gain.

8. Procédé selon la revendication 5, dans lequel un degré du coefficient de PLP est environ un 7^{ème} degré pour un taux d'échantillonnage de 8 kHz.

FIG.1
Related Art

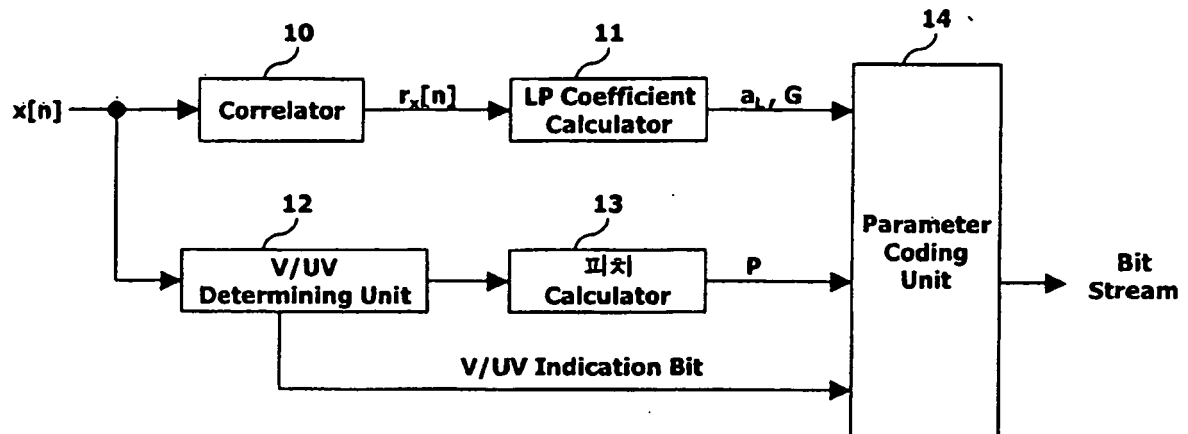


FIG.2

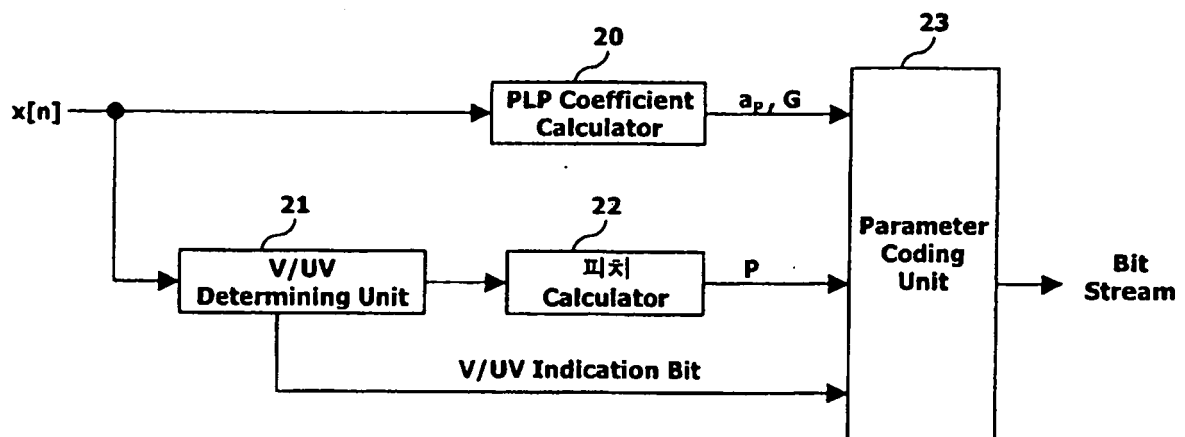
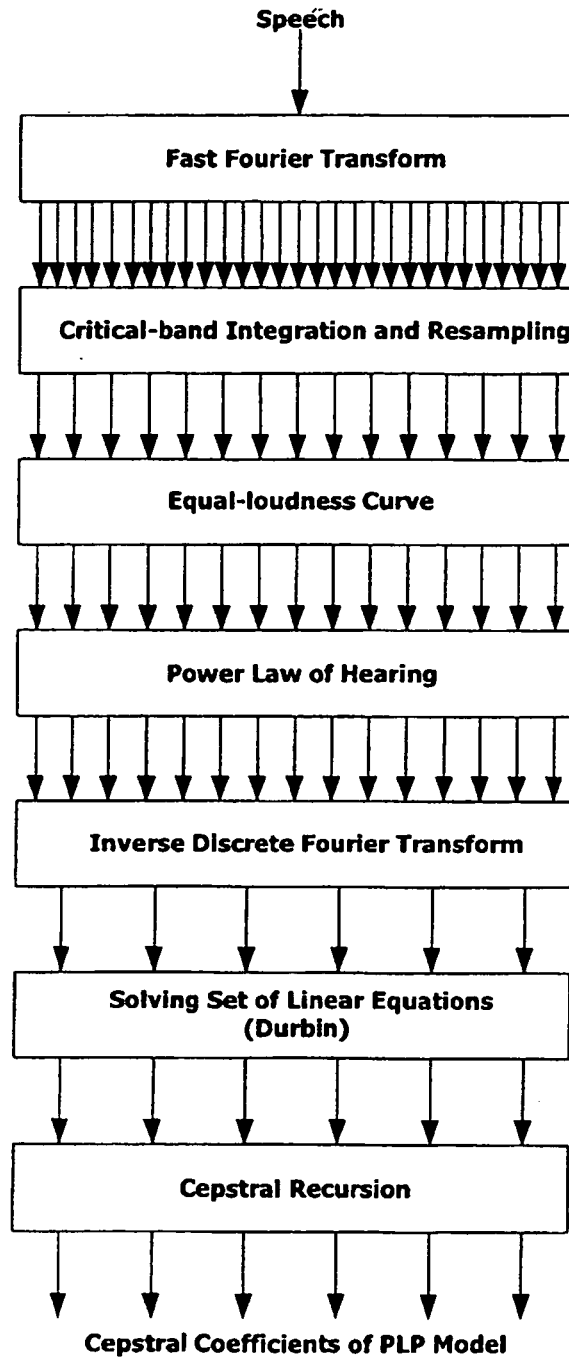


FIG.3



REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- Speech coding and synthesis. Elsevier, 1995 [0024]