(12) **UK Patent Application** (19) **GB** (11) **2 368 761** (13) **A**

(21) Application No 0026463.0

(22) Date of Filing 30.10.2000

(71) Applicant(s)
Motorola Inc
(Incorporated in USA - Delaware)
Third Floor, 1303 East Algonquin Road, Schaumburg,
Illinois 60196, United States of America

(72) Inventor(s)
Jonathan Alastair Gibbs
James Malcolm Hoskin

(74) Agent and/or Address for Service
Colin Treleven
European Intellectual Property Department, Midpoint,
Alencon Link, BASINGSTOKE, Hants, RG21 7PL,
United Kingdom

(51) INT CL$^7$
G10L 19/12 19/02 // G10L 101:02 101:10 101:12

(52) UK CL (Edition T )
H4R RPV R22P R22V
U1S S2204 S2215

(56) Documents Cited
GB 2300548 A          EP 0411655 A
WO 99/65017 A          WO 97/27578 A
WO 95/10760 A
Int. Conf. Acoustics, Speech & Signal Processing
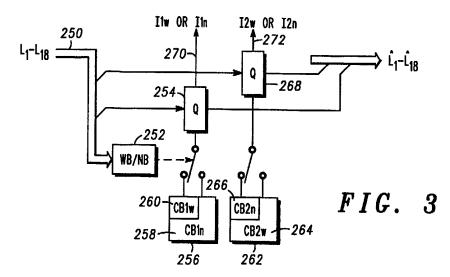(ICASSP-90) vol 1, pages 221-224

(58) Field of Search
UK CL (Edition S ) H4R RPBE RPV RPVA RPVS
INT CL$^7$ G10L 19/00 19/02 19/04 19/06 19/08 19/10
19/12 19/14
online: EPODOC,WPI,JAPIO,INSPEC,IEL

(54) Abstract Title
**Codec and methods for generating a vector codebook and encoding/decoding signals, e.g. speech signals**

(57) A speech codec (100, 200) for a speech communications unit includes a set of at least two embedded vector codebooks 256,262 capable of representing an input speech signal by a series of vectors. The codebooks 256, 262 share at least some common vectors of the series of vectors. The speech codec includes means for classifying 252 an input speech signal by selecting one of the codebooks in conjunction with at least a portion of at least one other embedded vector codebook to represent the input speech signal. The speech codec provides the advantage that at least two embedded vector codebooks, having overlapping vectors, can represent both wideband and narrowband signals. This provides for a reduced memory requirement of the codebook, in a memory constrained speech unit. Methods of generating such a combined vector codebook and encoding /decoding speech signals related to the combined vector codebook are disclosed. Application to video codecs is also envisaged.
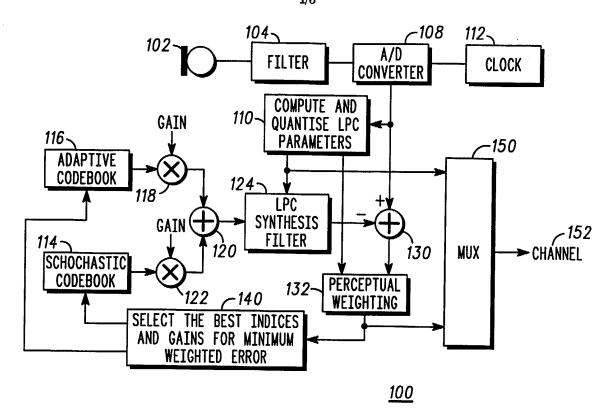


*FIG. 3*

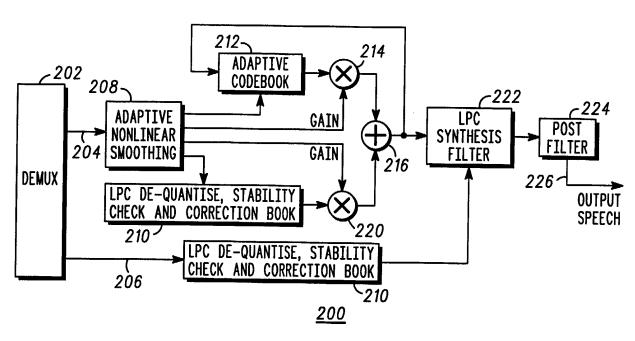GB 2 368 761 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

**FIG. 1**



**FIG. 2**

I1w OR I1n     I2w OR I2n

*250*

$L_1-L_{18}$

*270*

*272*

Q   *268*

*254*

Q

$\hat{L}_1-\hat{L}_{18}$

*252*

WB/NB

*266*

*260*— CB1w    CB2n

*258*— CB1n    CB2w —*264*

*256*      *262*

*FIG. 3*

| I1w | 0 ... 0 |
|---|---|
| *302* | *304* |

| I2w CORE | I2w XTRA |
|---|---|
| *306* | *308* |

*300* →

| I1n CORE | I1n XTRA |
|---|---|
| *310* | *312* |

| I2n | 0 ... 0 |
|---|---|
| *314* | *316* |

*322*    *324*    *326*    *328*

*320* →

| WB/NB BIT | I1w OR I1n CORE | I2w OR I2n CORE | I1n XTRA OR I2w XTRA |
|---|---|---|---|

*FIG. 4*

XXXX000

XXXX010    XXXX001

*352*

XXXX110    XXXX011

XXXX100    XXXX111

XXXX101

*350*

YYYY000

YYYY010    YYYY001

*356*

YYYY110    YYYY011

YYYY100    YYYY111
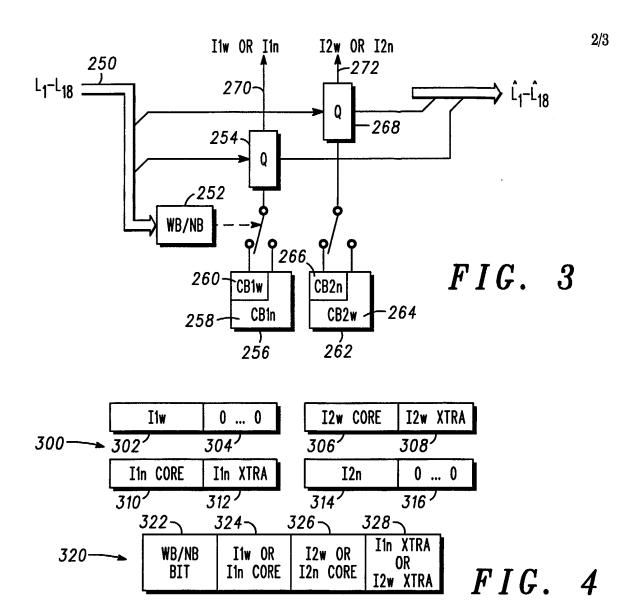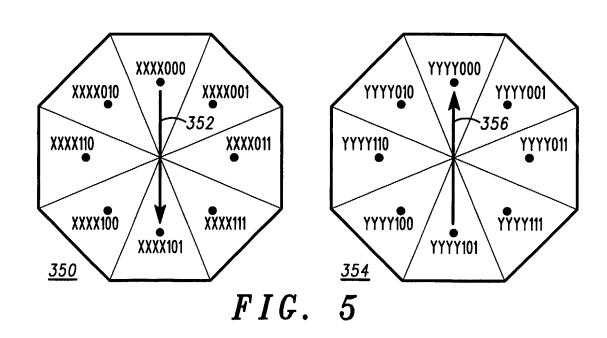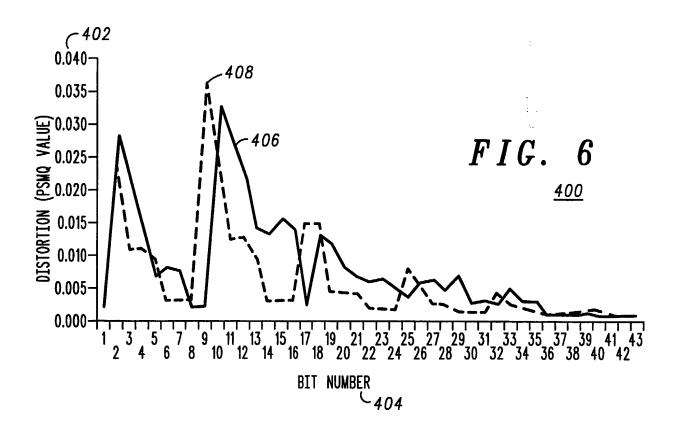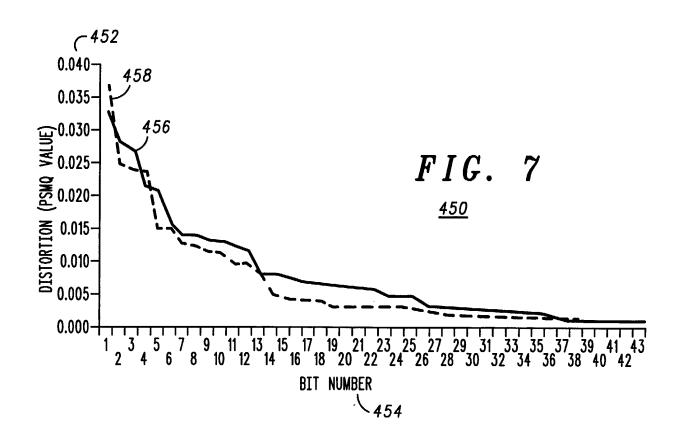
YYYY101

*354*

*FIG. 5*

*FIG. 6*

*400*



*FIG. 7*

*450*

Speech Codec And Methods For Generating A Vector Codebook
And Encoding/Decoding Speech Signals

5

## Field of the Invention

This invention relates to speech coding and methods of
optimising the performance of speech codecs in

10   communications systems.  The invention is applicable to,
but not limited to, speech codecs that accommodate
wideband and narrowband speech signals without
compromising the overall performance of the speech codec
quantiser.

15

## Background of the Invention

Many present day voice communications systems, such as

20   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

25   systems a speech encoder converts the analogue speech
pattern into a suitable digital format for
transmission and a speech decoder converts a received
digital speech signal into an audible analog speech
pattern.

30

As frequency spectrum for such voice communications
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

5 capacity of the speech patterns as much as possible, by use of compression techniques, without losing fidelity.

A popular solution in speech coding technology is the

10 application of vector quantisation (VQ). A prime incentive in using VQ can be found in Shannon's rate distortion theory, as known to those skilled in the art, which states that better performance can always be achieved by coding vectors instead of scalars.

15

The process of vector quantisation is to represent an input vector as a member of a set of fixed vectors. This set of fixed vectors is known as the VQ codebook. The fixed vector in the VQ codebook which best

20 represents the input vector is found by exhaustively searching all members of the VQ codebook and selecting the fixed vector which gives the minimum distance measure (or Euclidean distance) between it and the input vector.

25

This procedure requires that every fixed vector in the VQ codebook is searched in order to find the best representation of the input vector. Consequently searching a full VQ codebook is computationally

30 expensive and memory hungry. Although VQ has been shown to be very attractive and efficient in many

areas of speech coding, it is not without its
drawbacks.

In the field of this invention it is known that wideband
5   speech codecs are likely to find application in telephone
conferencing.  Wideband speech codecs have an input
speech bandwidth covering the 50Hz to 7KHz range,
compared to narrowband or telephone-band codecs that have
an input speech bandwidth of 250Hz to 3.3KHz.
10

The consequence of this is that wideband speech codecs
will often be tandemmed with narrowband speech codecs.
Tandemming is a term which is used to describe the
situation where speech previously processed by one speech
15   encoder/decoder is processed by a second speech
encoder/decoder pair.

Furthermore, in these situations, the speech quality
requirement of such tandemmed codecs is to achieve
20   equivalence to the best narrowband codecs i.e. GSM
Enhanced full-rate codec (EFR).  It is therefore
appropriate to consider the performance of any wideband
line spectral frequency (LSF) VQ scheme in the presence
of narrowband speech.
25

To accommodate the processing of both wideband and
narrowband speech signals in a speech synthesis unit, one
option may be to use a classified VQ scheme with two sets
of codebooks: one to represent the wideband speech and
30   one to represent the narrowband speech.  In this case, a
respective codebook would be selected by a special "mode"
bit, where the mode bit indicates whether the subsequent

data bits represent a wideband or narrowband speech signal.

One problem that is inherent with this technique is that two individual codebooks are required, one arranged for wideband performance and one arranged for narrowband performance. Subsequent bit errors applied to the mode bit would result in large de-quantiser errors. To resolve any such mode bit error problems, heavy forward error correction (FEC) protection is required in the resulting speech codec.

One disadvantage associated with quantising wideband LSFs, when there is a possibility that narrowband codec tandemming may be present, is highlighted below in Table 1.

The representative codecs have been simulated with each predictor of the speech codec arranged to be an 18th order split vector quantiser, with the eighteen associated line spectral frequencies split into six groups of three bits each.

Table 1: 7KHz & 3KHz Spectral Distortion Results for the
1st order MA-PVQ 40 bit Quantisers.

| Quantiser Configuration | Wideband Performance | | | Narrowband Performance | | |
|---|---|---|---|---|---|---|
| | Mean SD (dB) | % Frames 3-5 dB | % Frames > 5 dB | Mean SD (dB) | % Frames 2-4 dB | % Frames > 4 dB |
| 6,8,7,7,6,6 | 1.418 | 0.887 | 0.019 | 1.292 | 5.754 | 0.058 |
| 6,7,7,7,7,6 | 1.417 | 1.012 | 0.010 | 1.384 | 8.047 | 0.087 |
| 7,7,7,7,6,6 | 1.418 | 0.916 | 0.019 | 1.301 | 6.062 | 0.077 |
| 7,8,7,7,6,5 | 1.389 | 0.771 | 0.019 | 1.202 | 4.163 | 0.058 |
| 6,8,8,7,6,5 | 1.396 | 0.684 | 0.019 | 1.231 | 4.597 | 0.029 |
| 7,7,7,7,7,5 | 1.387 | 0.810 | 0.010 | 1.300 | 6.081 | 0.077 |
| 7,7,8,7,6,5 | 1.397 | 0.723 | 0.019 | 1.240 | 4.857 | 0.058 |
| **7,8,8,7,6,4** | **1.379** | **0.646** | **0.019** | 1.137 | 3.094 | 0.039 |
| 7,9,9,6,5,4 | 1.453 | 0.896 | 0.029 | 1.033 | 1.744 | 0.000 |
| 7,9,9,7,5,3 | 1.431 | 0.752 | 0.000 | 1.017 | 1.677 | 0.010 |
| 8,8,8,7,6,3 | 1.390 | 0.723 | 0.010 | 1.073 | 2.274 | 0.029 |
| **8,9,8,7,6,2** | **1.419** | **0.964** | **0.010** | **0.993** | **1.484** | **0.019** |
| 8,8,8,8,5,3 | 1.419 | 0.867 | 0.019 | 1.067 | 2.101 | 0.029 |
| 8,9,8,7,5,3 | 1.428 | 0.848 | 0.000 | 0.995 | 1.503 | 0.029 |
| 8,9,8,8,5,2 | 1.447 | 1.224 | 0.010 | 0.986 | 1.378 | 0.019 |
| **8,9,9,7,5,2** | 1.459 | 1.245 | 0.010 | **0.947** | **1.051** | **0.010** |

5

The wideband and narrowband distortion figures are
presented in the form suggested by; "LSF Quantization in
Wideband Speech Coders", Proceedings of 1999 IEEE
10  Workshop on Speech Coding, pp 25-27, M. Ferhaoui & S. Van
Gerven and "Efficient Vector Quantization of LPC
Parameters at 24 Bits/Frame", IEEE Transactions on Speech
and Audio Processing, Vol 1, No 1, March 1993, K.K.
Paliwal & B.S. Atal.

Table 1 details the wideband and narrowband spectral distortion figures for a 40-bit first order moving average quantiser trained on 50:50 wideband:narrowband speech. As mentioned, the configuration column denotes the number of bits allocated to each of six, moving-average predictive split-vector quantisers, applied to LSFs 1-3, 4-6, 7-9, 10-12, 13-15 and 16-18 respectively.

It is worth noting that a wideband speech codec would typically be represented by an even distribution of bits allocated to each of the six split-vector quantisers, to provide an approximately even frequency response across the full range of the line spectral frequencies. In contrast, a narrowband speech codec would have an uneven distribution of bits associated with each quantiser, with more bits allocated to the lower frequencies of the LSFs.

It is clear that the best wideband performance is obtained by a first configuration (7,8,8,7,6,4), whilst the best narrowband performance is obtained with a different second configuration (8,9,9,7,5,2). Both configurations consist of the same total number of bits, but with different distributions of the numbers of bits per quantiser.

A compromise quantiser such as 8,9,8,7,6,2 provides substantially inferior performance to both of these. In addition, it is noteworthy that this dual quantiser approach would require a total of:

3 x (640 + 1444) words = 6252 words in memory.

Hence, a need exists to provide a bit-error robust speech codec, preferably complementary to popular, proven speech codecs such as embedded split VQ codecs, that can provide improved quantisation of wideband line spectral frequencies (lsfs) for wideband speech and also in narrowband tandemming.

The present invention aims to provide a speech codec and method of optimising a performance of the speech codec to at least alleviate some of the aforementioned disadvantages.

## Summary of the Invention

In a first aspect of the present invention, a speech coder for a speech communications unit in accordance with claim 1 is provided.

In a second aspect of the present invention, a speech communications unit adapted to include the speech coder of any one of claims 1 to 10 is provided.

In a third aspect of the present invention, a method of generating a speech vector codebook in a speech communications unit in accordance with claim 12 is provided.

In a fourth aspect of the present invention, a speech communications unit adapted to include a speech vector codebook generated, in accordance with any one of claims 12 to 20, is provided.

In a fifth aspect of the present invention, a method of encoding a speech signal, in accordance with claim 22, is provided.

In a sixth aspect of the present invention, a speech communications unit adapted to employ a speech encoding method in accordance with any one of claims 22 to 26 is provided.

In a seventh aspect of the present invention, a method of decoding a speech signal, in accordance with claim 28, is provided.

In an eighth aspect of the present invention, a speech communications unit adapted to employ a speech decoding method in accordance with any one of claims 28 to 31 is provided.

## Brief Description of the Drawings

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

FIG. 1 shows a block diagram of a code excited linear predictive speech encoder that can be adapted to support the various inventive concepts of a preferred embodiment of the present invention;

FIG. 2 shows a block diagram of a code excited linear predictive speech decoder that can be adapted to support

the various inventive concepts of a preferred embodiment
of the present invention;

FIG. 3 shows a 2-way split VQ codebook applied to
eighteen wideband line spectral frequencies (LSFs)
adapted to support the inventive concepts of the
preferred embodiments of the present invention;

FIG. 4 shows a preferred packing arrangement of bits for
the 2-way split VQ codebook of FIG. 3, adapted to support
the inventive concepts of the preferred embodiments of
the present invention;

FIG. 5 shows an octagon partitioned to reflect eight
separate locations for the 2-way split VQ codebook of
FIG. 3, in accordance with a preferred embodiment of the
present invention;

FIG. 6 shows a graph that demonstrates the error
resilience of a preferred embodiment of the present
invention in the presence of 10% bit errors applied on a
per-bit basis with natural ordering; and

FIG. 7 shows a further graph that demonstrates the error
resilience of a preferred embodiment of the present
invention in the presence of 10% bit errors applied on a
per-bit basis with ranked ordering.

## Description of Preferred Embodiments

Referring first to FIG. 1, a block diagram of a code excited linear predictive speech encoder 100, according to a preferred embodiment of the present invention, is shown. An acoustic input signal to be analysed is applied to speech coder 100 at microphone 102. The input signal is then applied to filter 104. Filter 104 will generally exhibit band-pass filter characteristics. However, if the speech bandwidth is already adequate, filter 104 may comprise a direct wire connection.

The analog speech signal from filter 104 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 analog-to-digital (A/D) converter 108, 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) via clock 112.

The digital output of A/D 108, which may be represented as input speech vector $s(n)$, is then applied to coefficient analyser 110. 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 110. The generated speech coder parameters may include the following: LPC parameters, long-term

predictor (LTP) parameters, excitation gain factor ($\sqrt{}$) (along with the best excitation codeword I). Such speech coding parameters are applied to multiplexer 150 and sent over the channel 152 for use by the speech synthesizer at

5 the decoder. The input speech vector s(n) is also applied to subtractor 130, the function of which is described later.

In the preferred embodiment of the invention, a specially

10 constructed family of embedded codebooks is used in order to best represent the LPC parameters of the input speech signal. Hence, coefficient analyser 110 has been adapted to incorporate the specially constructed family of embedded codebooks.

15

It is within the contemplation of the invention that any number of embedded codebooks would benefit from the inventive concepts described herein. The inventive concepts of the embedded codebook arrangement are further

20 described with reference to FIG. 3.

Within the conventional CELP encoder of FIG. 1, the codebook search controller 140 selects the best indices and gains from the adaptive codebook within block 116 and

25 the stochastic codebook within block 114 in order to produce a minimum weighted error in the summed chosen excitation vector used to represent the input speech sample. The output of the stochastic codebook 114 and the adaptive codebook 116 are input into respective gain

30 functions 122 and 118. The gain-adjusted outputs are then summed in summer 120 and input into the LPC filter 124, as is known in the art.

For each individual excitation vector $u_i(n)$, a reconstructed speech vector $s'_i(n)$ is generated for comparison to the input speech vector $s(n)$. Gain block

5      122 scales the excitation gain factor '$\sqrt{}$'. Such gain may be pre-computed by coefficient analyser 110 and used to analyse all excitation vectors, or may be optimised jointly with the search for the best excitation codeword I, generated by codebook search controller 140.

10

The scaled excitation signal $\sqrt{} u_i(n)$ is then filtered by the linear predictive coding filter 124, which preferably includes a long-term predictor (LTP) filter and a short-term predictor (STP) filter, to generate the

15     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 130.

20

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 132, utilising the weighting filter

25     parameters (WTP) generated by coefficient analyser 110. Perceptual weighting accentuates those frequencies where the error is perceptually more important to the human ear, and attenuates other frequencies.

30     An energy calculator function inside the codebook search controller 140 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
5    is then output over the channel as the best excitation
code I.

In the alternative, codebook search controller 140 may
determine a particular codeword that provides an error
10   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
15   speech coding for low-bit rate communications systems" by
A. M. Kondoz, published by John Wiley in 1994.

As mentioned earlier, the coefficient analyzer 110 has
also been adapted to employ at least some of the
20   inventive concepts of the present invention. To
accommodate vectors in either the wideband or narrowband
vector space, the coefficient analyser 110 is used to
train the quantisers and to determine whether the input
speech comprises wideband or narrowband speech.
25

The inventors of the present invention have recognised
the opportunity to use the same training data, or at
least very similar data, to train each of the quantisers.
The different sized quantisers cover much the same signal
30   vector space and hence a smaller quantiser is embedded
within the larger quantiser leading to a more compact
representation. In the present case, the storage may be

reduced from the dual quantiser approach requiring a total of 3 x (640 + 1444) words = 6252 words to 3 x 1600 = 4800 words.

5    In the preferred embodiment of the invention, it is desirable for the coefficient analyser 110 to send an additional mode bit to indicate which quantiser set (i.e. which of the specially constructed family of embedded codebooks) is being used.  The quantiser set will

10   preferably refer to a wideband or narrowband arrangement. The codebook index transmission is structured in order to minimise the effect of errors to this mode bit, as described later with respect to FIG. 4.  The consequence of such a careful structuring of the codebook index

15   transmission means that any bit error(s) in the mode bit have much less impact than in any two or more independent-codebook prior art approach.

     Referring now to FIG. 2, a block diagram of a code

20   excited linear predictive speech decoder 200 is shown, according to a preferred embodiment of the present invention.  The decoder functionality is substantially the reverse of that of the encoder.

25   The received multiplexed signal is input into de-multiplexer 202, which separates the excitation parameters 204 from the LPC parameters 206.

     For each block of speech, a set of linear predictive

30   coding (LPC) parameters were produced in the encoder by coefficient analyser 110 as described with reference to FIG. 1.

In the preferred embodiment of the present invention, the LPC parameters are input into an LPC de-quantiser, stability check and correction block 210 to obtain a
5    local stable version of the synthesis filter even in the presence of channel bit errors. The LPC de-quantise, stability check and correction block 210 has been adapted to encompass the inventive concepts contained herein.

10   The LPC de-quantise, stability check and correction block 210 receives the LPC parameters and mode bit sent from the corresponding encoder function. The LPC de-quantise function of block 210 includes the corresponding embedded codebook arrangement of the encoder, such that the
15   determination of the at least one mode bit can select the embedded codebook arrangement that best describes the encoded and transmitted speech signal. The LPC de-quantise, stability check and correction block 210 also controls the filter co-efficients of the LPC synthesis
20   filter 222 in order to reconstruct the transmitted speech vector $s'_i(n)$. The output from the LPC synthesis filter 222 is input to a post filter process 224, which subsequently outputs the reconstructed speech 226.

25   The excitation parameters 204 may include: excitation gain factor $\sqrt{}$ together with the best excitation codeword I, and are input into an adaptive non-linear smoothing function 208. The output from the adaptive non-linear smoothing function 208 provides the precise adaptive and
30   stochastic codebook indices and gains that form the excitation for the synthesis filter. As such, the outputs from the adaptive non-linear smoothing function

208 are input to stochastic codebook 218 and adaptive codebook 212.

The gain controls are input to adaptive codebook gain
5   block 214, which receives an output from the adaptive codebook 212, and stochastic codebook gain block 220, which receives an output from the stochastic codebook 218.   The output from the respective gain blocks 214, 220 are input to summing junction 216, whose output is fed
10  into the LPC synthesis filter 222 and fed back to the adaptive codebook 212, as known in the art.

It is again within the contemplation of the invention that any number of embedded codebooks would benefit from
15  the inventive concepts described herein.   The inventive concepts of the embedded codebook arrangement, applicable to either the encoder or decoder function are further described with reference to FIG. 3.

20  A more detailed 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. Kondoz, published by John Wiley in 1994.
25

The inventive concepts of the preferred embodiment of the invention are best illustrated with an example.   FIG. 3 shows a 2-way split VQ codebook applied to eighteen wideband line spectral frequencies (LSFs).   The number of
30  split VQ codebooks together with the number of LSFs have only been selected to more easily show the benefits of the inventive concepts of the present invention.

The input LSFs (L1-L18) 250 are quantised by first quantiser 254 and second quantiser 268 to derive estimates and the two binary indices "I1" 270 and "I2" 272 using a respective first embedded codebook 256 and second embedded codebook 262. The LSFs (L1-L18) 250 are fed into a mode-bit detector 252, that selects the respective embedded codebook to provide the most appropriate one for the speech signal presented.

The first embedded codebook 256 contains a first set of core entries, in this case appropriate for wideband speech 260, and additional entries appropriate for narrowband speech 258. The second embedded codebook 262 contains a second set of core entries, this time appropriate for narrowband speech 266, and additional wideband entries 264.

In each case, the core entries are always searched in each quantiser and are indexed by a set of core bits. Additional entries are searched, depending upon the mode, and a set of "extra" bits are formed. The codebook is structured such that when the full codebook is searched, the "extra" bits are effectively zero for the core entries. This is depicted in FIG. 4.

This arrangement of core bits provides for a constant sum of the bit allocations for each of the two modes, wideband or narrowband.

FIG. 4 shows the preferred packing of bits for the 2-way split VQ of FIG. 3. The configuration of the bit stream

320 comprises the mode bit 322 (indicating a wideband or
narrowband input signal) followed by the "I1" core bits
324 (either wideband or narrowband) and the "I2" core
bits.  Finally the "I1" narrowband extra bits or the "I2"

5    wideband extra bits complete the preferred packing
configuration.


Since one of the codebooks is fully searched, and for the
other codebook only the core entries are searched,

10   advantageously only one set of extra bits needs to be
sent. The extra bits for the partially searched codebook
will effectively be zero i.e. either 302, 306 and 308 or
310, 312, 314 need to be sent since either 304 or 316
will convey no information.

15

By packing the bits as shown in FIG. 4, the mode bit and
core bits are beneficially always in the same locations.
Hence, the impact of a mode bit error can be arranged to
result in much smaller errors in the two quantisers than

20   in prior art arrangements.


Clearly, in order to first generate the embedded vector
codebook arrangement described in FIGs 1-4, a series of
"test" input speech signals may be used, to obtain the

25   optimum set of vectors to represent all input speech
signals.


FIG. 5 shows two octagons 350 and 354, partitioned to
reflect eight separate locations identified by a

30   respective 3-bit address.  The two octagons 350, 354
individually each represent one of the two split VQ
codebooks of FIG. 3.

For simplicity, the example shows the case where three "extra" LSBs are used. The xxxx & yyyy represent core entry bit patterns for each of the respective embedded codebooks. For embedded entries, a potentially "non-zero extra" position will be appended instead of all zeros and the "extra" LSBs of the larger codebooks will be set to zero. As can be seen from error arrows 352 and 356, the maximum error for a (WB/NB) mode bit is equivalent to that of several LSB errors in each codebook.

In order to achieve this advantageous feature, the codebook entries of the embedded codebook, trained using relevant and appropriately varied speech patterns, must be interlaced regularly within the large codebook. In addition, index reassignment of the combined codebook must be performed such that LSB errors in the indices result in small perceptual distances. This may be arranged using a simulated annealing method as is well known to those skilled in the art.

A set of codebooks was derived and selectively searched. In order to determine which codebook configuration to search during each frame an appropriate narrowband speech indicator was employed.

Any reliable indicator may be used, but in this case the squared sum of the LSF VQ weights, w13-w18, corresponding to LSFs 13-18, (neglecting the split adjacent weighting) was filtered:

(first-order recursive filter = $0.1/(1-0.9*z^{-1})$)    (1)

and thresholded.

This provides a reliable indication of whether the input is narrowband or wideband speech, except during silence periods, when the bias is slightly in favour of wideband speech.

The results of the hybrid wideband/narrowband scheme, in accordance with a preferred embodiment of the invention, is shown in Table 2.

Table 2: 7KHz and 3KHz Spectral Distortion Results for the Hybrid 41-bit Quantiser.

| Quantiser Configuration | Wideband Performance | | | Narrowband Performance | | |
|---|---|---|---|---|---|---|
| | Mean SD (dB) | % Frames 3-5 dB | % Frames > 5 dB | Mean SD (dB) | % Frames 2-4 dB | % Frames > 4 dB |
| Hybrid 7,8,8,7,6,4 & 8,9,9,7,5,2 | 1.396 | 0.761 | 0.019 | 0.985 | 1.359 | 0.019 |

As can be seen by comparing the results of Table 2 with the values in Table 1, the hybrid scheme performs very well for both wideband and narrowband speech and offers only slightly degraded performance to the optimum in either configuration.

The graph 400 shown in FIG. 6 demonstrates the error resilience of the preferred embodiment of the invention in the presence of 10% bit errors, applied on a per-bit

basis measured using an objective distortion measure, such as the perceptual speech quality measure (PSQM value)., as defined by the ITU-T Recommendation P.861.

5    Graph 400 shows the bit error sensitivity profiles, this time for two 43-bit quantisers according to the embodiment.  The distortion 402 (PSQM value) is shown plotted against bit number 404 on a bit-by-bit basis.

10   The two quantisers shown are the hybrid 8,8,8,7,6,5 & 8,9,9,8,6,2 wideband/narrowband scheme 406, according to the preferred embodiment of the invention, and an 8,8,8,7,7,5 wideband-only scheme 408.

15   For the hybrid quantiser the mode bit is the first bit and then the other core quantiser bits (8,8,8,7,6,2) are presented MSB first for each quantiser in-turn, followed finally by the three extra bits as depicted in FIG. 4. For the 8,8,8,7,7,5 wideband-only scheme the bits are
20   presented in MSB first natural order.

     The overall performance of the new quantiser in the presence of bit errors can be seen to be only very slightly worse than the wideband-only scheme (see rank-
25   ordered sensitivities).  In particular, the graph highlights that the sensitivity of the mode bit is 33rd out of 43, i.e. near, but not quite at, the bottom of the rank-ordered results.

30   The explanation for the mode bit not being the least sensitive bit (as in the optimal case) positioned at the bottom of the rank ordering (see FIG. 7) is that when a

mode error occurs, several LSB changes (in the three-quantiser tables) occur which together are more significant than a single LSB change (bottom of the rank ordering). This clearly shows that the embedded structuring of the LSF VQ and bit stream has beneficially rendered the LSF VQ relatively immune to bit errors.

The graph 450 shown in FIG. 7 demonstrates the error resilience of the preferred embodiment of the invention where the error sensitivity profile is shown rank-ordered, as compared to a bit-by-bit basis as shown in FIG. 6.

Graph 450 shows the bit error sensitivity profiles, of the same two 43-bit quantisers, 456 and 458. The distortion 452 is shown plotted against re-ordered bit position 454 on a rank-ordered bit basis. In particular, the graph highlights that the two quantisers have broadly similar error sensitivity profiles and that the addition of the mode bit has not increased error sensitivity.

It will be understood that the bit-error robust embedded split vector quantiser for wideband line spectral frequencies (lsfs) in narrowband tandemming described above provides at least the following advantages.

The invention provides for a single speech codec codebook that can quantise both wideband and narrowband signals in a near optimal manner to that of two independently-optimised speech codec codebooks. This provides for a reduced memory requirement of the codebook, in a memory constrained speech unit.

The inventive concepts described herein find particular use in speech processing units that are flexible enough to cope with a variety of bandwidth constrained speech input signals, such as future third generation cellular telecommunications systems

It will, of course, be appreciated that the above description has been given by way of example only and that modifications in detail may be made within the scope of the present invention. For example, whilst the preferred embodiment discusses the application of the present invention to a split vector quantiser codebook, it is envisaged by the inventors that other codebooks and speech coder techniques can benefit from the inventive concepts contained herein.

The skilled addressee will equally appreciate that any number of line spectral frequencies can be accommodated, in a LSF codebook arrangement. Indeed, the present invention may be implemented outside of the line spectral frequency area, such as in video encoding and decoding.

It is within the contemplation of the invention that alternative quantiser configurations can benefit from such inventive concepts. Furthermore, the inventive concepts can be applied to any LPC order, with any bit-division relationship.

It is also within the contemplation of the present invention, that the inventive concepts contained herein can be equally employed in any classified overlapping

codebook arrangement, not necessarily limited to the overlapping arrangement between wideband and narrowband speech signals.

5     Thus, a bit-error robust speech codec has been provided that is complementary to popular, proven speech codecs such as embedded split VQ codecs.  The bit-error robust speech codec accommodates wideband line spectral frequencies in narrowband tandemming, and alleviates at

10   least some of the aforementioned disadvantages.

## Claims

1. A speech coder (100, 200) for a speech communications
unit comprising a set of at least two embedded vector
5   codebooks capable of representing an input speech signal
by a series of vectors, the speech coder (100, 200)
characterised by the at least two embedded vector
codebooks (256, 262) sharing at least some common vectors
of the series of vectors and means for classifying (252)
10   the input speech signal by selecting one of the at least
two embedded vector codebooks (256, 262) in conjunction
with at least a portion of said at least one other
embedded vector codebook of the at least two embedded
vector codebooks (256, 262) to represent the input speech
15   signal.

2. The speech coder of claim 1, wherein indices which
address individual vector entries within the at least two
embedded vector codebooks (256, 262) are assigned such
20   that distortion resulting from incorrect classification
is minimised.

3. The speech coder of claim 1 or claim 2, further
characterised by the speech codebook being a vector
25   quantisation codebook, wherein the at least two embedded
vector codebooks (256, 262) are classified as
predominantly wideband or narrowband embedded codebooks
having shared common vectors.

4. The speech coder of any one of preceding claims 1 to
3, the speech coder further characterised by the selected
one of the embedded vector codebooks (256, 262) providing
a coarse resolution and the at least a portion of at

5    least one other embedded vector codebook providing a fine
resolution to represent the speech signal input to the
speech coder.

5. The speech coder of any one of the preceding claims,
10   further characterised by quantisation means (110, 210)
operably coupled to the means for classifying (252) for
quantising a line spectral frequency input speech signal.

6. The speech coder of any one of the preceding claims,
15   further characterised by analysis means operably coupled
to the speech vector codebook for analysing an incoming
speech signal to determine a particular characteristic of
the speech signal.

20   7. The speech coder of claim 6, wherein the classifying
means includes means for generating or recovering at
least one mode bit to identify a characteristic of the
speech signal associated with wideband or narrowband
speech.
25

8. The speech coder of claim 7, wherein the input speech
signal comprises at least one core bit, the speech coder
further characterised by positioning means to position
said mode bit and said at least one core bit for both
30   first and second embedded vector codebooks (256, 262) in
the same vector space.

9. The speech coder according to any one of the preceding claims, further comprising appending means, the appending means .adapted to append at least one zero (304, 306) to a vector representation (300) to represent vector positions

5  for either said first or second embedded speech vector codebook (256, 262) in a combined embedded speech vector codebook.

10. The speech coder according to any one of the
10 preceding claims, further comprising index reassignment means (350) for re-arranging the vector positions of the first and second codebook vectors in the combined embedded speech vector codebook to minimise perceptual distance between said embedded codebook vectors.
15

11. A speech communications unit adapted to include the speech coder of any one of preceding claims 1 to 10.

12. A method of generating a speech vector codebook in a
20 speech communications unit, the method characterised by the steps of:

representing speech signals by a series of vectors; and

25 generating at least two embedded vector codebooks that share common vectors of the series of vectors, to generate said speech vector codebook, such that, in use, at least a portion of each of the at least two embedded vector codebooks is used to represent an input speech
30 signal.

13. The method of generating a speech codebook according to claim 12, wherein the step of generating at least two embedded vector codebooks includes the steps of:

5   generating a first embedded vector codebook of the at least two embedded vector codebooks to substantially represent narrowband speech signals; and

generating a second embedded vector codebook of the at 10   least two embedded vector codebooks to substantially represent wideband speech signals.

14. The method of generating a speech vector codebook according to claims 12 or 13, the method further 15   characterised by the step of:

quantising, by quantisation means, a line spectral frequency input speech signal.

20   15. The method of generating a speech vector codebook according to any one of claims 12 to 14, wherein the step of generating at least two embedded vector codebooks includes generating one of the embedded vector codebooks to provide a coarse resolution, and generating the at 25   least one other embedded vector codebook to provide a fine resolution, to represent a speech signal input to the speech coder.

16.  The method of generating a speech vector codebook according to any one of claims 12 to 15, the method further characterised by the step of:

5   analysing an incoming speech signal by analysis means in the speech encoder to determine a particular characteristic of the speech signal associated with wideband or narrowband signals; and

10  generating a mode bit (322) to represent the particular characteristic of the speech signal.

17. The method of generating a speech vector codebook according to claim 16, wherein the speech signals include
15  core speech bits (324, 326), the method further characterised by the step of:

positioning the at least one mode bit and core speech bits for both embedded vector codebooks in substantially
20  the same location of the speech vector codebook.

18. The method of generating a speech vector codebook according to any one of preceding claims 12 to 17, the method further characterised by the step of:
25

appending (304, 316) zeros to vector entries in either of the first or second embedded vector codebooks.

19. The method of generating a speech vector codebook
30  according to any one of preceding claims 12 to 18, the method further characterised by the step of:

performing index re-assignment of the speech vector
codebook to maintain a relatively small perceptual
distance between respective vector positions thereby
minimising distortion errors resulting from incorrect
5    classification in a speech encoding or decoding process.


20. The method of generating a speech vector codebook
according to any one of preceding claims 12 to 19, the
method further characterised by the step of:
10

interlacing said vector entries at substantially evenly
distributed positions within the combined speech vector
codebook.


15   21. A speech communications unit adapted to incorporate a
speech vector codebook generated in accordance with any
one of method claims 12 to 20.

22. A method of encoding a speech signal, the method
characterised by the steps of:

receiving a speech signal;

5

identifying a characteristic of the speech signal;

selecting one of the at least two embedded vector
codebooks, in conjunction with at least a portion of at
10 least one other embedded vector codebook of the at least
two embedded vector codebooks, to represent the input
speech signal based on the identified characteristic; and

transmitting information identifying said selected one
15 embedded vector codebook as a representation of said
received speech signal.

23. The method of encoding a speech signal according to
claim 22, wherein the step of identifying a
20 characteristic of the received speech signal encompasses
identifying whether the speech signal is associated with
wideband or narrowband speech.

24. The method of encoding a speech signal according to
25 claim 22 or 23, wherein the step of selecting includes
selecting one of the embedded vector codebooks to provide
a coarse resolution, and the at least a portion of at
least one other embedded vector codebook to provide a
fine resolution, to represent the speech signal input to
30 the speech coder.

25. The method of encoding a speech signal according to any one of claim 22 to 24, the method further characterised by the step of:

5   generating at least one mode bit (322) to identify said characteristic of the speech signal.

26. The method of encoding a speech signal according to any one of preceding claims 22 to 25, the method further

10  characterised by the step of:

appending (304, 316) at least one zero to a vector representation of the input speech signal to represent said vector position for either said first or second

15  embedded speech vector codebook in the speech vector codebook.

27. A speech communications unit adapted to employ a speech encoding method in accordance with any one of

20  method claims 22 to 26.

28. A method of decoding a speech signal, the method characterised by the steps of:

receiving a speech signal;

5

identifying a characteristic of the received speech signal; and

selecting one of at least two embedded vector codebooks,
10 in conjunction with at least a portion of at least one other embedded vector codebook of the at least two embedded vector codebooks, to represent the input speech signal based on the identified characteristic.

15 29. The method of decoding a speech signal according to claim 27, wherein the step of identifying a characteristic of the received speech signal encompasses identifying whether the speech signal is associated with wideband or narrowband speech.
20

30. The method of decoding a speech signal according to claim 28 or 29, the method further characterised by the step of:

25 recovering (322) at least one mode bit to identify said characteristic of the speech signal.

31. The method of decoding a speech signal according to any one of preceding claims 28 to 30, the method further characterised by the step of:

5   removing at least one zero appended to a vector representation of the input speech signal to represent a vector position in either said first or second embedded speech vector codebook.

10   32. A speech communications unit adapted to employ a speech decoding method in accordance with any one of method claims 28 to 31.

33. A speech coder substantially as hereinbefore
15   described with reference to, and/or as illustrated by, FIG. 1 or FIG. 2 of the drawings.

34. A method of generating a vector codebook in a speech coder substantially as hereinbefore described with
20   reference to, and/or as illustrated by, FIG. 3 or FIG. 4 or FIG. 5 of the drawings.

**Application No:** GB 0026463.0     **Examiner:** Martyn Dixon

**Claims searched:** all     **Date of search:** 26 April 2001

## Patents Act 1977
## Search Report under Section 17

**Databases searched:**

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

    UK Cl (Ed.S): H4R (RPBE,RPV,RPVA,RPVS)

    Int Cl (Ed.7): G10L (19/00,19/02,19/04,19/06,19/08,19/10,19/12,19/14)

Other:     Online: EPODOC,WPI,JAPIO,INSPEC,IEL

**Documents considered to be relevant:**

| Category | Identity of document and relevant passage | | Relevant to claims |
|---|---|---|---|
| A | GB 2300548 A | (Motorola) | |
| X | WO 99/65017 A | (Matsushita) see e.g. page 14, lines 12-17; page 20, lines 4-26; page 26, lines 23 *et seq* | 22,27,28, 32 |
| X | WO 97/27578 A | (Motorola) see e.g. fig 4 and page 14, line 3 to page 17, line 38 | 22,24,25, 27,28,30, 32 |
| X | WO 95/10760 A | (Comsat) see e.g. page 11, line 5 to page 14, line 7 | 22,27,28, 32 |
| X | EP 0411655 A | (Fujitsu) see figs 3 and 4 | 22,27,28, 32 |
| A | J A Naylor, "A neural network algorithm for enhancing delta modulation/LPC tandem connections", International Conference on Acoustics, Speech and Signal Processing, 1990 (ICASSP-90), vol 1, pages 221-224 | | |

| | | | |
|---|---|---|---|
| X | Document indicating lack of novelty or inventive step | A | Document indicating technological background and/or state of the art. |
| Y | Document indicating lack of inventive step if combined with one or more other documents of same category. | P | Document published on or after the declared priority date but before the filing date of this invention. |
| & | Member of the same patent family | E | Patent document published on or after, but with priority date earlier than, the filing date of this application. |