

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 April 2011 (21.04.2011)

(10) International Publication Number
WO 2011/046446 A1

(51) International Patent Classification:
H04N 7/50 (2006.01)

(21) International Application Number:
PCT/NO2010/000361

(22) International Filing Date:
13 October 2010 (13.10.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20093155 16 October 2009 (16.10.2009) NO

(71) Applicant (for all designated States except US): **TANDBERG TELECOM AS** [NO/NO]; Philip Pedersens vei 20, N-1366 Lysaker (NO).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **BJØNTEGAARD, Gisle** [NO/NO]; Haukeliveien 12B, N-1415 Oppegård (NO).

(74) Agents: **HOKSMO, Elin** et al.; **ONSAGERS AS, P.O.** Box 6963 St. Olavs plass, N-0130 Oslo (NO).

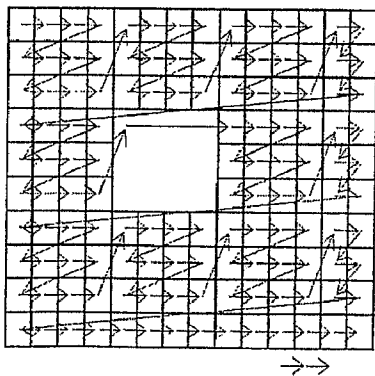
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, 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, 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: METHODS FOR VIDEO CODING AND DECODING



(57) Abstract: A method for scanning macroblocks in video compression and selecting alternatively sized large macroblocks accordingly. The scanning pattern is divided into scanning fragments with a size corresponding to different possible large macroblocks. This allows for varying the size of the selected macroblocks, based on minimizing distortion and/or bit rate consumption.

Figure 2

Methods for video coding and decoding

Technological Field

5

The exemplary embodiments described herein relate to the field of digital video compression, and more specifically variable macroblock sizes and the scanning order of macroblocks.

10

Background

Transmission of moving pictures in real-time is employed in several applications such as video conferencing, net meetings, TV broadcasting and video telephony.

15

However, representing moving pictures requires bulk information as digital video typically is described by representing each pixel in a picture with 8 bits (1 Byte). Such uncompressed video data results in large bit volumes, and cannot be transferred over conventional communication networks and transmission lines in real time due to limited bandwidth.

20

Thus, enabling real time video transmission requires a large extent of data compression. Data compression may, however, compromise the picture quality. Therefore, great efforts have been made to develop compression techniques allowing real time transmission of high quality video over bandwidth limited data connections.

25

30

Many video compression standards have been developed over the last 20 years. Many of those methods are standardized through ISO (the International Standards organization) or ITU (the International Telecommunications Union). In addition, a number of other proprietary methods have been developed. The main standardized methods are:

35

40 ITU: H.261, H.262, H.263, H.264

ISO: MPEG1, MPEG2, MPEG4/AVC)

In video compression systems, the main goal is to represent the video information with as little capacity as possible. Capacity is defined with bits, either as a constant value or as bits/time unit. In both cases, the main goal is to reduce the number of bits.

The first step in the coding process according to these standards is to divide the picture into square blocks of pixels, for instance 16x16. These blocks are typically denoted as macroblocks (or MBs for short). This is done for luminance information as well as for chrominance information. A scanning order of the MBs is established. A scanning order defines the encoding/decoding order of the MBs in a picture. A raster scan is typically used. This means that MBs are scanned as MB-lines from left to right and then the MB-lines from top to bottom. A raster scan order is illustrated in figure 1.

The following prediction process significantly reduces the amount of bits required for each picture in a video sequence to be transferred. It takes advantage of the similarity of parts of the sequence with other parts of the sequence, and produces a prediction for the pixels in the block. This may be based on pixels in an already coded/decoded picture (called inter prediction) or on already coded/decoded pixels in the same picture (intra prediction).

The prediction is mainly based on vectors representing movements. In a typical video sequence, the content of a present block M would be similar to a corresponding block in a previously decoded picture. If no changes have occurred since the previously decoded picture, the content of M would be equal to a block of the same location in the previously decoded picture. In other cases, an object in the picture may have been moved so that the content of M is more equal to a block of a different location in the previously decoded picture. Such movements are represented by motion vectors (V). As an example, a motion vector of (3;4) means that the content of M has moved 3 pixels to the left and 4 pixels

upwards since the previously decoded picture. For improved accuracy, the vector may also include decimals, requiring interpolation between the pixels.

5 Since the predictor part is known to both encoder and decoder, only the difference has to be transferred. This difference typically requires much less capacity for its representation. The difference between the pixels to be coded and the predicted pixels is often referred to as a residual.

10

The residual represented as a block of data (e.g. 4x3 pixels) still contains internal correlation. A well-known method of taking advantage of this is to perform a two dimensional block transform. In H.263 an 8x8 Discrete
15 Cosine Transform (DCT) is used, whereas H.264 uses a NxN (where N can be 4 or 8) integer type transform. This transforms NxN pixels into NxN transform coefficients and they can usually be represented by fewer bits than the pixel representation. Transform of a NxN array of pixels
20 with internal correlation will probably result in a 4x3 block of transform coefficients with much fewer non-zero values than the original 4x3 pixel block.

25

Direct representation of the transform coefficients is
25 still too costly for many applications. A quantization process is carried out for a further reduction of the data representation. Hence the transform coefficients undergo quantization. A simple version of quantisation is to divide parameter values by a number - resulting in a
30 smaller number that may be represented by fewer bits. This is the major tool for controlling the bit production and reconstructed picture quality. It should be mentioned that this quantization process has as a result that the reconstructed video sequence is somewhat different from
35 the uncompressed sequence. This phenomenon is referred to as "lossy coding". This means that the reconstructed pictures typically have lower quality than the original pictures. The output from the quantization process are integer numbers - which do not represent the original
40 transform coefficients correctly. These integers together with integers representing the side information

is coded in a lossless way and transmitted to the decoder.

5 Finally, a scanning within the MBs of the two dimensional transform coefficient data into a one dimensional set of data is performed, and the one dimensional set is further transformed according to an entropy coding scheme. Entropy coding implies lossless representation of the quantized transform coefficients. In depicting the
10 transform coefficients it is common to position the low frequency coefficient (or DC coefficient) in the upper left corner. Then the horizontal and vertical spatial frequency increase to the right and down. The scanning usually starts with the coefficient in the left upper
15 corner and follows a zig-zag pattern around the diagonal direction towards the lower right corner of the MB, but in other cases the entropy coding may be more efficient if "inverse scanning" (high to low frequency) is used.

20 The above steps are listed in a natural order for the encoder. The decoder will to some extent perform the operations in the opposite order and do "inverse" operations as inverse transform instead of transform and de-quantization instead of quantization.

25 In connection with the introduction of video formats of higher resolution in video conferencing, an increasing number of pixels will represent a picture of the same physical segment. Hence, an image section will contain an
30 increasing number of 16x16 MB. The probability of many adjacent MBs having the same characteristics, e.g. motion vectors and zero-transforms, is therefore increasing, and consequently also redundant data representation.

35 Summary of the invention

The invention has been set forth in the appended claims.

40 An exemplary embodiment described herein provides a method of video coding executed by an encoder apparatus, the method comprising:

arranging, at the encoder apparatus, residual pixels into macroblocks, the macroblocks comprising regular macroblocks, each having a size $p \times p$, and at least one of
5 first large macroblocks, each having a size $(k \times p) \times (l \times p)$, and

second large macroblocks, each having a size $(m \times p) \times (n \times p)$,

10

the arranging being determined by a predefined criterion for minimizing at least one of rate consumption or rate distortion, and wherein
p indicates a spatial size of a regular macroblock,
15 k, l, m and n are integer numbers,
m is divisible by k and
n is divisible by l;

20

scanning, at the encoder apparatus, said regular macroblocks in a first zig-zag scanning pattern within first rectangular picture fragments corresponding to the size of the first large macroblock or the second large macroblock; and

25

indicating, at the encoder apparatus, occurrences of regular macroblocks and at least one of first large macroblocks and second large macroblocks

30

by a block type code arranged in an order of which the regular macroblocks and at least one of the first large macroblocks and the second large macroblocks are scanned.

A corresponding inverse method for decoding is also described herein.

35

Furthermore, the present specification describes exemplary embodiments of an encoding apparatus, a decoding apparatus, and computer programs, implementing the methods.

40

Brief description of the drawings

In order to make the invention more readily understandable, the discussion that follows will refer to the accompanying drawings.

Figure 1 indicates the scanning pattern of Macro Blocks according to in a conventional way of video coding,

Figure 2 shows an example of a scanning pattern of macroblocks according to a first embodiment,

Figure 3 shows an example of a scanning pattern of macroblocks according to a second embodiment; and

Figure 4 shows an example of a computer system that is programmed as an encoder/decoder.

Detailed description of the present invention

Exemplary embodiments described herein provide a scanning order and a scanning scheme of macroblocks (or MBs for short) which is adjusted for coding of pictures of high resolution. Scanning in this context means the order of which the MBs within a video picture are coded or decoded. Raster scan is the conventional way of scanning MBs in video coding. As can be seen from figure 1, this means that MBs are scanned as MB-lines from left to right and then the MB-lines from top to bottom.

Figure 2 illustrates an example of a scanning pattern according to a first embodiment of the present invention. The scan starts with the upper left MB and proceeds horizontally towards the right. However, the scan does not proceed all the way to the end of the row. After the fourth MB in the upper line, the scan turns down to the leftmost MB in the row of MBs below, and proceeds towards the right. This is repeated in a zig-zag pattern within a 4x3 dimension rectangle of MB blocks. Such rectangles of MB blocks in which a zig-zag scanning pattern appears, may also be denoted as scanning fragments. After finishing the scan of these MBs, the scanning is performed equivalently on the right hand adjacent 4x3

dimension rectangle of MB blocks. When reaching the right picture border, the scanning proceeds to the leftmost 4x3 dimension rectangle of MB blocks below the already scanned blocks.

5 At the right or bottom edge of the picture, there may not be room for a full 4x3 fragment. In this case, the size of the fragments covering the remaining picture is adjusted accordingly as indicated in figure 2.

10 The scanning pattern according to the above description allows for the introduction of larger blocks of alternative sizes. For the scanning pattern in the example described above, a large macroblock (or LMB for short) with the size corresponding to the size of 4x3
15 regular macroblocks can be inserted in the picture without compromising with the overall scanning pattern. The LMBs can be positioned as a replacement of 4x3 MBs in what otherwise would have been a 4x3 dimension scanning fragment as described above. In the case of each MB
20 having a size of 16x16 pixels, the LMB would then have a size of 64x48 pixels. Smaller LMBs must be used at the right and bottom edge according to the size of the above-mentioned fragments covering the remaining picture is adjusted accordingly as indicated in figure 2.

25 The dimensions in the description of the first embodiment are of course only examples. A variety of other dimensions of MBs and LMBs are conceivable, but the size of the LMBs must correspond to the size of an integer
30 number MBs, the scanning pattern and hence the size of the scanning fragment. Consequently, other scanning patterns could be used. Generally, an LMB can have the size of $(m \cdot p) \times (n \cdot p)$ pixels, n and m being integer numbers, and p the pixel dimension of a regular
35 macroblock.

Figure 3 illustrates an example of an alternative scanning pattern according to a second embodiment of the present invention. The scan starts with the upper left MB
40 and proceeds towards the right. However, when reaching the second MB in the upper row, the scan turns down to the leftmost MB in the row of MBs below, and proceeds

towards the right to the second MB in the second row of MBs. As can be seen in figure 3, at this point, a zig-zag scan has been performed within a small fragment of 2x2 MBs. The scan then jumps up to the third MB in the upper
5 row and performs a corresponding zig-zag scan within the next small fragment of 2x2 MBs on the right hand side. After scanning these two small fragments the scanning jumps down to the leftmost MB on the third row, and perform a corresponding scanning on the two small
10 fragments just below, ending in the fourth MB in the fourth row. As can be seen, a group of blocks corresponding to the 4x4 scanning fragment as described in the first embodiment has been scanned, but in a different scanning pattern forming four small scanning
15 patterns each containing 2x2 MBs.

The second embodiment of the invention allows consequently for two alternative LMBs, a first LMB corresponding to the size of the small fragment (2x2
20 MBs), and a second LMB corresponding to a large fragment (4x4 MBs). In the case of each MB having a size of 16x16 pixels, the first LMB would then have a size of 32x32 pixels.

25 The dimensions in the description of the second embodiment above are of course only examples. A variety of other dimensions of MBs and first and second LMBs are conceivable, but the size of the LMBs must correspond to the size of an integer number MBs, the scanning pattern
30 and hence the size of the scanning fragment. The size of a second LMB must also correspond to the size of an integer number first LMBs. Consequently, other scanning patterns could be used. Generally, a second LMB can have the size of $(m \cdot p) \times (n \cdot p)$ pixels, n and m being integer
35 numbers, and p the pixel dimension of an MB. A first LMB can have the size of $(k \cdot p) \times (l \cdot p)$ pixels, k and l being integer numbers. The fact that the size of a second LMB must also correspond to the size of an integer number first LMBs implies further that m is divisible by k , and
40 n is divisible by l , without remainder.

In the example above, $k=2$, $l=2$, $m=4$, $n=4$ and $p=16$.

Further exemplary values for the size parameters are given below. It will readily be recognized that other possibilities also exist.

5 The size parameter k may be selected from the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$;

The size parameter l may be selected from the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$;

10

The size parameter m equals a positive integer multiplied by k , and in particular, a positive integer ≥ 2 multiplied by k . Thus, if $k=2$, m may be suitably selected from the set $\{4, 6, 8, 10, 12, 14, 18, 18, 20\}$. If $k=3$, m 15 may be suitably selected from the set $\{6, 9, 12, 15, 18, 21, 24, 27, 30\}$, and so on.

The size parameter n equals a positive integer multiplied by l , and in particular, a positive integer ≥ 2 20 multiplied by l . Thus, if $l=2$, n may be suitably selected from the set $\{4, 6, 8, 10, 12, 14, 18, 18, 20\}$. If $l=3$, n may be suitably selected from the set $\{6, 9, 12, 15, 18, 21, 24, 27, 30\}$, and so on.

25 The size parameter p may, e.g., be selected from the set $\{4, 8, 12, 16, 20, 24, 28, 32\}$

Regular macroblocks are, in general, square with size $p \times p$.

30

Large macroblocks of the first or second type, as described in the present disclosure, may also be square, but they are not necessarily square.

35 It will readily be understood that k equals 1 if the first large macroblock is square. Such an example has been explained with reference to figure 3, wherein $k=l=2$. If the first large macroblock is rectangular, but not square, k will be different from 1. Such an example has 40 been explained with reference to figure 2, wherein $k=4$ and $l=3$.

Likewise it will readily be understood that m equals n if the second large macroblock is square. Such an example has been described with reference to figure 3, wherein
5 $m=n=4$. If the second large macroblock is rectangular, but not square, m will be different from n .

Note the distinction between the following typographical symbols used in the present specification and claims: the
10 letter "1", used as a numerical (integer) size parameter for large macroblocks, and the fixed number (one) "1".

There is no limitation of the number of LMBs of different sizes, even if there are only two given in the
15 description above with reference to figure 3. In the embodiments described above, there are as many scanning fragment levels as there are LMB variants, and the scanning fragments have the same size as the respective LMB variants. A scanning fragment of a lower level must
20 further have a known scanning pattern within the next scanning fragment level, and the dimensions of a scanning fragment must be divisible by the dimensions of the a lower level scanning fragment. For example, in the case
25 of the second embodiment described above, there is a first scanning fragment level of 2×2 MBs, and a second scanning fragment level of 4×4 MBs. The scanning fragment level of 4×4 MBs is of a higher level than the scanning fragment level of 2×2 MBs, which has a scanning pattern within the fragment level of 4×4 MBs, and the dimensions
30 are divisible by each other. There could also accordingly be a third and a fourth scanning fragment following the same logic.

There are many ways of deciding if coding shall be based
35 on 16×16 MBs or a LMB of some size. One frequently used method is to make the decision based on a rate distortion measure. The rate distortion (or RD for short) measure over a block is typically:

40 $\text{Rate Distortion} = \text{Distortion} + \text{Rate} = D + R$
 $= \text{sum}(\text{pixel_errors}^2) + f(\text{bits_to_code_the_block})$.

Distortion (D) is most often calculated as the square_sum (indicated by pixel_errors²) of the reconstructed pixel errors. Rate (R) depends on the number of bits used to code the block.

5 Moreover, the reason that there are reconstructed pixel errors is that transform coefficients are not represented exactly but have been quantized. The quantization interval may be defined by q. A small q results in a small D, and a large number of bits. On the other hand, a large q results in a large D and a small number of bits. From this the function f() is of the form:

15 $f(\text{bits_to_code_the_block}) = k \cdot q \cdot \text{bits_to_code_the_block}$, where k is a constant not very different from 1, e.g. $k \in [0.5, 1.5]$, or $k \in [0.8, 1.2]$.

For the pixel area of a size corresponding to a LMB, RD_{LMB} and RD_{MB} are calculated. If $RD_{MB} < RD_{LMB}$ the coding is based on 16x16 MBs. Otherwise, the coding is based on LMB. Similarly, if there are more block sizes (more LMB sizes) the RD values are calculated and compared to arrive at the overall solution with the lowest Rate_Distortion measure (RD)."

25 The description above is for a full RD decision. This is one of many methods. In some situations the described procedure may be found too implementation costly. In such cases the distortion as described above may be replaced by some other measure of distortion - either of the reconstructed pixels or prediction errors. The last may typically come from motion estimation where a Sum of Absolute Differences (SAD) between pixels to be coded and already decoded pixels are calculated. Such a distortion measure alone or together with a number of bits used for coding may be used instead of a full RD measure as described above. The decision may also be based on minimizing the number of bits used (bit rate consumption) only. However, the decision methods are all referred to as rate distortion measure in this application.

40 Occurrences of MBs and different types of LMBs are

signalled in a block type code as bits in the bit stream between encoder and decoder. The combination of the order of the bit codes sent, the actual scanning order and the run code will uniquely indicate positions of MBs and different LMBs. Run coding is explained further below. Normally, coding will have a result that for most of the MBs there is no need to send any information. This is mainly because there has been no change in the MB since the last picture. An MB that does not need to transmit any information is often called a skipped block. The decoder will already know how to reconstruct the pixels in a skipped or "not coded" MB. For other MBs there will be bits transmitted in order to reconstruct the pixels on the decoder. Such a MB is referred to as "coded".

This is a situation where "run coding" is used in many areas. Instead of indicating whether each MB is "coded" or "not coded", it is more efficient to use a code only for the "coded" MBs. This code contains information of how many "not coded" MBs there were since the last "coded" MB. Hence the "run" from the last coded MB is coded. In combination with the block type code indicating MB or a LMB variant, the further scan pattern will be given, and the run code will have that pattern as an assumption.

Introducing LMBs provides the possibility of performing a transform operation on larger pixel areas when appropriate for example by transforming all the pixels of one luminance/chrominance component in an LMB with a size and dimension of the transform matrix corresponding to the size and dimension of the LMB. Likewise, a common motion vector or set of motion vectors can represent larger pixel areas, i.e. a whole LMB. The transform operation and motion vector estimate is preferably performed according to the standards, but then on a block size larger than conventionally.

All the transform coefficients resulting in the LMB transform may be quantized and entropy coded, but more realistically, a subset of for instance the 8x8 lowest frequency coefficients could be used. In this case, a

method of quantizing and coding 8x8 transform coefficients otherwise used for 8x8 block transform can be reused.

5 The above description is directed to encoding, but the corresponding applies also for decoding, but in an inverse manner.

10 A decoder will receive the indications of regular MB and LMB in a separate code arranged in the order of which they occur in the scanning pattern logics used at the encoder side. The actual positions and occurrences of MBs and LMBs in the picture can then be derived from the run code which states the run (distance) between non-skipped blocks, the block type indications, the received order of
15 which, and the zig-zag scanning pattern which in turn is given from the block type indications in combination with knowing the scanning pattern logics on the encoder side.

20 Fig. 4 illustrates an example of video conferencing encoder/decoder apparatus that is configured to implement the encoding and/or decoding embodiments described supra. The algorithms described supra may be written in computer executable instructions, which may be executed by the machine shown in Fig. 4. The encoder/decoder apparatus
25 may be embodied in a video conference endpoint or may be a stand alone device.

The computer system 1201 includes a bus 1202 or other communication mechanism for communicating information,
30 and a processor 1203 coupled with the bus 1202 for processing the information. The computer system 1201 also includes a main memory 1204, such as a random access memory (RAM) or other dynamic storage device (e.g., dynamic RAM (DRAM), static RAM (SRAM), and synchronous
35 DRAM (SDRAM)), coupled to the bus 1202 for storing information and instructions to be executed by processor 1203. In addition, the main memory 1204 may be used for

storing temporary variables or other intermediate information during the execution of instructions by the processor 1203. The computer system 1201 further includes a read only memory (ROM) 1205 or other static storage device (e.g., programmable ROM (PROM), erasable PROM (EPROM), and electrically erasable PROM (EEPROM)) coupled to the bus 1202 for storing static information and instructions for the processor 1203.

The computer system 1201 also includes a disk controller 1206 coupled to the bus 1202 to control one or more storage devices for storing information and instructions, such as a magnetic hard disk 1207, and a removable media drive 1208 (e.g., floppy disk drive, read-only compact disc drive, read/write compact disc drive, compact disc jukebox, tape drive, and removable magneto-optical drive). The storage devices may be added to the computer system 1201 using an appropriate device interface (e.g., small computer system interface (SCSI), integrated device electronics (IDE), enhanced-IDE (E-IDE), direct memory access (DMA), or ultra-DMA).

The computer system 1201 may also include special purpose logic devices (e.g., application specific integrated circuits (ASICs)) or configurable logic devices (e.g., simple programmable logic devices (SPLDs), complex programmable logic devices (CPLDs), and field programmable gate arrays (FPGAs)).

The computer system 1201 may also include a display controller 1209 coupled to the bus 1202 to control a display 1210, for displaying information and/or video to a user. The computer system includes input devices, such as a keyboard 1211 and a pointing device 1212, for interacting with a computer user and providing

information to the processor 1203. The pointing device 1212, for example, may be a mouse, a trackball, or a pointing stick for communicating direction information and command selections to the processor 1203 and for
5 controlling cursor movement on the display 1210. In addition, a printer may provide printed listings of data stored and/or generated by the computer system 1201.

The computer system 1201 performs a portion or all of the
10 processing steps of the exemplary embodiments discussed supra in response to the processor 1203 executing one or more sequences of one or more instructions contained in a memory, such as the main memory 1204. Such instructions may be read into the main memory 1204 from another
15 computer readable medium, such as a hard disk 1207 or a removable media drive 1208. One or more processors in a multi-processing arrangement may also be employed to execute the sequences of instructions contained in main memory 1204.

20 The computer system 1201 includes at least one computer readable storage medium or memory for holding instructions programmed according to the teachings of the invention and for containing data structures, tables,
25 records, or other data described herein. Examples of computer readable media are compact discs, hard disks, floppy disks, tape, magneto-optical disks, PROMs (EPROM, EEPROM, flash EPROM), DRAM, SRAM, SDRAM, or any other magnetic medium, compact discs (e.g., CD-ROM), or any
30 other optical medium, punch cards, paper tape, or other physical medium with patterns of holes.

The processing instructions, or a computer program, programmed according to the teachings of any of the

methods described herein, may be tangibly embodied on a computer readable storage medium, or in a memory.

Alternatively, the instructions or the program may be embodied on a propagated signal, which may be

5 transmitted, forwarded or received by any network element communicatively operating in a computer or communications network, such as computers, servers, terminals, clients, video conference endpoints, multipoint control units, routers, bridges, hubs, switches, or gateways.

10

Stored on any one or on a combination of computer readable storage media, the exemplary embodiments include software for controlling the computer system 1201, for driving a device or devices for implementing the

15 invention, and for enabling the computer system 1201 to interact with a human user (e.g., video conference participant). Such software may include, but is not limited to, device drivers, operating systems,

development tools, and applications software. Such

20 computer readable media further includes the computer readable storage medium of the present invention for performing all or a portion (if processing is distributed) of the processing performed in implementing the invention.

25

The computer code devices may be any interpretable or executable code mechanism, including but not limited to scripts, interpretable programs, dynamic link libraries (DLLs), Java classes, and complete executable programs.

30 Moreover, parts of the processing of the present invention may be distributed for better performance, reliability, and/or cost.

The computer system 1201 also includes a communication interface 1213 coupled to the bus 1202. The communication interface 1213 provides a two-way data communication coupling to a network link 1214 that is
5 connected to, for example, a local area network (LAN) 1215, or to another communications network 1216 such as the Internet. For example, the communication interface 1213 may be a network interface card to attach to any packet switched LAN. As another example, the
10 communication interface 1213 may be an asymmetrical digital subscriber line (ADSL) card, an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of communications line. Wireless
15 links may also be implemented. In any such implementation, the communication interface 1213 sends and receives electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

20
The network link 1214 typically provides data communication through one or more networks to other data devices. For example, the network link 1214 may provide a connection to another computer through a local network
25 1215 (e.g., a LAN) or through equipment operated by a service provider, which provides communication services through a communications network 1216. The local network 1214 and the communications network 1216 use, for example, electrical, electromagnetic, or optical signals
30 that carry digital data streams, and the associated physical layer (e.g., CAT 5 cable, coaxial cable, optical fiber, etc). The signals through the various networks and the signals on the network link 1214 and through the communication interface 1213, which carry the digital

data to and from the computer system 1201 maybe implemented in baseband signals, or carrier wave based signals. The baseband signals convey the digital data as unmodulated electrical pulses that are descriptive of a stream of digital data bits, where the term "bits" is to be construed broadly to mean symbol, where each symbol conveys at least one or more information bits. The digital data may also be used to modulate a carrier wave, such as with amplitude, phase and/or frequency shift keyed signals that are propagated over a conductive media, or transmitted as electromagnetic waves through a propagation medium. Thus, the digital data may be sent as unmodulated baseband data through a "wired" communication channel and/or sent within a predetermined frequency band, different than baseband, by modulating a carrier wave. The computer system 1201 can transmit and receive data, including program code, through the network(s) 1215 and 1216, the network link 1214 and the communication interface 1213. Moreover, the network link 1214 may provide a connection through a LAN 1215 to a mobile device 1217 such as a personal digital assistant (PDA) laptop computer, or cellular telephone.

The present invention is useful especially with large picture formats and result in both lower bit usage and better decoding fidelity than conventional methods using MB size of 16x16 pixels only.

Numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.

CLAIMS

1. A method of video coding executed by an encoder apparatus, the method comprising:

5 arranging, at the encoder apparatus, residual pixels into macroblocks, the macroblocks comprising regular macroblocks, each having a size $p \times p$, and at least one of

10 first large macroblocks, each having a size $(k \times p) \times (l \times p)$, and

second large macroblocks, each having a size $(m \times p) \times (n \times p)$,

15 the arranging being determined by a predefined criterion for minimizing at least one of rate consumption or rate distortion, and wherein

p indicates a spatial size of a regular macroblock,
20 k , l , m and n are integer numbers,
 m is divisible by k and
 n is divisible by l ;

scanning, at the encoder apparatus, said regular macroblocks in a first zig-zag scanning pattern within
25 first rectangular picture fragments corresponding to the size of the first large macroblock or the second large macroblock; and

30 indicating, at the encoder apparatus, occurrences of regular macroblocks and at least one of first large macroblocks and second large macroblocks
by a block type code arranged in an order of which the regular macroblocks and at least one of the first large
35 macroblocks and the second large macroblocks are scanned.

2. The method according to claim 1, wherein
40 if the first rectangular picture fragments correspond to the size of the first large macroblock, then a scanning order of the first rectangular picture fragments corresponds to a second zig-zag scanning

pattern within second rectangular picture fragments corresponding to the size of the second large macroblock.

5 3. The method according to claim 1 or 2, further comprising:

calculating a rate distortion as a square sum of reconstructed pixel errors and a rate depending on the number of bits used to code a block, and

10

determining, based on the rate distortion, whether coding will be based on the regular macroblock, the first large macroblock, or the second large macroblock.

15

4. The method according to claim one of the claims 1-3, further comprising:

20

transforming a first large macroblock by a $(k \times p) \times (l \times p)$ transform matrix, resulting in a corresponding first large macroblock of transform coefficients; and

25

transforming a second large macroblock by a $(m \times p) \times (n \times p)$ transform matrix, resulting in a corresponding second large macroblock of transform coefficients.

30

5. The method according to one of the claims 1-4, wherein the first or second large macroblock is represented by a set of motion vectors.

35

6. The method according to one of the claims 1-5, further comprising:

quantizing only $i \times i$ upper left transform coefficients of a large macroblock of transform coefficients; and

40

setting not-quantized transform coefficients of the large macroblock of transform coefficients to zero.

7. The method according to claim 2,
further comprising:

5 starting the second zig-zag scanning pattern in a
first rectangular picture fragment in an upper left
corner of one certain second rectangular picture fragment
and ending the scanning pattern in a first rectangular
picture fragment in a lowest right corner of the one
certain second rectangular picture fragment.

10

8. The method according to one of claims 1 to 6,
further comprising:

15 if the first rectangular picture fragments
correspond to the size of the second large macroblock,

starting the first zig-zag pattern in a leftmost and
first regular macroblock in an upper and first row of
regular macroblocks in a rectangular picture fragment,

20

heading horizontally towards the right along the
upper and first row of regular macroblocks to an m'th
regular macroblock of the upper and first row of regular
macroblocks,

25

turning to a leftmost and first regular macroblock
in a second row of regular macroblocks proceeding to the
m'th regular macroblock of the second row of regular
macroblocks, and

30

continuing accordingly until the m'th regular
macroblock in the n'th row of regular macroblocks is
reached.

35

9. The method according to one of the claims 1 to
6, further comprising:

if the first rectangular picture fragments correspond to
the size of the first large macroblock,

40

starting the first zig-zag pattern in a leftmost and
first macroblock in an upper and first row of regular

macroblocks in a rectangular picture fragment,

heading horizontally towards the right along the upper and first row of regular macroblocks to a k 'th regular macroblock of the upper and first row of regular macroblocks,

turning to a leftmost and first regular macroblock in a second row of regular macroblocks proceeding to the k 'th regular macroblock of the second row of regular macroblocks, and

continuing accordingly until the k 'th regular macroblock in the l 'th row of regular macroblocks is reached.

10. The method according to one of the claims 1 to 9, wherein p equals 16, m and n equals 4, and k and l equals 2.

11. A method of video decoding, executed by a decoding apparatus, for transforming transform coefficients in a coded video picture to residual pixels, the method comprising:

decoding, at the decoding apparatus, a run code resulting in non-skipped block runs;

decoding, at the decoding apparatus, a block type code resulting in block type indications of macroblocks of transform coefficients, the macroblocks comprising regular macroblocks having size pxp , and at least one of

first large macroblocks having size $(k*p)x(l*p)$, and

second large macroblocks having size $(m*p)x(n*p)$,

wherein p indicates the spatial size of a regular macroblock,

k , l , m and n are integer numbers,

m is divisible by k and

5 n is divisible by l ;

deriving, at the decoding apparatus, a zig-zag scanning pattern from the type block indications; and

10

determining, at the decoding apparatus, occurrences and positions of the regular macroblock and at least one of the first large macroblocks and the second large macroblocks in the coded video picture from the block type indications.

15

12. The method according to claim 11, further comprising:

20

transforming a regular macroblock by a $p \times p$ transform matrix resulting in a corresponding regular macroblock of residual pixels;

25
30

transforming a first large macroblock by a $(k \times p) \times (l \times p)$ transform matrix resulting in a corresponding first large macroblock of residual pixels; and

transforming a second large macroblock by a $(m \times p) \times (n \times p)$ transform matrix resulting in a corresponding second large macroblock of residual pixels.

30

13. A computer program, tangibly embodied on a computer readable storage medium, or in a memory, or on a propagated signal, the computer program comprising processing instructions which, when executed by a processing device in an encoding apparatus, causes the encoding apparatus to perform a method as set forth in one of the claims 1-10.

35

14. A computer program, tangibly embodied on a computer readable storage medium, or in a memory, or on a propagated signal, the computer program comprising processing instructions which, when executed by a

40

processing device in a decoding apparatus, causes the decoding apparatus to perform a method as set forth in one of the claims 11-12.

5 15. A video encoding apparatus, comprising a processor, input/output circuits and a memory, the memory containing a computer program as set forth in claim 13.

10 16. A video decoding apparatus, comprising a processor, input/output circuits and a memory, the memory containing a computer program as set forth in claim 14.

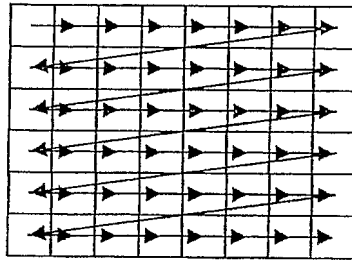


Figure 1

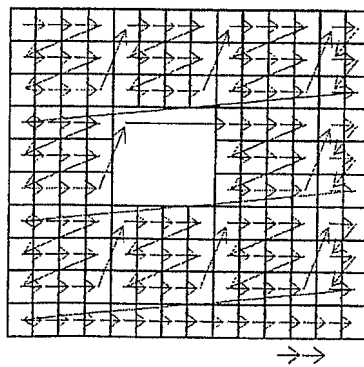


Figure 2

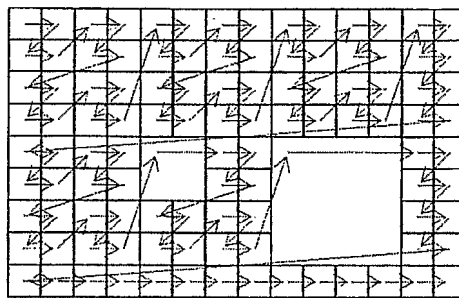


Figure 3

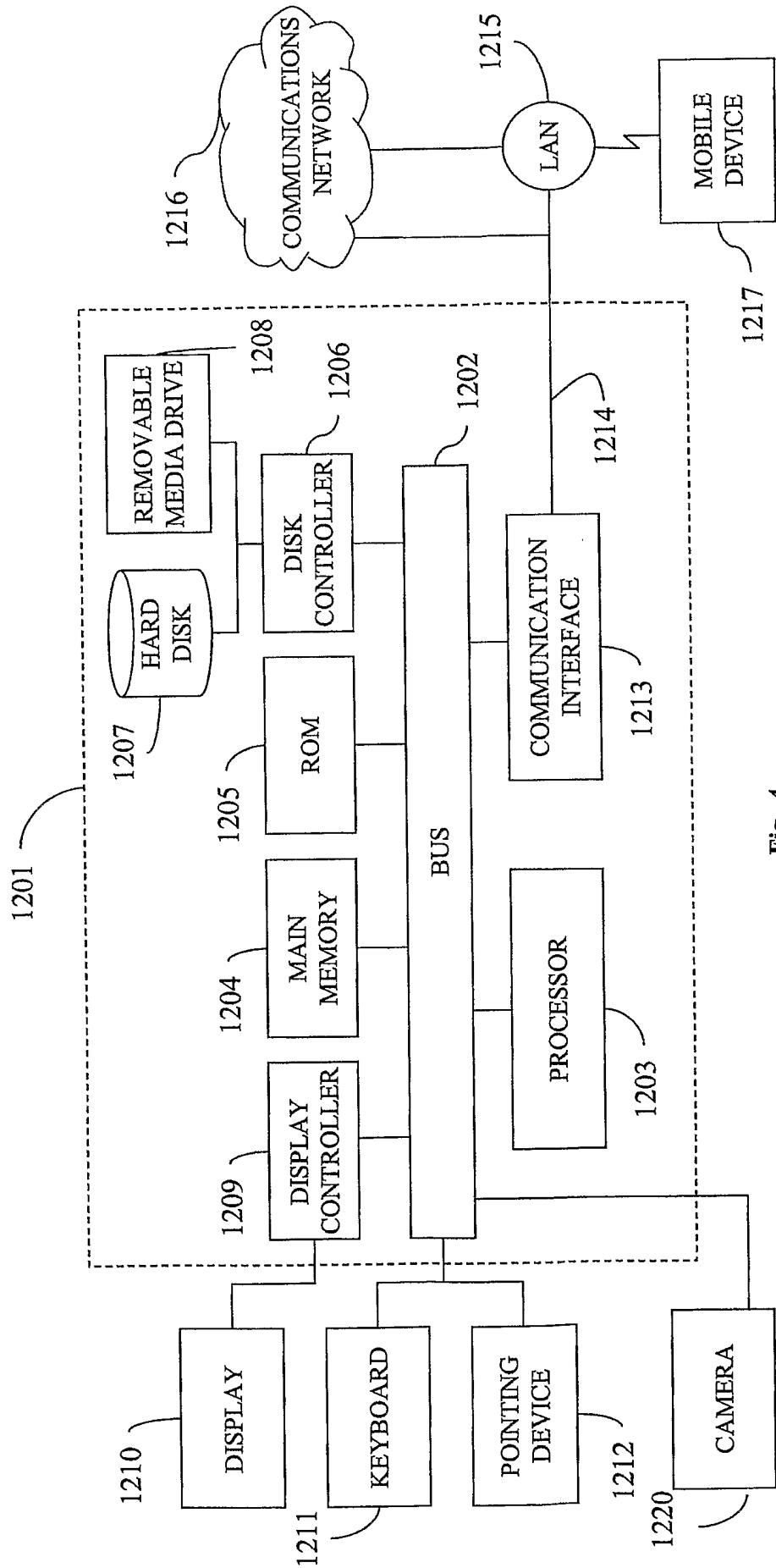


Fig. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/NO2010/000361

A. CLASSIFICATION OF SUBJECT MATTER

IPC: H04N 7/50 (2006.01)

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)

IPC H04N

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

SE, DK, NO, FI

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

WPI, EPODOC, NPL, XPETSI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/027192 A2 (THOMSON LICENSING [FR]; DIVORRA ESCODA OSCAR [US]; YIN PENG [US]) 6 March 2008 (2008-03-06)	1,3-4,8-16,
Y		2
Y	EP 1 956 847 A1 (MATSUSHITA ELECTRIC IND CO LTD [JP]) 13 August 2008 (2008-08-13)	2
A	JP 2004 253826 A (MATSUSHITA ELECTRIC IND CO LTD [JP]) 9 September 2004 (2004-09-09)	1-16

 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 application or patent 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
04/01/2011

Date of mailing of the international search report

10/01/2011

Name and mailing address of the ISA/
Nordic Patent Institute
Helgeshøj Allé 81, DK-2630 Taastrup,
Facsimile No. +45 4350 8008Authorized officer
Andreas Haffner

Telephone No. +47 22 38 73 50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NO2010/000361**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: 5,7
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
It is not clear in claim 5 how a macroblock can be represented by a set of motion vectors. An offset/movement is typically represented by a motion vector. Referring the special definition of zig-zag-scanning, claim 7 seems to disclose the same as claim 2. (PCT Art. 6)

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/NO2010/000361

Patent document Cited in search report	Publication	Patent family member(s)	Publication date
WO 2008027192 A	2008-03-06	EP 2055108 A CN 101507280 A JP 2010502102 T KR 20090057225 A US 2009196517 A	2009-05-06 2009-08-12 2010-01-21 2009-06-04 2009-08-06
EP 1956847 A	2008-08-13	CN 101292537 A US 2009122864 A WO 2007055158 A	2008-10-22 2009-05-14 2007-05-18
JP 2004253826	2004-09-09	NONE	