



(51) International Patent Classification:

H04N 7/26 (2006.01) *H04N* 7/36 (2006.01)
H04N 7/34 (2006.01) *H04N* 7/50 (2006.01)

(21) International Application Number:

PCT/US2009/063537

(22) International Filing Date:

6 November 2009 (06.11.2009)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

12/291,570 12 November 2008 (12.11.2008) US

(71) Applicant (for all designated States except US): **ALCATEL-LUCENT USA INC.** [US/US]; 600-700 Mountain Avenue, Murray Hill, NJ 07974-0636 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MCGOWAN, James, William** [US/US]; 114 Provincetown Court, Flemington, NJ 08822 (US). **RECCHIONE, Michael, Charles** [US/US]; 565 Passaic Avenue, Nutley, NJ 07110 (US).

(74) Agent: **BROWN, Kenneth M.**; Alcatel-Lucent USA Inc., Docket Administrator - Room 2F-192, 600-700 Mountain Avenue, Murray Hill, New Jersey 07974-0636 (US).

(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, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

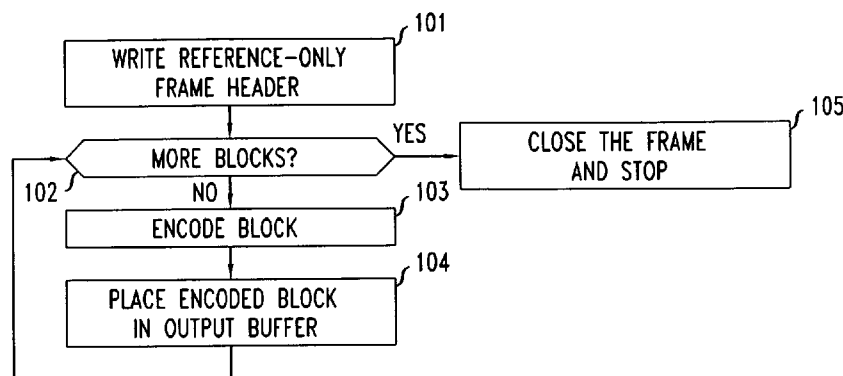
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, 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))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: VIDEO CODING USING VIRTUAL REFERENCE DATA

FIG. 1



(57) Abstract: A video codec is provided in which virtual reference data (e.g., virtual reference blocks) are advantageously employed by a video coder and decoder to significantly improve coding efficiency as compared to prior art codecs. Specifically, a "Virtual reference," as used herein, is defined as a group of pixels (e.g., a block) that is used as reference material for encoding portions of the video signal (e.g., a motion-compensated inter-predicted block), but that does not comprise or represent any portion of the actual video sequence to be displayed. For example, it may advantageously be determined that portions of the video sequence may be efficiently represented as a motion-compensated block of pixel values which are generated by the video coder based on the blocks being predicted but not based on any actual blocks in any of the actual video frames or slices.

J. McGowan 28-19

- 1 -

VIDEO CODING USING VIRTUAL REFERENCE DATA

Field of the Invention

5 The present invention relates generally to the field of video compression and more particularly to an improved method and apparatus for performing block-based video coding and decoding.

Background of the Invention

10 Most major standard video codecs (*i.e.*, coders and decoders) achieve a data compression advantage over still image coding techniques by using a block-based, motion-compensated prediction scheme. Such schemes are fully familiar to those of ordinary skill in the art and include, for example, the well known video coding standards ITU-T/H.261, H.262, H.263, and H.264, and the corresponding ISO/IEC-
15 11172 (MPEG-1), ISO/IEC 13818 (MPEG-2), ISO/IEC 14496 (MPEG-4/ASP), and ISO/IEC 14496-10 (MPEG-4/AVC).

 Block-based schemes treat (typically rectangular) regions of pixels (*i.e.*, picture elements) as a single unit, and typically encode such units using one of two broad methods. The first method is to independently code the block without reference
20 to other blocks at all (*i.e.*, “intra-coding”), or to code a block relative to some surrounding pixel values (*i.e.*, “intra-prediction”). The second method is to code a block by providing a pointer to a similar, previously coded block which is used as a *reference*, and to further provide an error block that indicates the differences between the reference and the block to be coded (*i.e.*, inter-coding). This latter method works
25 extremely well because video frames tend to be very similar to other nearby video frames, with many of the changes therebetween due to portions of the scene moving between frames – hence, the term “motion-compensated prediction” is used to indicate that these reference block pointers are compensating for object or camera motion. Note that inter-coding is generally more efficient than intra-coding. (The
30 efficiency of a codec in this context refers to the ability of a codec to represent an image with a minimum of distortion by using as few bits as possible.)

J. McGowan 28-19

- 2 -

In recent years, many efficiency gains have been realized by providing increasingly better ways of specifying a reference block, including “bi-directional prediction,” which uses multiple reference blocks (for example, one from a previous-in-time video frame and one from a future-in-time video frame – even though each of these video frames must necessarily have been previously decoded before the given video frame can be decoded), or by using variable size reference blocks. However, even with the best of the current video coding techniques, the bandwidth required for video transmission systems such as, for example, HDTV (High Definition Television), and for wireless video applications still exceeds the capabilities of these coding techniques and limits either the quality or the quantity of the resultant video.

Summary of the Invention

We have recognized that there exist certain limitations in current video coding techniques which limit the possible efficiency thereof. The first such limitation is that the reference blocks used in current video coders must necessarily be part of a previously decoded video frame (which may, in fact, represent either a previous-in-time and/or a later-in-time video frame). That is, the reference block must necessarily be based on an actual part of the video as presented in an actual video frame. The second such limitation is that reference blocks must be updated *en masse*, as part of a slice or frame. (Those of ordinary skill in the art will recognize that a frame refers to a single still-image picture in a video sequence, whereas a “slice” simply refers to some subset of a frame.) That is, in many standard video codec implementations, such as H.264, the encoder can save multiple slices or frames to use as references, but ultimately, some finite limit is chosen to fix the storage required at the decoder.

In accordance with various illustrative embodiments of the present invention, a video codec is provided in which *virtual reference data* (e.g., virtual reference blocks) are advantageously employed by a video coder and decoder to significantly improve coding efficiency as compared to prior art codecs. Specifically, a “virtual reference,” as used herein, is defined as a group of pixels (e.g., a block) that is used as reference material for encoding portions of the video signal (e.g., a motion-compensated inter-predicted block), but that does not comprise or represent any

J. McGowan 28-19

- 3 -

portion of the actual video sequence to be displayed. For example, it may advantageously be determined that portions of the video sequence may be efficiently represented as a motion-compensated block of pixel values which are generated by the video coder based on the blocks being predicted, but do not represent any actual
5 block in any of the actual video frames or slices.

Specifically, an illustrative embodiment of the present invention provides a method and apparatus for generating an encoded video signal from an original video signal, the original video signal comprising a sequence of video frames, the encoded video signal for use in a subsequent video display of said original video signal, the
10 method or apparatus comprising steps or machine-executable code for: generating virtual reference data based on a portion of the original video signal, wherein the generated virtual reference data does not represent any portion of any individual frame of the original video signal which is to be displayed in said subsequent video display thereof; encoding said generated virtual reference data and incorporating said
15 encoded virtual reference data into the encoded video signal; incorporating into the encoded video signal an indication that said encoded virtual reference data incorporated therein comprises data which does not represent any portion of any individual frame of the original video signal to be displayed in said subsequent video display thereof; encoding one or more portions of the original video signal which are
20 to be displayed in said subsequent video display thereof, wherein said encoding of said one or more portions of said original video signal which are to be displayed are based on and specifically reference a portion of said encoded virtual reference data; and incorporating said encoded portions of said original video signal which are to be displayed into said encoded video signal.

25 In addition, another illustrative embodiment of the present invention provides a method and apparatus for decoding an encoded video signal to generate a decoded video signal for use in a subsequent video display of an original video signal, the original video signal comprising a sequence of video frames, the encoded video signal having been generated from the original video signal, the method or apparatus
30 comprising steps or machine-executable code for: identifying in said encoded video signal an indication that encoded virtual reference data has been incorporated therein,

J. McGowan 28-19

- 4 -

wherein said encoded virtual reference data does not represent any portion of any individual frame of the original video signal which is to be displayed in said subsequent video display thereof; decoding from said encoded video signal said encoded virtual reference data incorporated therein; and decoding from said encoded
5 video signal one or more encoded portions of the original video signal which are to be displayed in said subsequent video display thereof, wherein said encoded portions of said original video signal which are to be displayed are based on and specifically reference a portion of said encoded virtual reference data.

10 **Brief Description of the Drawings**

Figure 1 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with a first illustrative embodiment of the present invention.

Figure 2 shows a flow chart of a method of decoding an encoded video signal
15 with use of virtual reference data in accordance with a first illustrative embodiment of the present invention.

Figure 3 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with a second illustrative embodiment of the present invention.

Figure 4 shows a flow chart of a method of decoding an encoded video signal
20 with use of virtual reference data in accordance with a second illustrative embodiment of the present invention.

Figure 5 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with a third illustrative embodiment of the
25 present invention.

Figure 6 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with a third illustrative embodiment of the present invention.

Figure 7 shows a flow chart of a method of encoding a video signal with use
30 of virtual reference data in accordance with a fourth illustrative embodiment of the present invention.

J. McGowan 28-19

- 5 -

Figure 8 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with a fourth illustrative embodiment of the present invention.

Figure 9 shows a flow chart of an illustrative method for generating coefficient values, which may be employed in connection with various ones of the methods of encoding a video signal with use of virtual reference data in accordance with various illustrative embodiments of the present invention.

Detailed Description of the Preferred Embodiments

10 Introduction

To clarify the operation of the illustrative embodiments described in detail below, consider the advantage of encoding a virtual reference in terms of minimizing encoded bits. Note first that typical state-of-the-art encoders represent blocks to be encoded in terms of previously encoded frames (or slices) by specifying the error between a pointed-to reference block and the block to be encoded. This error is typically encoded using entropy coding (which is fully familiar to those of ordinary skill in the art) or some similar method that reduces the required number of bits (typically to less than one bit of error per pixel).

In the H.264 standard encoder, for example, the Context-Adaptive Variable Length Coding (CAVLC) entropy coder (which is also familiar to those skilled in the art) is used as a lossless compression method well suited for block-based video coding. In a typical case, the error is a quantized difference between the discrete cosine transform (DCT) coefficients of the predicted pixels and the DCT coefficients of the actual pixels. (The use of DCT coefficients in video coding is also fully familiar to those of ordinary skill in the art.) In general, the encoding is more efficient if there are many differences which are equal to zero, but it is still highly efficient if the (absolute value of the) difference of select non-zero terms is equal to 1. An occasional absolute difference greater than 1 in these select terms breaks the efficiency of the entropy coder and requires a disproportionately large number of bits to encode.

Consider, for example, the case where one of these select terms is equal to 2.

J. McGowan 28-19

- 6 -

In the H.264 coder, for example, one would need to either accept the cost of coding the additional bits, distort the video by changing the absolute error to a 0 or a 1, or increase the quantization level of the error by, for example, dividing by 2 (which is identical to dropping the least significant bit of the error) so that the offending value of a 2 becomes a 1, and (unfortunately) all the previous 1 values become 0 values. Adopting the first choice is to accept lower efficiency, and adopting either the second or the third choices is to increase the distortion (*i.e.*, to decrease the quality of the resultant video). This creates a classic trade-off between bit-rate versus quality in a video codec. In summary, two blocks are sent (a first block for the previously encoded frame or slice, and a second block for the frame or slice being coded, *inter alia* in terms of the previously encoded frame or slice), but the second block may be coded in a less efficient manner than is, in theory, possible.

In accordance with the principles of the present invention, and in accordance with various illustrative embodiments thereof, the coding efficiency associated with the above example could be significantly improved by sending *virtual* reference data having the term in question set one bit higher than in the first block above. Thus, the first block above could advantageously use the virtual reference data with a -1 "error" (*i.e.*, the value of the term in question in the first block would be one less than the value specified in the virtual reference data), and the second block above could advantageously use the virtual reference data with a +1 "error" (*i.e.*, the value of the term in question in the second block would be one more than the value specified in the virtual reference data). In this case, a virtual block is sent, and two regular (*i.e.*, "real") blocks (*i.e.*, the first block above and the second block above) are also sent. The net coding, however, will be more efficient in many cases, since the virtual block can be carefully tuned to minimize the "error" values for the two "real" blocks. Effectively, although the coding of the virtual block and first regular block are likely to slightly exceed (in bits required) the coding of the first block for a prior art coder such as H.264, the second regular block is likely to be coded at such a better efficiency that there is a net reduction in the total number of required bits.

30

A first illustrative embodiment of the present invention

J. McGowan 28-19

- 7 -

In accordance with a first illustrative embodiment of the present invention, an entire frame (or slice) of virtual reference blocks is advantageously coded and transmitted in the coded bit stream to the decoder. These coded virtual reference blocks are then advantageously decoded and stored for future use as reference blocks, without ever being displayed as part of the output video (since they do not represent any part of the actual video signal itself). The coding of the virtual reference material may be otherwise identical to normal codec operation, with intra-coded blocks ("I-blocks"), bi-directionally-coded blocks ("B-blocks"), and predictive-coded ("P-blocks") each used where appropriate (*i.e.*, where most advantageous).

In particular, in accordance with the principles of the present invention, the virtual reference blocks are advantageously generated in such a manner so as to maximize the overall efficiency of the codec, whereas prior art encoders must necessarily choose reference blocks from existing blocks (*i.e.*, blocks which represent an actual portion of the video). One method for generating virtual reference blocks in accordance with an illustrative embodiment of the present invention is described below in connection with Figure 9 herein.

Figure 1 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with the first illustrative embodiment of the present invention. Flowchart box 101 writes a frame (or slice) header indicating that this is a virtual reference frame (or slice). Flowchart box 102 decides whether there are more video blocks in the virtual reference frame (or slice) to be encoded. If so, flowchart box 103 encodes a next video block, flowchart box 104 places the encoded video block in an output buffer, and control proceeds back to flowchart box 102. If not, flowchart box 105 closes the frame (or slice) and the flow terminates. In this manner, one or more virtual reference blocks have been advantageously generated and encoded for subsequent use in, for example, the predictive coding of actual video blocks by the encoder.

Figure 2 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with the first illustrative embodiment of the present invention. Flowchart box 201 identifies that the frame (or slice) to be decoded is a virtual reference (only) frame (or slice). Flowchart box 202 decides

J. McGowan 28-19

- 8 -

whether there are more blocks to be decoded. If so, flowchart box 203 decodes a next video block, flowchart box 204 places the video block directly in a reference frame buffer, and control proceeds back to flowchart box 202. If not, flowchart box 205 closes the frame (or slice) and flow terminates. In this manner, one or more virtual
5 reference blocks have been advantageously decoded and stored for subsequent use in, for example, the predictive decoding of actual video blocks by the decoder.

A second illustrative embodiment of the present invention

In accordance with a second illustrative embodiment of the present invention,
10 what will be referred to herein as “Dual blocks” or D-blocks” are introduced. As defined herein, a D-block is advantageously composed of three parts: a header, a primary block and a secondary block. The header is a flag indicating the type of reference to follow. For example, if the header is a “DBLOCK_NO” flag (which may, for example, be represented by a 0 bit), it indicates that there is no virtual
15 reference block, which therefore indicates a conventional encoding. In this case the flag may be merely removed by the decoder, and the block may be decoded as it is in conventional decoders (*e.g.*, as an I-block, a P-block, or a B-block). If, on the other hand, the header is a “DBLOCK_YES” flag (which may, for example, be represented by a 1 bit), it indicates that two blocks will be encoded (and will immediately follow the
20 header). First, what will be referred to herein as the “primary” block will be encoded in the bits immediately following the initial flag, and this will advantageously comprise the encoding of a virtual reference block. Following the encoding of the virtual reference block will be an encoding of what will be referred to herein as the “secondary” block, which will comprise an encoding of the actual block representing
25 the current location in the video (*i.e.*, a “normal video block” in a conventional coder). Either of the two blocks (*i.e.*, the primary block and the secondary block) may be an I-block, a P-block, or a B-block, independent of one another. In particular, the secondary block may advantageously be coded as a P-block or a B-block which references (*i.e.*, depends upon) the primary (virtual reference) block.

30 Figure 3 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with the second illustrative embodiment of the

J. McGowan 28-19

- 9 -

present invention. Flowchart box 301 determines whether the video block should be represented as a D-block. If it is advantageous to do so, flowchart box 302 writes a DBLOCK_YES flag into the output buffer as the header. Flowchart box 303 then encodes a virtual reference video block and writes it into the output buffer. Flowchart box 304 encodes a normal (*i.e.*, actual) video block and writes it into the output buffer, and flow terminates at flowchart box 305. Advantageously, the normal video block may be encoded with reference to the virtual reference video block encoded in flowchart box 303. If flowchart box 301 determines that it is not advantageous to represent the video block as a D-block, flowchart box 306 writes a DBLOCK_NO flag into the output buffer as the header, and flow proceeds to flowchart box 304 to encode the video block in a conventional manner.

Figure 4 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with the second illustrative embodiment of the present invention. Flowchart box 401 examines the DBLOCK flag in the header of the encoded data for the video block. If the flag is equal to DBLOCK_YES, flowchart box 402 decodes the first (primary) video block, which comprises a virtual reference block, and writes the decoded virtual reference block directly into the reference frame buffer. Flowchart box 403 then decodes the next (secondary) video block, which represents the actual video block to be displayed, and writes the decoded video block directly into a video frame output buffer. Flow then terminates at flowchart box 404. If, on the other hand, the flag from flowchart box 401 is equal to DBLOCK_NO, flow proceeds directly to flowchart box 403 where the (actual) video block is decoded and written to the video frame output buffer.

A third illustrative embodiment of the present invention

In accordance with a third illustrative embodiment of the present invention, the virtual reference block embedded in the D-block in accordance with the second illustrative embodiment of the present invention (as described above) advantageously contains an additional motion vector to place (*i.e.*, locate within the frame or slice) the primary block. In accordance with this third illustrative embodiment, this may be indicated by an extended header. For example, following a DBLOCK_YES flag (see

J. McGowan 28-19

- 10 -

above), a DBLOCK_RELOCATED_YES flag may be used to indicate that a D-block is to be relocated, whereas a DBLOCK_RELOCATED_NO flag may be used to indicate no such relocation (and thus, the scheme as described above in connection with the second illustrative embodiment of the present invention). Note that a leading
5 DBLOCK_NO may be advantageously handled in an identical manner as is described above in connection with the second illustrative embodiment of the present invention, and is advantageously not followed by the extended header.

When the DBLOCK_RELOCATED_YES flag is present, the first motion vector of the secondary block may be advantageously used to relocate the primary
10 virtual reference block, and the secondary block must therefore be either a P-block or a B-block. For example, a secondary P-block following an extended header may be advantageously parsed to decode the motion vector, and the decoded motion vector may then be advantageously applied to the primary block, such that the primary block is decoded and then replaces the block pointed to by the motion vector.

15 Note that an encoder will advantageously employ either the technique of the second illustrative embodiment of the present invention or the technique of the third illustrative embodiment of the present invention at any given time, since the second header bit needs to be properly interpreted by the corresponding decoder. In accordance with various illustrative embodiments of the invention, this may be
20 achieved by (a) only using one of these two techniques throughout the encoding of the entire video signal; (b) updating in a frame (or slice) header information specifying which of the techniques (*i.e.*, "modes") is to be used until a subsequent such update is provided; or (c) switching between modes using some parameter setting sequence in the syntax of the coder.

25 Figure 5 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with the third illustrative embodiment of the present invention. First, flowchart box 501 determines whether the video block should be represented as a D-block. If it is advantageous to do so, flowchart box 502 writes a DBLOCK_YES flag into the output buffer, and then, flowchart box 503
30 determines whether it is advantageous for the video block to be a relocated D-block. If it is, flowchart box 504 writes a DBLOCK_RELOCATED_YES flag into the

J. McGowan 28-19

- 11 -

output buffer, and flowchart box 505 encodes an appropriate motion vector and writes it into the output buffer. (The motion vector may be determined with use of conventional motion-compensation techniques fully familiar to those of ordinary skill in the art.) Then, flowchart box 506 encodes a virtual reference block and writes it to the output buffer, and flowchart box 507 encodes the (actual) video block, omitting the motion vector for the P-block or the first motion vector of a B-block, which is now advantageously assumed to be the motion vector from flowchart box 505 (*i.e.*, the motion vector from the virtual reference block). Flow terminates at flowchart box 508.

10 If, however, flowchart box 503 determines that the video block should not be a relocated D-block, then flowchart box 509 writes a DBLOCK_RELOCATED_NO flag into the output buffer. Flowchart box 510 then encodes the virtual reference block and places it in the output buffer, and flowchart box 511 encodes the normal block and places it in the output buffer as well. Flow is then terminated at flowchart box 508. Finally, if flowchart box 501 had determined that the video block should not be a D-block, then flowchart box 512 writes a DBLOCK_NO flag into the output buffer, and flow proceeds to flowchart box 511 to encode the normal block and place it in the output buffer. In accordance with an alternative illustrative embodiment of the present invention, all video blocks in flowchart box 507 may be encoded normally, without the encoder making any assumptions about the inter-coded blocks pointing to the virtual reference block encoded in flowchart box 506.

Figure 6 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with the third illustrative embodiment of the present invention. Flowchart box 601 examines the DBLOCK flag in the video block. If the flag is determined to be DBLOCK_YES, flowchart box 602 examines the DBLOCK_RELOCATED flag in the video block. If the flag is determined to be DBLOCK_RELOCATED_YES, then flowchart box 603 decodes the provided motion vector and flowchart box 604 decodes the provided virtual reference block. Flowchart box 605 then places the virtual reference block in the reference frame and location as indicated by the motion vector decoded in flowchart box 603, and flowchart box 606 then decodes the provided (actual) video block without use of a

J. McGowan 28-19

- 12 -

motion vector. As indicated above, if the video block is an inter-coded block, the first motion vector is advantageously assumed to be the block decoded in flowchart box 603 (*i.e.*, the virtual reference block is the virtual reference block decoded in flowchart box 604). Flowchart box 606 also advantageously writes the decoded
5 (actual) video block directly into a video frame output buffer, and then flow terminates at flowchart box 607.

If, on the other hand, the DBLOCK_RELOCATED flag is determined in flowchart box 602 to be DBLOCK_RELOCATED_NO, flow proceeds to flowchart box 608 which decodes a virtual reference block and places it in the reference buffer.
10 Flow then proceeds to flowchart box 609, which decodes a (normal) video block and places it in the output buffer. If, however, flowchart box 601 had determined that the DBLOCK flag in the video block is DBLOCK_NO, then flow would have proceeded directly to flowchart box 609 for the decoding of a (normal) video block. In accordance with an alternative illustrative embodiment of the present invention
15 corresponding to the alternative illustrative embodiment described above in connection with the illustrative encoder of Fig. 5, all video blocks in flowchart box 606 may be decoded normally, with the decoder making no assumptions about the inter-coded blocks pointing to the reference from flowchart box 603.

20 **A fourth illustrative embodiment of the present invention**

In accordance with a fourth illustrative embodiment of the present invention, *intra-coding* prediction modes, such as those employed by the H.264 standard codec (which is fully familiar to those of ordinary skill in the art) are advantageously modified. That is, whereas the previously described illustrative embodiments of the
25 present invention have employed virtual reference data (*i.e.*, virtual reference blocks) advantageously used in certain *inter-coding* prediction modes, this fourth illustrative embodiment employs virtual reference data which are advantageously used in certain *intra-coding* prediction modes. Note that the intra-coding prediction modes of the H.264 standard codec are fully familiar to those of ordinary skill in the art.

30 In particular, these intra-coding prediction modes function in a manner similar to that of inter-prediction modes such as motion compensation, but rather than using

J. McGowan 28-19

- 13 -

video blocks from other frames (or slices), added to an error block, to predict a video block to be coded, they use the values of surrounding pixels from other blocks in the same frame (or slice), added to an error block, to predict the video block to be coded. The surrounding pixels may, for example, comprise the column of pixels to the left of the block to be decoded, the row of pixels above the block to be coded, or a number of other such possibilities. In the H.264 standard codec, for example, there are eight such intra-coding prediction modes in total.

In accordance with this fourth illustrative embodiment of the present invention, the pixels used in a given prediction mode are advantageously encoded in the primary block of a D-block, such as is described in connection with the second illustrative embodiment of the present invention. (See above.) The specific prediction mode used in encoding a given block, which is conventionally specified in an intra-prediction block, is, in accordance with this fourth illustrative embodiment of the present invention, advantageously specified in the secondary block of the D-block. Note that the primary block need not be saved, since it is used only for the current block. The intra-coded pixels may be advantageously and efficiently encoded using standard methods, such as differential encoding, fully familiar to those of ordinary skill in the art.

Figure 7 shows a flow chart of a method of encoding a video signal with use of virtual reference data in accordance with the fourth illustrative embodiment of the present invention. Flowchart box 701 determines whether it would be advantageous to represent the current video block as an intra-predicted D-block. If so, flowchart box 702 writes a DBLOCK_YES flag into the output buffer and flowchart box 703 encodes virtual reference pixels and writes them into the output buffer. Flowchart box 704 then encodes the (normal) video block as an intra-predicted block using these virtual reference pixels, and writes the encoded video block into the output buffer. Flow terminates at flowchart box 705. If, on the other hand, flowchart box 701 determines that the current video block should not be represented as an intra-predicted D-block, flowchart box 706 writes a DBLOCK_NO flag into the output buffer, and flow proceeds to flowchart box 704, where the (normal) block is encoded and written into the output buffer in a conventional manner.

J. McGowan 28-19

- 14 -

Figure 8 shows a flow chart of a method of decoding an encoded video signal with use of virtual reference data in accordance with the fourth illustrative embodiment of the present invention. Flowchart box 801 examines the DBLOCK flag in the encoded video block. If the flag is determined to be DBLOCK_YES, flowchart box 802 decodes the provided set of virtual reference pixels. Flowchart box 803 then decodes the current video block using the virtual reference pixels, and writes the video block directly into a video frame output buffer. Flow then terminates at flowchart box 804. If, on the other hand, the flag from flowchart box 801 is determined to be DBLOCK_NO, flow proceeds directly to flowchart box 805, which decodes the block using reference pixels in a conventional manner. Flow then terminates at flowchart box 804.

Note that in accordance with various additional illustrative embodiments of the present invention, it may be advantageous to mix the techniques of the second illustrative embodiment described above, the third illustrative embodiment described above, and the fourth illustrative embodiment described above, by judiciously inspecting the primary and secondary blocks, and based thereupon, determining the which coding and decoding techniques are to be employed. Such judicious mixtures will be obvious to those skilled in the art based upon the disclosure provided herein.

An illustrative method for selecting coefficient values in virtual blocks

Figure 9 shows a flow chart of an illustrative method for generating frequency transform coefficient values, which may be employed in connection with various ones of the methods of encoding a video signal with use of virtual reference data in accordance with various illustrative embodiments of the present invention. (The use of frequency transform coefficients for coding video signals is fully familiar to those of ordinary skill in the art.) In particular, the figure shows an illustrative method which may be employed for generating a virtual reference block for use in the above-referenced illustrative embodiments of the invention. More specifically, a block of coefficient values – for example, DCT (Discrete Cosine Transform) or similar coefficients, fully familiar to those of ordinary skill in the art – may be generated by the illustrative method of Figure 9 and then used as the data for a virtual reference

J. McGowan 28-19

- 15 -

block in accordance with, for example, the illustrative encoding methods shown in Figures 1, 3 and 5.

Specifically, block encoding in accordance with the illustrative method of Figure 9 proceeds as follows: A DCT or similar block transform employed by the codec is advantageously performed on an $N \times M$ block of pixels, where, for example, $N = M = 16$. Such a step is specified in essentially all conventional MPEG and ITU block-based video codecs, and is fully familiar to those of ordinary skill in the art. This process is typically performed on a "Group of Pictures" (GOP), which is defined in conventional codecs as a set of frames (or slices) that are to be independently encoded without reference to any content outside the GOP. (Use of a GOP is also fully familiar to those of ordinary skill in the art). Conventional motion search procedures, implemented in essentially all such conventional encoders, can be then used to determine which coefficients belong to a "family", where a family is defined as a group of coefficients that move throughout at least some section of the GOP. (Note that, in terms of the images being encoded, this typically refers to a real-world object that is captured moving across the images, or that moves due to camera motion.) In accordance with certain illustrative embodiments of the present invention, a virtual reference block will be advantageously used to represent such a family. Ideally, the virtual reference block will have coefficients that minimize the total error of the P-blocks referencing it, in combination with the total encoded size of itself, and will be advantageously placed in the reference frame so as to minimize the encoded number of bits for all motion vectors pointing to that virtual reference block. A practical method for generating a virtual reference block which meets or is close to such an ideal in most cases is advantageously provided by the illustrative method shown in Figure 9.

Specifically, as shown in Figure 9, flowchart box 901 first selects the lowest modal value m – that is, if there is more than one mode, the lowest value is selected. (A "mode" of a set of values is defined in statistics theory to be a value which occurs most frequently in the set.) Using this number, m , will thereby result in the largest possible number of "0" entries in subsequent error blocks. However, in practice, the actual distribution of values in a set of transform coefficients representing a video

J. McGowan 28-19

- 16 -

block is often close to bimodal or is skewed, and therefore, using such a modal value may not always be optimal in reducing the total number of coefficients having an error with an absolute value less than or equal to 1 (*i.e.*, “small error” coefficients, as opposed to “large error” coefficients having an error with an absolute value greater than 1).

Thus, next, in flowchart box 902, N_m , which is defined herein to be the total number of large error coefficients (*i.e.*, those having an error with an absolute value greater than 1) assuming the use of the selected modal value m , is determined (*i.e.*, counted). Flowchart box 903 then computes N_{m-1} , which is defined herein to be the total number of large error coefficients assuming instead the use of a value of one less than the previously determined modal value – namely, using $m-1$ instead of m . Then, in flowchart box 904, N_{m-1} and N_m are compared, and if N_{m-1} is less than N_m , then flow proceeds to flowchart box 905, in which the value $m-1$ is selected and flow proceeds to flowchart box 910 where the process stops.

Otherwise (*i.e.*, if N_{m-1} is not less than N_m), flow proceeds to flowchart box 906, which computes N_{m+1} , which is defined herein to be the total number of large error coefficients assuming the use of a value of one above the previously determined modal value – namely, using $m+1$ instead of m (or $m-1$). Next, in flowchart box 907, N_{m+1} is compared to N_m , and if N_{m+1} is less than N_m , then flow proceeds to flowchart box 908 in which the value $m+1$ is selected and flow proceeds to flowchart box 910 where the process stops. Otherwise (*i.e.*, if N_{m+1} is not less than N_m), flowchart box 909 selects the originally determined modal value m , and again, flow proceeds to flowchart box 910 where the process stops.

Once all of the coefficients are selected, the block is advantageously placed in the virtual frame at a location that is the approximate average of its location through the frames. Since the blocks are allocated to non-overlapping places, the virtual frame is divided into a plurality of $N \times M$ sections, and so the virtual block may be advantageously placed in the closest one of these sections. If, however, the closet one of these sections is occupied, the block may instead be placed in one of the nearby open sections.

J. McGowan 28-19

- 17 -

Illustrative embodiments comprising a video encoder or a video decoder

In accordance with various illustrative embodiments of the present invention, video encoders and video decoders may be advantageously implemented using machine-executable code which implement the principles of the present invention.

- 5 For example, the above-described illustrative methods may be physically embodied in machine-executable code which may then be incorporated into an otherwise conventional memory, and such executable code may then be executed by a conventional associated processor. As used herein, the term "memory" is used without limitation and may comprise any type of data storage device which may hold
- 10 machine-executable code, including, for example, read-only memory (ROM), random access memory (RAM), mass storage devices (such as, for example, a disc), *etc.* Similarly, as used herein, the term "processor" is used without limitation and may comprise any device capable of executing program code supplied thereto in any possible manner.

15

Addendum to the detailed description

- It should be noted that all of the preceding discussion merely illustrates the general principles of the invention. It will be appreciated that those skilled in the art will be able to devise various other arrangements, which, although not explicitly
- 20 described or shown herein, embody the principles of the invention, and are included within its spirit and scope. In addition, all examples and conditional language recited herein are principally intended expressly to be only for pedagogical purposes to aid the reader in understanding the principles of the invention and the concepts contributed by the inventor to furthering the art, and are to be construed as being
- 25 without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention, as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. It is also intended that such equivalents include both currently known equivalents as well as equivalents developed in the future – *i.e.*, any
- 30 elements developed that perform the same function, regardless of structure.

J. McGowan 28-19

- 18 -

What is claimed is:

1. A method for generating an encoded video signal from an original video signal, the original video signal comprising a sequence of video frames, the encoded
5 video signal for use in a subsequent video display of said original video signal, the method comprising the steps of:

generating virtual reference data based on a portion of the original video signal, wherein the generated virtual reference data does not represent any portion of any individual frame of the original video signal which is to be displayed in said
10 subsequent video display thereof;

encoding said generated virtual reference data and incorporating said encoded virtual reference data into the encoded video signal;

incorporating into the encoded video signal an indication that said encoded virtual reference data incorporated therein comprises data which does not represent
15 any portion of any individual frame of the original video signal to be displayed in said subsequent video display thereof;

encoding one or more portions of the original video signal which are to be displayed in said subsequent video display thereof, wherein said encoding of said one or more portions of said original video signal which are to be displayed are based on
20 and specifically reference a portion of said encoded virtual reference data; and

incorporating said encoded portions of said original video signal which are to be displayed into said encoded video signal.

2. The method of claim 1 wherein said encoded video signal comprises a
25 block-based encoding of said original video signal, wherein said virtual reference data comprises one or more virtual reference blocks, and wherein said encoded portions of said original video signal which are to be displayed are encoded based on said encoded virtual reference data with use of a block-based motion compensated prediction scheme.

30

J. McGowan 28-19

- 19 -

3. The method of claim 2 wherein the step of generating the virtual reference data comprises:

identifying a set of video blocks comprising a video block from each of a plurality of video frames represented by said original video signal; and

5 generating the virtual reference data so as to minimize differences between said identified video blocks and the generated virtual reference data.

4. The method of claim 2 wherein said encoded virtual reference data and one encoded portion of said original video signal which is to be displayed are incorporated
10 into said encoded video signal together as an encoded dual block.

5. The method of claim 1 wherein the encoded video signal comprises a block-based encoding of said original video signal, wherein one of said portions of the original video signal which is to be displayed is comprised in a given frame of said
15 original video signal, wherein said virtual reference data is generated based on said given frame of said original video signal, and wherein said one of said portions of the original video signal which is to be displayed is encoded based on said encoded virtual reference data with use of an intra-coding prediction scheme.

20 6. A method for decoding an encoded video signal to generate a decoded video signal for use in a subsequent video display of an original video signal, the original video signal comprising a sequence of video frames, the encoded video signal having been generated from the original video signal, the method comprising the steps of:

25 identifying in said encoded video signal an indication that encoded virtual reference data has been incorporated therein, wherein said encoded virtual reference data does not represent any portion of any individual frame of the original video signal which is to be displayed in said subsequent video display thereof;

decoding from said encoded video signal said encoded virtual reference data
30 incorporated therein; and

J. McGowan 28-19

- 20 -

decoding from said encoded video signal one or more encoded portions of the original video signal which are to be displayed in said subsequent video display thereof, wherein said encoded portions of said original video signal which are to be displayed are based on and specifically reference a portion of said encoded virtual
5 reference data.

7. The method of claim 6 wherein said encoded video signal comprises a block-based encoding of said original video signal, wherein said encoded virtual reference data comprises one or more encoded virtual reference blocks, wherein said
10 encoded portions of said original video signal which are to be displayed are based on said encoded virtual reference data and have been encoded with use of a block-based motion compensated prediction scheme, and wherein the step of decoding from said encoded video signal said one or more encoded portions of the original video signal which are to be displayed comprises performing block-based motion compensation on
15 the decoded virtual reference data.

8. The method of claim 6 further comprising the step of storing said decoded virtual reference data in a reference frame buffer.

20 9. The method of claim 7 wherein said encoded virtual reference data and one encoded portion of said original video signal which is to be displayed have been incorporated into said encoded video signal together as an encoded dual block.

10. The method of claim 6 wherein the encoded video signal comprises a
25 block-based encoding of said original video signal, wherein one of said encoded portions of the original video signal which is to be displayed is comprised in a given frame of said original video signal, wherein said virtual reference data has been generated based on said given frame of said original video signal, and wherein said one of said encoded portions of the original video signal which is to be displayed has
30 been encoded based on said encoded virtual reference data with use of an intra-coding prediction scheme.

1/6

FIG. 1

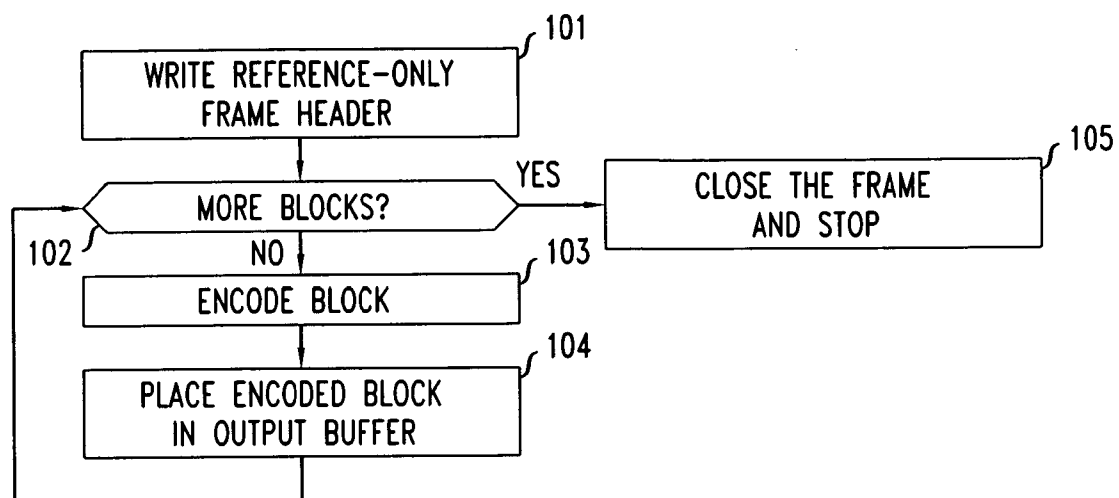
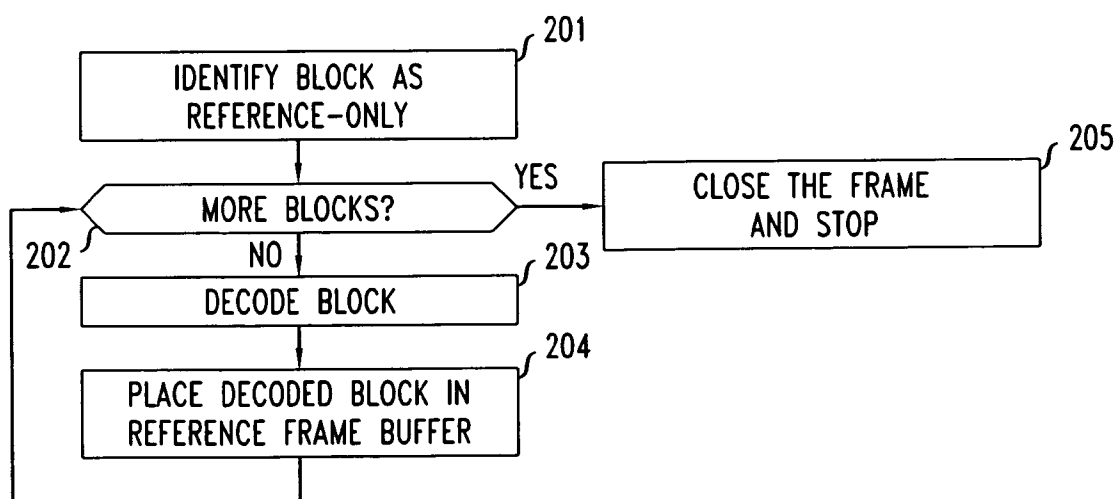


FIG. 2



2/6

FIG. 3

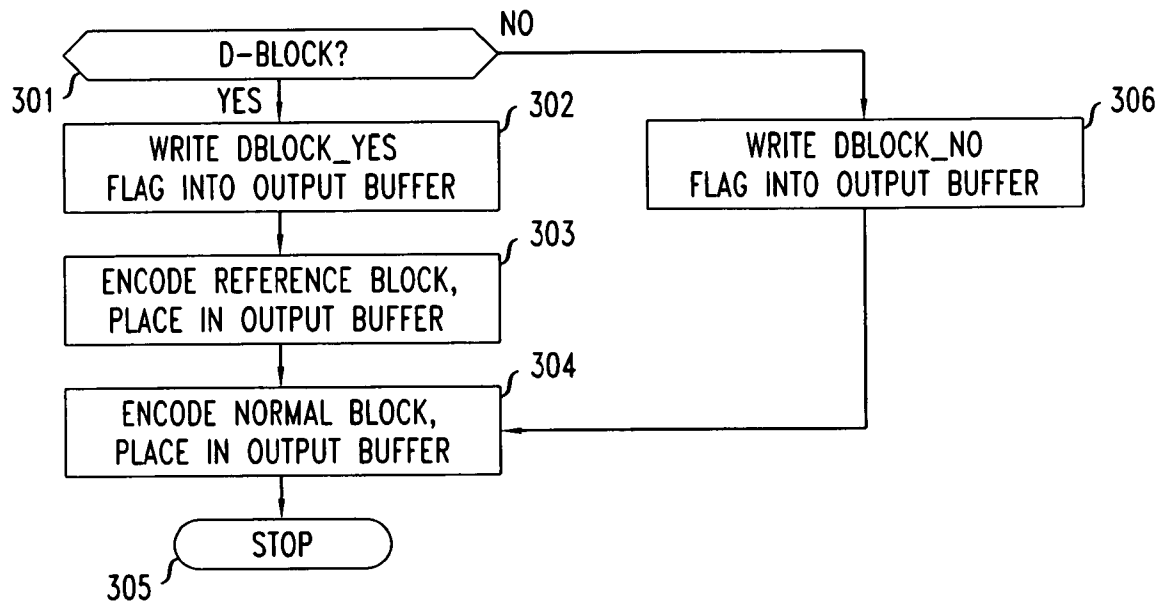
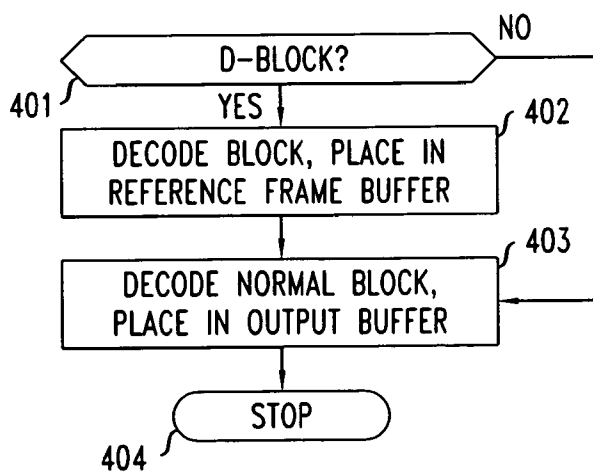
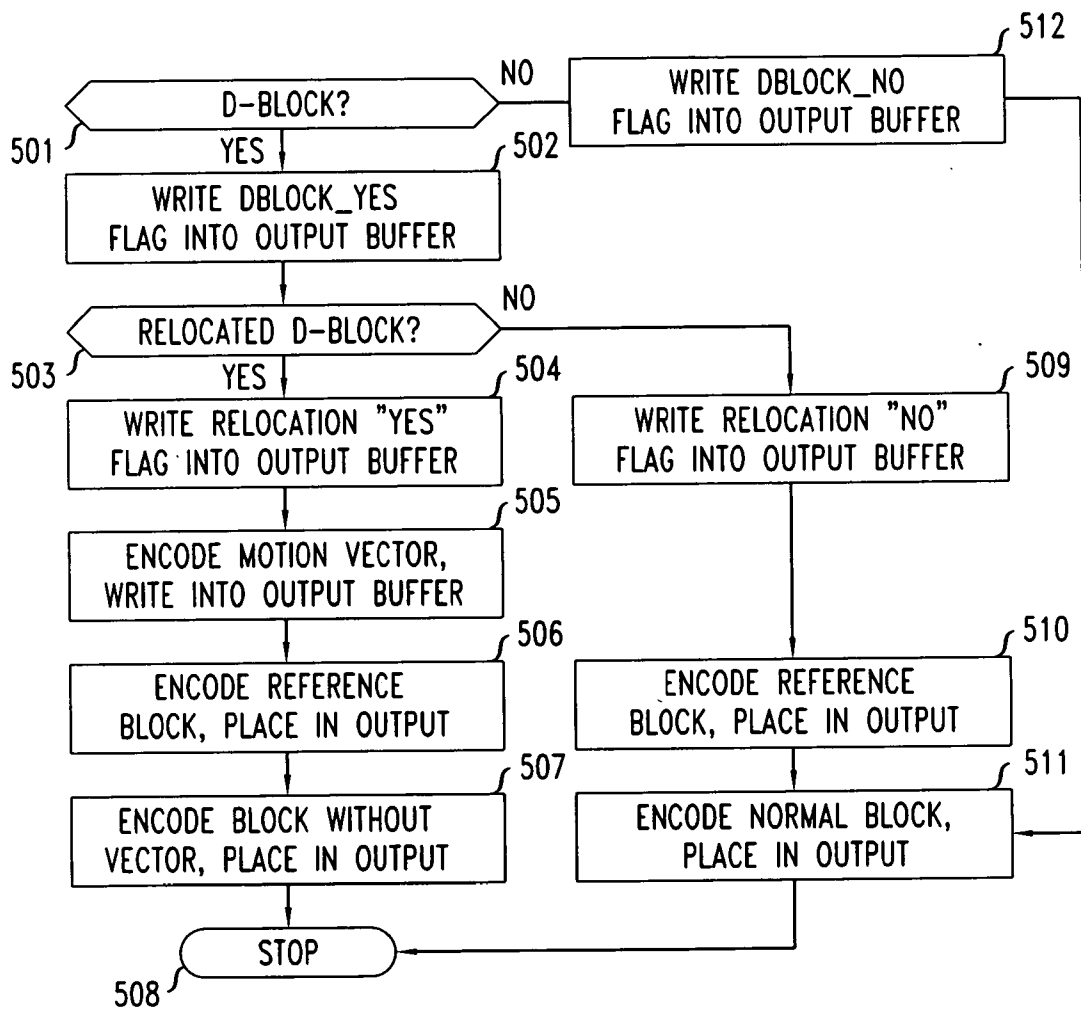


FIG. 4



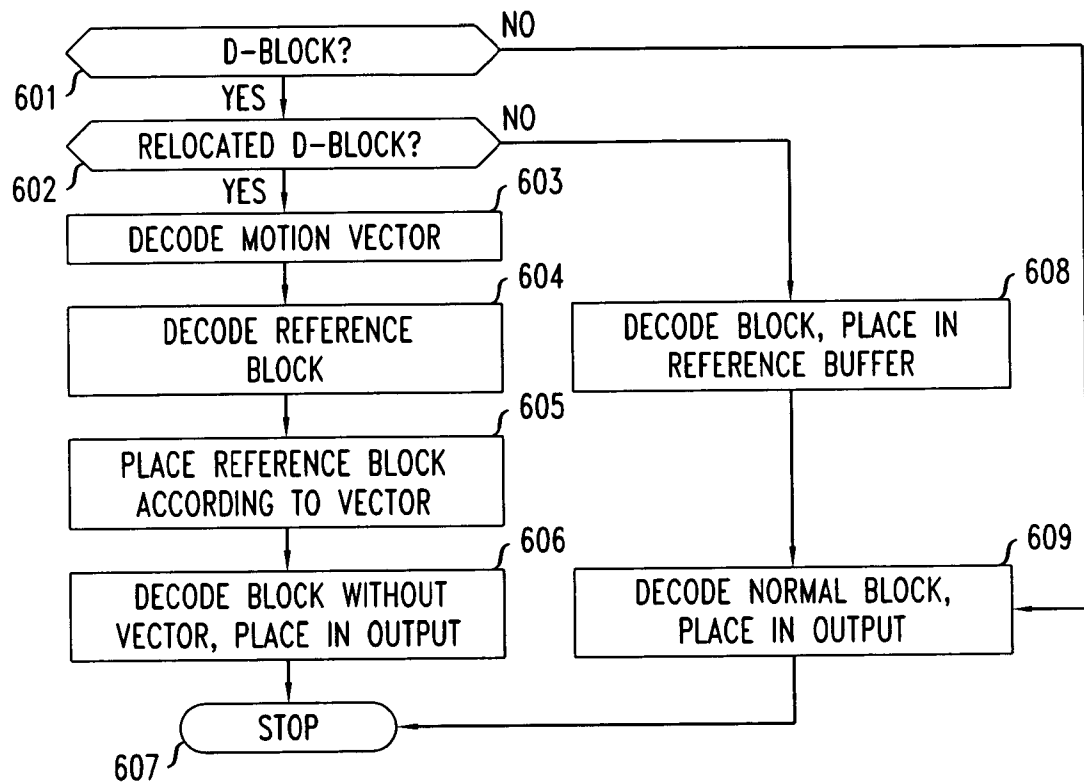
3/6

FIG. 5



4/6

FIG. 6



5/6

FIG. 7

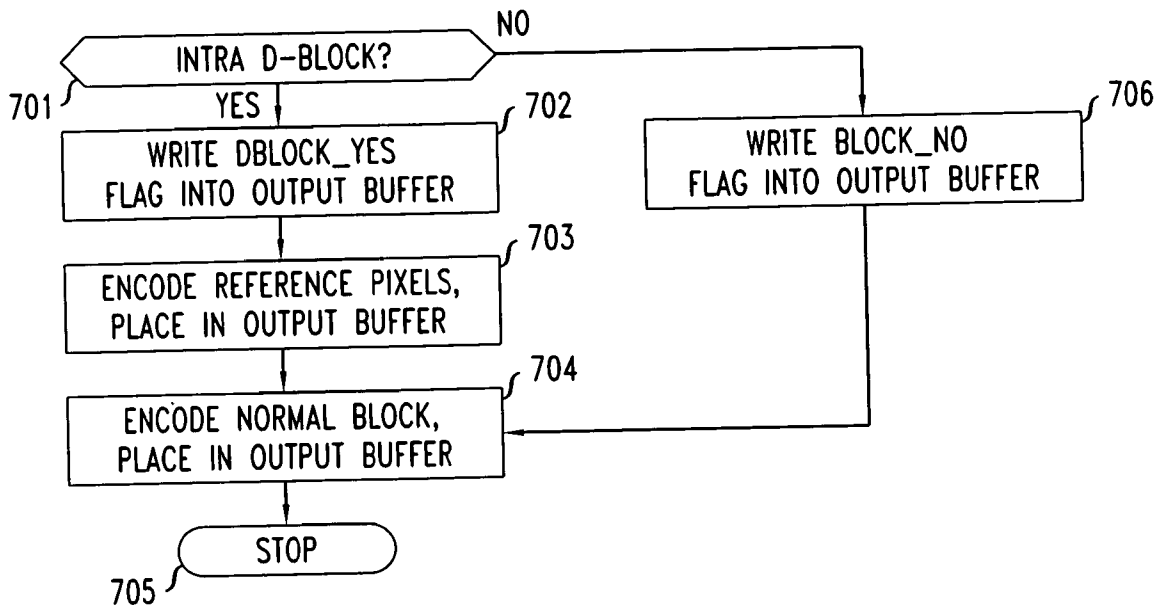
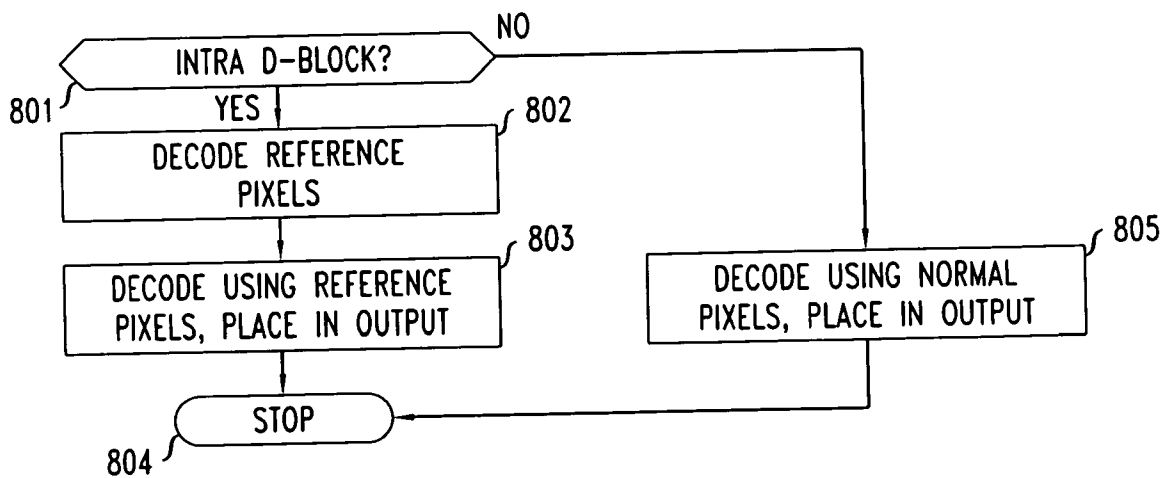
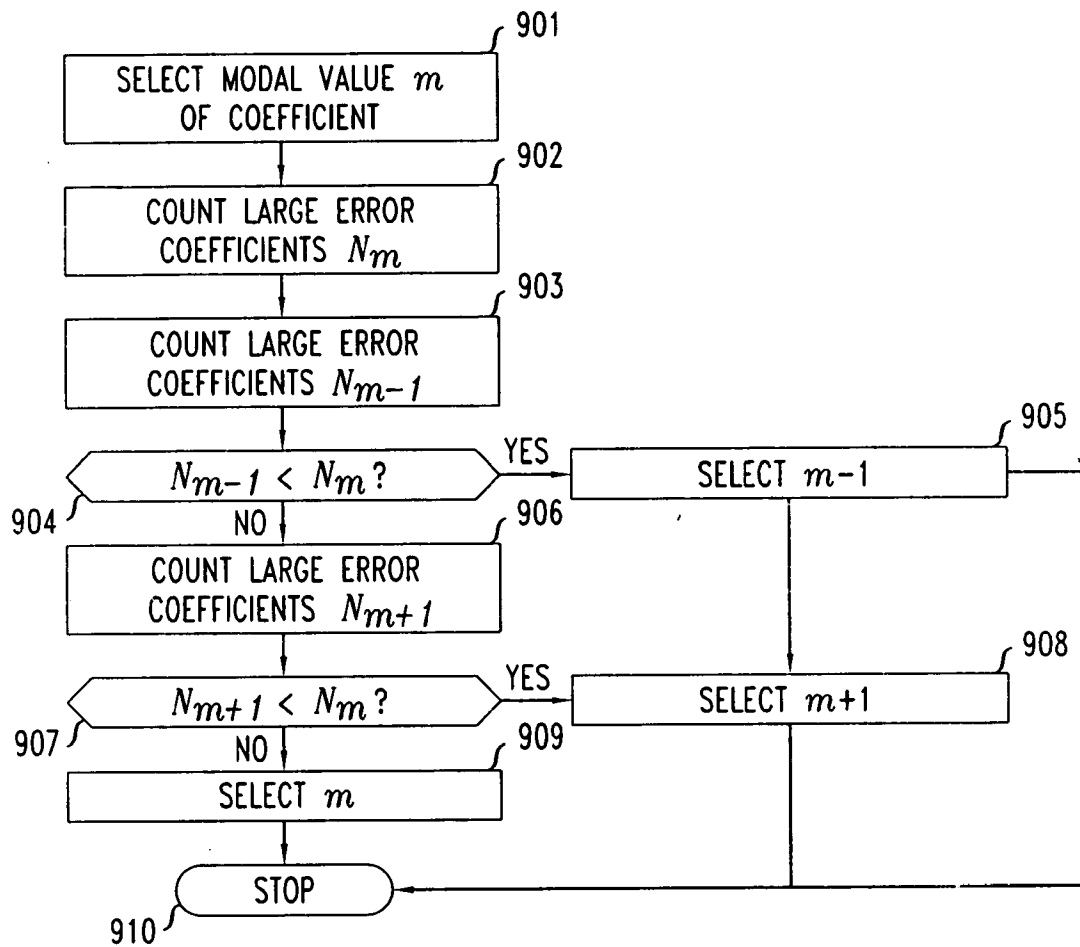


FIG. 8



6/6

FIG. 9



INTERNATIONAL SEARCH REPORT

International application No

PCT/US2009/063537

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N7/26 H04N7/34 H04N7/36 H04N7/50
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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/008331 A2 (THOMSON LICENSING [FR]; SU YEPING [US]; YIN PENG [US]; PANDIT PURVIN B) 17 January 2008 (2008-01-17) abstract page 1, line 13 - line 32 page 6, line 30 - page 8, line 26; figures 1,2	1-10
A	----- JIANPENG DONG AND NAM LING ET AL: "Adaptive Intra Prediction Padding to Improve Intra Motion Compensation" 25. PICTURE CODING SYMPOSIUM;24-4-2006 - 26-4-2006; BEIJING,, 24 April 2006 (2006-04-24), XP030080231 Sections 1 and 2 ----- -/--	1-10



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 docu-
ments, 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

1 April 2010

Date of mailing of the international search report

16/04/2010

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

Colesanti, Carlo

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2009/063537

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MATHIAS WIEN ET AL: "Integer Transforms for H.26L Using Adaptive Block Transforms" JOINT VIDEO TEAM (JVT) OF ISO/IEC MPEG & ITU-T VCEG(ISO/IEC JTC1/SC29/WG11 AND ITU-T SG16 Q6),, no. q15k24, 22 August 2000 (2000-08-22), XP030003117 Section "Rate-Distortion Optimization", page 4 -----	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2009/063537

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008008331 A2	17-01-2008	CN 101491099 A	22-07-2009
		EP 2041982 A2	01-04-2009
		JP 2009543521 T	03-12-2009
		KR 20090037888 A	16-04-2009