



(11) **EP 1 933 306 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
18.06.2008 Bulletin 2008/25

(51) Int Cl.:
G10L 19/12^(2006.01) G10L 19/14^(2006.01)

(21) Application number: **06025955.3**

(22) Date of filing: **14.12.2006**

(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**
Designated Extension States:
AL BA HR MK RS

(71) Applicant: **Nokia Siemens Networks GmbH & Co.
KG
81541 München (DE)**

(72) Inventor: **Beaugeant, Christophe
81735 München (DE)**

(54) **Method and apparatus for transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format**

(57) The method of the present invention comprises the following steps: receiving a first coded speech signal using the first CELP format and including at least a first pitch parameter; decoding the received first coded speech signal to a decoded pcm speech signal; detecting a voiced level of the pcm speech signal within a predetermined time window; determining, if the pcm speech signal is a voiced speech signal or an unvoiced speech signal dependent on at least a first parameter; if the pcm

speech signal is voiced, performing a closed loop search process which receives at least the first pitch parameter and estimates a second pitch parameter for the second CELP format dependent on at least the first pitch parameter; and if the pcm speech signal is unvoiced, copying the first pitch parameter as the second pitch parameter for the second CELP format.

EP 1 933 306 A1

Description

[0001] The present invention relates to a method and apparatus for transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format.

[0002] In the past, several types of networks have been developed, like mobile GSM, UMTS, CTMA, IP which all provide an alternative way to the known circuit switch network. The interconnection of all these networks relates to an interoperability problem regarding transmission of speech. Indeed, non-compatible speech standards were adopted in the different networks, although most of the codecs at medium rate (5-16,5 kbit/s) are based on the same model, namely the code excited linear prediction (CELP).

[0003] Many different formats of CELP coding are in use today. In order to successfully decode a CELP-coded speech signal, the decoder must employ the same CELP coding model, in the following referred as "format", as the encoder that produced the signal. When communications systems employing different CELP formats must share speech data, it is often desirable to convert the speech signal from one CELP coding format to another.

[0004] A known method to provide interconnectivity consists in decoding one standard compressed bitstream and to re-encode it into the other standard bitstream. This known method is called the tandem transcoding. In a tandem coding system an input CELP format is converted to an output CELP format. Therefore, tandem transcoding system includes an input CELP format decoder and an output CELP format encoder. The input format CELP decoder receives a speech signal that has been encoded using one CELP format. The decoder of the tandem coding system decodes the input coded speech signal to produce a pcm speech signal. The output CELP format encoder of the tandem coding system receives the decoded pcm speech signal and encodes it using the output CELP format to produce a compressed output signal in the output CELP format. The primary disadvantage of this approach is the perceptual degradation experienced by the speech signal in passing through multiple encoders and decoders. Further, the known tandem transcoding scheme suffers from the problems of complexity and delay.

[0005] Furthermore, smart transcoding solutions are internal known by the applicant, using the fact that the different standards are based on the CELP principle. Their aim is to reduce the complexity of the transcoding as many functions at the decoder of the transcoding system could be skipped. Further, they aim to decrease the delay and enhancing the quality or at least getting the same quality as the known transcoding scheme.

[0006] The basic idea of this internal smart transcoding solution of the applicant is to use the redundancy on the standard to avoid computing parameters that were already computed. For example, it is possible to use parameters already coded at the encoder of the sending

apparatus at the encoder of the transcoding system or apparatus to drive the re-encoding. One of these parameters mapped between the speech codecs is the pitch parameter.

[0007] According to an internal solution of the applicant, the pitch mapping is provided by copying the pitch or pitch parameter from the bitstream of a first codec to the encoder of a second codec.

[0008] The pitch estimation is done in two steps in standardized CELP coding. An open-loop search gives a first estimation of the pitch T_0 . In a second step by means of a closed loop search, a closed loop pitch T_{OP} is obtained as a refinement of the pitch parameter T_0 by a search in an interval $[T_0 - T_{LOW}; T_0 + T_{HIGH}]$.

[0009] A further enhanced internal solution is to provide a mapping skipping either the open-loop search or both the closed loop-search and the open-loop search dependent on predefined parameters. For example, the pitch parameter of the first codec $T_{OP}(A)$ is taken as the output of the open-loop search so that the closed loop search at the encoder of the second codec is done in an interval around $T_{op}(A)$. In the case where the closed loop search is also skipped, the pitch $T_{op}(A)$ is directly taken as the output of the closed loop search and is quantified at the encoder of the second codec.

[0010] More advanced approaches try to estimate more accurately the pitch or pitch parameter at the encoder of the second codec given the pitch computed by the first codec. Such approaches are for example known from "An Efficient transcoding algorithm for G.713.1 and EVCR speech coders", Kyung Tae Kim and al. IEEE 54th or from "A novel scheme from EVRC to G.729AB, Pankaj K. R., 37th Asilomar Conf. On Signals, Systems and Computers, 2003. Said advanced approaches could be called "pitch smoothing" method. Therein, the open-loop pitch computation at the encoder of the second codec is driven by the pitch parameter $T_{OP}(A)$ of the first codec. There is not a direct mapping, but a research of the open-loop pitch taking into account the difference between the pitch $T_{op}(A)$ and the one computed at the encoder of the second codec at the previous frame of the bitstream.

[0011] Further, in "Improvement issues on transcoding algorithms, for the flexible usage to the various pairs of speech codec, Jin-Kyu Choi and al, ICASSP 2004" an open-loop search at the encoder of the second codec is also driven by the pitch or pitch parameter $T_{op}(A)$, by limiting the closed loop-search in a restricted interval $(T_0 - T'_{LOW}; T_0 + T'_{HIGH})$ with $T'_{LOW} < T_{LOW}$ and $T'_{HIGH} < T_{HIGH}$. All the previous mentioned solutions work either at the encoder of the second codec on the output of the open-loop search or on the output of the closed loop search.

[0012] An object of the present invention is to provide an optimal compromise for a transcoding scheme between the quality of the transmission of the speech signal and the complexity of the generating of the pitch parameter.

[0013] The above mentioned object is solved by

means of a method with the features of claim 1 and/or by means of a transcoding apparatus of claim 12.

[0014] According to the present invention, a method of transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format is provided, wherein the method comprises the following steps:

- a) receiving a first encoded speech signal using the first CELP format and including at least a first pitch parameter;
- b) decoding the received first compressed speech signal to a decoded pcm speech signal;
- c) detecting a voiced level of the pcm speech signal within a predetermined time window;
- d) determining, if the pcm speech signal is a voiced speech signal or an unvoiced speech signal dependent on at least a first parameter;
- e) if the pcm speech signal is voiced, performing a closed loop search process which receives at least the first pitch parameter and estimates a second pitch parameter for the second CELP format dependent on at least the first pitch parameter; and
- f) if the pcm speech signal is unvoiced, copying the first pitch parameter as the second pitch parameter for the second CELP format.

[0015] Further, a transcoding apparatus for transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format is provided, wherein the transcoding apparatus comprises:

- a) a receiving means which receives the first speech signal encoded using the first CELP format and including at least the first pitch parameter;
- b) a decoding means which decodes the received first compressed speech signal to the decoded pcm speech signal;
- c) a first detecting means which detects the voiced level of the pcm speech signal within the predetermined time window;
- d) a first determining means which determines if the pcm speech signal is the voiced speech signal or the unvoiced speech signal dependent on at least the first parameter;
- e) a closed loop search means which performs a closed loop search, wherein the closed loop search means receives at least the first pitch parameter and

estimates a second pitch parameter for the second CELP format dependent on at least the first pitch parameter, if the pcm speech signal is voiced; and

f) a copying means which copies the first pitch parameter as the second pitch parameter for the second CELP format, if the pcm speech signal is unvoiced.

[0016] An experiment of the applicant has shown that an accurate pitch value is necessary at encoding during voiced periods to assure a good quality of the coded signal. During unvoiced periods the CELP encoders are less sensitive to a wrong estimation of the pitch.

[0017] Voiced and unvoiced characteristics are defined by the action of the vocal cords. The vocal cords vibrate for voiced sounds, but do not vibrate for unvoiced sounds. For example, all the vowels in English are voiced sounds. Some of the consonants such as "b", "d" are partially voiced. The beginning of the phoneme [b] or [d] is plosive, the end is voiced, while "p", "f" for instance are completely unvoiced.

[0018] The estimation of the pitch parameter at the encoder of the second codec is all the more accurate the open-loop search and the closed-loop search is done.

[0019] Definitions for the closed-loop search and the open-loop search can be found in the specification of the speech codec by ITU-T or 3GPP. It can be found for instance in "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec", release 6, 3GPP TS 26.090, v6.0.0 or in "Coding of Speech at 8 kbit/s using conjugate-structure algebraic-excited linear prediction", ITU-T recommendation G.729 03/1996.

[0020] The more the process of the open-loop search and the closed-loop search is applied, the more accurate and on the other hand the more complex the estimation of the pitch parameter is. According to the present invention, in voiced periods, the closed loop search is kept, wherein during unvoiced periods, the whole closed loop search can be principally skipped. Taking into account this inventive basic principal, in a transcoding scheme, the pitch parameter $T_{OP}(A)$ is taken as the output of the open loop at the encoder of the second codec and it is possible to skip the closed loop search depended on the energy level of the decoded speech signal and the voiced level of the speech signal. In the case of an unvoiced signal, the influence of the pitch is less so that an accurate estimation of the pitch parameter for the second CELP format is not needed according to the present invention. In a transcoding scheme of the present invention, in such a case the pitch parameter of the first codec can be used as the output of the closed loop of the encoder of the second codec.

[0021] Advantageously, according to the present invention, an optimal compromise between the quality of the transmission of the speech signal and the complexity of the generating of the pitch parameter of the encoder

of the second codec is provided.

[0022] Advantages, developments and improvements of the present invention are found in the subclaims.

[0023] According to an embodiment of the present invention, the method comprises the steps of:

a) receiving the first coded speech signal encoded using the first CELP format and including at least the first pitch parameter;

b) decoding the received first coded speech signal to the decoded pcm speech signal;

c) detecting the voiced level of the pcm speech signal within the predetermined time window;

d) determining, if the pcm speech signal is a voiced speech signal or an unvoiced speech signal dependent on at least the first parameter;

e) detecting an energy level of the pcm speech signal within the predetermined time window;

f) determining, if the energy level of the pcm speech signal is high or low dependent on at least a second parameter;

g) if the pcm speech signal is voiced and its energy level is high, performing the closed loop search process which receives at least the first pitch parameter and estimates the second pitch parameter for the second CELP format dependent on at least the first pitch parameter;

h) if the pcm speech signal is unvoiced or its energy level is low, copying the first pitch parameter as the second pitch parameter for the second CELP format.

[0024] According to an embodiment of the present invention, the transcoding apparatus comprises:

a) a receiving means which receives the first coded speech signal encoded using the first CELP format and including at least the first pitch parameter;

b) a decoding means which decodes the received first coded speech signal to the decoded pcm speech signal;

c) a first detecting means which detects the voiced level of the pcm speech signal within the predetermined time window;

d) a first determining means which determines if the pcm speech signal is the voiced speech signal or the unvoiced speech signal dependent on at least the first parameter;

e) a second detecting means which detects an energy level of the pcm speech signal within the predetermined time window;

5 f) a second determining means which determines if the energy level of the pcm speech signal is high or low dependent on at least the second parameter;

10 g) a closed loop search means which performs a closed loop search, wherein the closed loop search means receives at least the first pitch parameter and estimates a second pitch parameter for the second CELP format dependent on at least the first pitch parameter, if the pcm speech signal is voiced and its energy level is high; and

15 h) a copying means which copies the first pitch parameter as the second pitch parameter for the second CELP format, if the pcm speech signal is unvoiced or its energy level is low.

[0025] According to the present invention, the energy of the signal is an important factor determining if the quality needs to be optimal or not at the encoder of the second codec. Artefacts are principally more acceptable on signals of low energy than on high energy signals. Indeed, low signals are less audible and can be principally more degraded than energetic signals. Accordingly, according to the present invention an accurate estimation of the pitch parameter at the encoder of the second codec is principally only applied for high energy signals.

[0026] An advantage of the present invention is to provide an adaptive compromise between the closed loop search and the open loop search depending on the first pitch parameter of the first compressed speech signal encoded using the first CELP format and depending on its energy level.

[0027] According to a further embodiment of the present invention, the method further comprises the step of encoding the decoded pcm speech signal using the second CELP format to a second coded speech signal including at least a second pitch parameter.

[0028] According to a further embodiment, the closed loop search process is performed in a restricted interval $[T_{op}(A) - T'_{LOW}; T_{op}(A) + T'_{HIGH}]$ around the first pitch parameter ($T_{op}(A)$), wherein T'_{LOW} the signals a preselected lower pitch threshold value and T'_{HIGH} a preselected upper pitch threshold value.

[0029] According to a further embodiment, the lower and the upper pitch first threshold values are preselected the greater the detected voiced level is.

[0030] According a further embodiment, the first parameter and/or the second parameter are provided as a predetermined threshold value, respectively.

55 **[0031]** According to a further embodiment, the first CELP format is provided by a first codec and the second CELP format is provided by a second codec which is different to the first codec. Suitable examples for the first

codes and the second codec are the following:

1. AMR: Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec (Release 6), 3GPP TS 26.090, v6.0.0:
2. AMR-WB/G.722.2: Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB), ITU T Recommendation G.722.2, 07/2003.
3. G.729: Coding of Speech at 8 kbit/s using conjugate-structure algebraic-excited linear prediction, ITU-T Recommendation G.729 03/1996.
4. Annexes of G.729: Annex A, B, D, E, J.
5. G.723.1: Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, ITU-T recommendation G.723.1, 03/1996.
6. EVRC: Enhanced Variable Rate Codec (EVRC), 3GPP2 C.S0014-0, Version 1.0, 2003.
7. VMR-WB: Source-Controlled Variable-Rate Multimode 2 Wideband Speech Codec (VMR-WB) 3GPP2 C.S.0052-0, Version 1.0, 2004.

[0032] According to a further embodiment, the first codec and the second codec are selected from the group of AMR, AMR-WB/G.722.2, G.729, ANNEXES OF G.729, G.723.1, EVRC AND VMR-WB.

[0033] According to a further embodiment, the first coded speech signal includes at least the first pitch parameter T_{op} (A) and an additional parameter set comprising a linear prediction code (LPC) parameter and/or at least one fixed gain parameter and/or at least one adaptive gain parameter and/or one adaptive code-book parameter.

[0034] According to a further embodiment, the voiced level of the pcm speech signal is detected by means of using a variability of the first pitch parameter at a predetermined frame or for the predetermined time and/or by means of using at least one parameter of the additional parameter set.

[0035] According to a further embodiment, the energy level of the pcm speech signal is detected by means of using the fixed gain parameter of the first coded speech signal and/or by means of computing the energy level of the decoded pcm speech signal.

[0036] Exemplary embodiments of the invention are illustrated in the drawings and explained in more detail in the description below.

[0037] In the figures:

Figure 1: shows a schematic flow diagram of a first embodiment of the method of the present invention;

figure 2: shows a schematic flow diagram of a second embodiment of the method of the present invention;

5 figure 3: shows a diagram illustrating the pitch parameter over the time;

figure 4: shows a schematic block diagram of an embodiment of the transcoding apparatus of the present invention;

10 figure 5: shows a schematic block diagram of the transcoding apparatus of figure 4 coupled between two terminal units; and

15 figure 6: shows a detailed schematic block diagram of the pitch parameter providing device of figure 4.

20 **[0038]** In the figures, identical reference symbol designate identical or functionally identical elements.

[0039] Figure 1 shows a schematic flow diagram of a first embodiment of the method for transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format. In the following, the inventive method for transcoding the speech signal is explained by means of the schematic flow diagram of figure 1 referring to the block diagram of figure 4. The method of transcoding the speech signal of the present invention comprises the method steps S1-S8:

Method step S1:

35 **[0040]** A first coded speech signal DS1 is received, in particular the frame of a speech signal. Said coded speech signal DS1 is encoded using the first CELP format and includes at least a first pitch parameter T_{op} (A). Preferably, the first coded speech signal DS1 includes at least the first pitch parameter T_{OP} (A) and an additional parameter set. The additional parameter set comprises a linear prediction code (LPC) parameter and/or at least one fixed gain parameter and/or at least one adaptive gain parameter and/or at least one adaptive code-book parameter.

Method step S2:

40 **[0041]** The received first coded speech signal DS1 is decoded to a decoded pcm speech signal AS2.

Method step S3:

45 **[0042]** A voiced level VL of the pcm speech signal AS2 is detected within a predetermined time window T. Particularly, the voiced level VL of the pcm speech signal AS2 is detected by means of using a variability of the first pitch parameter T_{OP} (A) at the predetermined frame or

for the predetermined time window T. Alternatively, the voiced level VL can be detected by means of using at least one parameter of said additional parameter set.

Method step S4:

[0043] It is determined, if the pcm speech signal AS2 is a voiced speech signal or an unvoiced speech signal dependent on at least a first parameter P1. The first parameter P1 can be provided as a predetermined threshold value.

Method step S5:

[0044] Preferably, an energy level EL of the pcm speech signal AS2 is detected within the predetermined time window T. Preferably, the energy level EL of the pcm speech signal AS2 is detected by means of using the fixed gain parameter of the first coded speech signal DS1 and/or by means of computing the energy level of the decoded pcm speech signal AS2.

Method step S6:

[0045] Preferably, it is determined, if the energy level EL of the pcm speech signal AS2 is high or low dependent on at least a second parameter P2. Particularly, the second parameter P2 is provided as a predetermined threshold value.

Method step S7:

[0046] If the pcm speech signal AS2 is voiced and preferably its energy level EL is high, a closed loop search process is performed which receives at least the first pitch parameter $T_{OP}(A)$ and estimates a second pitch parameter $T_{OP}(B)$ for the second CELP format dependent on at least the first pitch parameter $T_{OP}(A)$.

[0047] Preferably, the closed loop search process is performed in a restricted interval $[T_{OP}(A)-T'_{LOW}; T_{OP}(A)+T'_{HIGH}]$ around the first pitch parameter $T_{OP}(A)$, wherein T'_{LOW} designates a pre-selected lower pitch threshold value and T'_{HIGH} designates an upper pitch threshold value.

[0048] Particularly, the lower and the upper pitch threshold values T'_{HIGH} , T'_{LOW} are pre-selected the greater the detected voiced level VL is. Typically, the variations of T'_{HIGH} , T'_{LOW} are between 1 and 3 samples for codecs like AMR, G729, EVCR. Indeed, for such codecs, the maximal range is the complete closed-loop search made in the interval $[T_{OP}(A)-T_{LOW}; T_{OP}(A)+T_{HIGH}]$ where $T_{LOW}=3$, $T_{HIGH}=3$

Method step S8:

[0049] If the pcm speech signal AS2 is unvoiced or preferably its energy level EL is low, the first pitch parameter $T_{OP}(A)$ is copied as second pitch parameter T_{OP}

(B) for the second CELP format. Typically, a preferable value for energy level EL on a frame is energy level EL < 20 dB.

[0050] Figure 2 is a schematic flow diagram of a second embodiment of the method of the present invention. The second embodiment of figure 2 comprises the method steps S1-S8 as shown in figure 1 and as explained above. Further, the second embodiment of the method of the present invention of figure 2 comprises the additional method step S9.

Method step S9:

[0051] The decoded pcm speech signal AS2 is encoded using the second CELP format to a second coded speech signal DS2 including at least the second pitch parameter $T_{OP}(B)$.

[0052] Figure 3 is a diagram showing the pitch parameter T_{OP} over the time t. Figure 3 shows that the amplitude of the pitch parameter T_{OP} is low in the voiced periods VP. On the other hand, the pitch parameter T_{OP} is high in the unvoiced periods UP.

[0053] Figure 4 is a schematic block diagram of an embodiment of the transcoding apparatus 1 of the present invention.

[0054] The transcoding apparatus 1 of figure 4 is adapted to execute the method of figure 1, respectively, of figure 2.

[0055] Therefore, the transcoding apparatus 1 comprises a receiving means 2, a decoding means 3, a first detecting means 4, a first determining means 5, a second detecting means 6, a second determining means 7 and a pitch parameter providing means 8 comprising at least a closed loop search means 8a and a copying means 8b (see figure 6).

[0056] The receiving means 2 is adapted to receive the first coded speech signal DS1 encoded using the first CELP format and including at least the first pitch parameter $T_{OP}(A)$.

[0057] The decoding means 3 is adapted to decode the received first coded speech signal DS1 to provide an pcm speech signal AS2.

[0058] The first detecting means 4 is adapted to detect the voiced level VL of the pcm speech signal AS2 within the predetermined time window T.

[0059] The first determining means 5 is adapted to determine, if the pcm speech signal AS2 is a voiced speech signal or an unvoiced speech signal dependent on at least the first parameter P1.

[0060] The second detecting means 6 is adapted to detect an energy level EL of the pcm speech signal AS2 within the predetermined time window T.

[0061] The second determining means 7 is adapted to determine, if the energy level EL of the pcm speech signal AS2 is high or low dependent on at least the second parameter P2.

[0062] The closed loop search means 8a is adapted to perform a closed loop search. The closed loop search

means 8a receives at least the first pitch parameter T_{OP} (A) and estimates a second pitch parameter T_{OP} (B) for the second CELP format dependent on at least the first pitch parameter T_{OP} (A), if the pcm speech signal AS2 is voiced and its energy level EL is high (see figure 6).

[0063] The copying means 8b is adapted to copy the first pitch parameter T_{OP} (A) as the second pitch parameter T_{OP} (B) for the second CELP format, if the pcm speech signal AS2 is unvoiced or its energy level EL is low (see figure 6).

[0064] The closed loop search means 8a and the copying means 8b are shown in detail in figure 6. The pitch parameter providing means 8 comprises further a decision means 8c. The decision means 8c receives the signals EL' and VL'. EL' designates the detection result of the second determining means 7. For example, if the energy level EL is greater than the second parameter P2 which is a threshold value, the decision signal EL' is high. On the other hand, if the energy level EL is smaller or equal to the second parameter P2, the decision signal EL' is low. Further, the signal VL' is the decision signal of the first detecting means 4. If, for example, the voiced level VL is greater than the first parameter P1 which is a threshold value, the decision signal VL' is high. On the other hand, the decision signal VL' is low.

[0065] Particularly, the closed loop search means 8a performs the closed loop search in a restricted interval $[T_{OP}(A) - T_{LOW}; T_{OP}(A) + T_{HIGH}]$ around the first pitch parameter $T_{OP}(A)$.

[0066] Figure 5 shows a schematic block diagram of the transcoding apparatus 1 coupled between two terminal units 11 and 12. A first terminal unit 11 comprises an encoding means 12 and a sending means 13. The encoding means 12 receives a first pcm speech signal AS1 and encodes said first pcm speech signal AS1 using the first CELP format to a first coded speech signal DS1. The sending means 13 receives the first coded speech signal DS1 encoded with the first CELP format and sends it to the transcoding apparatus 1.

[0067] The second terminal unit 14 comprises a receiving means 15 and a decoding means 16. The receiving means 15 receives the second coded speech signal DS2 encoded with the second CELP format. The receiving means 15 transfers the received second coded speech signal DS2 to the decoding means 16 for decoding. The decoding means 16 works with the second CELP format.

[0068] Although the present invention has been explained on the basis of particular exemplary embodiments, it is not restricted thereto, but rather can be modified in any desired manner without departing from the basic principle of the invention.

[0069] In particular, the present invention is not limited to the use of one transcoding apparatus between the terminal units, but there could be also provided a lot of different transcoding apparatuses, wherein neighbouring transcoding apparatuses which are coupled to each other work on the same CELP format.

Claims

1. Method of transcoding a speech signal from a first code excited linear prediction (CELP) format to a second code excited linear prediction (CELP) format, comprising:

- a) receiving a first coded speech signal (DS1) encoded using the first CELP format and including at least a first pitch parameter (T_{op} (A));
- b) decoding the received first coded speech signal (DS1) to a decoded pcm speech signal (AS2);
- c) detecting a voiced level (VL) of the pcm speech signal (AS2) within a predetermined time window (T);
- d) determining, if the pcm speech signal (AS2) is a voiced speech signal or an unvoiced speech signal dependent on at least a first parameter (P1);
- e) if the pcm speech signal (AS2) is voiced, performing a closed loop search process which receives at least the first pitch parameter (T_{op} (A)) and estimates a second pitch parameter (T_{op} (B)) for the second CELP format dependent on at least the first pitch parameter (T_{op} (A));
- f) if the pcm speech signal (AS2) is unvoiced, copying the first pitch parameter (T_{op} (A)) as the second pitch parameter (T_{op} (B)) for the second CELP format.

2. The method of claim 1, comprising:

- a) receiving the first coded speech signal (DS1) encoded using the first CELP format and including at least the first pitch parameter (T_{op} (A));
- b) decoding the received first coded speech signal (DS1) to the decoded pcm speech signal (AS2);
- c) detecting the voiced level (VL) of the pcm speech signal (AS2) within the predetermined time window (T);
- d) determining, if the pcm speech signal (AS2) is a voiced speech signal or an unvoiced speech signal dependent on at least the first parameter (P1);
- e) detecting an energy level (EL) of the pcm speech signal (AS2) within the predetermined time window (T);
- f) determining, if the energy level (EL) of the pcm speech signal (AS2) is high or low dependent on at least a second parameter (P2);
- g) if the pcm speech signal (AS2) is voiced and its energy level (EL) is high, performing the closed loop search process which receives at least the first pitch parameter (T_{op} (A)) and estimates the second pitch parameter (T_{op} (B)) for the second CELP format dependent on at least

- the first pitch parameter ($T_{op}(A)$);
 h) if the pcm speech signal (AS2) is unvoiced or its energy level (EL) is low, copying the first pitch parameter ($T_{op}(A)$) as the second pitch parameter ($T_{op}(B)$) for the second CELP format.
3. The method of claim 1 or 2, comprising further the step of encoding the decoded pcm speech signal (AS2) using the second CELP format to a second coded speech signal (DS2) including at least the second pitch parameter ($T_{op}(B)$).
 4. The method of claims 1, 2 or 3, wherein the closed loop search process is performed in a restricted interval [$T_{op}(A) - T'_{LOW}$; $T_{op}(A) + T'_{HIGH}$] around the first pitch parameter ($T_{op}(A)$), wherein T'_{LOW} designates a preselected lower pitch threshold value and T'_{HIGH} designates an upper pitch threshold value.
 5. The method of claim 4, wherein the lower and the upper pitch threshold values (T'_{HIGH} , T'_{LOW}) are preselected the greater the detected voiced level (VL) is.
 6. The method of claim 1 or one of claims 2 to 5, wherein the first parameter (P1) and/or second parameter (P2) are provided as a predetermined threshold value, respectively.
 7. The method of claim 1 or one of claims 3 to 6, wherein the first CELP format is provided by a first codec (3, 12) and the second CELP format is provided by a second codec (9, 16) which is different to the first codec (3, 12).
 8. The method of claim 7, wherein the first codec (3, 12) and the second codec (9, 16) are selected from the group of AMR, AMR-WB/G.722.2, G.729, Annexes of G.729, G.723.1, EVRC and VMR-WB.
 9. The method of claim 1 or one of claims 2 to 8, wherein the first coded speech signal (DS1) includes at least the first pitch parameter ($T_{op}(A)$) and an additional parameter set comprising a linear prediction code (LPC) parameter and/or at least one fixed gain parameter and/or at least one adaptive gain parameter and/or at least one adaptive code-book parameter.
 10. The method of claim 9, wherein the voiced level (VL) of the pcm speech signal (AS2) is detected by means of using a variability of the first pitch parameter ($T_{op}(A)$) at a predetermined frame or for the predetermined time window (T) and/or by means of using at least one parameter of the additional parameter set.
 11. The method of claim 9 or 10, wherein the energy level (EL) of the pcm speech signal (AS2) is detected by means of using the fixed gain parameter of the first coded speech signal (DS1) and/or by means of computing the energy level of the decoded pcm speech signal (AS2).
 12. Transcoding apparatus (1) for executing the method of claim 1, comprising:
 - a) a receiving means (2) which receives the first coded speech signal (DS1) encoded using the first CELP format and including at least the first pitch parameter;
 - b) a decoding means (3) which decodes the received first coded speech signal (DS1) to the decoded pcm speech signal (AS2);
 - c) a first detecting means (4) which detects the voiced level (VL) of the pcm speech signal (AS2) within the predetermined time window (T);
 - d) a first determining means (5) which determines if the pcm speech signal (AS2) is the voiced speech signal or the unvoiced speech signal dependent on at least the first parameter (P1);
 - e) a closed loop search means (8a) which performs a closed loop search, wherein the closed loop search means (8a) receives at least the first pitch parameter ($T_{op}(A)$) and estimates the second pitch parameter ($T_{op}(B)$) for the second CELP format dependent on at least the first pitch parameter ($T_{op}(A)$), if the pcm speech signal (AS2) is voiced; and
 - f) a copying means (8b) which copies the first pitch parameter ($T_{op}(A)$) as the second pitch parameter for the second CELP format, if the pcm speech signal (AS2) is unvoiced.
 13. Transcoding apparatus (1) for executing the method of claim 2 or one of claims 3 to 11, comprising:
 - a) a receiving means (2) which receives the first coded speech signal (DS1) encoded using the first CELP format and including at least the first pitch parameter;
 - b) a decoding means (3) which decodes the received first coded speech signal (DS1) to the decoded pcm speech signal (AS2);
 - c) a first detecting means (4) which detects the voiced level (VL) of the pcm speech signal (AS2) within the predetermined time window (T);
 - d) a first determining means (5) which determines if the pcm speech signal (AS2) is the voiced speech signal or the unvoiced speech signal dependent on at least the first parameter

(P1) ;

e) a second detecting means (6) which detects an energy level (EL) of the pcm speech signal (AS2) within the predetermined time window (T);

f) a second determining means (7) which determines if the energy level (EL) of the pcm speech signal (AS2) is high or low dependent on at least the second parameter (P2);

g) a closed loop search means (8a) which performs a closed loop search, wherein the closed loop search means (8a) receives at least the first pitch parameter ($T_{op}(A)$) and estimates a second pitch parameter ($T_{op}(B)$) for the second CELP format dependent on at least the first pitch parameter ($T_{op}(A)$), if the pcm speech signal (AS2) is voiced and its energy level (EL) is high; and

h) a copying means (8b) which copies the first pitch parameter ($T_{op}(A)$) as the second pitch parameter for the second CELP format, if the pcm speech signal (AS2) is unvoiced or its energy level (EL) is low.

5

10

15

20

25

30

35

40

45

50

55

FIG 1

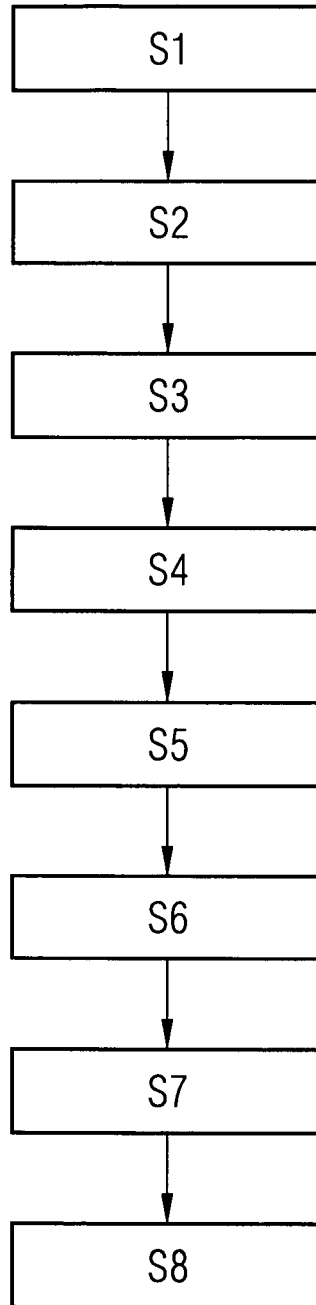


FIG 2

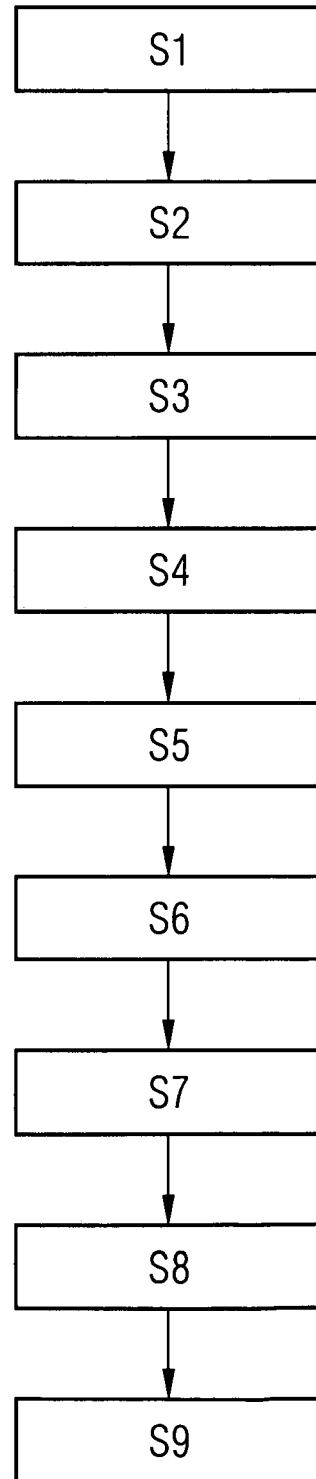


FIG 3

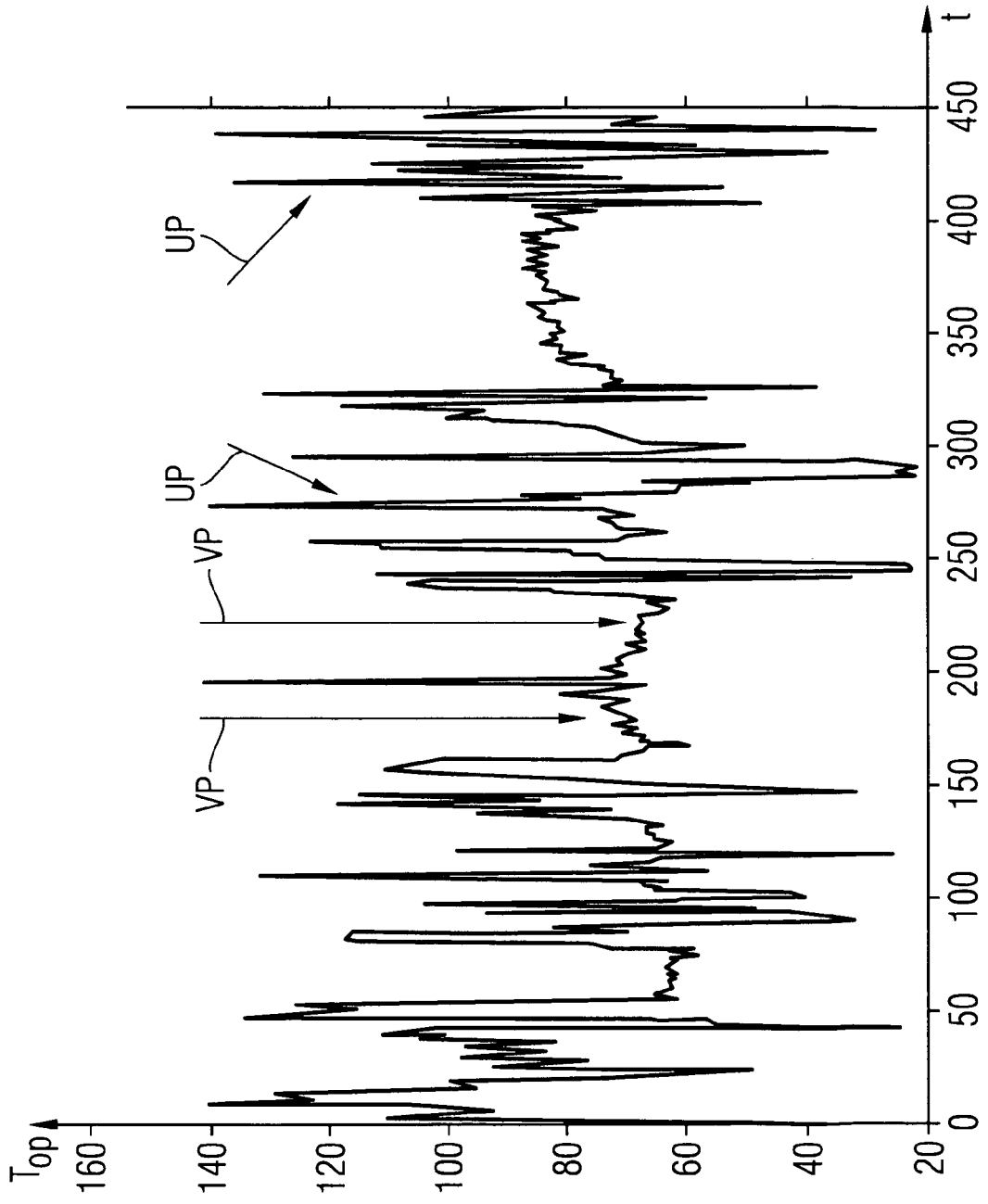


FIG 4

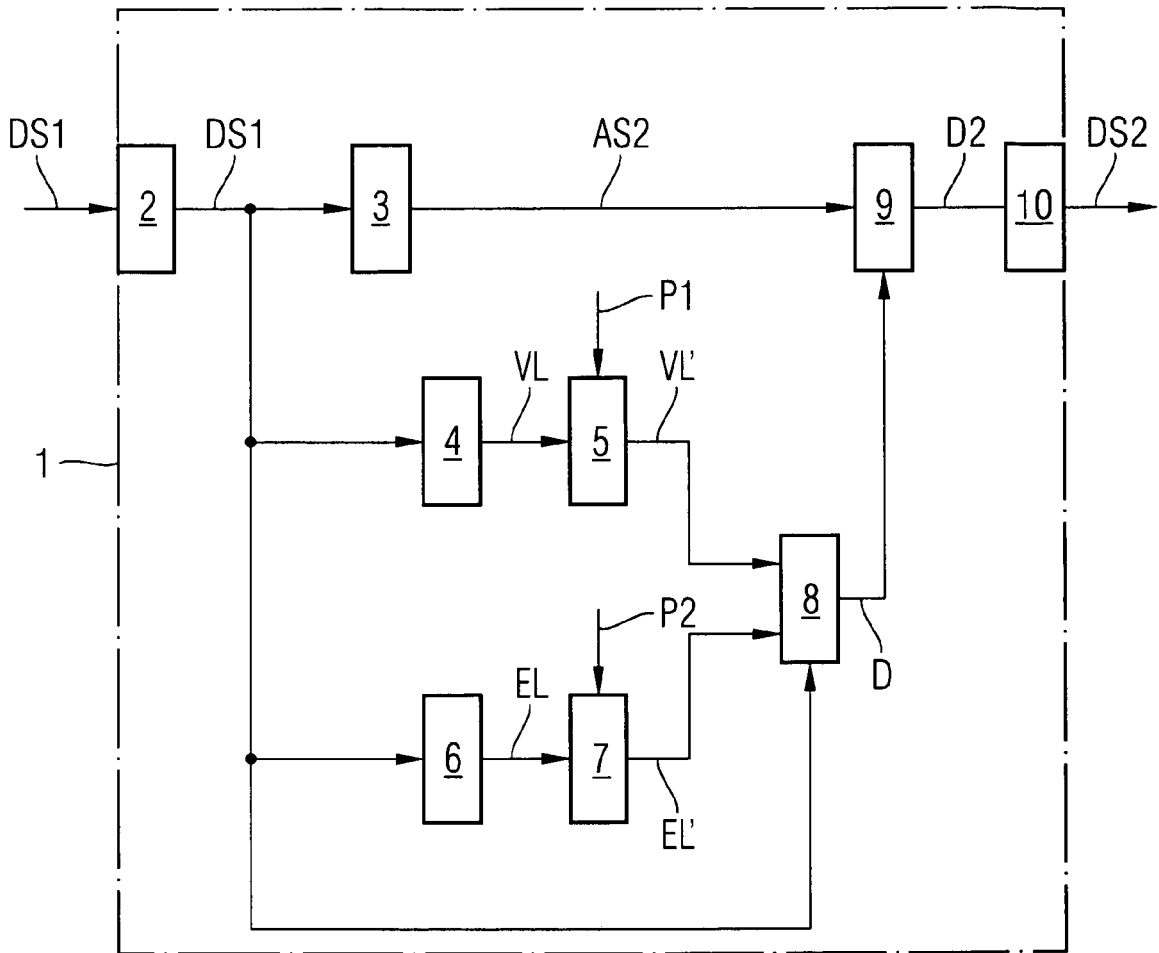


FIG 5

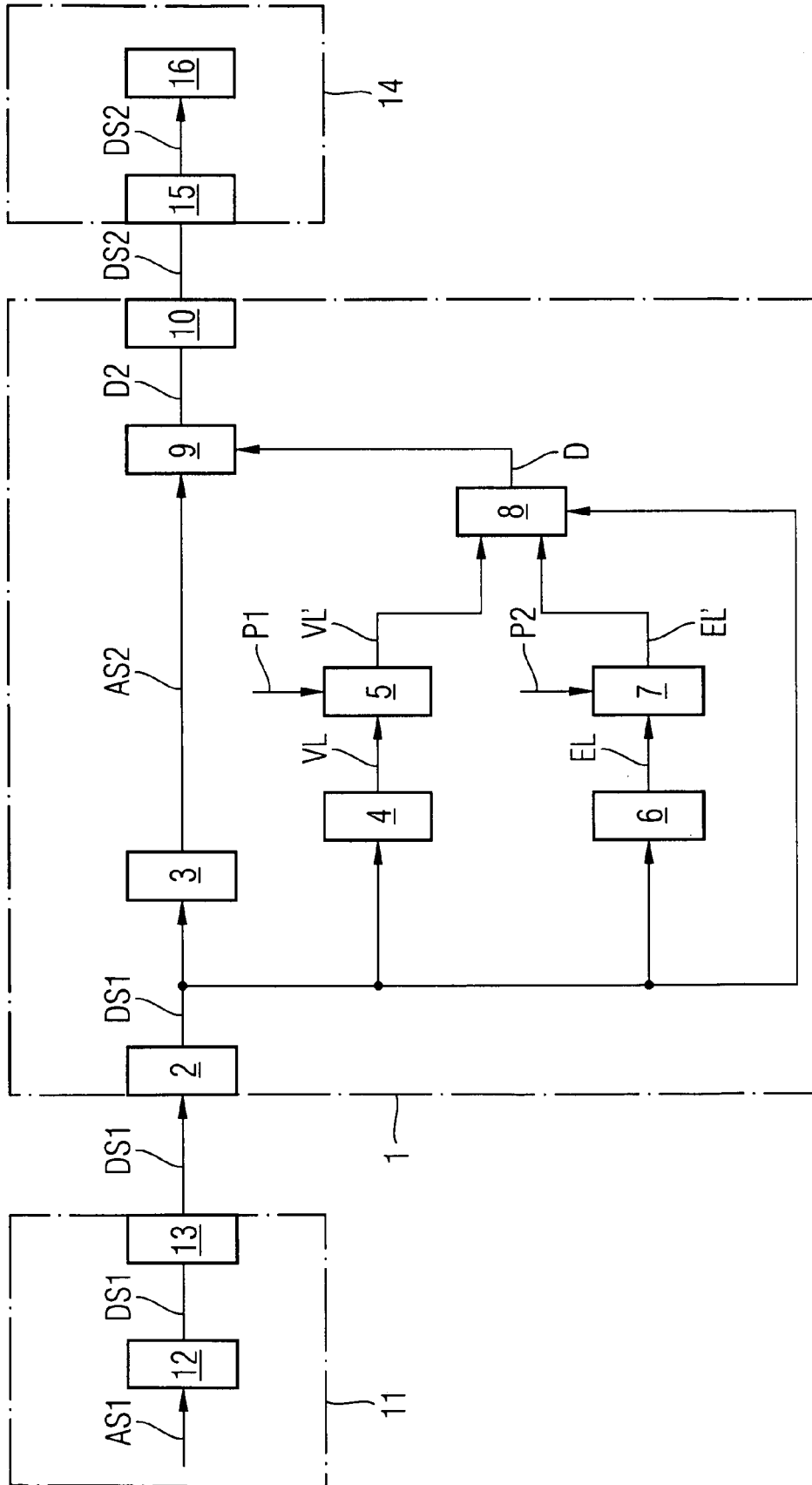
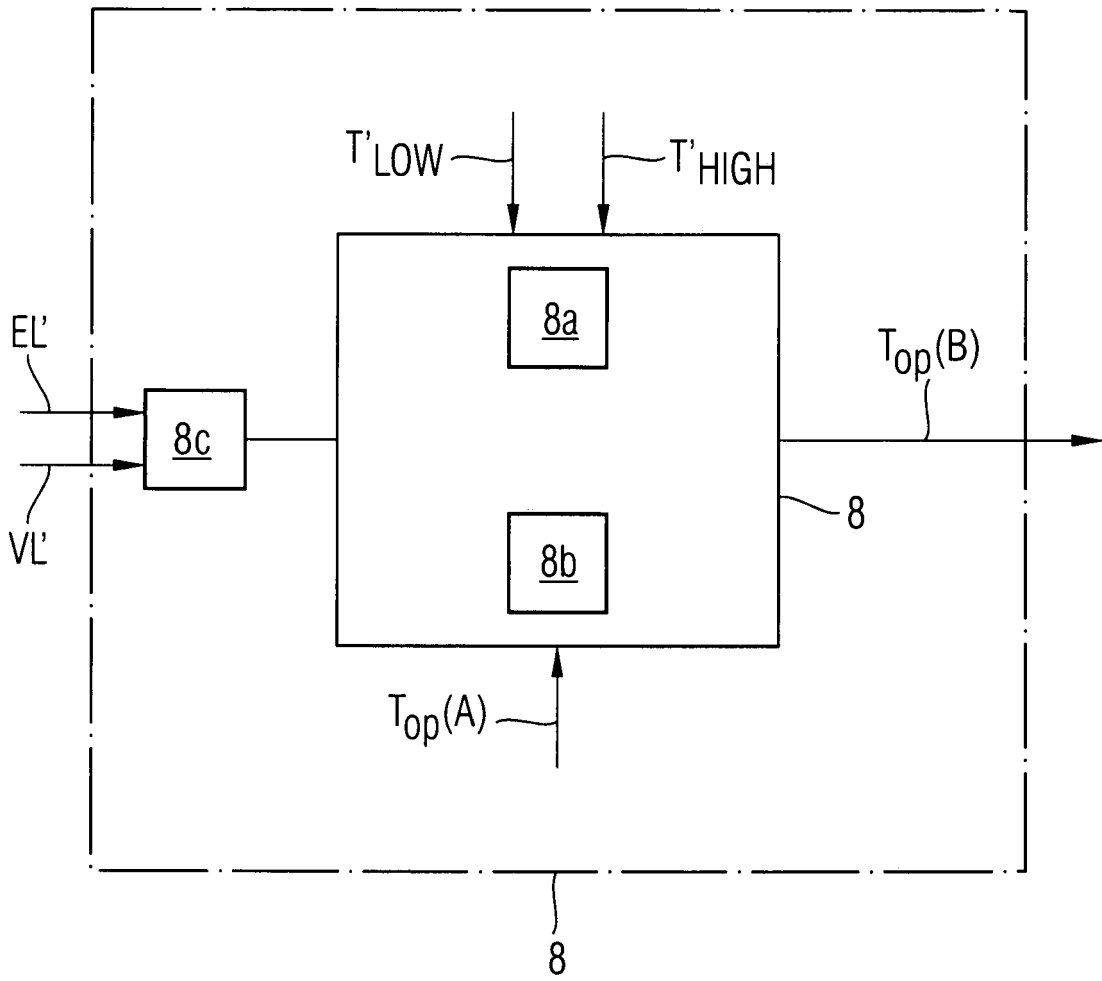


FIG 6





DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
Y	DALWON JANG ET AL: "A Novel Rate Selection Algorithm for Transcoding CELP-type Codec and SMV" EUROSPEECH 2003, September 2003 (2003-09), page 2865, XP007007018 Geneva, CH * the whole document *	1-13	INV. G10L19/12 G10L19/14
D,Y	JIN-KYU CHOI ET AL: "Improvement issues on transcoding algorithms : for the flexible usage to the various pairs of speech codec" IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 2004. PROCEEDINGS. (ICASSP '04)., 17 May 2004 (2004-05-17), - 21 May 2004 (2004-05-21) pages 269-272, XP010717617 MONTREAL, QUEBEC, CANADA ISBN: 0-7803-8484-9 * abstract * * section 2.2 Pitch estimation *	1-13	TECHNICAL FIELDS SEARCHED (IPC)
A	WO 03/058407 A (DILITHIUM NETWORKS INC [US]; JABRI MARWAN A [AU]; WANG JIANWEI [AU]; G) 17 July 2003 (2003-07-17) * abstract; figures 8,10,11 * * paragraph [0103] *	1,4,5, 7-9,12	G10L
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 18 April 2007	Examiner Quélavoine, Régis
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

2

EPO FORM 1503 03.02 (P04C01)



DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
D,A	PANKAJ K R ED - MATTHEWS M B (ED) INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS: "A novel transcoding scheme from EVRC to G.729AB" CONFERENCE RECORD OF THE 37TH. ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS, & COMPUTERS. PACIFIC GROOVE, CA, NOV. 9 - 12, 2003, ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 2. CONF. 37, 9 November 2003 (2003-11-09), pages 533-536, XP010702678 ISBN: 0-7803-8104-1 * abstract * * section "C. Pitch Parameter Transformation" *	1,4,5, 7-9,12	
D,A	KYUNG TAE KIM ET AL: "An efficient transcoding algorithm for G.723.1 and EVRC speech coders" VTC FALL 2001. IEEE 54TH. VEHICULAR TECHNOLOGY CONFERENCE. PROCEEDINGS. ATLANTIC CITY, NJ, OCT. 7 - 11, 2001, IEEE VEHICULAR TECHNOLOGY CONFERENCE, NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 4. CONF. 54, 7 October 2001 (2001-10-07), pages 1561-1564, XP010562224 ISBN: 0-7803-7005-8 * abstract * * section III *	1,4,5, 7-9,12	TECHNICAL FIELDS SEARCHED (IPC)
----- -/--			
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 18 April 2007	Examiner Quélavoine, Régis
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

2
EPO FORM 1503 03 82 (P04C01)



DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	<p>GHENANIA M ET AL: "TRANSCODAGE INTELLIGENT A FAIBLE COMPLEXITE EXTRE LES CODEURS UIT-T G.729 ET 3GPP NB-AMR (12.2 KBIT/S)" CORESA. COMPRESSION ET REPRESENTATION DES SIGNAUX AUDIOVISUELS, 25 May 2004 (2004-05-25), pages 85-88, XP001199662 * section 3.2 "Conversion du pitch" * * abstract *</p> <p style="text-align: center;">-----</p>	1,4,7-9,12	
			TECHNICAL FIELDS SEARCHED (IPC)
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 18 April 2007	Examiner Quélavoine, Régis
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p>		<p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>	

2
EPO FORM 1503 03 82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 06 02 5955

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

18-04-2007

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03058407 A	17-07-2003	AU 2003207498 A1	24-07-2003
		CN 1701353 A	23-11-2005
		EP 1464047 A2	06-10-2004
		JP 2005515486 T	26-05-2005

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

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

- An Efficient transcoding algorithm for G.713.1 and EVCR speech coders. **KYUNG TAE KIM**. 37th Asilomar Conf. On Signals, Systems and Computers. IEEE, 2003 **[0010]**
- **JIN-KYU CHOI**. ICASSP, 2004 **[0011]**