

Internal distribution code:

- (A) [] Publication in OJ
(B) [] To Chairmen and Members
(C) [X] To Chairmen
(D) [] No distribution

D E C I S I O N
of 6 September 2005

Case Number: T 0770/03 - 3.4.01

Application Number: 96202584.7

Publication Number: 0751496

IPC: G10L 19/06

Language of the proceedings: EN

Title of invention:

Speech coding method and apparatus for the same

Patentee:

Nippon Telegraph and Telephone Corporation

Opponent:

Telefonaktiebolaget LM Ericsson (publ)

Headword:

-

Relevant legal provisions:

EPC Art. 52(1), 56, 100(a)

Keyword:

"Inventive step (yes) "

Decisions cited:

-

Catchword:

-



Case Number: T 0770/03 - 3.4.01

D E C I S I O N
of the Technical Board of Appeal 3.4.01
of 6 September 2005

Appellant: Telefonaktiebolaget LM Ericsson (publ)
(Opponent) Telefonvägen 30
S-126 25 Stockholm (SE)

Representative: Kribbler, Klaus-Dieter
Ericsson GmbH
Ericsson Allee 1
D-52134 Herzogenrath (DE)

Respondent: Nippon Telegraph and Telephone Corporation
(Proprietor of the patent) 19-2 Nishi-Shinjuku 3-chome Shinjuku-ku
Tokyo 163-19 (JP)

Representative: Audier, Philippe André
Brevalex
3, rue du Docteur Lancereaux
F-75008 Paris (FR)

Decision under appeal: Decision of the Opposition Division of the
European Patent Office posted 26 May 2003
rejecting the opposition filed against European
patent No. 0751496 pursuant to Article 102(2)
EPC.

Composition of the Board:

Chairman: B. Schachenmann
Members: R. Bekkering
M. Rognoni

Summary of Facts and Submissions

- I. The appellant (opponent) lodged an appeal against the decision of the opposition division, dispatched on 26 May 2003, rejecting the opposition against European patent No. 0 751 496. The notice of appeal was received on 21 July 2003, the appeal fee being paid on the same day, and the statement setting out the grounds of appeal was received on 30 September 2003.
- II. Opposition had been filed against the patent as a whole, based on Article 100(a) EPC on the grounds of lack of novelty and inventive step (Articles 52(1), 54, 56 EPC).
- III. In the appeal proceedings reference was *inter alia* made to the following documents:
- E2: A. Kataoka et al, "8kbit/s low-delay speech coder based on CELP coding", Institute of Electronics, Information and Communication Engineers (IEICE) Technical Report, vol. 91, No. 471, 19 February 1992, Tokyo, Japan, SP91-119 (in Japanese) and corresponding English translation.
- E3: J. Chen et al, "Gain-Adaptive Vector Quantization with Application to Speech Coding", IEEE Transactions on Communications, vol. COM-35, No. 9, September 1987, pages 918 to 930
- E4: G.S. Kang et al, "Application of line-spectrum pairs to low-bit-rate speech encoders" Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Tampa,

Florida (USA), 26 to 29 March 1985, pages 244
to 247

E5: US-A-4 975 956

E5a: P. Kroon et al, "A Class of Analysis-by-Synthesis
Predictive Coders for High Quality Speech Coding
at Rates Between 4.8 and 16 kbits/s", IEEE Journal
on Selected Areas in Communications, vol. 6, No. 2,
February 1988, pages 353 to 363

E5b: R. Steele, "Mobile Radio Communications", IEEE
Press, New York (USA), 1992, pages 237 to 245

E10: N. Sugamura et al, "Speech analysis and synthesis
methods developed at ECL in NTT - from LPC to LSP",
Speech Communications 5, Elsevier Science
Publishers, 1986, pages 199 to 215

- IV. Oral proceedings, requested by the appellant as an
auxiliary measure, were held on 6 September 2005.
- V. The appellant requested that the decision under appeal
be set aside and the patent revoked.
- VI. The respondent requested that the appeal be dismissed.
- VII. Independent claims 1 and 6 of the patent as granted
read as follows:

*"1. A speech coding method comprising at least four
flows of steps;*

a first flow comprising:

a first step for forming a vector from speech signals comprising a plurality of samples as a unit of frame operation, and storing said vector as a speech input vector;

a second step for sequentially checking, one frame at a time, the amplitude of each speech input vector, and compressing said amplitude when the absolute value of said amplitude exceeds a predetermined value;

a third step for conducting linear prediction analysis and calculating an LPC coefficient for each speech input vector outputted by means of said second step;

a fourth step for converting each LPC coefficient calculated in said third step into an LSP parameter;

a fifth step for quantizing said LSP parameter by means of using a vector quantizing process;

a sixth step for converting said quantized LSP parameter into a quantized LPC coefficient;

a seventh step for synthesizing a synthetic speech vector based on a driving vector supplied from the exterior, and said quantized LPC coefficient;

an eighth step for calculating distortion data by means of subtracting said synthetic speech vector outputted by means of said seventh step from said speech input vector outputted by means of said second step;

a ninth step for weighting said distortion data calculated by means of said eighth step; and

a tenth step for calculating the distortion power of said distortion data with regard to each distortion data weighted by means of said ninth step;

a second flow comprising:

an eleventh step for selecting one pitch period vector from among a plurality of pitch period vectors;

a twelfth step for selecting one noise waveform vector from among a plurality of noise waveform vectors;

a thirteenth step for calculating a prediction gain for each noise waveform vector selected by means of said twelfth step;

a fourteenth step for multiplying said prediction gain calculated by means of said thirteenth step by said noise waveform vector selected by means of said twelfth step;

a fifteenth step for respectively multiplying a gain selected from among a plurality of gains by said pitch period vector selected by means of said eleventh step, and an output vector of said fourteenth step; and

a sixteenth step for adding two multiplication results obtained by means of said fifteenth step, and supplying said addition result to said seventh step as said driving vector;

a third flow for selecting a value which will minimize said distortion power calculated by means of said tenth step when selecting a pitch period vector according to said eleventh step, selecting a noise waveform vector according to said twelfth step, and selecting a gain according to said fifteenth step; and

a fourth flow for encoding processed information obtained by means of said structural means into bit series, adding as necessary error correctional coding, and then transmitting said encoded bit series;

wherein said thirteenth step comprises calculating said prediction gain by means of conducting linear

prediction analysis based on the power of an output vector of said fourteenth step multiplied by a gain during processing of said fifteenth step for the current frame, and the power of an output vector of said fourteenth step multiplied by a gain during the processing of said fifteenth step for a past frame."

"6. A speech coding apparatus comprising:

a buffer (22) for forming a vector from speech signals comprising a plurality of samples as a unit of frame operation, and storing said vector as a speech input vector;

an amplitude limiting means (23) for sequentially checking, one frame at a time, the amplitude of each speech input vector stored in said buffer (22), and compressing said amplitude when the absolute value of said amplitude exceeds a predetermined value;

an LPC analyzing means (24) for conducting linear prediction analysis and calculating an LPC coefficient for each speech input vector outputted by means of said amplitude limiting means (23);

an LPC parameter converting means for converting each LPC coefficient calculated by means of said LPC analyzing means (24) into an LSP parameter;

a vector quantizing means for quantizing said LSP parameter by means of using a vector quantizing process;

an LPC coefficient converting means for converting said quantized LSP parameter into a quantized LPC coefficient;

a synthesizing means (26) for synthesizing a synthetic speech vector based on a driving vector supplied from the exterior, and said quantized LPC coefficient;

a distortion data calculating means (33) for calculating distortion data by means of subtracting said synthetic speech vector outputted by means of said synthesizing means (26) from said speech input vector outputted by means of said amplitude limiting means (23);

a perceptual weighting means (34) for weighting said distortion data obtained by means of said distortion data calculating means (33);

a distortion power calculating means (35) for calculating the distortion power of said distortion data with regard to each distortion data weighted by means of said perceptual weighting means (34);

a pitch period vector searching means (27) for storing a plurality of pitch period vectors, and for selecting one pitch period vector from among said plurality of stored pitch period vectors;

a noise waveform vector searching means (28) for storing a plurality of noise waveform vectors, and for selecting one noise waveform vector from among said plurality of stored noise waveform vectors;

a gain adapting means (29) for calculating a prediction gain for each noise waveform vector selected by means of said noise waveform vector searching means (28);

a prediction gain multiplying means (30) for multiplying said prediction gain calculated by means of said gain adapting means (29) by said noise waveform vector selected by means of said noise waveform vector searching means (28);

a gain multiplying means (31) for storing a plurality of gains, and for respectively multiplying a gain selected from among said plurality of stored gains by said pitch period vector selected by means of said

pitch period vector searching means (27) and an output vector of said prediction gain multiplying means (30);

an adding means (32) for adding two multiplication results obtained by means of said gain multiplying means (31), and supplying said addition result to said synthesizing means (26) as said driving vector;

a control means for selecting a value which will minimize said distortion power calculated by means of said distortion power calculating means (35) when selecting a pitch period vector by means of said pitch period vector searching means (27), selecting a noise waveform vector by means of said noise waveform vector searching means (28), and selecting a gain by means of said gain multiplying means (31); and

a code outputting means (36) for encoding processed information obtained by means of said structural means into bit series, adding as necessary error correctional coding, and then transmitting said encoded bit series;

wherein said gain adapting means (29) calculates said prediction gain by means of conducting linear prediction analysis based on the power of an output vector of a prediction gain multiplying means (30) multiplied by a gain during the processing of gain multiplying means (31) for the current frame, and the power of an output vector of a prediction gain multiplying means (30) multiplied by a gain during the processing of gain multiplying means (31) for a past frame."

VIII. The appellant argued in substance that the subject-matter of claim 1, and of independent claim 6, was rendered obvious by the teaching of document E2 in combination with the teachings of documents E3 and E10 (or any one of documents E4, E5, E5a, E5b) and the

general knowledge of the skilled person. In particular, the subject-matter of claim 1 differed from the method known from document E2 in that it involved amplitude compression of the input signal, generally known, the use of LSP quantization, already suggested in document E10 (and E4, E5, E5a, E5b), and a different gain setting structure, as such already suggested in document E3. Since these features were directed at solving distinct problems, they could be treated separately for the purpose of assessing inventive step.

- IX. The respondent submitted that synergy was present between the features of claim 1, so that a separation as suggested by the appellant was not permissible. Furthermore, an amplitude compression of the speech input vector as claimed was nowhere disclosed. The filter disclosed in document E2 differed since the LPC predictor acted on the output of the synthesis filter, rather than the speech input vector as claimed, and the use of LSP quantization was not suggested. Moreover, document E3 failed to disclose a multiplication with a gain selected from among a plurality of gains. Accordingly, the subject-matter of claim 1 as granted, and of independent claim 6 for the same reasons, involved an inventive step.

Reasons for the Decision

1. The appeal complies with the requirements of Articles 106 to 108 and Rule 64 EPC and is therefore admissible.

2. *Inventive step*

2.1 Having regard to the subject-matter of claim 1, the closest prior art consists of a conventional code excited linear prediction coding method such as the LD-CELP method disclosed in document E2 (see figure 1) and the prior art CELP, VSELP and LD-CELP methods acknowledged in the patent specification (see page 2, line 15 to page 3, line 46; figures 15, 16 and 17).

2.2 In particular, the method known from document E2 involves the following steps:

- forming a vector from speech signals comprising a plurality of samples as a unit of frame operation, and storing said vector as a speech input vector (see translation of document E2, page 5, lines 14 to 15, according to which an input speech vector ("input voice signal vector X") is used in the calculation of the error d) (cf. "first step" of claim 1);
- synthesising a speech vector using an excitation vector from a codebook and the LPC coefficients for the filter (cf. "seventh step");
- calculating distortion data by subtracting the synthesised vector from the input speech vector (see document E2, figure 1 and translation, page 5, lines 9 to 12) (cf. "eighth step");
- perceptually weighting the distortion (see document E2, figure 1, "perceptual weighting") (cf. "ninth step");

- calculating the distortion power (see document E2, figure 1, "minimum error") (cf. "tenth step");
- selecting a pitch period vector from among a plurality of pitch period vectors (see document E2, figure 1, "pitch candidates" and translation, sentence bridging pages 4 and 5) (cf. "eleventh step");
- selecting a noise waveform vector from among a plurality of noise waveform vectors (see document E2, figure 1, "codebook" and translation, sentence bridging pages 4 and 5) (cf. "twelfth step");
- calculating a prediction gain for each noise waveform vector selected from the codebook (see document E2, figure 1, "gain adapter") (cf. "thirteenth step");
- multiplying said prediction gain by said noise waveform vector selected from the codebook (see document E2, figure 1, "gain") (cf. "fourteenth step");
- multiplying a pitch gain by said pitch period vector from the codebook (see document E2, figure 1, "pitch gain") (cf. part of the "fifteenth step");
- adding the resulting pitch period and noise excitation vectors and supplying the result to the synthesis filter (cf. "sixteenth step"); and

- encoding the processed information into bit series, adding as necessary error correctional coding, and then transmitting said encoded bit series (see document E2, translation, page 6, lines 2 to 5) (cf. "fourth flow" of claim 1).

2.3 Not disclosed in document E2 are in substance:

- the second step of claim 1 in suit,
- the third to sixth step,
- part of the fifteenth step, and the last feature of the claim, further specifying the thirteenth step.

These differences of the claimed method with respect to the method known from document E2, in particular, the use of LSP quantization (see the third to sixth step), the use of gain codebooks (see the fifteenth step and last feature of claim 1), as well as the overdrive prevention on the input signal (see second step), all contribute to an improved speech quality, and coding and transmission efficiency.

Accordingly, the objective problem to be solved in the patent in suit may be seen as improving the speech coding in terms of efficiency and speech quality. This problem as such has to be considered obvious for a skilled person working on the technical field of speech coding at issue.

2.4 The claimed solution consists of a number of measures as listed above.

Regarding the above second step, as held by the appellant, indeed the compression of the amplitude of an input signal in case it exceeds a given threshold, thereby preventing overdriving the input stage of a circuit, is a conventional measure and must be considered obvious in the present context to a skilled person in the technical field at issue.

2.5 As far as the above third to sixth steps are concerned, it is noted that document E2 indeed uses backward prediction for determining the LPC coefficients used in the synthesis filter (see figure 1), whereas according to claim 1 in suit forward prediction is used acting on the speech input vector (cf. "third step"). These prediction schemes, however, are well-known alternatives (see also the prior art discussed in the patent in suit, figures 15, 16 and 17, respectively), so that the use of forward prediction must be considered an obvious design option. Furthermore, both the use of LSP parameters and the fact that LSP parameters provide for a more efficient quantization is known from the prior art, see document E10 (see "Introduction" and page 207, left-hand column, last paragraph to right-hand column, second paragraph), as well as document E4 (see abstract and page 245, penultimate paragraph) also indicating the advantage of converting the LSP parameters in prediction coefficients, document E5a (see page 359, right-hand column, first paragraph) and E5b (see chapter 3.5.1.1.4.2). Similarly, document E5 (see figures 1A, 1B and corresponding description) discloses the quantization of LPC coefficients using LSP parameters with, furthermore, the conversion of the LSP parameters

back into LPC coefficients for use in the synthesis module of the decoder. Furthermore, the prior art CELP and VSELP methods disclosed in the patent specification (see figures 15 and 16 and corresponding description) already include the steps of converting the LPC coefficients to LSP parameters, quantizing the LSP parameters by means of vector quantization and converting the quantized LSP parameters into quantized LPC coefficients.

- 2.6 Having regard to the computation of the gain as defined in part of the fifteenth step and the last feature of the claim, the appellant argued that the claimed method steps were obvious in the light of document E3, in particular the teaching concerning figure 3.

Document E3 (see abstract, "II. Basic structure" and figure 3) concerns gain-adaptive vector quantization in the context of speech coding. Figure 3 shows the basic structure of backward gain-adaptive vector quantization. An input speech vector is normalised by means of division by the predicted gain. The normalised vector is then encoded in a VQ encoder and the index of the best matching code vector from the codebook is sent to the receiver. Furthermore, for the purpose of the gain prediction, a quantized input vector is reconstructed by multiplying the quantized code vector by the predicted gain and fed to the gain predictor in which the predicted gain is computed based on the previous quantized input vectors. The predicted gain need not be sent to the receiver as it can be generated by the duplicated gain predictor in the receiver.

However, contrary to what has been argued by the appellant, a substitution of the gain setting steps in document E2 by the above teaching of document E3 would not lead to the gain setting steps as claimed. It already appears doubtful that a skilled person would consider this teaching of document E3 as being of any relevance to the gain setting steps of document E2. In E2 a vector from a noise codebook is multiplied by a suitable gain in order to provide a fitting noise component of the excitation vector. Document E3, on the other hand, is concerned with the vector quantization of an input speech vector. Moreover, document E3 fails to provide a multiplication step involving a gain selected from a plurality of gains (cf. the fifteenth step of claim 1) such as from a codebook, the only multiplication with a gain being the one for reconstructing the quantized input vector as discussed above.

In the claimed method, by calculating the prediction gain by linear prediction analysis based on the power of the noise waveform vector multiplied by the prediction gain and a gain selected from e.g. a codebook containing a plurality of gains of current and past frames, the gain selected from e.g. a codebook will generally be close to unity. Therefore, the variation in this gain will be small, allowing for e.g. a small codebook size and thus for a small number of encoding bits to be transmitted. As a result, improved speech quality can be provided with a high coding and transmission efficiency.

Accordingly, document E3 cannot be held to render the above claimed gain setting steps obvious.

- 2.7 In view of the above, the submission of the respondent that indeed some synergy is present between the various features distinguishing the claimed subject-matter from the method known from document E2 listed above, is of no consequence and, therefore, need not further be addressed.
- 2.8 For the reasons given above, in the board's opinion the cited prior art cannot be held to render the claimed method obvious. Therefore, an inventive step has to be recognised for the subject-matter of claim 1.
- 2.9 The same applies to independent claim 6, which is directed to a corresponding speech coding apparatus. The subject-matter of claim 6, thus, also involves an inventive step.
- 2.10 The remaining claims 2 to 5 and 7 to 10 are dependent on either claim 1 or 6 and provide further preferred features of the speech coding method and apparatus, respectively. The subject-matter of these claims, therefore, involves an inventive step as well.
3. Consequently, the grounds of opposition invoked by the appellant do not prejudice the maintenance of the patent as granted.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar:

The Chairman:

R. Schumacher

B. Schachenmann