

(43) International Publication Date  
19 July 2012 (19.07.2012)(51) International Patent Classification:  
**H04N 7/64** (2006.01)      **H04N 7/36** (2006.01)  
**H04N 7/26** (2006.01)(21) International Application Number:  
PCT/EP2012/050390(22) International Filing Date:  
11 January 2012 (11.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
1100462.9      12 January 2011 (12.01.2011)      GB(71) Applicant (for all designated States except US): **CANON KABUSHIKI KAISHA** [JP/JP]; 30-2 Shimomaruko 3-chome, Ohta-ku, Tokyo 146-8501 (JP).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GISQUET, Christophe** [FR/FR]; 18 Avenue de Cucillé, F-35000 Rennes (FR). **LAROCHE, Guillaume** [FR/FR]; 6 Rue François Lanno, F-35700 Rennes (FR).(74) Agent: **HITCHING, Peter Matthew**; 3 The Square, Stockley Park, Uxbridge Middlesex UB11 1ET (GB).

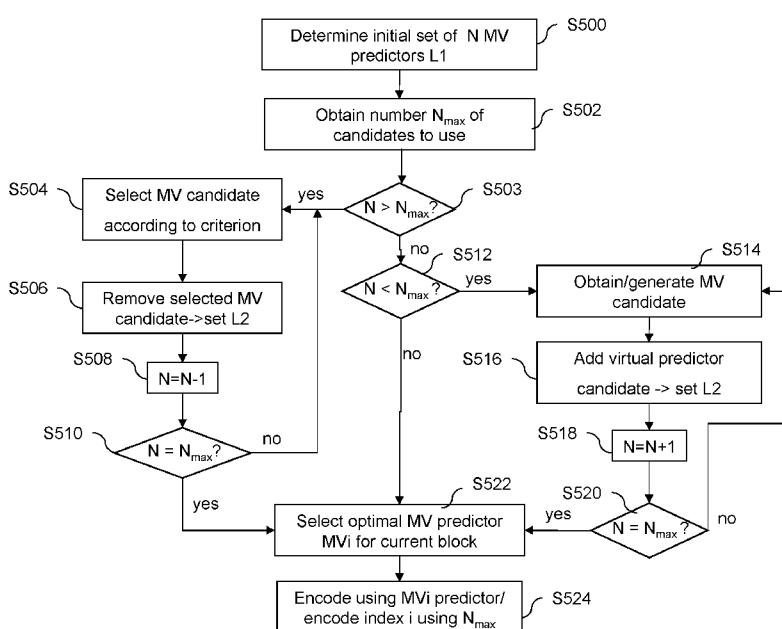
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

— with international search report (Art. 21(3))

(54) Title: VIDEO ENCODING AND DECODING WITH IMPROVED ERROR RESILIENCE



(57) **Abstract:** A sequence of digital images is encoded into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion. A target number (Nmax) of motion information predictors is obtained. Using the target number a set of motion information predictors is generated having controlled diversity. A motion information predictor for the image portion to encode is selected from the generated set of motion information predictors. The target number is signaled in the bitstream, together with information relating to the selected motion information predictor.

Figure 5

**5    Video encoding and decoding with improved error resilience****10            Field of the invention**

The invention relates to a method and device for encoding a sequence of digital images and a method and device for decoding a corresponding bitstream.

15            The invention belongs to the field of digital signal processing, and in particular to the field of video compression using motion compensation to reduce spatial and temporal redundancies in video streams.

**20            Description of the prior-art**

Many video compression formats, for example H.263, H.264, MPEG-1, MPEG-2, MPEG-4, SVC, use block-based discrete cosine transform (DCT) and motion compensation to remove spatial and 25 temporal redundancies. They can be referred to as predictive video formats. Each frame or image of the video signal is divided into slices which are encoded and can be decoded independently. A slice is typically a rectangular portion of the frame, or more generally, a portion of a frame or an entire frame. Further, each slice is divided into 30 macroblocks (MBs), and each macroblock is further divided into blocks, typically blocks of 8x8 pixels. The encoded frames are of two types: temporal predicted frames (either predicted from one reference frame called P-frames or predicted from two reference frames called B-frames) and non temporal predicted frames (called Intra frames or I- 35 frames).

Temporal prediction consists in finding in a reference frame, either a previous or a future frame of the video sequence, an image portion or reference area which is the closest to the block to encode. This step is known as motion estimation. Next, the difference between 5 the block to encode and the reference portion is encoded (motion compensation), along with an item of motion information relative to the motion vector which indicates the reference area to use for motion compensation.

In order to further reduce the cost of encoding motion 10 information, it has been proposed to encode a motion vector by difference from a motion vector predictor, typically computed from the motion vectors of the blocks surrounding the block to encode.

In H.264, motion vectors are encoded with respect to a median predictor computed from the motion vectors situated in a causal 15 neighbourhood of the block to encode, for example from the blocks situated above and to the left of the block to encode. Only the difference, also called residual motion vector, between the median predictor and the current block motion vector is encoded.

The encoding using residual motion vectors saves some 20 bitrate, but necessitates that the decoder performs the same computation of the motion vector predictor in order to decode the value of the motion vector of a block to decode.

Recently, further improvements have been proposed, such as using a plurality of possible motion vector predictors. This method, 25 called motion vector competition, consists in determining between several motion vector predictors or candidates which motion vector predictor minimizes the encoding cost, typically a rate-distortion cost, of the residual motion information. The residual motion information comprises the residual motion vector, i.e. the difference between the 30 actual motion vector of the block to encode and the selected motion vector predictor, and an item of information indicating the selected motion vector predictor, such as for example an encoded value of the index of the selected motion vector predictor.

In the High Efficiency Video Coding (HEVC) currently in the course of standardization, it has been proposed to use a plurality of motion vector predictors as schematically illustrated in figure 1: 3 so-called spatial motion vector predictors  $V_1$ ,  $V_2$  and  $V_3$  taken from blocks 5 situated in the neighbourhood of the block to encode, a median motion vector predictor computed based on the components of the three spatial motion vector predictors  $V_1$ ,  $V_2$  and  $V_3$  and a temporal motion vector predictor  $V_0$  which is the motion vector of the co-located block in a previous image of the sequence (e. g. block of image N-1 located at the 10 same spatial position as block 'Being coded' of image N). Currently in HEVC the 3 spatial motion vector predictors are taken from the block situated to the left of the block to encode ( $V_3$ ), the block situated above ( $V_2$ ) and from one of the blocks situated at the respective corners of the block to encode, according to a predetermined rule of availability. This 15 motion vector predictor selection scheme is called Advanced Motion Vector Prediction (AMVP). In the example of figure 1, the vector  $V_1$  of the block situated above left is selected.

Finally, a set of 5 motion vector predictor candidates mixing spatial predictors and temporal predictors is obtained. In order to 20 reduce the overhead of signaling the motion vector predictor in the bitstream, the set of motion vector predictors is reduced by eliminating the duplicated motion vectors, i.e. the motion vectors which have the same value. For example, in the illustration of figure 1,  $V_1$  and  $V_2$  are equal, and  $V_0$  and  $V_3$  are also equal, so only two of them should be kept 25 as motion vector prediction candidates, for example  $V_0$  and  $V_1$ . In this case, only one bit is necessary to indicate the index of the motion vector predictor to the decoder.

A further reduction of the set of motion vector predictors, based on the values of the predictors, is possible. Once the best motion 30 vector predictor is selected and the motion vector residual is computed, it is possible to further eliminate from the prediction set the candidates which would have not been selected, knowing the motion vector residual and the cost optimization criterion of the encoder. A sufficient

reduction of the set of predictors leads to a gain in the signalling overhead, since the indication of the selected motion vector predictor can be encoded using fewer bits. At the limit, the set of candidates can be reduced to 1, for example if all motion vector predictors are equal,

5 and therefore it is not necessary to insert any information relative to the selected motion vector predictor in the bitstream.

To summarize, the encoding of motion vectors by difference with a motion vector predictor, along with the reduction of the number of motion vector predictor candidates leads to a compression gain.

10 However, as explained above, for a given block to encode, the reduction of the number of motion vector predictor candidates is based on the values taken by the motion vector predictors of the set, in particular the values of the motion vectors of the neighbouring blocks and of the motion vector of the co-located block. Also, the decoder  
15 needs to be able to apply the same analysis of the set of possible motion vector predictors as the encoder, in order to deduce the amount of bits used for indicating the selected motion vector predictor and to be able to decode the index of the motion vector predictor and finally to decode the motion vector using the motion vector residual received.  
20 Referring to the example of figure 1, the set of motion vector predictors of the block 'being coded' is reduced by the encoder to  $V_0$  and  $V_1$ , so the index is encoded on one single bit. If the block 'Co-located' of image N-1 is lost during transmission, the decoder cannot obtain the value of  $V_0$ , and therefore cannot find out that  $V_0$  and  $V_3$  are equal. Therefore,  
25 the decoder cannot find how many bits were used for encoding the index of the motion vector predictor for the block 'Being coded', and consequently the decoder cannot correctly parse the data for the slice because it cannot find where the index encoding stops and the encoding of video data starts.

30 Therefore, the fact that the number of bits used for signalling the motion vector predictors depends of the values taken by the motion vector predictors makes the method very vulnerable to transmission errors, when the bitstream is transmitted to a decoder on a lossy

communication network. Indeed, the method requires the knowledge of the values of the motion vector predictors to parse the bitstream correctly at the decoder. In case of packet losses, when some motion vector residual values are lost, it is impossible for the decoder to

5 determine how many bits were used to encode index representing the motion vector predictor, and so it is impossible to parse the bitstream correctly. Such an error may propagate causing the decoder's de-synchronization until a following synchronization image, encoded without prediction, is received by the decoder.

10 It would be desirable to at least be able to parse an encoded bitstream at a decoder even in case of packet losses, so that some re-synchronization or error concealment can be subsequently applied.

15 It was proposed, in the document JCTVC-C166r1, 'TE11: Study on motion vector coding (experiment 3.3a and 3.3c)' by K. Sato, published at the 3<sup>rd</sup> meeting of the Joint Collaborative Team on Video Coding (JTC-VC) of Guangzhou, 7-15 of October 2010, to use only the spatial motion vector predictors coming from the same slice in the predictor set. This solution solves the problem of parsing at the decoder in case of slice losses. However, the coding efficiency is significantly 20 decreased, since the temporal motion vector predictor is no longer used. Therefore, this solution is not satisfactory in terms of compression performance.

25 SUMMARY OF THE INVENTION

It is desirable to address one or more of the prior art drawbacks.

According to one aspect of the present invention there is 30 provided a method of encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion, the method comprising: obtaining for an image portion to encode a target number of

motion information predictors; generating using said target number a set of motion information predictors; selecting a motion information predictor for said image portion to encode from said generated set of motion information predictors, and signaling in said bitstream said target

5 number and information relating to the selected motion information predictor.

According to another aspect of the present invention there is provided a method of decoding a bitstream comprising an encoded sequence of digital images, at least one portion of an image being

10 encoded by motion compensation with respect to a reference image, the method comprising: obtaining from said bitstream a target number of motion information predictors for an image portion to decode; generating using said target number a set of motion information predictors; and determining a motion information predictor for said

15 image portion to decode from the generated set of motion information predictors.

According to another aspect of the present invention there is provided a device for encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion

20 compensation with respect to a reference image portion, the device comprising: means for obtaining a target number (Nmax) of motion information predictors; means for generating using said target number a set of motion information predictors; and means for selecting a motion information predictor for said image portion to encode from said

25 generated set of motion information predictors, and signaling in said bitstream said target number and information relating to the selected motion information predictor.

According to another aspect of the present invention there is provided a device for decoding a bitstream comprising an encoded sequence of digital images, at least one portion of an image being

30 encoded by motion compensation with respect to a reference image, the device comprising: means for obtaining from said bitstream a target number (Nmax) of motion information predictors; means for generating

using said target number a set of motion information predictors; and means for determining a motion information predictor for an image portion to decode from said generated set of motion information predictors.

5 Further aspects of the present invention provide computer programs which, when run on a computer, cause the computer to carry out a method for encoding a digital video signal as described above and a method for decoding a bitstream as described above. In each case the program may be stored on a computer-readable storage medium.

10 It is also desirable to provide a method allowing correct parsing at the decoder even in the case of a bitstream corrupted by transmission losses while keeping good compression efficiency.

15 To that end, the invention relates to method of encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion. The method comprises, for at least one image portion to encode, the steps of:

-obtaining a target number of motion information predictors to be used for said image portion to encode, and

20 -generating a set of motion information predictors consisting of said target number of motion information predictors, each motion information predictor of the generated set being different from any other motion information predictor of the generated set.

25 Advantageously, the method of the invention allows the systematic determination of a target number of motion information predictors to be used for encoding motion information, such as a motion vector, associated with an image portion, and the compression is advantageously improved by generating a set of motion information predictors which are all different from one another. The potential overhead of using a fixed target number of motion information predictors is compensated by the variety of predictors selected which helps improving the compression rate. The target number of different motion information predictors is determined and fixed independently of

the actual values of the items of motion information, such as motion vectors, selected as motion information predictors for the current image portion to encode.

An embodiment of the present invention is effective when the  
5 number of motion information predictors that is initially generated is a priori unknown, for example as when AMVP is used. For example, if reduction of an initial set is carried out, and the number of initial predictors removed by the reduction process is a priori unknown, an embodiment of the present invention can be used to ensure that the  
10 final set of motion information predictors consists of the target number of motion information predictors.

According to an embodiment, the encoding method further comprises the steps of:

-selecting a motion information predictor for said image  
15 portion to encode from said generated set of motion information predictors, and

-encoding an item of information representative of said selected motion information predictor based upon said target number obtained.

20 Advantageously, a motion information predictor can be selected for a current block to encode and the selected motion vector predictor can be encoded depending on the number of motion information predictors determined. The number of motion information predictors can be systematically retrieved by the decoder, so that the  
25 encoded bitstream can be systematically parsed at a decoder even in case of losses.

According to an embodiment, the item of information representative of said selected motion vector predictor is an index of the selected motion vector predictor in the generated set of motion  
30 information predictors, and the index is encoded on a number of bits dependent upon said target number obtained.

According to an embodiment, in the step of obtaining a target number of motion information predictors to be used for said image

portion to encode, said target number is set equal to a predetermined value for any image portion to encode of the sequence of digital images.

The advantage of this embodiment is that the target number 5 of motion information predictors can be easily obtained, without any supplementary computation or signaling overhead, at both the encoder or the decoder.

According to another embodiment, in the step of obtaining a target number of motion information predictors to be used for said 10 image portion to encode, said target number is determined, for a given image portion to encode, depending upon an encoding information of said given image portion to encode.

Advantageously, such an encoding information can be an encoding parameter, such as for example, if the images are divided into 15 variable size macroblocks for processing, the size of the macroblock to which the image portion to encode belongs. Such an encoding information may also be for example an encoding mode associated with the image portion to encode.

According to yet another aspect, the invention relates to a 20 device for encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion. The device comprises, for at least one image portion to encode:

-means for obtaining a target number of motion information 25 predictors to be used for said image portion to encode, and

-means for generating a set of motion information predictors consisting of said target number of motion information predictors, each motion information predictor of the generated set being different from any other motion information predictor of the generated set.

30 According to yet another aspect, the invention also relates to a computer program product that can be loaded into a programmable apparatus, comprising sequences of instructions for implementing a method for encoding a sequence of digital images as briefly described

above, when the program is loaded into and executed by the programmable apparatus. Such a computer program may be transitory or non transitory. In an implementation, the computer program can be stored on a non-transitory computer-readable carrier medium.

5 The particular characteristics and advantages of the device for encoding a sequence of digital images, of the storage means and of the computer program product being similar to those of the digital video signal encoding method, they are not repeated here.

According to yet another aspect, the invention also relates to  
10 a method for decoding a bitstream comprising an encoded sequence of digital images, at least one portion of an image being encoded by motion compensation with respect to a reference image. For at least one said image portion to decode, the method comprises the steps of:

15 -obtaining a target number of motion information predictors to be used for said image portion to decode, and  
-generating a set of motion information predictors consisting of said target number of motion information predictors, each motion information predictor of the generated set being different from any other motion information predictor of the generated set.

20 The method for decoding a bitstream has the advantage of allowing determining a target number of motion information predictors and using such a number of different motion information predictors. The target number of motion information predictors can be systematically retrieved, and consequently the bitstream can be parsed systematically, 25 even in case of transmission errors. A further advantage is that in all cases, the parsing of the bitstream is simple, and in particular simpler than with prior art methods which adaptively reduce the number of motion information predictors instead of using a predetermined target number that can be obtained by the decoder.

30 According to an embodiment, the method further comprises a step of decoding an item of information representative of a selected motion information predictor for said image portion to decode based upon said target number obtained.

Advantageously, in case an encoding dependent on the number of motion information predictors has been applied at the encoder, the item of information representative of the selected motion information predictor for said image portion to decode can be 5 systematically decoded, even in case of transmission errors.

According to yet another aspect, the invention also relates to a device for decoding a bitstream comprising an encoded sequence of digital images, at least one portion of an image being encoded by motion compensation with respect to a reference image portion. The 10 device comprises, for at least one said image portion to decode:

-means for obtaining a target number of motion information predictors to be used for said image portion to decode, and  
-means for generating a set of motion information predictors consisting of said target number of motion information predictors, each 15 motion information predictor of the generated set being different from any other motion information predictor of the generated set.

According to yet another aspect, the invention also relates to an information storage means that can be read by a computer or a microprocessor, this storage means being removable, and storing 20 instructions of a computer program for the implementation of the method for decoding a bitstream as briefly described above.

According to yet another aspect, the invention also relates to a computer program product that can be loaded into a programmable apparatus, comprising sequences of instructions for implementing a 25 method for decoding a bitstream as briefly described above, when the program is loaded into and executed by the programmable apparatus.

Such a computer program may be transitory or non transitory. In an implementation, the computer program can be stored on a non-transitory computer-readable carrier medium.

30 The particular characteristics and advantages of the device for decoding a bitstream, of the storage means and of the computer program product being similar to those of the decoding method, they are not repeated here.

According to another aspect of the present invention there is provided a method of encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion, the method

5 comprising generating a set of motion information predictors having controlled diversity and selecting a motion information predictor for said image portion to encode from said generated set of motion information predictors.

Controlled diversity means that the motion information predictors

10 of the set are different from one another but one or more of them are statistically likely to be close to the actual motion information so that a residual (difference between the actual motion information and the predictor concerned) is small and hence efficiently compressible.

The method may comprise:

15 generating first motion information predictors;

identifying one or more first motion information predictors as seed predictors;

generating one or more second motion information predictors based on the seed predictor(s); and

20 forming said set of motion information predictors from the first and/or second motion information predictors.

In this case, the first motion information predictors can be predictors expected statistically to give good results in terms of compression efficiency. The second motion information predictors,

25 being based on the first predictors, can then be used to explore in an organized or systematic way other predictors in the predictor space neighbouring the first predictors. Such predictors may also be expected to give good results and the more the predictors that are examined the greater the chance of finding a good match to the actual motion

30 information.

In one embodiment, at least one first motion information predictors is identified as such a seed predictor based on an importance of the first motion information predictor concerned.

The importance may be dependent on a number of times the first motion information predictor concerned appears among the first motion information predictors. The greater the number of times the more important the predictor is considered to be and the more likely it is to be 5 used in the set. As well as looking for identical predictors (duplicates) it can also be effective to look for close matches, too.

Alternatively the importance may be dependent on a measure of how representative the first motion information predictor concerned is of the first motion information predictors as a whole. For example, if the 10 first motion information predictors are averaged, the difference or distance between the average predictor and a given first motion information predictor is a measure of how representative the given predictor is of the first motion information predictors as a whole.

One way of controlling the diversity is to generate at least one 15 said second motion information predictor by adding or subtracting an offset from one of said seed predictors. The offset may be fixed. It could also be a pseudo-random value as long as the same seed value is available to the decoder as to the encoder. If the seed predictors are vectors, it is also possible to control the diversity by adding to the seed 20 predictor another vector, e.g. of fixed magnitude and predetermined direction relative to the direction of the seed predictor.

A plurality of said second motion information predictors may be generated based on the same said seed predictor. If the motion information predictors are vectors each having X and Y components, 25 the plurality of second motion information predictors can be obtained by adding and/or subtracting offsets to/from one or both said components of the same said seed predictor. For example, the same offset can be added to and subtracted from the same seed predictor. If the seed predictor is a vector having X and Y components, there are a 30 number of permutations of adding/subtracting offsets to/from one or both of the X and Y components of the same seed predictor. This is an efficient way of generating controlled diversity without a large processing burden.

Another way of controlling the diversity is to generate a plurality of second motion information predictors by forming average of different pairs (or other combinations) of first motion information predictors. For example, if the first motion information predictors are V1, V2 and V3,

5 three second motion information predictors could be formed from the averages of V1 & V2, V2 & V3 and V3 & V1. It would also be possible to form different weighted combinations of the same first motion information predictors as different second motion information predictors.

The first motion information predictors may be or include motion information predictors each associated with an image portion having a predetermined spatial and/or temporal relationship with the image portion being encoded. For example the motion information predictors used in AMVP may be first motion information predictors. These are a good source of seed predictors.

15 In another embodiment the method comprises:  
generating first motion information predictors;  
checking the differences between the generated first motion information predictors; and  
excluding from said set of motion information predictors one or  
20 more first motion information predictors based on the differences.

By taking into account the differences between the first motion information predictors it is possible to control the diversity of the motion information predictors of the set. It is not necessary in this case to identify seed predictors among the first motion information predictors  
25 and generate second motion information predictors based on the seed predictors. This can be effective, for example, if a sufficiently high number of first motion information predictors are initially available.

For example, a first motion information predictor having the smallest difference from another first motion information predictor can  
30 be removed, as a way of controlling the diversity. The process can be repeated again, as necessary, to successively remove the less diverse predictors.

It is important to note that in this aspect of the invention a number of motion information predictors in said set can be variable.

However, as in the other basic aspect of the present invention the number of motion information predictors in said set can be

5 predetermined at least for a given image portion to encode or even for all image portions (a target number). This makes it possible not only to achieve controlled diversity among the predictors of the set but also to solve the parsing problem noted in the introduction.

This aspect of the invention also provides a corresponding 10 decoding method and corresponding encoding and decoding devices, as well as programs which cause the encoding and decoding.

According to yet another aspect, the invention relates to a method of encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation

15 with respect to a reference image portion. The method comprises, for at least one image portion to encode, the steps of:

-obtaining a first set of motion vector predictors to be used for said image portion to encode, and

-generating a second set of motion vector predictors from 20 said first set of motion vector predictors, each motion vector predictor of the generated set being different from any other motion vector predictor of the generated second set of motion vector predictors, wherein at least one motion vector predictor of the second set is computed from a selected motion vector predictor of the first set.

25 Advantageously, the second set of motion vector predictors generated is used for encoding the motion vector associated with the portion of image to encode. The second set of motion vector predictors comprises a variety of different motion vector predictors, which are generated (and possibly selected) so as to enhance the compression 30 efficiency.

According to an embodiment, a motion vector predictor of the first set is selected in a selecting step according to an importance value.

According to an embodiment, the encoding method comprises a step of computing an importance value associated with each motion vector predictor of the first set.

According to an embodiment, a motion vector predictor of the 5 first set is selected in a selecting step according to a distance among the motion vector predictors of the first set.

The various embodiments for selecting a motion vector predictor to generate further additional or virtual motion vector predictors allow applying a controlled diversity selection, which has the 10 advantage of improving the compression efficiency. Indeed, the use of motion vector predictors computed from important motion vector predictors of the initial set allows to more accurately represent the motion vector of the current image portion to encode. Again, it is not essential to have a fixed or target number of predictors in the final set.

15 According to an embodiment, the motion vector predictors of the first set of motion vector predictors are motion vectors associated with image portions to encode of the image being encoded and/or of a reference image. The first set may be made up of, or include, the predictors used in AMVP.

20

#### BRIEF DESCRIPTION OF THE DRAWINGS

25 Other features and advantages will appear in the following description, which is given solely by way of non-limiting example and made with reference to the accompanying drawings, in which:

- Figure 1, already described, illustrates schematically a set of motion vector predictors used in a motion vector prediction scheme;

30 - Figure 2 is a diagram of a processing device adapted to implement an embodiment of the present invention;

- Figure 3 is a block diagram of an encoder according to an embodiment of the invention;

- Figure 4 illustrates a block diagram of a decoder according to an embodiment of the invention;

- Figure 5 details the determination of a set of motion vector predictors according to a first embodiment;

5 - Figure 6 details the determination of a set of motion vector predictors according to a second embodiment;

- Figure 7 illustrates schematically a second set of motion vector predictors,

10 - Figure 8 illustrates schematically motion vectors in a coordinates system.

## 15 DETAILED DESCRIPTION OF THE EMBODIMENTS

**Figure 2** illustrates a diagram of a processing device 1000 adapted to implement one embodiment of the present invention. The apparatus 1000 is for example a micro-computer, a workstation or a 20 light portable device.

The apparatus 1000 comprises a communication bus 1113 to which there are preferably connected:

-a central processing unit 1111, such as a microprocessor, denoted CPU;

25 -a read only memory 1107 able to contain computer programs for implementing the invention, denoted ROM;

-a random access memory 1112, denoted RAM, able to contain the executable code of the method of the invention as well as the registers adapted to record variables and parameters necessary for 30 implementing the method of encoding a sequence of digital images and/or the method of decoding a bitstream; and

-a communication interface 1102 connected to a communication network 1103 over which digital data to be processed are transmitted.

5 Optionally, the apparatus 1000 may also have the following components:

-a data storage means 1104 such as a hard disk, able to contain the programs implementing the invention and data used or produced during the implementation of the invention;

10 -a disk drive 1105 for a disk 1106, the disk drive being adapted to read data from the disk 1106 or to write data onto said disk;

-a screen 1109 for displaying data and/or serving as a graphical interface with the user, by means of a keyboard 1110 or any other pointing means.

15 The apparatus 1000 can be connected to various peripherals, such as for example a digital camera 1100 or a microphone 1108, each being connected to an input/output card (not shown) so as to supply multimedia data to the apparatus 1000.

20 The communication bus affords communication and interoperability between the various elements included in the apparatus 1000 or connected to it. The representation of the bus is not limiting and in particular the central processing unit is able to communicate instructions to any element of the apparatus 1000 directly or by means of another element of the apparatus 1000.

25 The disk 1106 can be replaced by any information medium such as for example a compact disk (CD-ROM), rewritable or not, a ZIP disk or a memory card and, in general terms, by an information storage means that can be read by a microcomputer or by a microprocessor, integrated or not into the apparatus, possibly removable and adapted to store one or more programs whose execution enables the method of 30 encoding a sequence of digital images and/or the method of decoding a bitstream according to the invention to be implemented.

The executable code may be stored either in read only memory 1107, on the hard disk 1104 or on a removable digital medium

such as for example a disk 1106 as described previously. According to a variant, the executable code of the programs can be received by means of the communication network 1103, via the interface 1102, in order to be stored in one of the storage means of the apparatus 1000

5 before being executed, such as the hard disk 1104.

The central processing unit 1111 is adapted to control and direct the execution of the instructions or portions of software code of the program or programs according to the invention, instructions that are stored in one of the aforementioned storage means. On powering

10 up, the program or programs that are stored in a non-volatile memory, for example on the hard disk 1104 or in the read only memory 1107, are transferred into the random access memory 1112, which then contains the executable code of the program or programs, as well as registers for storing the variables and parameters necessary for implementing the

15 invention.

In this embodiment, the apparatus is a programmable apparatus which uses software to implement the invention. However, alternatively, the present invention may be implemented in hardware (for example, in the form of an Application Specific Integrated Circuit or

20 ASIC).

**Figure 3** illustrates a block diagram of an encoder according to an embodiment of the invention. The encoder is represented by connected modules, each module being adapted to implement, for example in the form of programming instructions to be executed by the

25 CPU 1111 of device 1000, a corresponding step of a method implementing an embodiment of the invention.

An original sequence of digital images  $i_0$  to  $i_n$  301 is received as an input by the encoder 30. Each digital image is represented by a set of samples, known as pixels.

30 A bitstream 310 is output by the encoder 30.

The bitstream 310 comprises a plurality of encoding units or slices, each slice comprising a slice header for encoding values of

encoding parameters used to encode the slice and a slice body, comprising encoded video data.

The input digital images are divided into blocks (302), which blocks are image portions and may be of variable sizes (e.g. 4x4, 8x8, 5 16x16, 32x32). A coding mode is selected for each input block. There are two families of coding modes, spatial prediction coding or Intra coding, and temporal prediction coding or Inter coding. The possible coding modes are tested.

Module 303 implements Intra prediction, in which the given 10 block to encode is predicted by a predictor computed from pixels of the neighbourhood of said block to encode. An indication of the Intra predictor selected and the difference between the given block and its predictor is encoded if the Intra coding is selected.

Temporal prediction is implemented by modules 304 and 15 305. Firstly a reference image among a set of reference images 316 is selected, and a portion of the reference image, also called reference area, which is the closest area to the given block to encode, is selected by the motion estimation module 304. The difference between the selected reference area and the given block, also called a residual 20 block, is computed by the motion compensation module 305. The selected reference area is indicated by a motion vector.

An information relative to the motion vector and the residual block is encoded if the Inter prediction is selected. To further reduce the bitrate, the motion vector is encoded by difference with respect to a 25 motion vector predictor. A set of motion vector predictors, also called motion information predictors, is obtained from the motion vectors field 318 by a motion vector prediction and coding module 317.

Advantageously, the set of motion vector predictors used to select a best motion vector predictor to encode a current motion vector 30 is generated as explained in more detail hereafter with respect to figures 5 and 6. For a given current block to encode, a predetermined number  $N_{max}$  of motion vector predictors is set, and consequently the index of the selected motion vector predictor, which is an item of

information representative of the selected motion vector predictor, can be encoded using a predetermined number of bits. This predetermined number of bits can be also retrieved by the decoder even in case of losses, therefore it is ensured that the decoder will be able to parse the 5 bitstream even in case of errors or losses. The  $N_{max}$  motion vector predictors are selected according to various embodiments to be all different from one another so as to enhance the compression efficiency.

The selection of the predetermined number  $N_{max}$  of motion vector predictors and of the corresponding number of bits to encode the 10 index of the motion vector predictor can be applied either for the entire sequence, or for a group of images of the sequence, or at the block level depending on an encoding parameters such as the block size or the encoding mode. For example, a first predetermined number of motion vector predictors  $N_{max1}$  can be used for the blocks encoded 15 using Inter prediction for which a residual block is encoded, and a second predetermined number motion vector predictors  $N_{max2}$  can be used for the blocks encoded using the SKIP mode, for which only a motion vector is encoded, but no residual block. The respective numbers of motion vector predictors  $N_{max1}$  and  $N_{max2}$  can be for 20 example signaled in the bitstream by inserting them in a header, such as the slice header, or in any appropriate metadata field.

The encoder 30 further comprises a module of selection of the coding mode 306, which uses an encoding cost criterion, such as a rate-distortion criterion, to determine which is the best mode among the 25 spatial prediction mode and the temporal prediction mode. A transform 307 is applied to the residual block, the transformed data obtained is then quantized by module 308 and entropy encoded by module 309. Finally, the encoded residual block of the current block to encode is inserted in the bitstream 310, along with the information relative to the 30 predictor used. For the blocks encoded in 'SKIP' mode, only a reference to the predictor is encoded in the bitstream, without any residual block.

The encoder 30 further performs the decoding of the encoded image in order to produce a reference image for the motion

estimation of the subsequent images. The module 311 performs inverse quantization of the quantized data, followed by an inverse transform 312. The reverse motion prediction module 313 uses the prediction information to determine which predictor to use for a given block and

- 5 the reverse motion compensation module 314 actually adds the residual obtained by module 312 to the reference area obtained from the set of reference images 316. Optionally, a deblocking filter 315 is applied to remove the blocking effects and enhance the visual quality of the decoded image. The same deblocking filter is applied at the decoder, so
- 10 that, if there is no transmission loss, the encoder and the decoder apply the same processing.

**Figure 4** illustrates a block diagram of a decoder according to an embodiment of the invention. The decoder is represented by connected modules, each module being adapted to implement, for  
15 example in the form of programming instructions to be executed by the CPU 1111 of device 1000, a corresponding step of a method implementing an embodiment of the invention.

The decoder 40 receives a bitstream 401 comprising encoding units, each one being composed of a header containing  
20 information on encoding parameters and a body containing the encoded video data. As explained with respect to figure 3, the encoded video data is entropy encoded, and the motion vector predictors' indexes are encoded, for a given block, on a predetermined number of bits. The received encoded video data is entropy decoded (402), dequantized  
25 (403) and then a reverse transform (404) is applied.

In particular, when the received encoded video data corresponds to a residual block of a current block to decode, the decoder also decodes motion prediction information from the bitstream, so as to find the reference area used by the encoder.

**30** The module 410 applies the motion vector decoding for each current block encoded by motion prediction, comprising determining the number  $N_{max}$  of motion vector predictors used and retrieving the motion vector predictor index encoded on a number of bits dependent on  $N_{max}$ .

Similarly to module 317 of figure 3, motion vector decoding module 410 generates a set of  $N_{max}$  motion vector predictors. The embodiments explained hereafter with respect to figures 5 and 6 apply similarly. If the bitstream is received without losses, the decoder generates exactly the

- 5 same set of motion vector predictors as the encoder. In case of losses, it may not be possible to generate the set of motion vector predictors and therefore to correctly decode the motion vector associated with the current block. However, the parsing of the bitstream is always possible, even in case of losses, since the number of bits used to encode the
- 10 index of the motion vectors predictor can be systematically retrieved by the decoder.

Once the index of the motion vector predictor for the current block has been obtained, if no losses have occurred, the actual value of the motion vector associated with the current block can be decoded and

- 15 used to apply reverse motion compensation (406). The reference area indicated by the decoded motion vector is extracted from a reference image (408) to finally apply the reverse motion compensation 406.

In case an Intra prediction has been applied, an inverse Intra prediction is applied by module 405.

- 20 Finally, a decoded block is obtained. A deblocking filter 407 is applied, similarly to the deblocking filter 315 applied at the encoder. A decoded video signal 409 is finally provided by the decoder 40.

**Figure 5** details the generation of the set of motion vector predictors or motion vector candidates in a first embodiment of the 25 present invention. All the steps of the algorithm represented in figure 5 can be implemented in software and executed by the central processing unit 1111 of the device 1000.

Figure 5 represents a flowchart applied for a given current block to encode, which has an associated motion vector designating a 30 reference area in a reference image.

An initial set of motion vector predictors L1 is firstly obtained in step S500. The set L1 is composed of  $N$  candidates. In an embodiment, the initial set of motion vector predictors comprises the

motion vector candidates selected according to the motion vector prediction scheme AMVP already described with reference to figure 1, for example vectors  $V_0$  to  $V_3$  of figure 1 and the median vector computed from  $V_1$ ,  $V_2$  and  $V_3$ . Accordingly,  $N$  is a maximum of 5.

5 Preferably, the initial set of motion vector predictors  $L1$  contains only motion vectors different from one another. Taking the example of figure 1, only motion vectors  $V_0$ ,  $V_1$  and the median vector should be kept in  $L1$ , and the number of motion vector predictors is  $N=3$ .

10 In an alternative embodiment, any other scheme for selecting motion vectors already computed and computing other motion vectors from available ones (i.e. average, median etc) to form the initial set of motion vector predictors  $L1$  can be applied.

15 In a yet another alternative embodiment, the initial set of motion vector predictors  $L1$  is empty and  $N=0$ .

20 In the following step S502, the target number  $N_{max}$  of candidate motion vector predictors to use is obtained.  $N_{max}$  may either be predetermined for the entire sequence of digital images to encode, for example  $N_{max}=4$ , or may be selected according to encoding parameters of the current block to encode or of the encoding unit (for example the slice) to which the current block to encode belongs.

25 For example, the number  $N_{max}$  may be dependent on the size of the transform applied to the macroblock to which the current block being processed belongs: e.g.  $N_{max}=p$  for a transform applied on a block of  $2^{p+1} \times 2^{p+1}$ .

30 In an embodiment, the number of bits  $k$  to use for encoding the index of the motion vector predictor is directly related to  $N_{max}$ :  $k=INT_{sup}(\log_2(N_{max}))$ , where  $INT_{sup}(x)$  is the integer number immediately following the value  $x$ . Preferably,  $N_{max}$  should be a power of 2,  $N_{max}=2^k$ , so as to use all the indexes that can be encoded on  $k$  bits. Advantageously, all the bits used for signaling are used so as to be able to designate a variety of motion vector predictors and to improve the compression.

Alternatively, any type of encoding of the indexes representative of the motion vector predictors can be used, after the number of different motion vector predictors  $N_{max}$  has been determined. In particular, any type of entropy encoding, such as Huffman encoding 5 or arithmetic encoding can be used. Also, the indexes may be also encoded using a prefix type code, such as a Rice-Golomb or a unary code.

Next, it is tested in step S503 whether the number  $N$  of motion vector predictors of set  $L1$  is higher than  $N_{max}$ .

10 In case of positive answer, test S503 is followed by step S504 of selection of a motion vector predictor candidate from  $L1$ , followed by the removal of the selected motion vector predictor candidate from  $L1$  in step S506 to form a modified set of motion vector predictors  $L2$ .

15 The selection of step S504 is applied according to a removal criterion, for example a distance criterion. For example, the set  $L1$  comprises motion vectors predictors  $\{V_1, \dots, V_N\}$ , wherein each motion vector predictor is a vector represented by its components or coordinates on the X-axis and the Y-axis in a coordinate system, as 20 represented in figure 8.

In an embodiment, step S504 determines the distance between each vector  $V_k$  and each other vector  $V_n$  of  $L1$  other than  $V_k$  itself, by computing the distance as:  $d(V_k, V_n) = |V_{kx} - V_{nx}| + |V_{ky} - V_{ny}|$ , where  $|a|$  represents the absolute value of  $a$ . Taking the example of 25 figure 8, vector  $V$  has the coordinates (3,2), vector  $V'$  has the coordinates (4,2) and  $V''$  has the coordinates (3,3). In this example,  $d(V, V') = d(V, V'') = 1$ , so  $V'$  and  $V''$  are at the same distance from vector  $V$ , whereas  $d(V', V'') = 2$ .

Any other type of metric for the calculation of the distances 30 can be alternatively applied.

The minimal distance found  $d(V_p, V_q)$  indicates the two closest vectors  $V_p, V_q$ , among the set  $L1$ , and therefore one of these

two vectors is selected for removal. The selection of one of these two vector can be based on the distance of each one of them to the remaining motion prediction vectors in the modified set L1: the vector between  $V_p$  and  $V_q$  which has the smallest distance to another vector of

5 the set L1 is selected for removal.

The effect is to ensure that the distance between the remaining vectors in the modified set L2 is maximal, so as to allow the use of motion vector predictors as varied or diverse as possible.

After the removal of the selected vector, the value of N is  
 10 decreased (S508), and then N is compared to  $N_{max}$  (S510). If the value of N has not reached yet  $N_{max}$  (answer 'no' to test S510) steps S504 to S510 are repeated. Otherwise, if N has reached  $N_{max}$ , step S510 is followed by step S522 described hereafter.

If the answer to test S503 is 'no', then test S512 checks  
 15 whether N is lower than  $N_{max}$ . In case of negative answer, i.e. if  $N=N_{max}$ , test S512 is followed by step S522 described hereafter.

If the answer to test S512 is 'yes', i.e. if N is strictly lower than  $N_{max}$ , then test S512 is followed by step S514 of obtaining or generating an additional motion vector predictor candidate. Indeed,  
 20 starting from the initial set of motion vector predictors L1, it is possible to add other candidates as motion vector predictors, in a predetermined order, to form a modified set of motion vector predictors L2. Taking the example of figure 7, the motion vectors of blocks 710, 720, 730 and 740 can be added as possible motion vector predictors. Further, the 2  
 25 predictors among 770, 760, 750 which were not selected in S500 can be added as possible motion vector predictors.

For each potential motion vector predictor candidate MV, it is checked whether the motion vector predictor MV is different from all the motion vector predictor candidates already stored in the set L2.

30 If each potential motion vector candidate considered, as for example the motion vectors of blocks 710, 720, 730, 740, 750, 760 and 770 of figure 7, is equal to a motion vector predictor of set L2, new 'virtual' motion vector predictor candidates are computed in step S514.

Such motion vector predictor candidates are called virtual because they are not motion vectors of other blocks of the current image or of the reference image. The virtual motion vector predictors are computed from existing motion vector predictors, for example by

5 adding offsets. For example, from a motion vector MV of set L2 of coordinates ( $MV_x$ ,  $MV_y$ ), it is possible to compute four virtual motion vector predictors by adding/subtracting an offset off to its coordinates:  $MV' (MV_x \pm off, MV_y \pm off)$ . Typically, off may be set equal to 1 or 2.

Alternatively, other modifications of the components of the

10 motion vector predictor MV to obtain virtual motion vector predictors can be applied, so as to obtain diverse motion vector predictors starting from the existing motion vector predictors in order to increase the compression efficiency.

For example, the components of the motion vector MV may

15 be modified independently, using respectively two values offx and offy, and either offx or offy may be set to 0.

In an embodiment, both offx and offy are proportional to the corresponding component:  $offx = aMV_x$  and  $offy = bMV_y$  with a and b typically smaller than 0.5. If necessary, the modified coordinates  $MV_x$

20 and  $MV_y$  are rounded to the closest integer value, so as to represent a displacement on the grid of pixels.

In an embodiment, a supplementary motion vector, of predetermined norm, is added to the vector MV, the supplementary vector having the same direction as motion vector MV, as represented

25 in figure 8: supplemental vector 850 is added to vector 820.

In yet another alternative embodiment, a variance of the motion vectors of the set L2 is computed:

$$var = \sum_{MV \in L2} (MV_x - \overline{MV}_x)^2 + (MV_y - \overline{MV}_y)^2 \quad \text{where } \overline{MV}_x \text{ represents the}$$

mean value of the  $MV_x$  coordinates of the vectors of L2 and  $\overline{MV}_y$

30 represents the mean value of the  $MV_y$  coordinates of the vectors of L2. Then, the offset off is selected by comparing the calculated value var to a predetermined threshold T. T can be equal to  $50/L2$ . If var is lower

than T, the value off is small, for example off=1; if var is higher than T, off is set to a larger value, for example off=3. Further, in this embodiment also, a differentiated value offx or offy can be computed for each component.

5 One motion vector predictor obtained in step S514 is added to the set of motion vector predictors L2 in step S516, and the number N is increased by 1 (step S518).

Next, it is checked in step S520 if N is equal to  $N_{max}$ . In case of negative answer, steps S514 to S520 are repeated.

10 In case of positive answer, the determined target number of motion vector predictor candidates  $N_{max}$  has been reached, and step S520 is followed, at the encoder, by the step S522 of selection of an optimal motion vector predictor for the current block from set L2. For example, a rate-distortion optimization criterion is applied to select the 15 optimal motion vector predictor  $MV_i$  to encode the motion vector of the current block.

At the encoder, the motion residual, i.e. the difference between the motion vector of the current block and the selected motion vector predictor is encoded, as well as an indication of the motion 20 vector predictor selected in step S524. For example, the index i of the selected motion vector predictor  $MV_i$  of L2 is encoded, using  $k=INT_{sup}(\log_2(N_{max}))$  bits.

Alternatively, an entropy encoding of the index i may be applied.

25 In yet another alternative, the index i can be encoded using a prefix type code, such as the Rice-Golomb code, in which each value i is encoded using i '1's followed by a '0'.

The algorithm of figure 5 can also be implemented by a decoder to generate the set of motion vector predictor or motion vector 30 candidates for a given block, without steps S522 and S524.

At the decoder, the index i of the selected motion vector predictor  $MV_i$  for the given block to decode is obtained from the bitstream, knowing  $N_{max}$  and therefore the number of bits k on which the

index  $i$  has been encoded. The steps S500 to S518 are similarly implemented to obtain the set of motion vector predictors L2, so that the index  $i$  decoded from the bitstream designates the motion vector predictor actually used by the encoder.

5           In case of losses during transmission, since the number  $N_{max}$  can be systematically retrieved by the decoder, the received bitstream can be systematically parsed to extract the index  $i$  designating the selected motion vector predictor, even if, depending on the packets lost, the complete set of motion vector predictors L2 may not be obtained at  
10           the decoder.

15           **Figure 6** details the generation of the set of motion vector predictors or motion vector candidates in a second embodiment of the present invention. All the steps of the algorithm represented in figure 6 can be implemented in software and executed by the central processing unit 1111 of the device 1000.

Figure 6 represents a flowchart applied for a given current block to encode, which has an associated motion vector designating a reference area in a reference image.

20           Firstly, similarly to step S502 of figure 5, the target number  $N_{max}$  of motion vector predictor candidates to use is determined in step S600.

In an embodiment,  $N_{max}$  is of the form  $2^k$ , so that each index value that can be coded on  $k$  bits corresponds to a possible motion vector predictor.

25           For example, in order to use all motion vector predictors of the AMVP scheme which proposes 5 motion vector predictors, 3 bits are necessary to encode the index of a motion vector predictor. In this case, preferably  $N_{max}=2^3=8$ .

30           An initial set of motion vector predictor candidates L1 is obtained in step S602. For example, the initial set of  $N=5$  motion vector predictors of AMVP is selected.

A reduction process is applied on the initial set of motion vector predictors to eliminate duplicates, so as to obtain a reduced set

of motion vector predictors containing  $N_1$  elements. Preferably, the number of duplicates of each remaining vector after the reduction process is recorded and stored in a memory for a subsequent use in step S612 described hereafter.

5 It is next checked (test S606) whether  $N_1$  is higher than or equal to  $N_{max}$ , the target number of motion vector predictors. It may be pointed out that a positive outcome to this test only occurs if the algorithm starts with a first set of motion vector predictors a greater number of motion vectors than  $N_{max}$ . In case of positive answer, step  
10 10 S606 is followed by step S630 of selection of the first  $N_{max}$  motion vector predictor candidates of the set L1 to form the set of motion vector predictors L2.

In case of negative answer, i.e. if  $N_1$  is lower than  $N_{max}$ , the set of motion vector predictors has to be complemented with additional  
15 motion vector predictors.

A second set of motion vector predictor candidates L1' is obtained in step S608.

The second set of motion vector predictors L1' is composed of the remaining motion vector predictors of the first set L1 and of  
20 additional motion vectors, for example corresponding to the motion vectors of the block 710, 720, 730 and 740 of the reference image as represented on figure 7. Further, the 2 predictors among 770, 760, 750 which were not selected in S600 can be added as possible motion vector predictors. Each motion vector predictor of the set L1' has a  
25 corresponding index.

Next, a reduction process is applied to the second set of motion vector predictors in step S610 to obtain a reduced second set of motion vector predictors L1" of  $N_2$  vectors. The reduction process eliminates the duplicates, so that all motion vector predictors of L1" are  
30 different from one another. The number of duplicates of each vector kept in L1" is recorded and stored in a memory for a subsequent use in step S612 described hereafter.

It is then checked in step S628 whether the number of motion vector predictors  $N_2$  is higher than or equal to  $N_{max}$ . In case of positive answer, step S628 is followed by step S630 already described.

In case of negative answer, it is necessary to add more  
5 motion vector predictors to the reduced second set of motion vector predictors  $L_1''$  to obtain the final set of  $N_{max}$  motion vector predictors.

Following the test S628, in case of negative answer, an importance value is assigned to each remaining motion vector predictor candidate of the reduced second set of motion vector predictors  $L_1''$  in  
10 step S612.

In an alternative embodiment, step S612 follows directly test S606, in case of negative answer to test S606.

The importance value is computed in this embodiment as the  
15 number of duplicates of a given motion vector predictor, using the number of duplicates of a given motion vector predictor computed and stored during steps S604 and S610. Taking the example of figure 1, two vectors,  $V_0$  and  $V_3$ , are equal, so vector  $V_0$  has an importance value equal to 2.

In an alternative embodiment, the importance value can be  
20 computed as a function of the distance to a representative vector of the set of vectors considered, such as the average value of the vectors of the set or the median of the vectors of the set. Then, the importance may be computed as the inverse of the distance of a given vector of the set  $V_n$  to the representative vector: the closer a vector  $V_n$  is to the  
25 representative vector of the set, the higher the importance of  $V_n$ .

Then, the  $N_2$  remaining motion vector predictor candidates are ordered in step S614 according to an order of decreasing importance value. If several motion vector predictors have the same importance value, they can be ordered according to the increasing  
30 order of their indexes.

The re-ordered motion vector predictors are re-assigned increasing indexes  $\{V_0, V_1, \dots, V_{N_2-1}\}$ .

In the following step S616, a variable  $n$  is initialized to 0 and a variable  $N$  is initialized to  $N_2$ , which is the current number of motion vector predictors in the re-ordered set.

Next, in step S618 following step S616, virtual motion vector predictor candidates are added to the re-ordered set. In this embodiment, the virtual motion vector predictors are computed from the remaining motion vector predictors, ordered according to their importance. The motion vector predictor of index  $n$  of the re-ordered set,  $V_n$  of coordinates  $(V_{n_x}, V_{n_y})$ , is considered. The following list of 8 virtual motion vector predictors, defined by their coordinates, can be computed from  $V_n$ , by adding successively +off and -off to one or the two coordinates of  $V_n$ :  $\{(V_{n_x}+\text{off}, V_{n_y}), (V_{n_x}-\text{off}, V_{n_y}), (V_{n_x}+\text{off}, V_{n_y}+\text{off}), (V_{n_x}+\text{off}, V_{n_y}-\text{off}), (V_{n_x}-\text{off}, V_{n_y}+\text{off}), (V_{n_x}-\text{off}, V_{n_y}-\text{off}), (V_{n_x}, V_{n_y}+\text{off}), (V_{n_x}, V_{n_y}-\text{off})\}$ .

Any alternative computation of the virtual motion vector predictors starting from the motion vector predictor  $V_n$ , in particular the alternatives described above with respect to step S514 of figure 5, may be used.

This list of virtual motion vector predictors is added to the current set of motion vector predictors.

The duplicates are eliminated in step S620.

The value  $N$  is updated in step S622 to the remaining number of motion vector predictors after the removal of potential duplicates.

Next, it is checked whether  $N$  is higher than or equal to  $N_{\max}$  in step S624. In case of negative answer, step S624 is followed by step S634 of increasing of the value  $n$  by 1, and steps S618 to S624 are repeated.

In case of positive answer to step S624, sufficient motion vector predictors have been obtained. Step S624 is followed by step S630 of selection of the first  $N_{\max}$  motion vector candidates to constitute the final set of motion vector predictors  $L_2$  of  $N_{\max}$  vectors.

At the encoder, step S630 is followed by step S632, analogous to step S522 of figure 5, of selection of an optimal motion vector predictor MV<sub>i</sub> among the set of motion vector predictors for the current block, according to a predetermined criterion such as a rate-  
5 distortion criterion.

Step S632 is followed by a step S634 of encoding the motion vector of the current block using the motion vector predictor MV<sub>i</sub>, similar to step S524 of figure 5. For example, the index i of the motion vector predictors MV<sub>i</sub> is encoded using k bits, k being computed from N<sub>max</sub>,  
10 k=INT<sub>sup</sub>(log<sub>2</sub>(N<sub>max</sub>)).

Alternatively, an entropy encoding of the index i may be applied.

In yet another alternative, the index i can be encoded using a prefix type code, such as the Rice-Golomb code, in which each value i  
15 is encoded using i '1's followed by a '0'.

The algorithm of figure 6 can also be implemented by a decoder to generate the set of motion vector predictors or motion vector candidates for a given block, except steps S632 and S634 are omitted on the decoder side.

20 At the decoder, the index i of the selected motion vector predictor MV<sub>i</sub> for the given block to decode is obtained from the bitstream, knowing N<sub>max</sub> and therefore the number of bits k on which the index i has been encoded. The steps S600 to S630 are similarly implemented to obtain the set of motion vector predictors L<sub>2</sub>, so that the  
25 index i decoded from the bitstream designates the motion vector predictor actually used by the encoder.

In case of losses during transmission, since the number N<sub>max</sub>  
30 can be systematically retrieved by the decoder, the received bitstream can be systematically parsed to extract the index i designating the selected motion vector predictor, even if, depending on the packets lost, the complete set of motion vector predictors L<sub>2</sub> may not be obtained at the decoder.

The embodiments described above are based on block partitions of input images, but more generally, any type of image portions to encode or decode can be considered, in particular rectangular portions or more generally geometrical portions.

5 Other alternative embodiments may be envisaged, such as for example starting with a large set of motion vector predictor candidates, comprising a number of vectors  $N$  larger than the target number of motion vector predictors determined  $N_{max}$ , and applying a clustering type algorithm to reduce the set of vectors. For example, a

10 Voronoï partition can be applied to reduce the set to the  $N_{max}$  most representative vectors of the set according to a predefined distance between the vectors of the set.

15 More generally, any modification or improvement of the above-described embodiments, that a person skilled in the art may easily conceive should be considered as falling within the scope of the invention.

The present application claims priority from United Kingdom Patent Application No. 1100462.9 filed on 12 January 2011, the entire content of which is incorporated herein by reference.

CLAIMS

5

1. A method of encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion, the method comprising: obtaining for an image portion to encode a target number (Nmax) of motion information predictors; generating using said target number a set of motion information predictors having controlled diversity; and selecting a motion information predictor for said image portion to encode from said generated set of motion information predictors, and signaling in said bitstream said target number and information relating to the selected motion information predictor.

2. A method as claimed in claim 1, wherein said target number is inserted in a header included in said bitstream.

20 3. A method as claimed in claim 1, wherein said target number is inserted in a slice header included in said bitstream.

25 4. A method as claimed in any preceding claim, comprising: obtaining a first target number (Nmax1) of motion information predictors applicable in a first encoding mode; obtaining a second target number (Nmax2) of motion information predictors applicable in a second encoding mode; and signaling the first target number in said bitstream when the first encoding mode is applied and signaling the second target number in said bitstream when the second encoding mode is applied.

30 5. A method as claimed in any preceding claim, wherein generating said set of motion information predictors comprises: obtaining (S500; S602, S604) an initial set of motion information predictors; testing (S512; S606) whether the number (N; N1) of motion

information predictors in said initial set is lower than the obtained target number (Nmax) and, if so, adding (S514, S516; S608-S630) one or more motion information predictors to said initial set.

- 5                 6. A method as claimed in claim 5, wherein the motion information predictors of said initial set are actual motion information predictors, having motion vectors obtained from image portions of said image being encoded or of a reference image, and potential motion information predictors for addition include one or more further such
- 10                10. actual motion information predictors and also include one or more virtual motion information predictors not having motion vectors obtained from image portions of said image being encoded or of a reference image.
- 15                15. 7. A method as claimed in claim 5 or 6, wherein at least one said virtual motion information predictor is computed from an existing motion information predictor.
- 20                20. 8. A method as claimed in claim 7, wherein a supplementary vector is added to a motion vector of an existing motion information predictor, the supplementary vector having a predetermined direction relative to the direction of the motion vector of the existing motion information predictor.
- 25                25. 9. A method as claimed in claim 8, wherein the magnitude of the supplementary vector is dependent on the magnitude of the motion vector of the existing motion information predictor.
- 30                30. 10. A method as claimed in claim 8 or 9, wherein the supplementary vector has components (aMVx, bMVy) proportional to respective corresponding components of the motion vector of the existing motion information predictor.

11. A method as claimed in any one of claims 5 to 9, comprising eliminating duplicates from said initial set (S604). .

12. A method of decoding a bitstream comprising an  
5 encoded sequence of digital images, at least one portion of an image  
being encoded by motion compensation with respect to a reference  
image, the method comprising: obtaining from said bitstream a target  
number (Nmax) of motion information predictors for an image portion to  
decode; generating using said target number a set of motion  
10 information predictors having controlled diversity; and determining a  
motion information predictor for said image portion to decode from the  
generated set of motion information predictors.

13. A method according to claim 12, further comprising  
15 decoding an item of information representative of a selected motion  
information predictor for said image portion to decode.

14. A method according to claim 13, further comprising  
retrieving said selected motion information predictor from said  
20 generated set of motion information predictors using said decoded item  
of information.

15. A method as claimed in any one of claims 12 to 14,  
wherein said target number is obtained from a header included in said  
25 bitstream.

16. A method as claimed in any one of claims 12 to 14,  
wherein said target number is obtained from a slice header included in  
said bitstream.

30

17. A method as claimed in any one of claims 12 to 16,  
comprising: obtaining from said bitstream a first target number (Nmax1)  
of motion information predictors when a first encoding mode is applied;

and obtaining from said bitstream a second target number (Nmax2) of motion information predictors when a second encoding mode is applied.

18. A method as claimed in any one of claims 12 to 17,  
5 wherein generating said set of motion information predictors comprises:  
obtaining (S500; S602, S604) an initial set of motion information predictors; testing (S512; S606) whether the number (N; N1) of motion information predictors in said initial set is lower than the obtained target number (Nmax) and, if so, adding (S514, S516; S608-S630) one or  
10 more motion information predictors to said initial set.

19. A method as claimed in claim 18, wherein the motion information predictors of said initial set are actual motion information predictors, having motion vectors obtained from image portions of said image being decoded or of a reference image, and potential motion information predictors for addition include one or more further such actual motion information predictors and also include one or more virtual motion information predictors not having motion vectors obtained from image portions of said image being decoded or of a  
15 reference image.  
20

20. A method as claimed in claim 19, wherein at least one said virtual motion information predictor is computed from an existing motion information predictor.

25  
21. A method as claimed in claim 19 or 20, wherein a supplementary vector is added to a motion vector of an existing motion information predictor, the supplementary vector having a predetermined direction relative to the direction of the motion vector of the existing  
30 motion information predictor.

22. A method as claimed in claim 21, wherein the magnitude of the supplementary vector is dependent on the magnitude of the motion vector of the existing motion information predictor.

23. A method as claimed in claim 21 or 22, wherein the supplementary vector has components (aMVx, bMVy) proportional to respective corresponding components of the motion vector of the 5 existing motion information predictor.

24. A method as claimed in any one of claims 18 to 23, comprising eliminating duplicates from said initial set (S604).

10

25. A device for encoding a sequence of digital images into a bitstream, at least one portion of an image being encoded by motion compensation with respect to a reference image portion, the device comprising:

15 means for obtaining a target number (Nmax) of motion information predictors;

means for generating using said target number a set of motion information predictors having controlled diversity; and

20 means for selecting a motion information predictor for said image portion to encode from said generated set of motion information predictors, and signaling in said bitstream said target number and information relating to the selected motion information predictor.

26. A device for decoding a bitstream comprising an 25 encoded sequence of digital images, at least one portion of an image being encoded by motion compensation with respect to a reference image, the device comprising:

means for obtaining from said bitstream a target number (Nmax) of motion information predictors;

30 means for generating using said target number a set of motion information predictors having controlled diversity; and

means for determining a motion information predictor for an image portion to decode from said generated set of motion information predictors.

27. A computer program which, when run on a computer, causes the computer to carry out a method for encoding a digital video signal according to any one of claims 1 to 11 or a method  
5 for decoding a bitstream according to one of the claims 12 to 24.

28. A computer-readable storage medium storing a program according to claim 27.

1/8

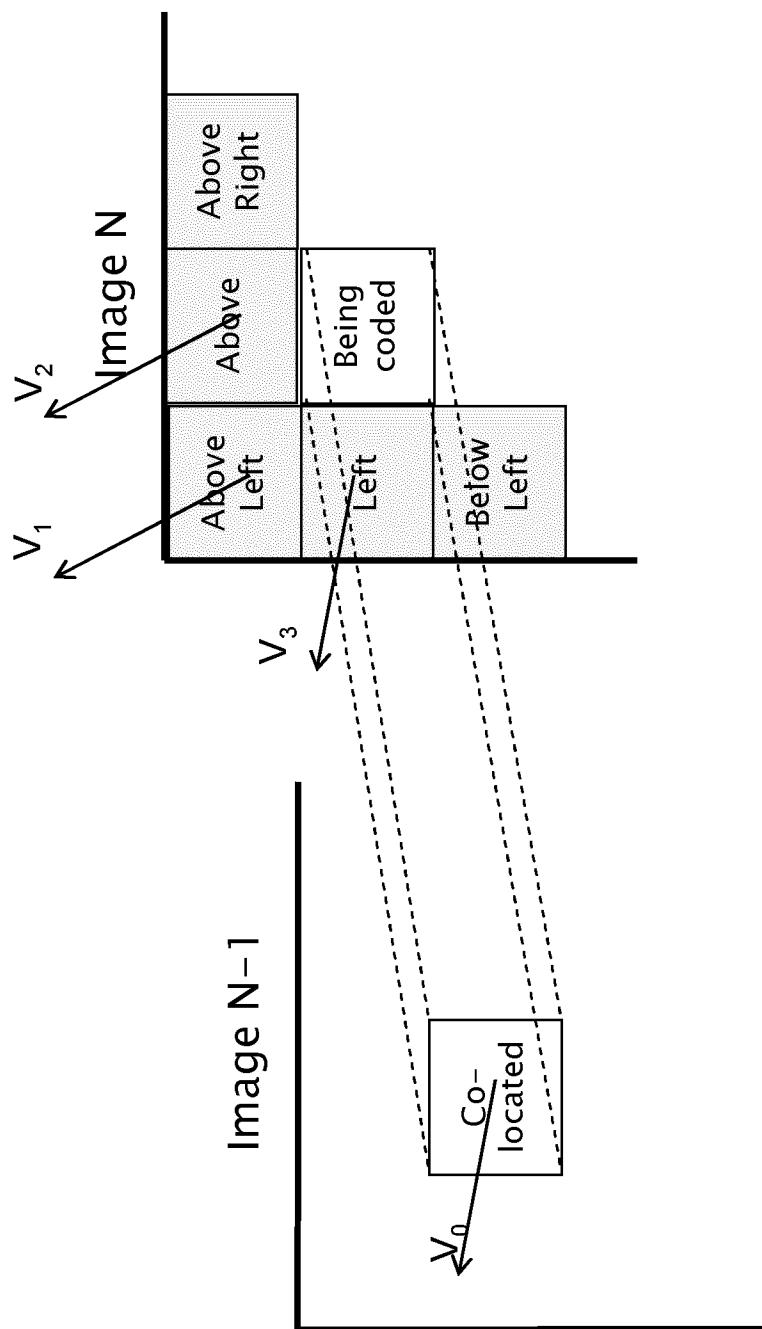


Figure 1

2/8

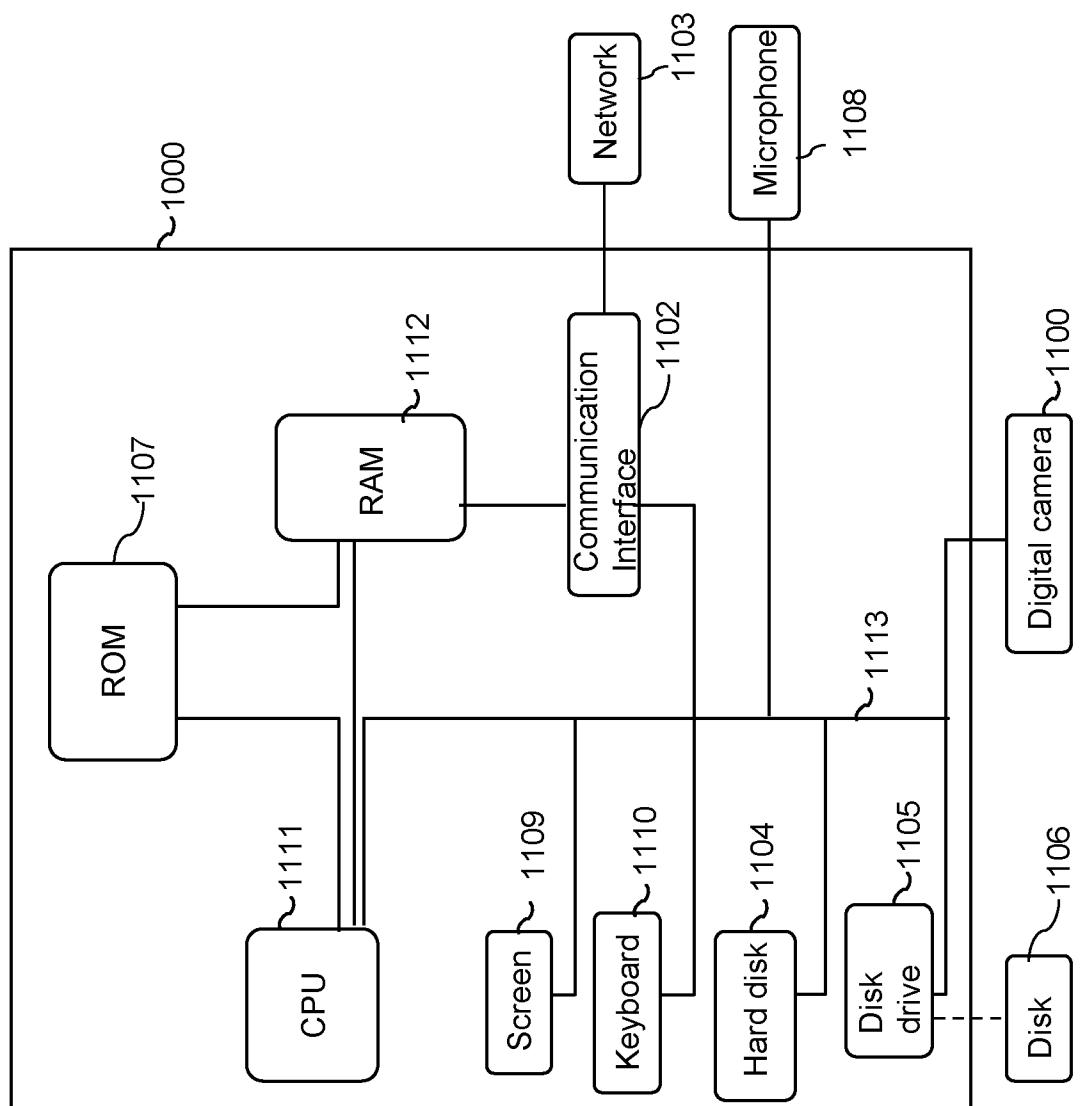


Figure 2

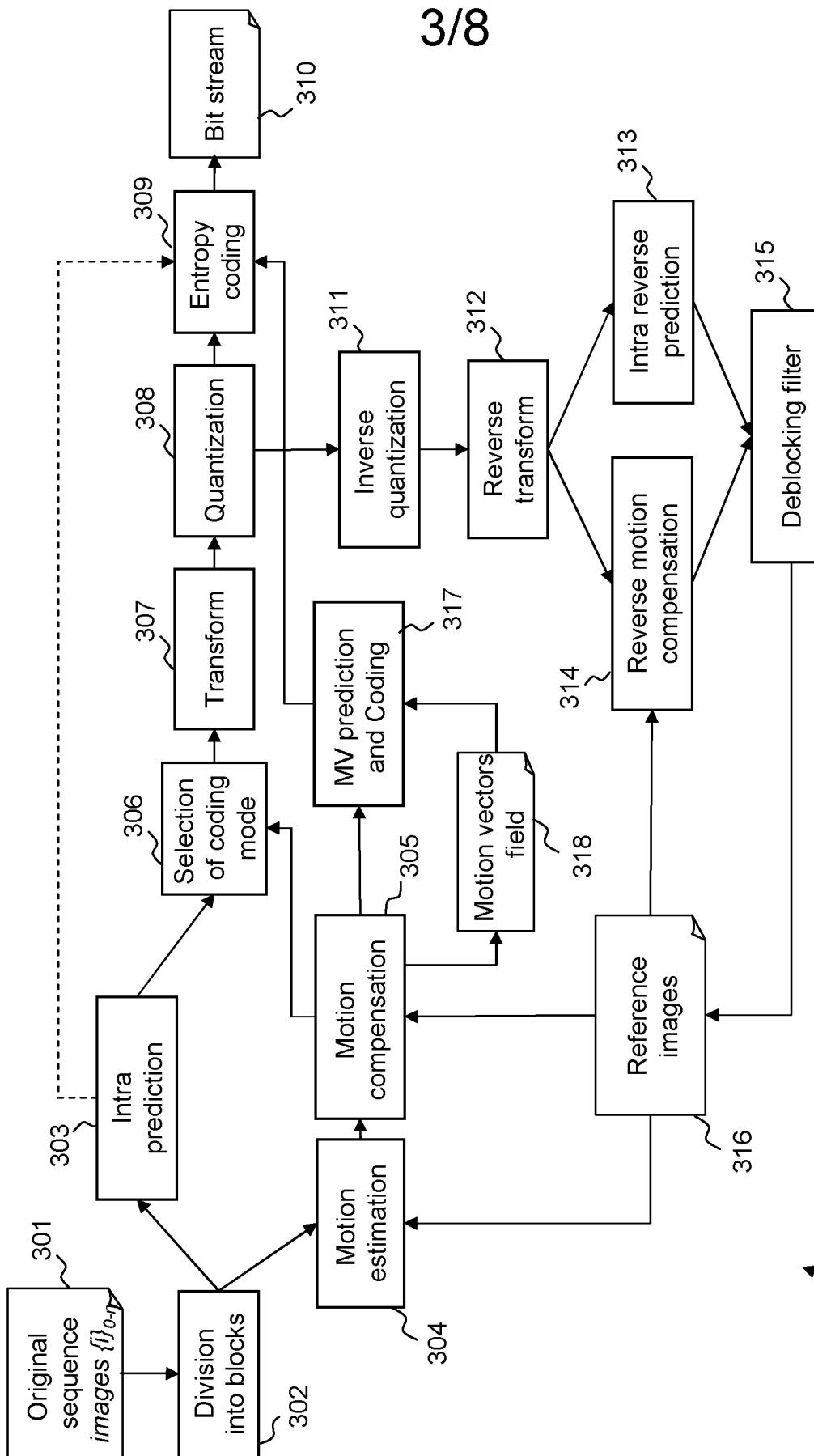


Figure 3

4/8

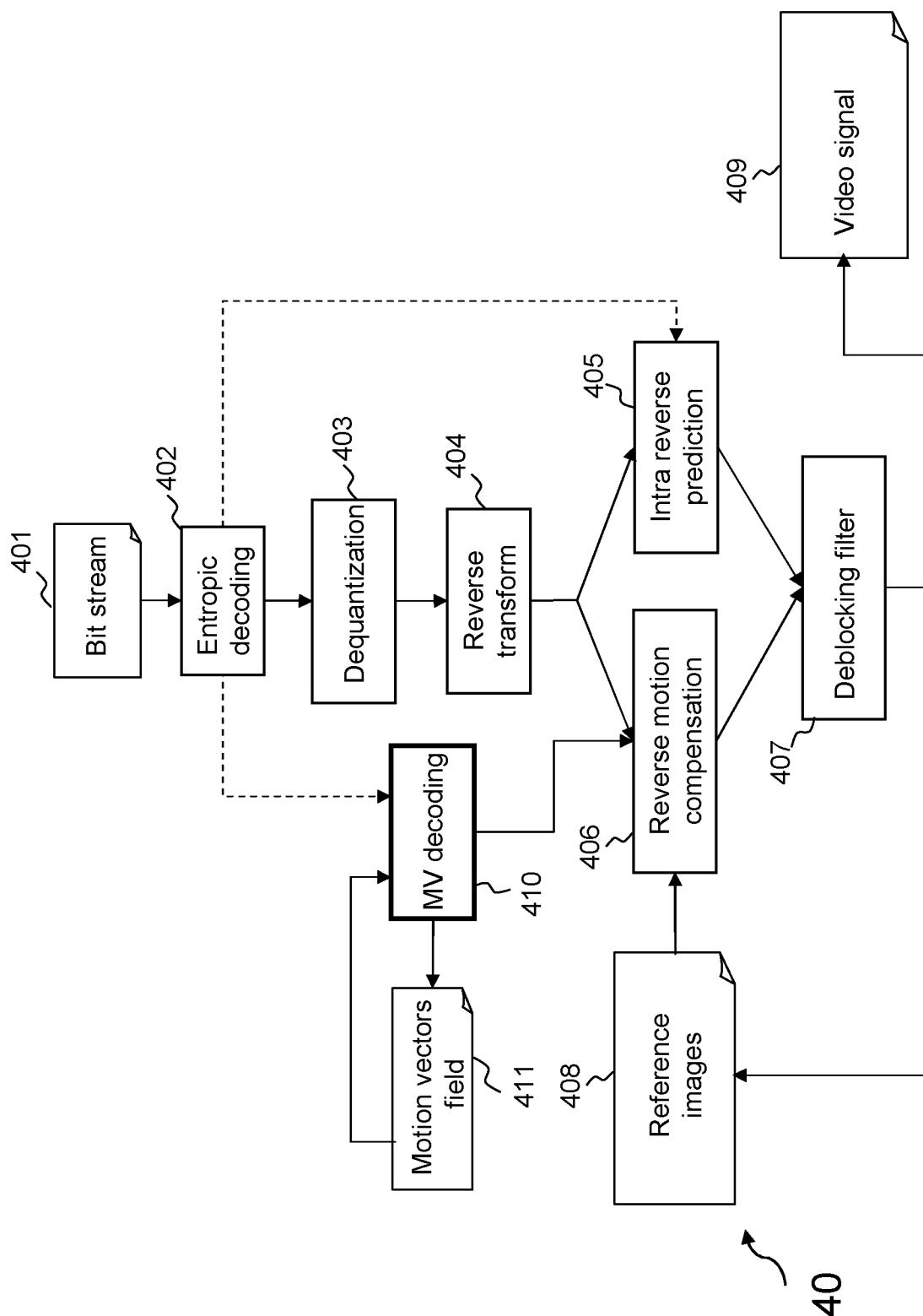


Figure 4

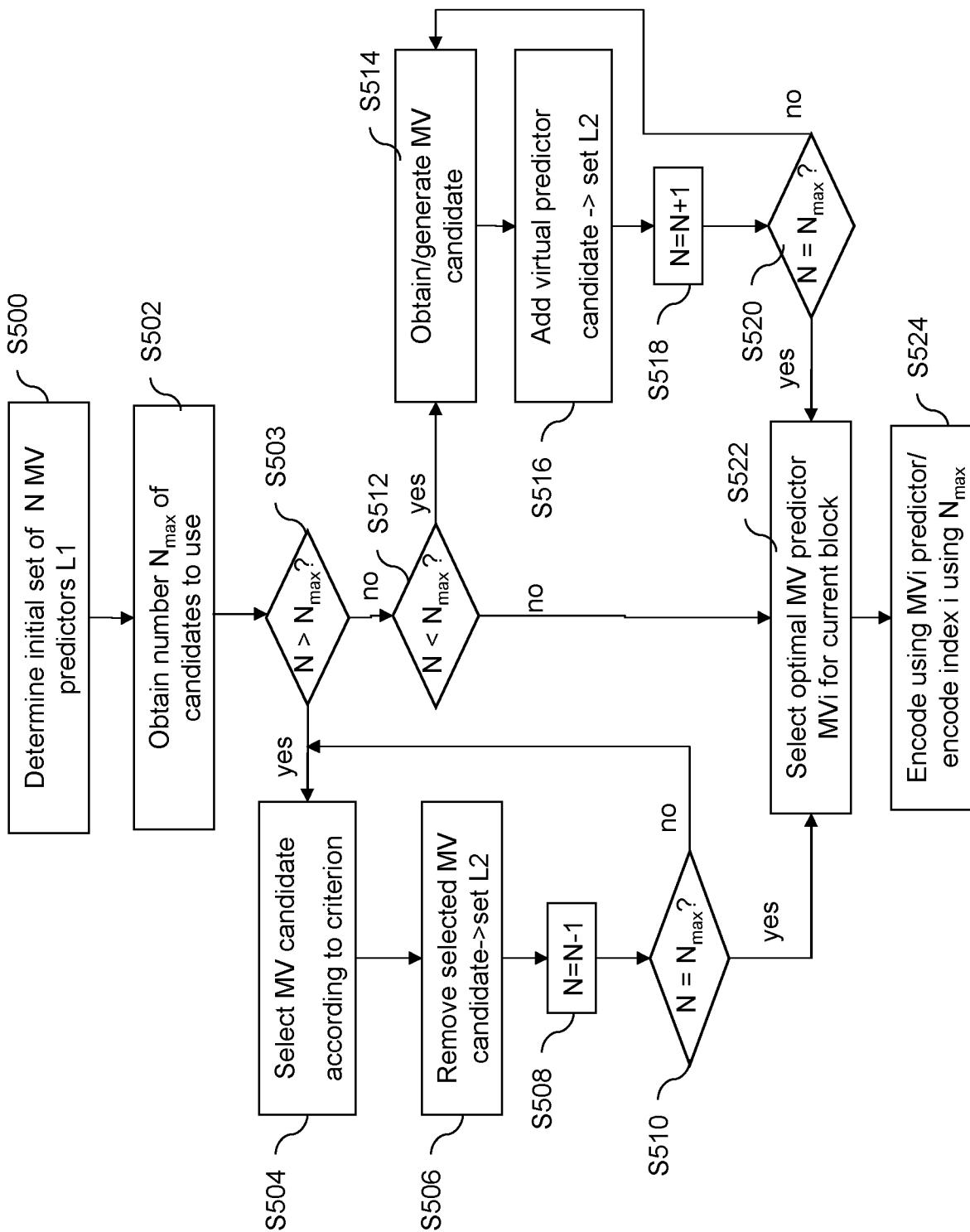


Figure 5

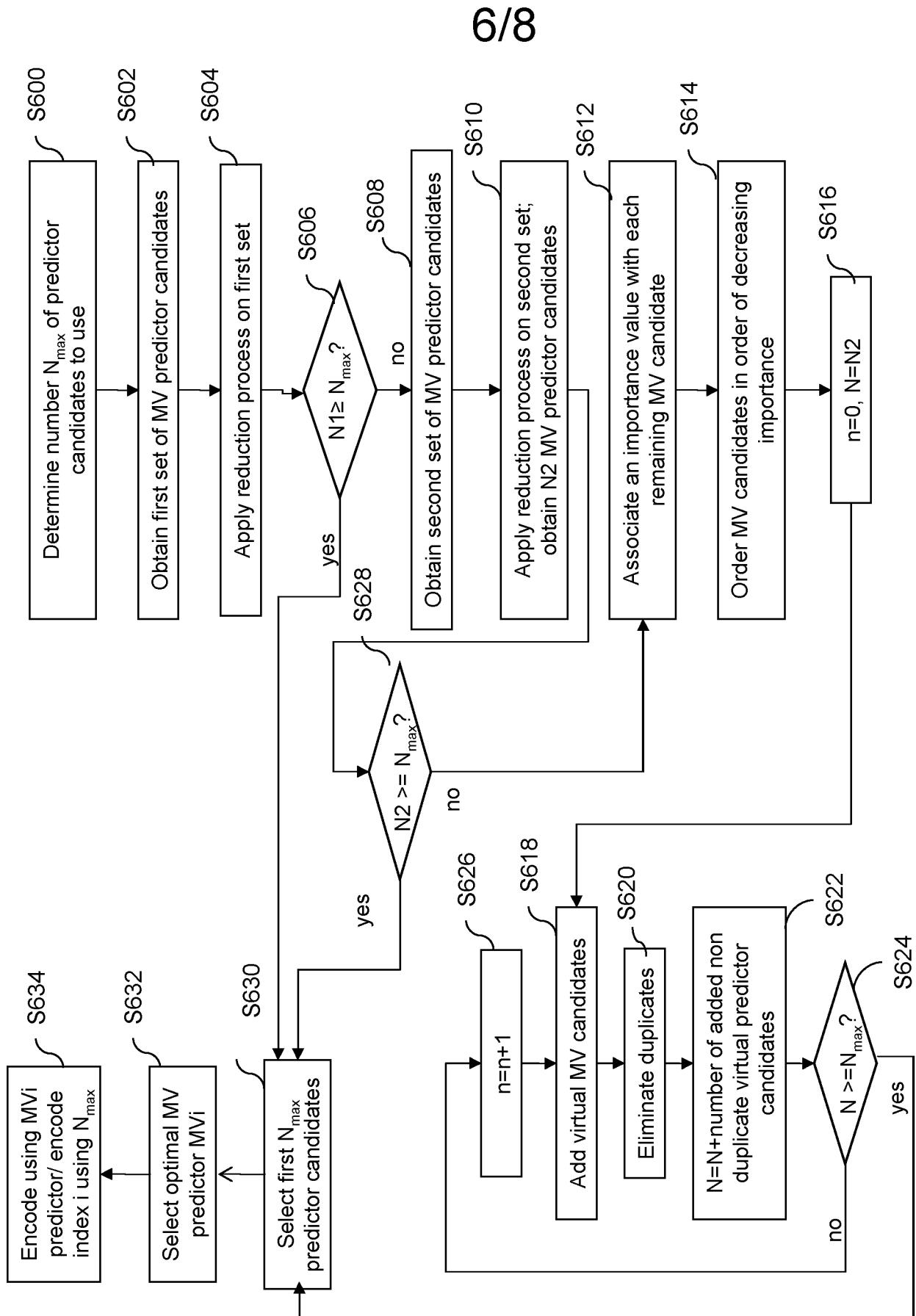


Figure 6

7/8

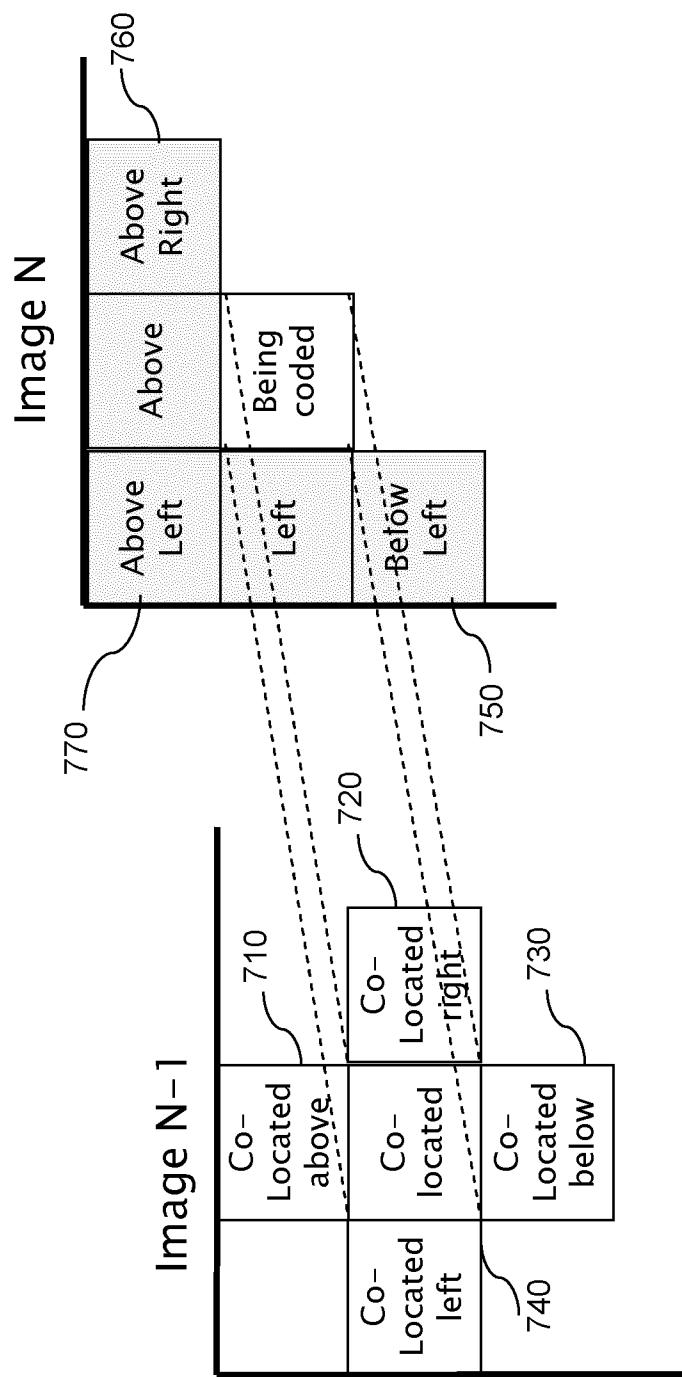


Figure 7

8/8

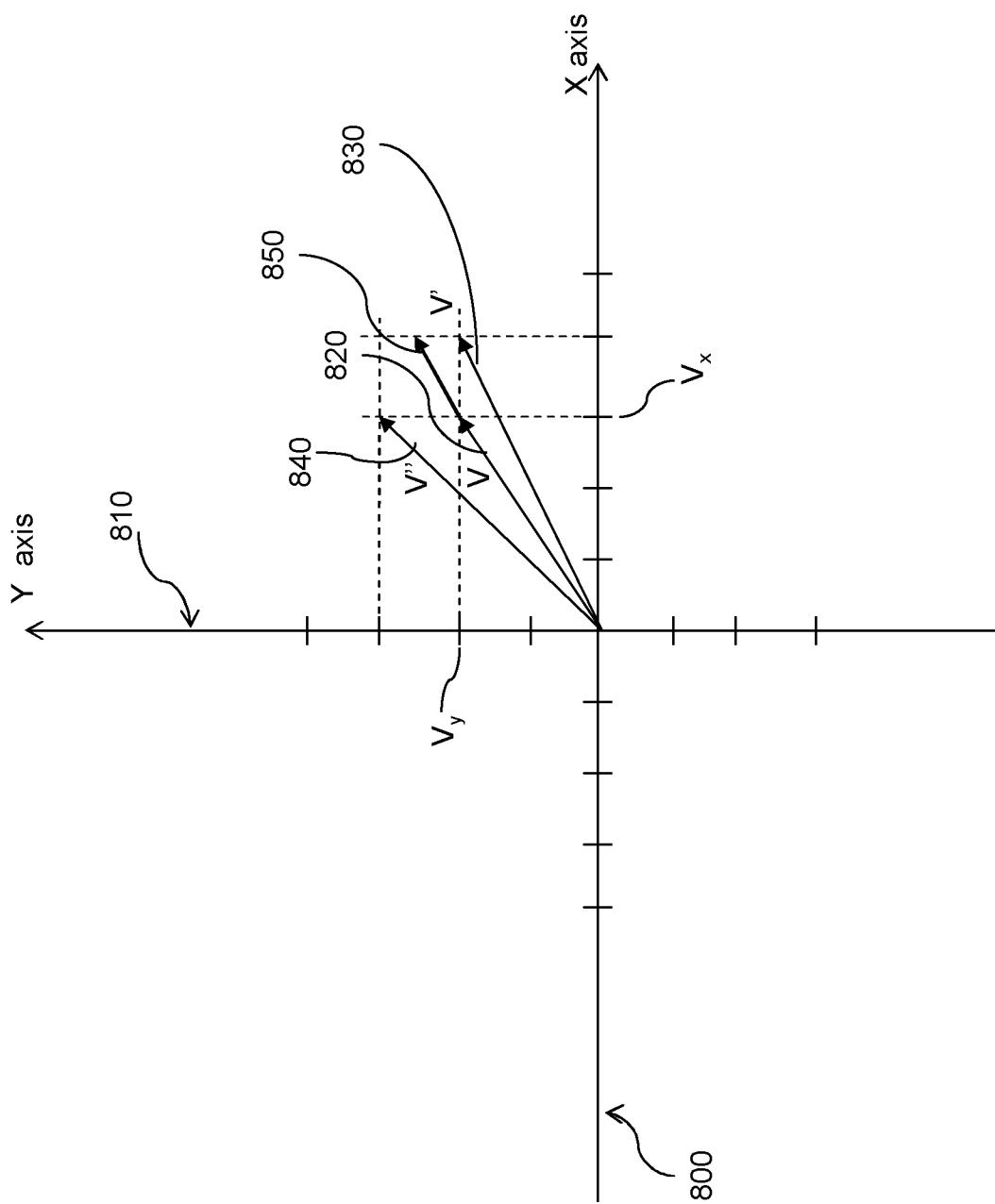


Figure 8

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2012/050390

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. H04N7/64 H04N7/26 H04N7/36  
ADD.

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
H04N

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

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

EPO-Internal, COMPENDEX, INSPEC, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>LAROCHE G ET AL: "RD Optimized Coding for Motion Vector Predictor Selection", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 18, no. 9, 1 September 2008 (2008-09-01), pages 1247-1257, XP011231739, ISSN: 1051-8215, DOI: 10.1109/TCSVT.2008.928882</p> <p>abstract</p> <p>section I. "Introduction"</p> <p>section III.A. "Competition-Based MV Coding"</p> <p>section III.B. "Competition-Based Skip Mode"</p> <p>page 1252, lines 7-13</p> <p>-----</p> <p style="text-align: center;">-/-</p>	1-7, 11-20, 24-28
Y		8,21

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 March 2012	02/04/2012
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  Stoufs, Maryse

## INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/050390

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 2009/115901 A2 (NOKIA CORP [FI]; HALLAPURO ANTTI OLLI [FI]; UGUR KEMAI [FI]; LAINEMA J) 24 September 2009 (2009-09-24)</p> <p>figures 2, 5a, 5b, 6, 7, 8 page 3, lines 1-12 page 9, lines 17-27 page 10, lines 24-34 page 11, line 3 - page 12, line 14</p> <p>-----</p>	1-3,5-7, 11,12, 15,16, 18-20, 24-28
Y	<p>BERIC A ET AL: "A technique for reducing complexity of recursive motion estimation algorithms", SIGNAL PROCESSING SYSTEMS, 2003. SIPS 2003. IEEE WORKSHOP ON 27 - 29 AUG. 2003, PISCATAWAY, NJ, USA, IEEE, 27 August 2003 (2003-08-27), pages 195-200, XP010661014, ISBN: 978-0-7803-7795-0</p> <p>page 198, left-hand column, last paragraph - right-hand column, line 3</p> <p>section 3.4 "Pruning the motion vector candidate set"</p> <p>-----</p>	8,21
A	<p>JINGJING DAI ET AL: "Motion Vector Coding Based on Optimal Predictor Selection", 15 December 2009 (2009-12-15), ADVANCES IN MULTIMEDIA INFORMATION PROCESSING - PCM 2009, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 1040 - 1047, XP019134851, ISBN: 978-3-642-10466-4</p> <p>section 2.1 "Adaptive Predictor Candidate Set"</p> <p>section 2.2 "Optimal Predictor Selection"</p> <p>-----</p>	1-28
X,P	<p>WO 2011/095260 A1 (ERICSSON TELEFON AB L M [SE]; RUSERT THOMAS [SE]; ANDERSSON KENNETH [S] 11 August 2011 (2011-08-11) the whole document</p> <p>-----</p>	1-7, 11-20, 24-28
X,P	<p>GUILLAUME LAROCHE ET AL: "Robust solution for the AMVP parsing issue", 5. JCT-VC MEETING; 96. MPEG MEETING; 16-3-2011 - 23-3-2011; GENEVA; (JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16); URL: HTTP://WFTP3.IITU.INT/AV-ARCH/JC TVC-SITE/,, no. JCTVC-E219, 10 March 2011 (2011-03-10), XP030008725, ISSN: 0000-0007</p> <p>section 2.2 "Tool 2: Replacement of redundant candidates"</p> <p>-----</p>	1-8, 11-21, 24-28
1		-/-

## INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/050390

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	<p>CHEN J ET AL: "MVP index parsing with fixed number of candidates", 6. JCT-VC MEETING; 97. MPEG MEETING; 14-7-2011 - 22-7-2011; TORINO; (JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16 ); URL: <a href="HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/">HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/</a>, no. JCTVC-F402, 12 July 2011 (2011-07-12), XP030009425, figure 1 section 2.1 "Alternative candidates in AMVP mode" section 2.2. "Alternative candidates in Merge mode" -----</p>	1-8, 11-21, 24-28

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No  
PCT/EP2012/050390

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 2009115901	A2	24-09-2009	EP 2266318 A2	29-12-2010
			US 2009304084 A1	10-12-2009
			WO 2009115901 A2	24-09-2009
<hr/>				
WO 2011095260	A1	11-08-2011	US 2011194608 A1	11-08-2011
			US 2011194609 A1	11-08-2011
			WO 2011095259 A1	11-08-2011
			WO 2011095260 A1	11-08-2011
<hr/>				