



**(43) International Publication Date**  
**19 February 2004 (19.02.2004)**

## PCT

**(10) International Publication Number**  
**WO 2004/015690 A1**

- (51) **International Patent Classification<sup>7</sup>:** **G10L 19/06**,  
H04L 1/00

(21) **International Application Number:**  
PCT/EP2003/005076

(22) **International Filing Date:** 12 May 2003 (12.05.2003)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**  
0217729.3 31 July 2002 (31.07.2002) GB

(71) **Applicant (for all designated States except US):** **MOTOROLA INC** [US/US]; 1303 E.Algonquin Road, Schaumburg, IL 60196 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants (for US only):** **GIBBS, Jonathan, Alastair** [GB/GB]; Motorola Limited, Jays Close, Viabes Industrial Estate, Basingstoke, Hampshire RG22 4PD (GB). **AFTELAK, Stephen** [GB/GB]; Motorola Limited, Jays Close, Viabes Industrial Estate, Basingstoke, Hampshire RG22 4PD (GB).

(74) **Agent:** **MCCORMACK, Derek**; Motorola European Intellectual, Property Operations, Midpoint, Alencon Link, Basingstoke, Hampshire RG21 7PL (GB).

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**  
— with international search report

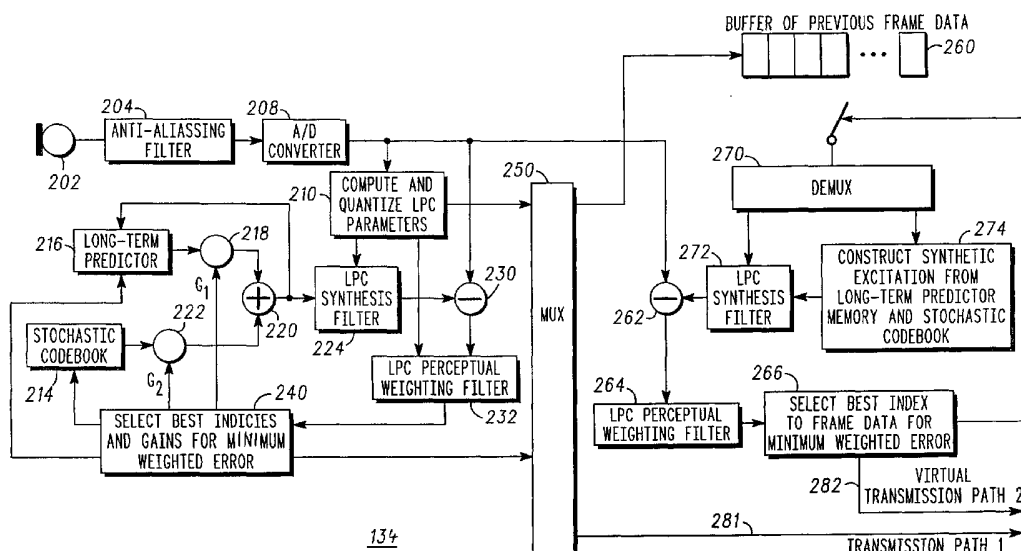
*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

- (54) Title:** SPEECH COMMUNICATION UNIT AND METHOD FOR ERROR MITIGATION OF SPEECH FRAMES



- (57) Abstract:** A speech communication unit (100) comprising a speech encoder (134) capable of representing an input speech signal, the speech encoder (134) comprising a transmission path (281) for transmitting a number of speech frames to a speech decoder, the speech encoder (134) characterised by a virtual transmission path (282) for transmitting one or more references for a number of speech frames transmitted in the transmission path (281) wherein the one or more references relate to an alternative speech frame within the number of speech frames transmitted on the transmission path (281) to be used as a replacement frame when a frame is received in error. The speech communication unit provides at least the advantage that a more accurate replacement frame mechanism is provided, thereby reducing the risk of undesirable artefacts being audible in recovered speech frame.

SPEECH COMMUNICATION UNIT AND METHOD FOR ERROR MITIGATION  
OF SPEECH FRAMES

Field of the Invention

5

This invention relates to speech coding and methods for improving the performance of speech codecs in speech communication units. The invention is applicable to, but not limited to, error mitigation in speech codecs.

10

Background of the Invention

Many present day voice communications systems, such as the global system for mobile communications (GSM) cellular telephony standard and the TERrestrial Trunked Radio (TETRA) system for private mobile radio users, use speech-processing units to encode and decode speech patterns. In such voice communications systems a speech encoder in a transmitting unit converts the analogue speech pattern into a suitable digital format for transmission. A speech decoder in a receiving unit converts a received digital speech signal into an audible analogue speech pattern.

25 As frequency spectrum for such wireless voice communication systems is a valuable resource, it is desirable to limit the channel bandwidth used by such speech signals, in order to maximise the number of users per frequency band. Hence, a primary objective in the use of speech coding techniques is to reduce the occupied capacity of the speech patterns as much as possible, by use of compression techniques, without losing fidelity.

In the context of voice and data communication systems, a further approach is to provide substantially less protection on speech signals, when compared to comparable data signals. This approach leads to comparably more errors within speech packets than data packets, as well as increased risk of losing whole speech packets.

In speech decoders, it is common for error mitigation techniques to be used, for example to improve the performance of the speech communication unit in the event of:

- (i) Too many bit errors being present within a received speech frame; or
- (ii) A data packet (which may include speech information) within an Internet Protocol (IP) based network being lost.

'Bad-frame' mitigation techniques are needed to minimise the audible effect of frames received in error, where 'received in error' is taken here to mean either received with errors or not received at all. These techniques reproduce an estimate of the missing speech frame, rather than injecting either silence or noise into the decoded speech. Such techniques typically involve exploiting the statistical static properties of speech. A single frame in error is usually adequately estimated by replacing it with similar parameters including energy, pitch, spectrum and voicing from the previous frame. However, speech is not truly stationary e.g. speech onsets and plosives are very short events. Hence, this simple 'replacement' technique sometimes leads to unnatural, and therefore undesirable, artefacts.

In an ideal world it would be preferable to interpolate the data from either side of a transmission break, i.e. take data following the bad-frame sequence, as well as before, and interpolate therebetween. However, such an approach is unacceptable in voice communication systems as it introduces undesirable delay.

If several bad frames are received then the energy of the speech signals is often reduced to zero after a few frames. Often a 'voicing' parameter is included because it is useful to change what is repeated dependent upon whether the speech is voiced or not. In principle, for voiced speech, it is preferable to just repeat the periodic component. In contrast, for unvoiced speech, it is preferable to generate a similar audio spectrum and similar energy without making it too periodic.

The inventors of the present invention have recognised and appreciated the limitations in using such a simple 'replacement' frame mechanism as a bad-frame mitigation strategy. In particular, they have recognised that only on rare occasions is the replacing frame a truly suitable frame. Furthermore, if a number of frames are received in error, which may frequently occur on a poor quality wireless communication link, then the replacement frame mechanism is even less acceptable.

Hence, a need has arisen for provision of an improved error mitigation technique when using such speech codecs, to alleviate at least some of the aforementioned disadvantages.

#### Summary of the Invention

In a first aspect of the present invention, a speech communication unit is provided, in accordance with Claim 1.

- 5 In a second aspect of the present invention, a speech communication unit is provided, in accordance with Claim 11.

- 10 In a third aspect of the present invention, a method of performing bad-frame error mitigation in a voice communication unit is provided, in accordance with Claim 13.

- 15 In a fourth aspect of the present invention, a speech communication unit is provided, in accordance with any Claim 14.

- 20 In a fifth aspect of the present invention, a wireless communication system is provided, in accordance with any Claim 15.

Further aspects of the present invention are defined in the dependent Claims.

- 25 In summary, the present invention aims to provide a communication unit, comprising a speech codec and method of performing bad-frame error mitigation that at least alleviate some of the aforementioned disadvantages associated with current bad-frame error mitigation  
30 techniques. This is primarily achieved by transmitting speech frames on a transmission path and using a reference/pointer that is transmitted on a virtual transmission path to indicate alternative replacement speech frames to be used by a speech decoder, should a

speech frame on the transmission path be received in error. By utilising an additional virtual transmission path, ideally with different error statistics e.g. separate FEC scheme, the reference/pointer will not be  
5 subject to the same errors as the speech frame it is referencing. Furthermore, a buffering technique is used in the encoder to select an alternative speech frame, from a number of previously transmitted speech frames, which exhibits similar characteristics to the selected  
10 speech frame to be referenced.

#### Brief Description of the Drawings

Exemplary embodiments of the present invention will now  
15 be described, with reference to the accompanying drawings, in which:

FIG. 1 shows a block diagram of a wireless communication unit containing a speech coder adapted to support the  
20 various inventive concepts of a preferred embodiment of the present invention;

FIG. 2 shows a block diagram of a code excited linear predictive speech coder adapted to support the various  
25 inventive concepts of a preferred embodiment of the present invention;

FIG. 3 shows a use of a reference mechanism indicated by an alternative virtual transmission path, whereby  
30 replacement frames are selected from a number of other frames, in accordance with the preferred embodiments of the present invention; and

FIG. 4 shows an enhanced use of an alternative virtual transmission path, to address multiple errors occurring in the main transmission path, in accordance with the preferred embodiments of the present invention.

5

#### Description of Preferred Embodiments

Turning now to FIG. 1, there is shown a block diagram of a wireless subscriber unit, hereinafter referred to as a mobile station (MS) 100 adapted to support the inventive concepts of the preferred embodiments of the present invention. The MS 100 contains an antenna 102 preferably coupled to a duplex filter, antenna switch or circulator 104 that provides isolation between a receiver and a transmitter chain within the MS 100.

As known in the art, the receiver chain typically includes scanning receiver front-end circuitry 106 (effectively providing reception, filtering and intermediate or base-band frequency conversion). The scanning front-end circuit is serially coupled to a signal processing function 108. An output from the signal processing function is provided to a suitable output device 110, such as a speaker via a speech-processing unit 130.

The speech-processing unit 130 includes a speech encoding function 134 to encode a user's speech into a format suitable for transmitting over the transmission medium. The speech-processing unit 130 also includes a speech decoding function 132 to decode received speech into a format suitable for outputting via the output device (speaker) 110. The speech-processing unit 130 is operably coupled to a memory unit 116, and a timer 118

via a controller 114. In particular, the operation of the speech-processing unit 130 has been adapted to support the inventive concepts of the preferred embodiments of the present invention. In particular, the

5 speech-processing unit 130 has been adapted to select a replacement speech frame from a number of previously transmitted speech frames. The speech processing unit 130, or signal processor 108, then initiates transmission of a reference/pointer signal (indicating the selected

10 replacement speech frame) in an alternative virtual transmission path to the primary transmission path. The adaptation of the speech-processing unit 130 is further described with regard to FIG. 2.

15 For completeness, the receiver chain also includes received signal strength indicator (RSSI) circuitry 112 (shown coupled to the scanning receiver front-end 106, although the RSSI circuitry 112 could be located elsewhere within the receiver chain). The RSSI circuitry

20 is coupled to a controller 114 for maintaining overall subscriber unit control. The controller 114 is also coupled to the scanning receiver front-end circuitry 106 and the signal processing function 108 (generally realised by a DSP). The controller 114 may therefore

25 receive bit error rate (BER) or frame error rate (FER) data from recovered information. The controller 114 is coupled to the memory device 116 for storing operating regimes, such as decoding/encoding functions and the like. A timer 118 is typically coupled to the controller

30 114 to control the timing of operations (transmission or reception of time-dependent signals) within the MS 100.



In the context of the present invention, the timer 118 dictates the timing of speech signals, in the transmit (encoding) path and/or the receive (decoding) path.

- 5 As regards the transmit chain, this essentially includes an input device 120, such as a microphone transducer coupled in series via speech encoder 134 to a transmitter/modulation circuit 122. Thereafter, any transmit signal is passed through a power amplifier 124
- 10 to be radiated from the antenna 102. The transmitter/modulation circuitry 122 and the power amplifier 124 are operationally responsive to the controller, with an output from the power amplifier coupled to the duplex filter or circulator 104. The
- 15 transmitter/modulation circuitry 122 and scanning receiver front-end circuitry 106 comprise frequency up-conversion and frequency down-conversion functions (not shown).
- 20 Of course, the various components within the MS 100 can be arranged in any suitable functional topology able to utilise the inventive concepts of the present invention. Furthermore, the various components within the MS 100 can be realised in discrete or integrated component form,
- 25 with an ultimate structure therefore being merely an arbitrary selection.

- It is within the contemplation of the invention that the preferred buffering or processing of speech signals can
- 30 be implemented in software, firmware or hardware, with preferably a software processor (or indeed a digital signal processor (DSP)), performing the speech processing function.

Referring now to FIG. 2, a block diagram of a code excited linear predictive (CELP) speech encoder 134 is shown, according to the preferred embodiment of the present invention. An acoustic input signal to be analysed is applied to speech coder 134 at microphone 202. The input signal is then applied to filter 204. Filter 204 will generally exhibit band-pass filter characteristics. However, if the speech bandwidth is already adequate, filter 204 may comprise a direct wire connection.

The analogue speech signal from filter 204 is then converted into a sequence of N pulse samples, and the amplitude of each pulse sample is then represented by a digital code in analogue-to-digital (A/D) converter 208, as known in the art. The sampling rate is determined by sample clock (SC). The sample clock (SC) is generated along with the frame clock (FC).

The digital output of A/D 208, which may be represented as input speech vector  $s(n)$ , is then applied to coefficient analyser 210. This input speech vector  $s(n)$  is repetitively obtained in separate frames, i.e., blocks of time, the length of which is determined by the frame clock (FC), as is known in the art.

For each block of speech, a set of linear predictive coding (LPC) parameters is produced in accordance with a preferred embodiment of the invention by coefficient analyser 210. The generated speech coder parameters may include the following: LPC parameters, long-term predictor (LTP) parameters, excitation gain factor ( $G_2$ ) (along with the best stochastic codebook excitation codeword I). Such speech coding parameters are applied

to multiplexer 250 and sent over the channel for use by the speech synthesizer at the decoder. The input speech vector  $s(n)$  is also applied to subtractor 230, the function of which is described later.

5

Within the conventional CELP encoder of FIG. 2, the codebook search controller 240 selects the best indices and gains from the adaptive codebook within block 216 and the stochastic codebook within block 214 in order to  
10 produce a minimum weighted error in the summed chosen excitation vector used to represent the input speech sample. The output of the stochastic codebook 214 and the adaptive codebook 216 are input into respective gain functions 222 and 218. The gain-adjusted outputs are  
15 then summed in summer 220 and input into the LPC filter 224, as is known in the art.

Firstly, the adaptive codebook or long-term predictor component is computed  $l(n)$ . This is characterised by a  
20 delay and a gain factor ' $G_1$ '.

For each individual stochastic codebook excitation vector  $u_i(n)$ , a reconstructed speech vector  $s'_i(n)$  is generated for comparison to the input speech vector  $s(n)$ . Gain  
25 block 222 scales the excitation gain factor ' $G_2$ ' and summing block 220 adds in the adaptive codebook component. Such gain may be pre-computed by coefficient analyser 210 and used to analyse all excitation vectors, or may be optimised jointly with the search for the best  
30 excitation codeword  $I$ , generated by codebook search controller 240.

The scaled excitation signal  $G_1 l(n) + G_2 u_i(n)$  is then filtered by the linear predictive coding filter 224,

which constitutes a short-term predictor (STP) filter, to generate the reconstructed speech vector  $s'_i(n)$ .

The reconstructed speech vector  $s'_i(n)$  for the  $i$ -th excitation code vector is compared to the same block of input speech vector  $s(n)$  by subtracting these two signals in subtractor 230.

The difference vector  $e_i(n)$  represents the difference between the original and the reconstructed blocks of speech. The difference vector is perceptually weighted by weighting filter 232, utilising the weighting filter parameters (WTP) generated by coefficient analyser 210. Perceptual weighting accentuates those frequencies where the error is perceptually more important to the human ear, and attenuates other frequencies.

An energy calculator function inside the codebook search controller 240 computes the energy of the weighted difference vector  $e'_i(n)$ . The codebook search controller compares the  $i$ -th error signal for the present excitation vector  $u_i(n)$  against previous error signals to determine the excitation vector producing the minimum error. The code of the  $i$ -th excitation vector having a minimum error is then output over the channel as the best excitation code  $I$ .

A copy of the scaled excitation  $G_{1l}(n) + G_2 u_I(n)$  is stored within the Long Term Predictor memory of 216 for future use.

In the alternative, codebook search controller 240 may determine a particular codeword that provides an error signal having some predetermined criteria, such as meeting a predefined error threshold.

A more detailed description of the functionality of a typical speech encoding unit can be found in "Digital speech coding for low-bit rate communications systems" by  
5 A. M. Kondo, published by John Wiley in 1994.

In the preferred embodiment of the invention, an error mitigation technique has been applied to the speech frames following the multiplexer 250. The invention  
10 makes use of an alternative, preferably parallel, virtual transmission path 282 that is used to send a pointer to a previously encoded speech frame sent from the encoder on the main transmission path 281.

15 In the context of the present invention, the expression 'virtual' is defined as a transmission path that is provided from the encoder to the decoder in addition to the primary transmission path that supports the speech communication. The 'virtual' transmission path, may be  
20 located within the same bit-stream, or within the same time frame or multi-frame in a time division multiplexed scheme, or via a different communication route, for example in a VoIP system. By utilising an additional virtual transmission path, ideally with different error  
25 statistics e.g. a separate FEC scheme, the reference/pointer will not be subject to the same errors as the speech frame that it is referencing.

One notable difference to known encoding arrangements is  
30 that there is a second minimisation section following the multiplexing operation. Such circuitry assesses the speech parameter data held in the buffer and select the one that is closest to the current speech frame.

- In an enhanced embodiment, the parallel virtual transmission path uses different forward error correction (FEC) protection from that used in the main transmission path by the speech coder. In this manner, by using an independent FEC path, the speech data packet suffers from different error statistics. This difference between the main and parallel virtual transmission paths helps improve robustness to errors.
- 10 The multiplexer 250 outputs data packets/frames into a buffer 260 that holds previously multiplexed frames. A de-multiplexer 270 accesses the buffered frames of the multiplexed signal held in the buffer 260. In this regard, the de-multiplexer 270 separates the excitation parameters 274 from the LPC parameters 272. Note that
- 15 the memory of the long-term predictor used to generate the excitation parameters must be the same as the long-term predictor 216 at the start of the frame.
- 20 For each block of multiplexed speech, a set of linear predictive coding (LPC) parameters for current frames and previous frames are therefore produced. In the preferred embodiment of the present invention, each set of quantized LPC parameters and excitation parameters form
- 25 reconstructed speech vectors  $s'_j(n)$  for the  $j$ -th previous frame of buffered data. These are compared to previously buffered speech vectors  $s(n)$  by subtracting these two signals in subtractor 262.
- 30 The difference vector  $e_j(n)$  represents the difference between the original and the previously buffered blocks of speech. The difference vector is perceptually weighted by LPC weighting filter 264. As indicated, perceptual weighting accentuates those frequencies where

the error is perceptually more important to the human ear, and attenuates other frequencies.

An energy calculator function inside a codebook search  
5 controller 266 computes the energy of the weighted  
difference vector  $e'_j(n)$ . The codebook search controller  
266 compares the  $j$ -th error signal for the present  
excitation vector  $u_j(n)$  against previous error signals to  
10 determine the excitation vector producing the minimum  
error. The codebook search controller 266 then selects  
the 'best index to frame data' to provide the minimum  
weighted error. The encoder then transmits, to the  
decoder, a 'pointer' to the previous frame determined as  
15 providing the minimum weighted error between itself and  
the respective speech frame in the main transmission  
path.

In essence, the speech frame that is referenced (ideally  
differentially in time or frame number from the current  
20 transmitted frame) constitutes the frame within a certain  
moving window of speech that most closely resembles, in a  
perceptually weighted error sense, the frame that was  
encoded by the encoder. It therefore represents the best  
match (pointer) to the current frame for use in the error  
25 mitigation procedure if the frame is received in error.  
This representation, or pointer, is described in more  
detail with respect to FIG. 3.

Referring now to FIG. 3, a buffer timing diagram 300 is  
30 shown illustrating the preferred process of the present  
invention. The timing diagram illustrates frame 0 310 as  
having been received at a speech decoder and determined  
as being in error. The decoder has then accessed the  
alternative virtual transmission path to determine the

most appropriate frame to replace frame 0 310. As shown in FIG. 3 the alternative virtual transmission path has included a pointer to frame -4 320 as a preferred replacement of frame 0 310. By replacing frame 0 310 with frame -4 320, there is minimal effect on speech quality in the speech decoding process.

The inventors of the present invention have appreciated, and utilised, the fact that the immediately preceding frames were (typically) all spoken by the same talker, i.e. the speech frames will exhibit similar pitch and formant positions. Therefore, it is highly likely that a similar previous speech frame can be found to the current speech frame.

In accordance with the preferred embodiment of the present invention, the minimum perceptual error is found by evaluating the weighted Segmental signal-to-noise (SEGSNR) or average weighted SNR for each of the buffered frames, given the sets of parameters for each frame within the memory. Preferably, a segment is defined at the speech codec sub-frame level.

$$SEGSNR = \frac{1}{Subframes} \sum_{i=1}^{Subframes} 10 \log \left( \frac{Energy\ of\ Perceptually\ Weighted\ Speech}{Energy\ of\ Perceptually\ Weighted\ Error} \right) \quad [1]$$

This determination is performed in the encoder. In cases where there is a small pitch error, it is envisaged that wildly different SEGSNR values may result. This is because the source speech and the buffered signal can quickly move out of phase. Hence, in an enhanced embodiment of the present invention, it is proposed to search around the pitch period for the buffered frames,



say +/-5%, using sub-sample resolution (usually  $\frac{1}{3}$  or  $\frac{1}{4}$  samples) and take the highest SEGSNR value.

In a yet further enhancement of the present invention, if  
5 that frame itself was received in error, then a frame  
that was used to mitigate against that frame's bad  
reception will itself be the best source of speech  
information for the current frame received in error, as  
shown in FIG. 4. Hence, FIG. 4 illustrates a timing  
10 diagram indicating how multiple errors are handled. The  
data from frame 0 410 is known to be in error. The  
proposed error mitigation process employs an alternative  
virtual transmission path that has indicated data frame -  
4 420 as a suitable replacement. However, data frame -4  
15 420 is determined to be in error. In which case, a  
pointer indicates data from frame -6 430 as a frame that  
is the most similar frame to the corrupted frame -4 420.  
Therefore, frame -6 450 is used to replace frame -4 420  
and is suitable for replacing frame -1 410. In this  
20 manner, multiple frame errors can be handled, to overcome  
the problem of out-of-memory references.

This may result in references (pointers) eventually  
leading out of, what is effectively, a storage window.  
25 However, this does not need to be a problem if the  
erroneous values within the window are updated by  
removing the need for multiple references.

Alternatively, if replaced frames are stored within the  
30 buffer, then previously when frame -4 420 was the current  
frame, it would have been replaced by frame -6 430 (then  
frame -2) within the buffer, so that the buffer always  
contains only useable data.

In summary, a reference or a pointer is transmitted to the decoder in an alternative bit stream to the primary bit stream. The reference or pointer indicates a previously transmitted frame that best matches the  
5 currently transmitted frame. The reference or pointer is preferably transmitted in a parallel bit stream. If the frame is received in error at the speech decoder, the reference or pointer is used in the frame replacement error mitigation process. Hence, frame mitigation has  
10 been enhanced by extending the known immediately preceding or immediately succeeding frame replacement mechanism to any frame from a number of frames. In this regard, the number of frames used in the process is only limited by the buffering/storage mechanism and/or the  
15 processing power required to determine the minimum weighted error frame.

As indicated, the buffering/storage process of the speech parameters of the speech coder is performed over a number  
20 of frames. For example, in the context of a GSM enhanced full rate (EFR) codec, of <12 kb/sec, the storage for three seconds of speech is only 5 Kbytes. The most difficult task is therefore identifying the closest frame match from the one hundred and fifty possible frames.  
25 Hence, in one embodiment of the present invention, the aforementioned minimum-weighted error selection technique may be applied to subsets of parameters or to parameters derived from the synthesised speech, rather than all of the parameters of a speech coder frame. In other words,  
30 the LPC filter parameters (LSFs) and energy of the synthesised speech frame (a derived speech parameter from the synthesised speech computed in both the encoder and decoder) may be referenced (or pointed to) rather than

the precise coder parameters in order to save on memory and comparison processing.

In this regard, since a speech frame includes many parameters, the proposed technique can be applied in principle to any number of them. Examples of such parameters, in a CELP coder, include the following:

- (i) Line Spectral Pairs (LSPs) that represent the LPC parameters;
- (ii) Long-term predictor (LTP) lag for subframe-1;
- (iii) LTP Gain for subframe-1;
- (iv) Codebook Index for subframe-1;
- (v) Codebook Gain for subframe-1;
- (vi) Long-term predictor lag for subframe-2;
- (vii) LTP Gain for subframe-2;
- (viii) Codebook Index for subframe-2;
- (ix) Codebook Gain for subframe-2;
- (x) Long-term predictor lag for subframe-3;
- (xi) LTP Gain for subframe-3;
- (xii) Codebook Index for subframe-3;
- (xiii) Codebook Gain for subframe-3;
- (xiv) Long-term predictor lag for subframe-4;
- (xv) LTP Gain for subframe-4;
- (xvi) Codebook Index for subframe-4; or
- (xvii) Codebook Gain for subframe-4.

It is within the contemplation of the invention that a pointer could be sent referencing the set of LSPs from previous frames to match those of the current frame, rather than the whole set of parameters. Alternatively, it would be possible to have a pointer for each of a number of the above parameters.

In a wireless communication system, the parallel virtual transmission path preferably consists of transmitting a block coded reference word (where seven bits would be sufficient to support a 128-frame buffer, equating to approximately 2.5 seconds) within the unprotected bits of the data payload. This could be encoded with a BCH block code of 15 bits (with an equivalent rate of 75 bits/sec) providing up to 2-bit error correction.

Alternatively, it is envisaged that the alternative virtual transmission path may provide a combination of error correction and error detection functions. Error detection would be useful since poor reception of the reference could lead to bad mitigation. In the event of a badly received reference word, the scheme could default to the previous frame repetition. The 75 bits/sec of channel rate would only reduce the gross bit-rate of the GSM full-rate channel from 22.8 Kbits/sec. to 22.725 Kbits/sec., which would result in an insignificant loss of sensitivity.

In an alternative embodiment, such as a voice over an Internet Protocol (VoIP) communication link, the alternative virtual transmission path may be achieved by sending multiple packet streams. Although, in this context, it is desirable that the total traffic does not increase substantially since this is likely to increase the packet dropping probabilities.

A preferred mechanism would be to send the references to previous frames as described above, only where transitions occur and the speech is non-stationary. When the speech is stationary, and when conventional techniques will work relatively well, the references are

not sent. In this way the packet network is not unduly overloaded, but the majority of the performance gains are achieved. The degree of how static a speech signal becomes can be generated as a variable that can be  
5 adjusted to improve the reproduced quality in the event of a lost packet.

The decoder functionality is substantially the reverse of that of the encoder (without the additional circuitry  
10 following the multiplexer), and is therefore not described here in detail. A description of the functionality of a typical speech decoding unit can also be found in "Digital speech coding for low-bit rate communications systems" by A. M. Kondo, published by  
15 John Wiley in 1994. At the decoder, the decoder follows the standard decoding process until it determines a bad frame. When a bad frame is detected, the decoder assesses the alternative virtual transmission path to determine the alternative frame indicated by the  
20 respective reference/pointer. The decoder then retrieves the 'similar' frame, as indicated by the reference/pointer transmission. The previously indicated frame is then used to replace the received frame, to synthesise the speech.

25 Advantageously, the inventive concepts herein described may be retrofitted to existing codecs by stealing bits from an already constructed FEC scheme.

30 It is within the contemplation of the invention that any speech processing circuit would benefit from the inventive concepts described herein.

It will be understood that the bad-frame error mitigation mechanism, as described above, provides at least the following advantages:

5 (i) A more accurate replacement frame mechanism is provided, thereby reducing the risk of undesirable artefacts being audible in recovered speech frames.

(ii) The alternative virtual transmission path may be retrofitted to existing codecs, for example, by stealing bits from an already constructed FEC scheme.

10 (iii) When references to previous frames are only sent where transitions occur and the speech is non-stationary, the existing bad-frame error mitigation techniques can be used, thereby minimising any additional data required in the present invention.

15 (iv) By cross-referencing the data received for a given frame with the frames referenced in this scheme, erroneously received parameters may be detected.

Whilst the preferred embodiment discusses the application  
20 of the present invention to a CELP coder, it is envisaged by the inventors that other any speech-processing unit, where transmission errors may occur, can benefit from the inventive concepts contained herein. The inventive concepts described herein find particular use in speech  
25 processing units for wireless communication units, such as universal mobile telecommunication system (UMTS) units, global system for mobile communications (GSM), TERrestrial Trunked RAdio (TETRA) communication units, Digital Interchange of Information and Signalling  
30 standard (DIIS), Voice over Internet Protocol (VoIP) units, etc.

Apparatus of the Invention:

A speech communication unit includes a speech encoder capable of representing an input speech signal. The speech encoder includes a transmission path for transmitting a number of speech frames to a speech  
5 decoder. The speech encoder further includes a virtual transmission path for transmitting one or more references for a number of speech frames transmitted in the transmission path. The one or more references relate to  
10 an alternative speech frame within the number of speech frames transmitted on the transmission path to be used as a replacement frame when a frame is received in error.

A speech communication unit, for example the above speech communication unit having a speech encoder, includes a  
15 speech decoder adapted to receive a number of speech frames on a transmission path and one or more alternative speech frame references on a virtual transmission path. The one or more references relate to an alternative speech frame within the number of speech frames received  
20 on the transmission path to be used as a replacement frame when a frame is received in error.

#### Method of the Invention:

25 A method of performing bad-frame error mitigation in a voice communication unit includes the step of transmitting, by a speech encoder in a speech communication unit, a number of speech frames on a transmission path to a speech decoder. The speech  
30 encoder transmits, on a virtual transmission path, one or more references for a number of speech frames transmitted in the transmission path. The one or more references relate to an alternative speech frame within the number of speech frames transmitted on the transmission path to

be used as a replacement frame when a frame is received in error.

In this manner, an improved replacement frame from a  
5 number of speech frames may be selected when a speech frame is received in error.

Thus, a bad-frame error mitigation technique, and associated speech communication units and circuits, have  
10 been described that substantially alleviate at least some of the aforementioned disadvantages with known error mitigation techniques.



## Claims

1. A speech communication unit (100) comprising a speech encoder (134) capable of representing an input  
5 speech signal, the speech encoder (134) comprising a transmission path (281) for transmitting a number of speech frames to a speech decoder, the speech encoder (134) characterised by a virtual transmission path (282) for transmitting one or more references for a number of  
10 speech frames transmitted in the transmission path (281), wherein the one or more references relate to an alternative speech frame within the number of speech frames transmitted on the transmission path (281) to be used as a replacement frame when a frame is received in  
15 error.

2. The speech communication unit (100) according to Claim 1, wherein the speech encoder (134) is further characterised by:  
20 a multiplexer (250) for multiplexing said number of speech frames;  
a buffer (260), operably coupled to said multiplexer (250), to store multiplexed speech data; and  
a processor (130,270), operably coupled to said  
25 buffer (260), for characterising a current speech frame in said buffer (260) and selecting an alternative speech frame that exhibits a similar characteristic to said speech frame wherein reference to said alternative speech frame is transmitted to the decoder in the virtual  
30 transmission path (282).

3. The speech communication unit (100) according to Claim 2, wherein said processor includes a de-multiplexer function (270) to access one or more speech frames in the

buffer (260) and separates excitation parameters (274) from LPC parameters (272) of the buffered speech frame to select a speech frame exhibiting a similar characteristic.

5

4. The speech communication unit (100) according to any preceding Claim, wherein the virtual transmission path (282) is contained within the same bit stream of the transmission path (281).

10

5. The speech communication unit (100) according to any preceding Claim, wherein said transmission path (281) employs a first forward error correction protection scheme and said virtual transmission path (282) employs a second forward error correction protection different from that used in the transmission path (281).

15

6. The speech communication unit (100) according to any of preceding Claims 2 to 5, wherein said processor (130, 266, 270) selects an alternative replacement frame to provide a minimum weighted error.

20

7. The speech communication unit (100) according to Claim 6, wherein said processor (130, 266, 270) determines a minimum weighted error by evaluating a weighted Segmental signal-to-noise (SEGSNR) or average weighted SNR for each of the buffered frames.

25

8. The speech communication unit (100) according to Claim 6 or Claim 7, wherein said processor (130, 266, 270) determines a minimum weighted error of a subset of speech coding parameters.

30

9. The speech communication unit (100) according to Claim 6, Claim 7 or Claim 8, wherein said processor (130, 266) searches substantially around a pitch period of said buffered speech frames, and selects a frame exhibiting  
5 the highest SEGSNR value.

10. The speech communication unit (100) according to any preceding Claim, wherein said alternative speech frame (320) is referenced to said current speech frame  
10 only where transitions occur and speech is non-stationary.

11. A speech communication unit (100) according to any preceding Claim, characterised by a speech decoder  
15 (132) adapted to receive a number of speech frames on a transmission path (281) and one or more alternative speech frame (320) references on a virtual transmission path (282), wherein the one or more references relate to  
20 speech frames received on the transmission path (281) to be used as a replacement frame when a frame is received in error.

12. The speech communication unit (100) according to  
25 Claim 11, wherein if said alternative speech frame (420) is received in error, then a frame (430) selected as the alternative frame for said alternative frame (420) received in error and is used in the replacement of the current speech frame (410) received in error as well as  
30 the alternative speech frame (420) received in error.

13. A method of performing bad-frame error mitigation in a voice communication unit (100), the method comprising the step of:

5 transmitting, by a speech encoder (134) in a speech communication unit (100), a number of speech frames on a transmission path (281) to a speech decoder; the method characterised by the step of:

10 transmitting, on a virtual transmission path (282), one or more references for a number of speech frames transmitted in the transmission path (281), wherein the one or more references relate to an alternative speech frame within the number of speech frames transmitted on the transmission path (281) to be used as a replacement frame when a frame is received in  
15 error.

14. A speech communications unit (100) adapted to perform the method steps according to Claim 13.

20 15. A wireless communication system adapted to support the use of a transmission path (281) and a virtual transmission path (282) in accordance with any preceding Claim.

25

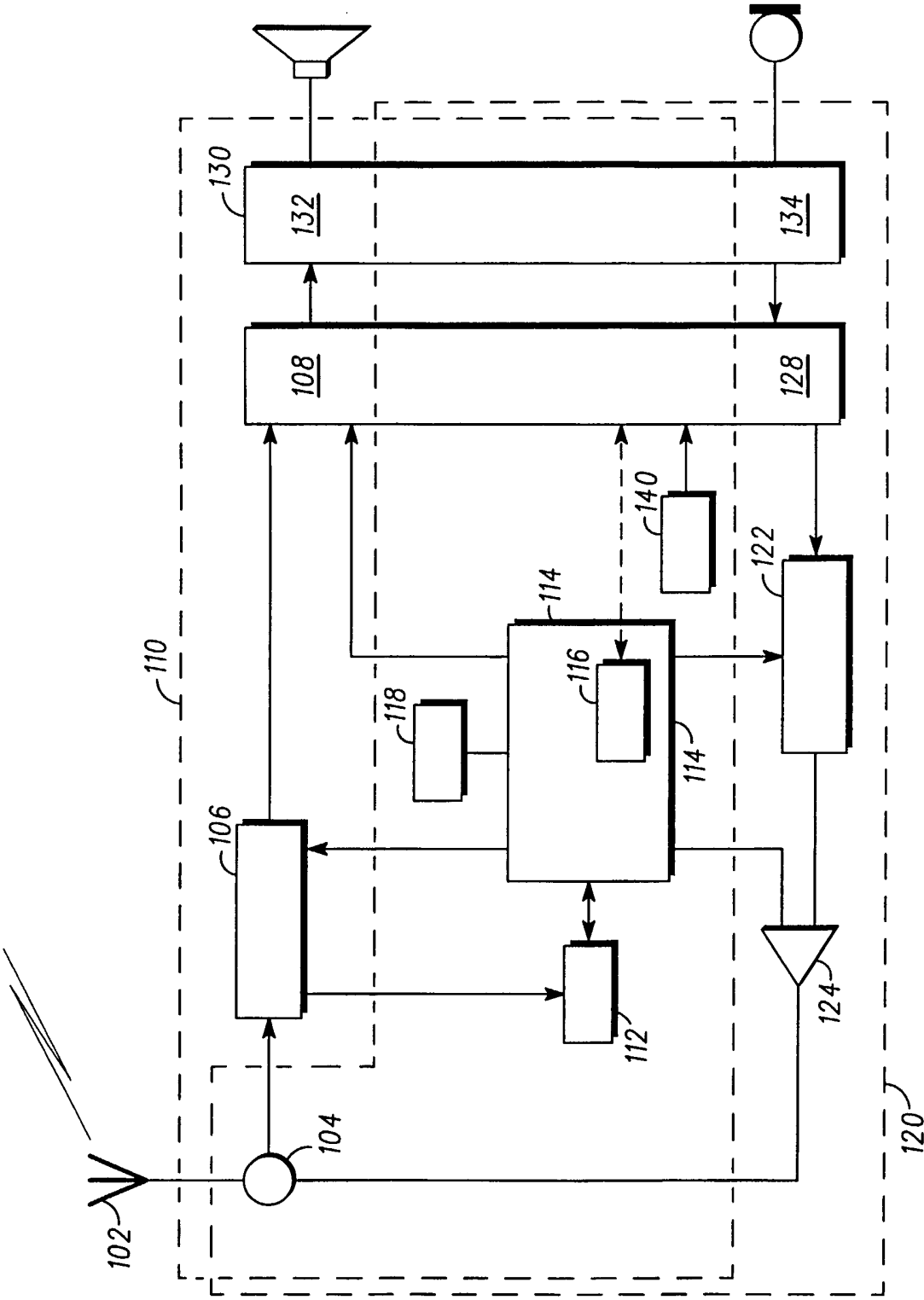
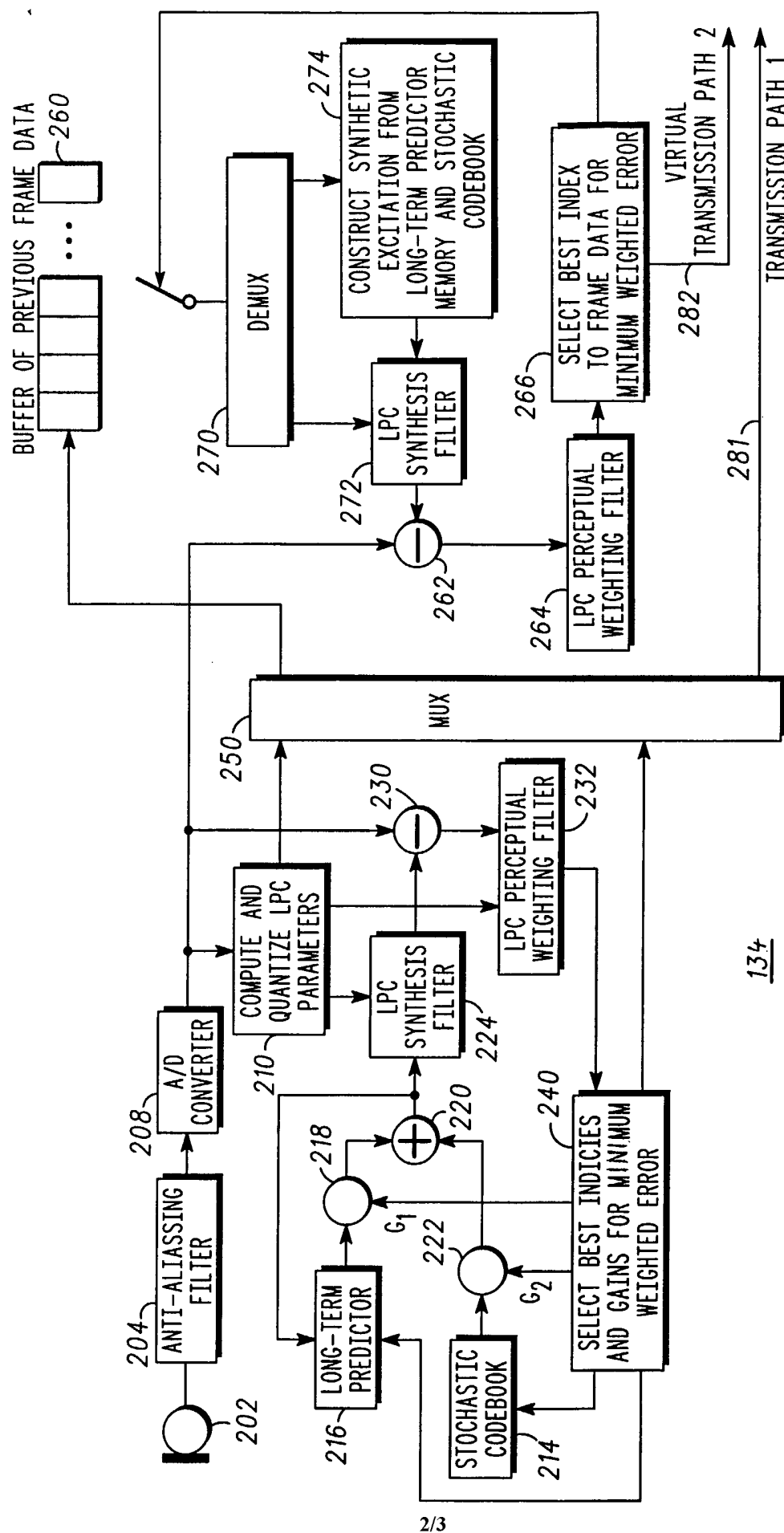


FIG. 1



134

FIG. 2

FIG. 3

300

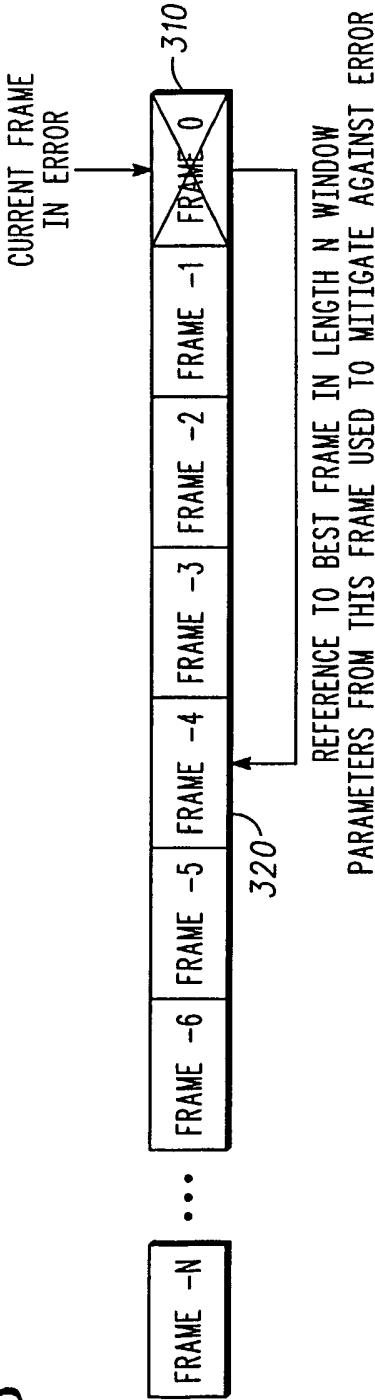
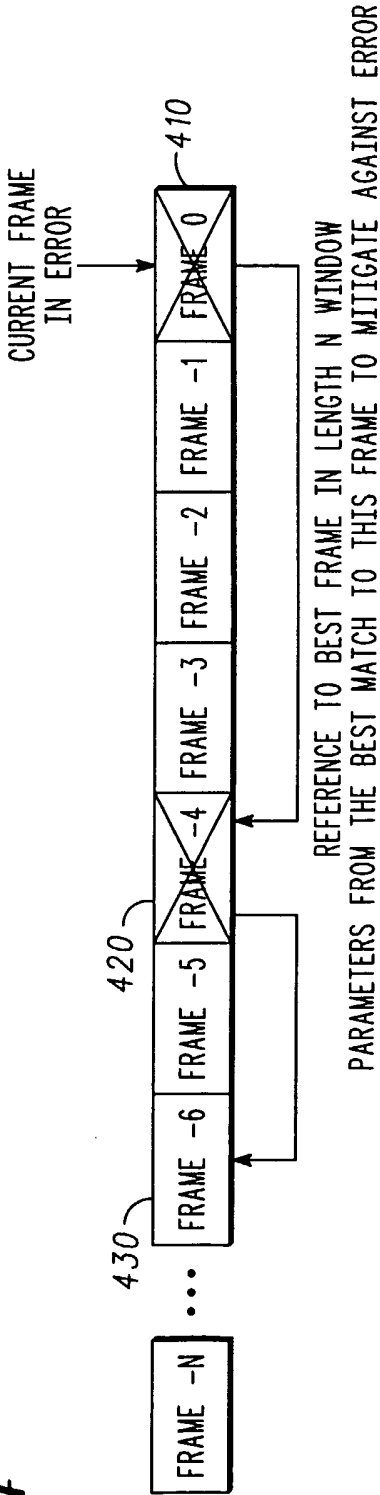


FIG. 4

400



## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/05076

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G10L19/06 H04L1/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G10L H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 917 835 A (BARRETT PHIL ET AL) 29 June 1999 (1999-06-29) abstract column 2, line 33 claims 1-4	1,10,13, 14
A	---	2-9,11, 12,15
A	WO 02 07061 A (CONEXANT SYSTEMS INC) 24 January 2002 (2002-01-24) abstract page 1, line 21 -page 2, line 15 claim 1 -----	1-15



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*G\* document member of the same patent family

Date of the actual completion of the international search

1 October 2003

Date of mailing of the international search report

09/10/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Reilly, D



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/05076

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5917835	A	29-06-1999	NONE		
-----					
WO 0207061	A	24-01-2002	AU	6627801 A	30-01-2002
			CN	1441950 T	10-09-2003
			EP	1301891 A2	16-04-2003
			WO	0207061 A2	24-01-2002
-----					