Datasheet for the decision
of 27 June 2019

Case Number: T 1690/13 - 3.5.04

Application Number: 01931769.2

Publication Number: 1282982

IPC: H04N7/36

Language of the proceedings: EN

Title of invention:
Method for encoding and decoding video information, a motion compensated video encoder and a corresponding decoder

Applicant:
Nokia Technologies Oy

Headword:

Relevant legal provisions:
EPC 1973 Art. 56, 84

Keyword:
Claims - clarity after amendment (yes)
Inventive step - (yes)

Decisions cited:
Catchword:
Decision of Technical Board of Appeal 3.5.04
of 27 June 2019

Appellant: Nokia Technologies Oy
Karaportti 3
02610 Espoo (FI)

Representative: Berggren Oy, Helsinki & Oulu
P.O. Box 16
Eteläinen Rautatiekatu 10A
00101 Helsinki (FI)

Decision under appeal: Decision of the Examining Division of the European Patent Office posted on 19 February 2013 refusing European patent application No. 01931769.2 pursuant to Article 97(2) EPC.

Composition of the Board:
Chairman C. Kunzelmann
Members: R. Gerdes
B. Müller
Summary of Facts and Submissions

I. The appeal is against the decision to refuse European patent application No. 01 931 769.2, published as international application WO 01/86962 A1.

II. The examining division refused the patent application on the grounds that claim 1 of the main request and each of the first to fourth auxiliary requests did not meet the requirements of Article 84 EPC and that the subject-matter of these claims lacked inventive step (Article 56 EPC) in view of document:


III. The applicant filed notice of appeal against this decision and, with its statement of grounds of appeal, submitted claims of a new main request and first to third auxiliary requests.

IV. The board issued a communication pursuant to Rule 100(2) EPC indicating that it disagreed with the objections in the decision under appeal. However, the board raised several objections as to lack of clarity and added subject-matter. It noted that as far as the claims could be understood, their subject-matter seemed to involve an inventive step.

V. The appellant replied by letter dated 12 April 2019 to the board's communication. It submitted amended claims according to a main request and a first auxiliary
request and withdrew its second and third auxiliary requests.

VI. In response to telephone conversations held on 29 May 2019 and 14 June 2019, the appellant submitted amended claims of the main request and amended description pages.

VII. The appellant's final requests are that the decision under appeal be set aside and that the case be remitted to the examining division with instructions to grant a patent on the basis of the claims of the main request filed on 17 June 2019 (labelled "MAIN REQUEST amended on 14 June 2019") or of the first auxiliary request filed with the letter dated 12 April 2019.

VIII. The independent claims of the main request read as follows:

"1. A method for encoding video information, comprising:
- defining (401) a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that specify possible ways in which a macroblock can be segmented into blocks,
- defining a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining (402) for each available macroblock segmentation (70, 71, 72, 83, 84, 85) at least one available prediction method as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a respective available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock
one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and
an associated method for deriving the prediction motion coefficient for said respective block of the macroblock from motion coefficients of said one or more associated prediction blocks,
- segmenting (403) a frame of video information into macroblocks,
- selecting (411) for a macroblock one of the available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), wherein the selection of the macroblock-segmentation - prediction-method pair is based on minimizing a cost function and
- segmenting the macroblock into blocks and deriving prediction motion coefficients for the blocks within said macroblock using the selected macroblock-segmentation - prediction-method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85).

... 26. A method for decoding encoded video information, characterized in that it comprises:
- defining (501) a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that specify possible ways in which a macroblock can be segmented into blocks,
- defining a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining (502) for each available macroblock segmentation (70, 71, 72, 83, 84, 85) at least one available prediction method as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a respective
available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and an associated method for deriving the prediction motion coefficient for said respective block of the macroblock from motion coefficients of said one or more associated prediction blocks, - receiving (503) information indicating at least a macroblock segmentation (70, 71, 72, 83, 84, 85) selected for a macroblock in said encoded video information, and - determining (505) a prediction method relating to the macroblock segmentation (70, 71, 72, 83, 84, 85) of said macroblock with reference to the defined macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) and deriving prediction motion coefficients for blocks within said macroblock using the determined prediction method.

... 32. An encoder for performing motion compensated encoding of video information, comprising: - means for receiving a frame of video information, - means (111) for segmenting the frame of video information into macroblocks, - means (112) for specifying a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that define possible ways in which a macroblock can be segmented into blocks, - means (122) for specifying a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining at least
one available prediction method for each available macroblock segmentation (70, 71, 72, 83, 84, 85) as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a respective available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and

- means (125) for selecting one macroblock-segmentation - prediction method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) among the available macroblock-segmentation - prediction method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), wherein the selection of the macroblock-segmentation - prediction-method pair is based on minimizing a cost function,

- means (123) for deriving macroblock-segmentation-specific prediction motion coefficients for blocks within said macroblock using the prediction method specified by the selected macroblock-segmentation - prediction-method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), and

... 34. A decoder for performing motion compensated decoding of encoded video information, comprising:
- input means for receiving encoded video information,
- means for defining a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that specify possible ways in which a macroblock can be segmented into blocks,
- means for specifying a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining at least one available prediction method for each available macroblock segmentation (70, 71, 72, 83, 84, 85) as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a respective available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and
- an associated method for deriving the prediction motion coefficient for said respective block of the macroblock from motion coefficients of said one or more associated prediction blocks,
- input means for receiving information indicating at least the macroblock segmentation (70, 71, 72, 83, 84, 85) selected for a macroblock,
- means (132) for determining the prediction-method relating to the segmentation of the macroblock with reference to the defined macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), and
- means (131) for deriving prediction motion coefficients for blocks within said macroblock using the determined prediction method.

...
39. A computer program element for performing motion compensated encoding of video information, comprising:
- means for receiving a frame of video information,
- means (111) for segmenting the frame of video information into macroblocks,
- means (112) for specifying a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that define possible ways in which a macroblock can be segmented into blocks,
- means (122) for specifying a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining at least one available prediction method for each available macroblock segmentation (70, 71, 72, 83, 84, 85) as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a given available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and
- an associated method for deriving the prediction motion coefficient for said respective block of the macroblock from the motion coefficients of said one or more associated prediction blocks,
- means (125) for selecting one macroblock-segmentation - prediction method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) among the available macroblock-segmentation - prediction method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) wherein the selection of the macroblock-segmentation - prediction-method pair is based on minimizing a cost function,
- means for segmenting a macroblock using the macroblock segmentation (70, 71, 72, 83, 84, 85) specified by the selected macroblock-segmentation - prediction-method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), and - means (125) for deriving macroblock-segmentation-specific prediction motion coefficients for blocks within said macroblock using the prediction method specified by the selected macroblock-segmentation - prediction-method pair (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85).

...  
41. A computer program element for performing motion compensated decoding of encoded video information, comprising:
- input means for receiving encoded video information,
- means for defining a certain number of available macroblock segmentations (70, 71, 72, 83, 84, 85) that specify possible ways in which a macroblock can be segmented into blocks,
- means for specifying a certain finite number of available macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85) by defining at least one available prediction method for each available macroblock segmentation (70, 71, 72, 83, 84, 85) as a predetermined subset of possible prediction methods (60A, 60B, 60C, 60D) in dependence of a given available macroblock segmentation (70, 71, 72, 83, 84, 85), wherein each available prediction method (60A, 60B, 60C, 60D) defines for each block of a macroblock
  one or more associated prediction blocks used in producing a prediction motion coefficient for a respective block of the macroblock, and
  an associated method for deriving the prediction motion coefficient for said respective block of the
macroblock from the motion coefficients of said one or more prediction blocks,
- input means for receiving information indicating at least the macroblock segmentation (70, 71, 72, 83, 84, 85) selected for a macroblock,
- means (132) for determining the prediction-method relating to the segmentation of the macroblock with reference to the defined macroblock-segmentation - prediction-method pairs (70A, 70B, 70C, 70D, 71A, 71B, 71C, 71D, 71E, 72A, 72B, 72C, 72D, 83, 84, 85), and
- means (131) for deriving prediction motion coefficients for blocks within said macroblock using the determined prediction method."

Further claims 2 to 25, 27 to 31, 33 and 35 to 38, 40 and 42 to 49 are dependent on claims 1, 26, 32, 34, 39 and 41, respectively. The wording of the claims of the auxiliary request is of no relevance to this decision.

IX. In the decision under appeal, the examining division held that then valid claim 1 of the main request was not clear (see Reasons, points 3.1.1.2 and 3.3.1). In particular, the examining division argued that:

(a) the term "subset" did not limit the ways in which a macroblock segmentation might occur, and

(b) it was unclear whether any specific criteria needed to be adhered to in order to obtain a subset.

(c) Claim 1 also did not solve the problem posed by the applicant. Such inconsistency between the claims and description cast doubt on the matter for which protection was sought.
Based on these objections, the examining division considered the "defining" step in claim 1 as unclear and technically meaningless. The examining division acknowledged that D3 did not disclose pairs of macroblock segmentations and prediction methods. However, since the claim allowed for the interpretation that a single prediction method was assigned to each macroblock segmentation, the pairing was - in the absence of any details regarding the limitations thereto - obvious (see decision under appeal, Reasons, points 3.2 and 3.3.2).

**Reasons for the Decision**

1. The appeal is admissible.

**The invention**

2. The present invention relates to motion compensated prediction in video coding.

2.1 Motion compensated prediction utilises the fact that, in a typical video sequence, image intensity values in a particular frame segment can be predicted using image intensity values of a segment in another already coded and transmitted frame, given the motion vector between these two segments. Hence, video data can be compressed by coding motion vector coefficients representing motion vectors instead of image intensity values (see present application as published, page 1, lines 4 and 5 and page 2, lines 16 to 30).

In order to further minimise the amount of information needed in sending the video frames to a decoder, the motion vector coefficients can themselves be predicted
from the coefficients of neighbouring segments, which have been previously coded. When this kind of motion prediction is used, a motion vector is expressed as a sum of a predicted motion vector and a refinement motion vector. Only the refinement motion vector coefficients need to be coded and transmitted to the decoder because the predicted motion vector can be predicted from the coefficients of the neighbouring segments in the decoder in the same way as in the encoder. This approach typically results in transmission bit rate savings (see page 5, lines 4 to 17).

Motion vectors are not computed for every pixel. Instead, the image is divided into macroblocks, which are further segmented in blocks. Motion vectors are then determined for each block. There may be many possible ways to segment a macroblock and many different methods of how to predict the motion coefficients of a specific block from those of neighbouring blocks. As a consequence, the calculation of an optimal segmentation together with an optimal prediction method is computationally complex (see page 6, lines 23 to 30 and page 7, lines 5 to 34).

2.2 The inventive idea of the application is to form pairs of macroblock segmentations and prediction methods, such that only a limited number of prediction methods are available for each macroblock segmentation. By restricting the number of possible prediction methods per macroblock segmentation to a subset of all possible pairs, the complexity of the encoding process is reduced. The search for an optimal combination is restricted to these allowed pairs, and the pair minimising a cost function is chosen for determining
the motion coefficients (see page 9, line 32 to page 10, line 23).

Main request, amendments (Article 123(2) EPC)

3. The claims of the present main request are based on the claims of the first auxiliary request underlying the decision under appeal. Compared with claim 1 of the latter request, present claim 1 has been reformulated to specify more clearly the step of defining pairs of macroblock segmentations and prediction methods. The reformulation of claim 1 is based on page 9, line 31 to page 10, line 14 of the application as published. The further independent claims, namely claims 26, 32, 34, 39 and 41, have been amended to correspond to claim 1. In addition, dependent claims 2 and 3 have been amended based on claims 2 and 3 of the application as published. Claims 28, 37 and 40 now refer to "at least two macroblock segmentation – prediction methods", whose basis is found in claim 28 as published.

Hence, the board finds that the claims of the appellant's main request do not contain subject-matter extending beyond the content of the application as filed and that they thus comply with Article 123(2) EPC.

Main request, clarity (Article 84 EPC 1973)

4. In the decision under appeal, the examining division argued that then valid claim 1 lacked clarity (see point IX above).

4.1 In particular, the examining division objected that
(a) the term "subset" did not limit the ways in which a macroblock segmentation might occur, and

(b) it was unclear whether any specific criteria needed to be adhered to in order to obtain a subset.

(c) Claim 1 also did not solve the problem posed by the applicants. Such inconsistency between the claims and description cast doubt on the matter for which protection was sought.

4.2 The interpretation of claim 1 by the examining division in (a) is tantamount to stating that the "defining" step in claim 1 is technically meaningless. This applies when subsets contain all possible pairs of macroblock segmentations and prediction methods. Furthermore, this argument implies that the step of "defining a certain finite number of ... pairs" is only a mental step, or at least, it is unclear what technical limitations are conveyed by the definition of pairs.

4.3 The board is of the opinion that this interpretation of claim 1 can be ruled out on a proper reading of the claim. It would not be technically meaningful to define subsets that contain all pairs.

4.4 More importantly, the board considers the step of "defining a certain finite number of ... pairs" to imply a technical limitation, which goes beyond a simple (mental) assignment of prediction methods to macroblock segmentations. The "available" pairs are used in the further method step of "selecting (411) for a macroblock one of the available macroblock-segmentation - prediction-method pairs". In order to be selectable, the available pairs have to be represented
using technical means in a corresponding encoder or encoding program. The "predetermined subset" of prediction methods in dependence on available macroblock segmentations could, for example, be represented in the form of a table mapping macroblock-segmentations to corresponding prediction methods in the internal storage of the encoder. Other representations are possible, but the predetermined subsets have to be stored in an encoder carrying out the method of claim 1 in order to retrieve the possible prediction methods for a certain macroblock segmentation.

4.5 Hence, the board considers objection (a) as not justified.

4.6 Regarding objection (b), the board is convinced that the skilled person in the technical field of video coding would have realised that possible pairs could be identified, for example, by carrying out statistical analyses to figure out which pairs are mostly used in an unrestricted encoder having no limitation on the pairs of macroblock segmentations and prediction methods. Hence, the skilled person would have been in a position to obtain meaningful subsets.

4.7 As a consequence of the above and contrary to objection (c), the subject-matter of claim 1 also solves the problem posed by the applicant.

4.8 Hence, the board is not convinced by the objections in the decision under appeal with respect to the then valid set of claims. The board also sees no other lack of clarity in the claims.
Main request, inventive step

5. It is common ground that D3 may be considered as the closest prior art with respect to the claimed subject-matter.

5.1 D3 relates to a low bit-rate video codec that employs multiple-level segmentation and affine motion compensation. The video codec of D3 uses variable block size motion compensation (see abstract and chapter II.C). D3 is based on the realisation that for objects undergoing coherent movement and being divided into a number of blocks, it is likely that neighbouring blocks have the same motion vectors. In order to reduce the data for image data transmission, it is sufficient to indicate which of the neighbouring blocks moves with the same motion as the current block or minimises an error criterion. If a suitable neighbouring block is detected, no motion vectors need to be transmitted. Instead, the location of that neighbouring block is coded. Furthermore, in a situation where no good match between the velocity of the neighbouring blocks and the current block can be detected, a reference velocity is determined using the median of the velocity of the first block to the left, the velocity of the first block on the top, and the velocity of the first block on the top-right corner. The velocity difference between the current block and the reference velocity is then encoded using variable length coding (see chapter III.B).

5.2 D3 does not disclose defining pairs of macroblock segmentations and prediction methods, as specified in claim 1. Instead, it compares the motion vectors of the current block with those of neighbouring blocks and - if need be - with the reference motion vector being
computed as the median of the motion vectors of neighbouring blocks. Consequently, D3 also does not disclose the selecting step of claim 1, which selects one of the available macroblock-segmentation prediction-method pairs.

5.3 The distinguishing features have the technical effect of reducing the complexity of the encoding process and increasing flexibility by allowing preselection of optimal pairs between macroblock segmentations and prediction methods.

5.4 Hence, these features solve the technical problem of how to reduce the complexity and increase the flexibility of the encoding process in motion compensated prediction.

5.5 None of the cited documents discloses a relationship being formed between macroblock segmentations and prediction methods, let alone organising them in pairs. Also, there is no indication on file that a person skilled in the art would form such a relationship in an obvious manner on the basis of common general knowledge. The realisation that a prediction method may be selected in dependence on the macroblock segmentation allows the determination of combinations of pairs of macroblock segmentations and prediction methods that achieve near optimal results in terms of the selected cost function and reduce the computational load on the encoder.

5.6 As a consequence of the above, the subject-matter of claim 1 involves an inventive step in view of the available state of the art.
5.7 The same reasoning applies to independent claim 26, which specifies a method for decoding using steps corresponding to those of the encoding method of claim 1. Similarly, independent claims 32 and 34 relating to an encoder and a decoder, respectively, specify the devices corresponding to the methods of claims 1 and 26. Claims 39 and 41 define computer program elements for encoding and decoding corresponding to the methods of claims 1 and 26.
Order

For these reasons it is decided that:

1. The decision under appeal is set aside.

2. The case is remitted to the examining division with the order to grant a patent in the following version:

Description:
- pages 1 to 7, 13 to 17 and 19 to 23 as published
- pages 8 and 10 to 12 filed by telefax on 24 February 2009
- pages 9, 18 and 24 filed in electronic form on 3 June 2019

Claims:
Nos. 1 to 49 filed in electronic form on 17 June 2019

Drawings:
Sheets 1/8 to 8/8 as published.

The Registrar: The Chairman:

K. Boelicke C. Kunzelmann

Decision electronically authenticated