(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 May 2018 (31.05.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/095313 A1

(51) International Patent Classification:
H04N 19/139 (2014.01)

(21) International Application Number:
PCT/CN2017/112140

(22) International Filing Date:
21 November 2017 (21.11.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/425,172 22 November 2016 (22.11.2016) US

(71) Applicant: **MEDIATEK INC.** [CN/CN]; No. 1, Dusing Rd. 1st, Science-Based Industrial Park, Hsin-Chu, Taiwan 300 (CN).

(72) Inventor: **HSU, Chih-Wei**; No. 1, Dusing Rd. 1st, Science-Based Industrial Park, Hsin-Chu, Taiwan 300 (CN).

(74) Agent: **BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.**; 16th Fl., Block A, Corporate Square, No. 35 Jinrong Street, Beijing 100033 (CN).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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,

(54) Title: METHOD AND APPARATUS FOR MOTION VECTOR SIGN PREDICTION IN VIDEO CODING

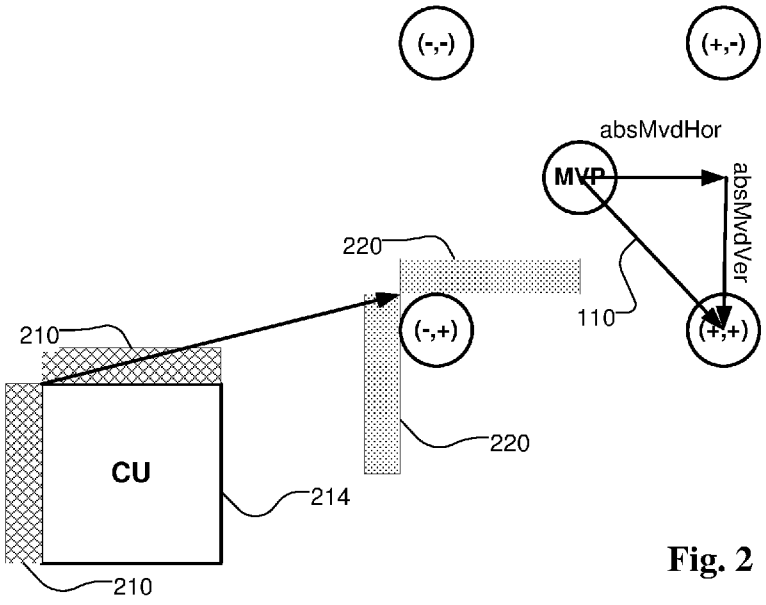


Fig. 2

(57) Abstract: A method and apparatus of motion vector reconstruction based on one MVP (motion vector predictor) and one MVD (motion vector difference) are disclosed. According to this method, input data associated with a video sequence including a current block in a current picture are received. Candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD are derived, where each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD. Template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position is evaluated. A target candidate MV position among all candidate reconstructed MV positions is selected. The current block is encoded or decoded using a target candidate MV corresponding to the target candidate MV position.

WO 2018/095313 A1

[Continued on next page]

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, KM, ML, MR, NE, SN, TD, TG).

Published:

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

METHOD AND APPARATUS FOR MOTION VECTOR SIGN PREDICTION IN VIDEO CODING

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present invention claims priority to U.S. Provisional Patent Application, Serial No. 62/425,172, filed on November 22, 2016. The U.S. Provisional Patent Application is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] The present invention relates to motion vector prediction for video coding. In particular, the present invention relates to motion vector sign derivation or prediction to improve coding efficiency of motion vector coding.

BACKGROUND

[0003] High-Efficiency Video Coding (HEVC) is a new international video coding standard developed by the Joint Collaborative Team on Video Coding (JCT-VC). HEVC is based on the hybrid block-based motion-compensated DCT-like transform coding architecture. The basic unit for compression, termed as coding unit (CU), is a 2Nx2N square block, and each CU can be recursively split into four smaller CUs until the predefined minimum size is reached. Each CU contains one or multiple prediction units (PUs).

[0004] To achieve the best coding efficiency of hybrid coding architecture in HEVC, there are two kinds of prediction modes for each PU, which are Intra prediction and Inter prediction. For Intra prediction modes, the spatial neighbouring reconstructed pixels can be used to generate the directional predictions. There are up to 35 directions in HEVC. For Inter prediction modes, the temporal reconstructed reference frames can be used to generate motion compensated predictions. There are three different modes, including Skip, Merge and Inter Advanced Motion Vector Prediction (AMVP) modes.

[0005] When a PU is coded in Inter AMVP mode, motion-compensated prediction is performed with transmitted motion vector differences (MVDs) that can be used together with Motion Vector Predictors (MVPs) for deriving motion vectors (MVs). To decide MVP in Inter AMVP mode, the

advanced motion vector prediction (AMVP) scheme is used to select a motion vector predictor among an AMVP candidate set including two spatial MVPs and one temporal MVP. Therefore, in AMVP mode, an MVP index for MVP and the corresponding MVDs are required to be encoded and transmitted. In addition, the Inter prediction direction to specify the prediction directions among bi-prediction, and uni-prediction related to list 0 (L0) and list 1 (L1) along with the reference frame index for each list should also be encoded and transmitted.

[0006] When a PU is coded in either Skip or Merge mode, no motion information is transmitted except for the Merge index of the selected candidate. That is because the Skip and Merge modes utilize motion inference methods (i.e., MV=MVP+MVD, where MVD is zero) to obtain the motion information from spatial neighbouring blocks (spatial candidates) or a temporal block (temporal candidate) located in a co-located picture where the co-located picture is the first reference picture in list 0 or list 1, which is signalled in the slice header. In the case of a Skip PU, the residual signal is also omitted. To decide the Merge index for the Skip and Merge modes, the Merge scheme is used to select a motion vector predictor among a Merge candidate set containing four spatial MVPs and one temporal MVP.

SUMMARY

[0007] A method and apparatus of motion vector reconstruction in a video coding system based on one MVP (motion vector predictor) and one MVD (motion vector difference) are disclosed. According to this method, input data associated with a video sequence including a current block in a current picture are received. An MVP from the input data for the current block is determined. Also, first magnitude and second magnitude of an MVD associated with the current block are determined. Candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD are derived, where each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD. A template for the current block is determined, where the template is formed from neighbouring reconstructed pixels of the current block. Template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position is evaluated. A target candidate MV position among all candidate reconstructed MV positions is selected based on one or more performance criteria including best template matching among all candidate reconstructed MV positions. The current block is encoded or decoded using a target candidate MV corresponding to the target candidate MV position.

[0008] The template matching is evaluated based on a pre-determined criterion comprising

SAD (sum of absolution differences) or SSD (sum of squared differences). The performance criteria further include other decoder side information comprising statistics of previous selected sign pair or previous MVP. In one embodiment, the template matching between the template for the current block and the corresponding template for each candidate reconstructed MV position is evaluated based on nearby integer pixel positions of the corresponding template if the MVP or the MVD has fractional-pel precision. In this case, if two corresponding templates for two candidate reconstructed MV positions fall into a same location, locations of said two corresponding templates are modified to become separate. Furthermore, the locations of the two corresponding templates can be modified to be on a left side and a right side of the MVP, or on an upper side and a lower side of the MVP.

[0009] The method may comprise pruning one or more candidate reconstructed MV positions based on distance between one MVP and one candidate reconstructed MV position, where if a specific candidate reconstructed MV position is closer to another MVP in a candidate MVP set than the MVP, the specific candidate reconstructed MV position is pruned and not evaluated for the template matching.

[0010] Whether to encode or decode the current block using the target candidate MV can be turned on or off explicitly by signalling a control flag. Alternatively, whether to encode or decode the current block using the target candidate MV can also be turned on or off implicitly. For example, whether to encode or decode the current block using the target candidate MV can be turned on or off based on a result of an exclusive OR operation applied to the first magnitude and the second magnitude of the MVD.

[0011] The selected sign pair corresponding to the target candidate MV position can be used as a predictor for encoding or decoding signs of the MVD. A prediction hit condition corresponding to the signs of the MVD being equal to the selected sign pair can be coded using a single bit or the shortest codeword. Other prediction conditions corresponding to the signs of the MVD being not equal to the selected sign pair can be coded using two or more bits, where the other prediction conditions include a horizontal flip condition, a vertical flip condition and a diagonal flip condition. The bins generated for prediction conditions associated with blocks in the current picture can be context coded.

[0012] In another embodiment, if any of the first magnitude and the second magnitude of the MVD is zero, a first single sign from the current block can be combined with a second single sign from another block or another prediction list to form a joint sign pair, and a prediction hit condition corresponding to both the first single sign and the second single sign being correctly predicted can be coded using a single bit or the shortest codeword. Other prediction conditions corresponding to at least one of the joint sign pair being not correctly predicted can be coded using two or more bits,

where the other prediction conditions include a first single sign hit and a second single sign miss condition, a first single sign miss and a second single sign hit condition and both the first single sign miss and the second single sign miss condition.

[0013] Aspects of the disclosure further provide a non-transitory computer readable medium storing program instructions for causing a processing circuit of an apparatus to perform a motion vector reconstruction method based on one MVP (motion vector predictor) and one MVD (motion vector difference). In some embodiment, input data associated with a video sequence including a current block in a current picture are received. An MVP from the input data for the current block is determined. Also, first magnitude and second magnitude of an MVD associated with the current block are determined. Candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD are derived, where each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD. A template for the current block is determined, where the template is formed from neighbouring reconstructed pixels of the current block. Template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position is evaluated. A target candidate MV position among all candidate reconstructed MV positions is selected based on one or more performance criteria including best template matching among all candidate reconstructed MV positions. The current block is encoded or decoded using a target candidate MV corresponding to the target candidate MV position.

BRIEF DESCRIPTION OF DRAWINGS

[0014] Fig. 1 illustrates an example of motion vector (MV) reconstruction based on the motion vector predictor (MVP) and magnitudes of motion vector differences (MVDs).

[0015] Fig. 2 illustrates an example of sign derivation for motion vector differences (MVDs) using template matching according to the present invention.

[0016] Fig. 3 illustrates an example of pruning one or more candidate sign positions using available information at the decoder side, such as the distance between a candidate sign position and motion vector predictors (MVP) including a current MVP and another MVP from the candidate MVP set.

[0017] Fig. 4 illustrates an exemplary flowchart of video coding system incorporate an embodiment of the present invention, where the signs of motion vector differences (MVDs) are derived or predicted using template matching.

DETAILED DESCRIPTION

[0018] The following description is of the best-contemplated mode of carrying out the invention. This description is made for the purpose of illustrating the general principles of the invention and should not be taken in a limiting sense. The scope of the invention is best determined by reference to the appended claims.

[0019] When a current block (or a prediction unit, PU) is coded in Inter AMVP mode, an MVP index and two MVDs for X and Y components are coded to represent the MVs. According to HEVC, for each MVD, magnitude and sign are coded separately using CABAC (Context-Adaptive Binary Arithmetic Coding) entropy coding. Due to different statistics, the magnitude part is always coded using context bin and the sign part is coded using bypass bin (i.e., with equal probability). In this invention, a scheme to code (i.e., encode/decode) the MVD signs more efficiently is disclosed.

[0020] When MVP is used, MVs are reconstructed using MVP and MVDs at the decoder side. With the signs of MVDs un-knowing, there are four possibilities for each MV, as shown in Fig. 1. As shown in Fig. 1, absMvdHor and absMvdVer are the absolute values (i.e., magnitudes) of the X and Y components of the MVD 110 respectively. As shown in this example, if the signs of both X and Y components of the MVD are positive, then the MV can be derived using MVP plus MVD (i.e., MV = MVP + MVD) as indicated in Fig. 1. If the decoder can infer the correct sign position, there is no need to code the signs in the bitstream. In order to perform the decoding process correctly, the same sign derivation process will be performed at the encoder side as well. In Fig. 1, the circle labelled as MVP indicates the position pointed to by the MVP. The four circles labelled with different sign pairs (i.e., (+, +), (+, -), (-, +), (-, -)) correspond to four possible sign positions pointed to by the reconstructed MV. In this disclosure, the MVD may also refer to a vector with MVD X component and MVD Y component for convenience.

[0021] In one embodiment, template matching method is used for the decoder to derive the correct sign position. An example of deriving the correct sign position is shown in Fig. 2, where the MV of a current block (214) is to be decoded. The neighbouring reconstructed pixels of current block 214 are used to form a template (210). The MVP can be derived at the decoder side according to MVP index. At the encoder side, the MVP will be selected from a candidate list and the index of the selected MVP is signalled in the bitstream. When the absMvdHor and absMvdVer are decoded at the decoder, the four possible positions of the reconstructed MV can be determined as shown by the four circles labelled with different sign pairs. The corresponding templates at each possible reconstructed MV position (also referred as “candidate reconstructed MV position” in this

disclosure) can be formed. In Fig. 2, the corresponding template (220) for possible position corresponding to (-,+) is shown. Similarly, the templates for three other possible positions can be determined. The corresponding templates are formed with the same number of pixels, shape and distance as the template for the current block in order to perform the template matching. For example, template 210 is matched with template 220. After the matching results for all four possible positions are determined, the best match position based on a predetermined criterion (e.g. SAD or SSD) can be inferred to be the final MVD sign position. Any other information available at the decoder side can also be used to help the sign deriving process and to improve the correctness. For example, the previous statistics of the sign position, MVP position etc. can be used.

[0022] In another embodiment, the templates for the four possible sign positions can be derived without using the exact MV precision. For example, if the MVs are coded using quarter-pixel precision, the template positions can be derived using the nearest integer position. However, the two integer-based template positions associated with the "-" and "+" signs in the X direction or the Y direction may fall into the same position, which may occur particularly for small absMvdHor or absMvdVer. If any two of these integer-based templates fall into a same position, constraints or rules can be applied to make sure the two templates are apart from each other. For example, if two templates fall into the same position, one should be located at the left hand side of the MVP and the other should be located at the right hand side of the MVP. Alternatively, one should be located at the upper side of the MVP and the other should be located at the lower side of the MVP.

[0023] In another embodiment, the available information at the decoder can be used to improve the correctness of the sign prediction derivation process. Fig. 3 illustrates an example, where there are a total of two MVPs (e.g. MVP1 and MVP2) and MVP1 is used for the current block by the encoder. In other words, MVP2 is in the candidate MVP set, but not selected by the encoder as the MVP for the current block. At the decoder side, one or more of the four possible MVD sign positions can be pruned according to the corresponding distances from a sign position to these two MVPs. As shown in this example, if the (-,+) position is used to reconstruct the MV, the resulted MV position will be closer to MVP2 position instead of MVP1 position. Therefore, the sign (-,+) position can be pruned since the encoder would favour a shorter distance (i.e., smaller MVD magnitude) between a sign position and the MVP position. If the (-,+) position is the correct position, the encoder would have selected MVP2 instead of MVP1. By pruning some of the sign positions, it can reduce the possibility of selecting a wrong sign position and also reduce the required computations for evaluating the template matching.

[0024] In another embodiment, the sign derivation process can be signalled using one explicit flag to turn on and off.

[0025] In another embodiment, the sign derivation process can be determined using implicit method to turn on and off. For example, to turn on the process when the XOR (i.e., exclusive OR operation) between the X component of the MVD and the Y component of the MVD results in an even number (i.e., "0"), and to turn off otherwise.

[0026] In another embodiment, the derived sign position can be used a predictor. The true MVD sign can be coded using residues. Table 1 illustrates an example of a codeword table to code the sign residues using the sign predictors. According to the codeword table, if the predictor inferred by the derived sign process matches the true MVD sign position, only one bin or bin (i.e., "0" in Table 1) needs to be coded. On the other hand, the original method will require two bins or bits (one for each sign) for the MVDs. If the predictor is wrong and is the true MVD sign position matches the diagonal sign position of the derived sign position (referred as Diagonal flipping for this condition), more bins or bits (three in this example) are needed. However, no matter how many bins or bits are used in this sign residue coding, coding sign will no longer use bypass bin with equal probability. Instead, it can use context coding with more efficient adaptive probability update. Table 1 also include horizontal flip prediction condition and vertical flip prediction condition. The horizontal flip prediction condition refers to the condition that the true MVD sign position matches the other sign position in the horizontal direction of the derived sign position. The vertical flip prediction condition refers to the condition that the true MVD sign position matches the other sign position in the vertical direction of the derived sign position. In the example as shown in Table 1, the predictor matching the true MVD sign position (i.e., a prediction hit condition corresponding to the signs of the MVD being equal to the selected sign pair) is coded using shortest codeword.

Table 1.

	Code word
Prediction hit	0
Horizontal flip	10
Vertical flip	110
Diagonal flip	111

[0027] In another embodiment, when any one of the X component and Y component of the MVD is zero, there exist only one MVD sign and two possible sign positions for the non-zero components. In that case, only one code word is needed to indicate if the prediction is hit or not (i.e., miss) in Table 1. The codeword table in Table 1 is still applicable. However, more efficient variation

may be applied. Two single sign residues that may come from different prediction lists or different blocks can be combined into one and then use the codeword table in Table 2. In this situation, the code word table will represent different semantics as shown in Table 2.

Table 2.

	Code word
Both hit	0
The first hit and the second miss	10
The first miss and the second hit	110
Both miss	111

[0028] In another embodiment, at the encoder side, the bits resulted from the sign residues can be taken into consideration during the mode decision in order to choose the best modes. Moreover, during the motion estimation for searching the MVs, it can also consider the sign residue bits for every MV position. In other words, the MV not only generates the better predictors for the current block, but also results in less coded bin for the MVD sign residues.

[0029] Fig. 4 illustrates an exemplary flowchart of video coding system incorporate an embodiment of the present invention, where the signs of motion vector differences (MVDs) are derived or predicted using template matching. The steps shown in the flowchart, as well as other flowcharts in this disclosure, may be implemented as program codes executable on one or more processors (e.g., one or more CPUs) at the encoder side and/or the decoder side. The steps shown in the flowchart may also be implemented based on hardware such as one or more electronic devices or processors arranged to perform the steps in the flowchart. According to this method, input data associated with a video sequence including a current block in a current picture are received in step 410. At the encoder side, the input data may correspond to pixel data associated with a video sequence. At the decoder side, the input data may correspond to video bitstream or compressed data for the video sequence. An MVP from the input data for the current block is determined in step 420. As is known in video coding, the encoder may selects the MVP from a candidate MVP set and signal the MVP index associated with the selected MVP. The decoder can selected the MVP from the candidate MVP set according to the MVP index. The first magnitude and second magnitude of an MVD associated with the current block are determined in step 430. The encoder can obtain the MVD from the difference between the current MV and the MVP. The magnitudes of the MVD are signalled in the bitstream. The decoder may recover the magnitudes of the MVD from the bitstream.

Candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD are derived in step 440, where each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD. This same step is performed at the encoder as well as the decoder. A template for the current block is determined in step 450, where the template is formed from neighbouring reconstructed pixels of the current block. Template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position is then performed in step 460. A target candidate MV position is selected among all candidate reconstructed MV positions based on one or more performance criteria including best template matching among all candidate reconstructed MV positions in step 470. Encoding or decoding is then applied to the current block using a target candidate MV corresponding to the target candidate MV position in step 480.

[0030] The flowchart shown is intended to illustrate an example of video coding according to the present invention. A person skilled in the art may modify each step, re-arranges the steps, split a step, or combine steps to practice the present invention without departing from the spirit of the present invention. In the disclosure, specific syntax and semantics have been used to illustrate examples to implement embodiments of the present invention. A skilled person may practice the present invention by substituting the syntax and semantics with equivalent syntax and semantics without departing from the spirit of the present invention.

[0031] The above description is presented to enable a person of ordinary skill in the art to practice the present invention as provided in the context of a particular application and its requirement. Various modifications to the described embodiments will be apparent to those with skill in the art, and the general principles defined herein may be applied to other embodiments. Therefore, the present invention is not intended to be limited to the particular embodiments shown and described, but is to be accorded the widest scope consistent with the principles and novel features herein disclosed. In the above detailed description, various specific details are illustrated in order to provide a thorough understanding of the present invention. Nevertheless, it will be understood by those skilled in the art that the present invention may be practiced.

[0032] Embodiment of the present invention as described above may be implemented in various hardware, software codes, or a combination of both. For example, an embodiment of the present invention can be one or more circuit circuits integrated into a video compression chip or program code integrated into video compression software to perform the processing described herein. An embodiment of the present invention may also be program code to be executed on a Digital Signal Processor (DSP) to perform the processing described herein. The invention may also

involve a number of functions to be performed by a computer processor, a digital signal processor, a microprocessor, or field programmable gate array (FPGA). These processors can be configured to perform particular tasks according to the invention, by executing machine-readable software code or firmware code that defines the particular methods embodied by the invention. The software code or firmware code may be developed in different programming languages and different formats or styles. The software code may also be compiled for different target platforms. However, different code formats, styles and languages of software codes and other means of configuring code to perform the tasks in accordance with the invention will not depart from the spirit and scope of the invention. Embodiment of the present invention as described above may be implemented in a video encoder and a video decoder. The components of the video encoder and video decoder may be implemented by hardware components, one or more processors configured to execute program instructions stored in a memory, or a combination of hardware and processor. For example, a processor executes program instructions to control receiving of input data associated with a video sequence including a current block in a current picture. The processor is equipped with a single or multiple processing cores. In some examples, the processor executes program instructions to perform functions in some components in the encoder and the decoder, and the memory electrically coupled with the processor is used to store the program instructions, information corresponding to the reconstructed images of blocks, and/or intermediate data during the encoding or decoding process. The memory in some embodiment includes a non-transitory computer readable medium, such as a semiconductor or solid-state memory, a random access memory (RAM), a read-only memory (ROM), a hard disk, an optical disk, or other suitable storage medium. The memory may also be a combination of two or more of the non-transitory computer readable medium listed above.

[0033] The invention may be embodied in other specific forms without departing from its spirit or essential characteristics. The described examples are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.

CLAIMS

1.A method of motion vector reconstruction in a video coding system based on one MVP (motion vector predictor) and one MVD (motion vector difference), the method comprising:

receiving input data associated with a video sequence including a current block in a current picture;

determining an MVP from the input data for the current block;

determining first magnitude and second magnitude of an MVD associated with the current block;

deriving candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD, wherein each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD;

determining a template for the current block, wherein the template is formed from neighbouring reconstructed pixels of the current block;

evaluating template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position;

selecting a target candidate MV position among all candidate reconstructed MV positions based on one or more performance criteria including best template matching among all candidate reconstructed MV positions; and

encoding or decoding the current block using a target candidate MV corresponding to the target candidate MV position.

2.The method of claim 1, wherein the template matching is evaluated based on a pre-determined criterion comprising SAD (sum of absolution differences) or SSD (sum of squared differences).

3.The method of claim 1, wherein said one or more performance criteria further include other decoder side information comprising statistics of previous selected sign pair or previous MVP.

4.The method of claim 1, wherein the template matching between the template for the current block and the corresponding template for each candidate reconstructed MV position is evaluated based on nearby integer pixel positions of the corresponding template if the MVP or the MVD has fractional-pel precision.

5.The method of claim 4, wherein if two corresponding templates for two candidate reconstructed MV positions fall into a same location, locations of said two corresponding templates are modified to become separate.

6.The method of claim 5, wherein said locations of said two corresponding templates are

modified to be on a left side and a right side of the MVP, or on an upper side and a lower side of the MVP.

7.The method of claim 1, further comprising pruning one or more candidate reconstructed MV positions based on distance between one MVP and one candidate reconstructed MV position, wherein if a specific candidate reconstructed MV position is closer to another MVP in a candidate MVP set than the MVP, the specific candidate reconstructed MV position is pruned and not evaluated for the template matching.

8.The method of claim 1, wherein whether to encode or decode the current block using the target candidate MV is turned on or off explicitly by signalling a control flag.

9.The method of claim 1, wherein whether to encode or decode the current block using the target candidate MV is turned on or off implicitly.

10.The method of claim 9, wherein whether to encode or decode the current block using the target candidate MV is turned on or off based on a result of an exclusive OR operation applied to the first magnitude and the second magnitude of the MVD.

11.The method of claim 1, wherein selected sign pair corresponding to the target candidate MV position is used as a predictor for encoding or decoding signs of the MVD.

12.The method of claim 11, wherein a prediction hit condition corresponding to the signs of the MVD being equal to the selected sign pair is coded using shortest codeword.

13.The method of claim 11, wherein a prediction hit condition corresponding to the signs of the MVD being equal to the selected sign pair is coded using a single bit.

14.The method of claim 13, wherein other prediction conditions corresponding to the signs of the MVD being not equal to the selected sign pair are coded using two or more bits, wherein said other prediction conditions include a horizontal flip condition, a vertical flip condition and a diagonal flip condition.

15.The method of claim 14, wherein bins generated for prediction conditions associated with blocks in the current picture are context coded.

16.The method of claim 1, wherein if any of the first magnitude and the second magnitude of the MVD is zero, a first single sign from the current block is combined with a second single sign from another block or from the current block in another prediction list to form a joint sign pair.

17.The method of claim 16, wherein a prediction hit condition corresponding to both the first single sign and the second single sign being correctly predicted is coded using a single bit or a shortest codeword.

18.The method of claim 16, wherein other prediction conditions corresponding to at least one of the joint sign pair being not correctly predicted are coded using two or more bits, wherein said other

prediction conditions include a first single sign hit and a second single sign miss condition, a first single sign miss and a second single sign hit condition and both the first single sign miss and the second single sign miss condition.

19.An apparatus for motion vector reconstruction in a video coding system based on one MVP (motion vector predictor) and one MVD (motion vector difference), the apparatus comprising one or more electronic devices or processors configured to:

receive input data associated with a video sequence including a current block in a current picture;

determine an MVP from the input data for the current block;

determine first magnitude and second magnitude of an MVD associated with the current block;

derive candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD, wherein each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD;

determine a template for the current block, wherein the template is formed from neighbouring reconstructed pixels of the current block;

evaluate template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position;

select a target candidate MV position among all candidate reconstructed MV positions based on one or more performance criteria including best template matching among all candidate reconstructed MV positions; and

encode or decode the current block using a target candidate MV corresponding to the target candidate MV position.

20.A non-transitory computer readable medium storing program instructions causing a processing circuit of an apparatus to perform a motion vector reconstruction method based on one MVP (motion vector predictor) and one MVD (motion vector difference), and the method comprising:

receiving input data associated with a video sequence including a current block in a current picture;

determining an MVP from the input data for the current block;

determining first magnitude and second magnitude of an MVD associated with the current block;

deriving candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD, wherein each candidate reconstructed MV position is derived as a vector sum

of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD;

determining a template for the current block, wherein the template is formed from neighbouring reconstructed pixels of the current block;

evaluating template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position;

selecting a target candidate MV position among all candidate reconstructed MV positions based on one or more performance criteria including best template matching among all candidate reconstructed MV positions; and

encoding or decoding the current block using a target candidate MV corresponding to the target candidate MV position.

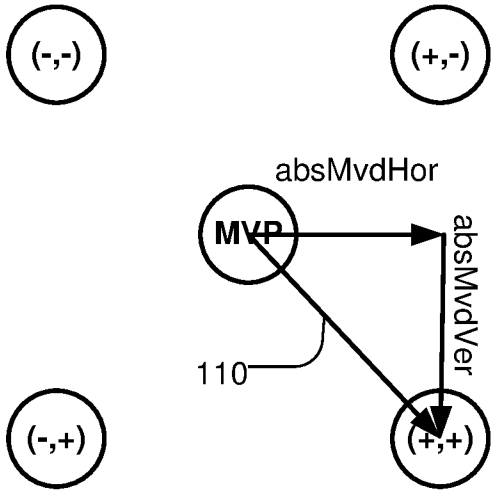

(-,-)
(+,-)
absMvdHor
MVP
absMvdVer
110
(-,+)
(+,+)

Fig. 1

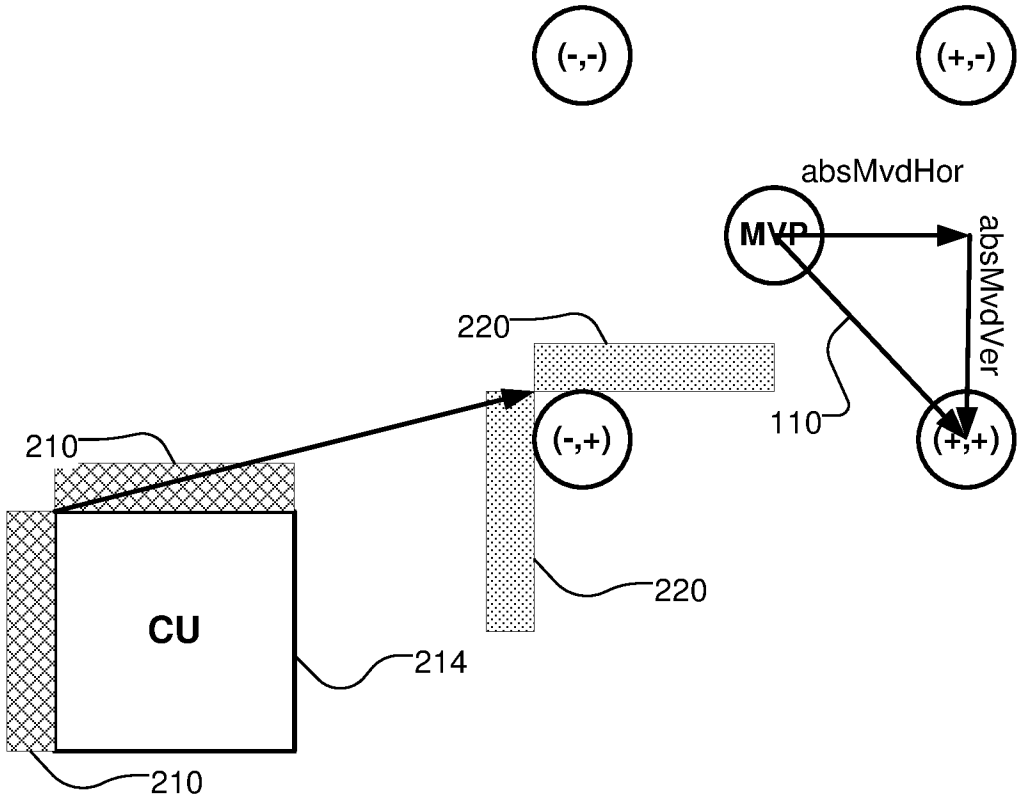

(-,-)
(+,-)
absMvdHor
MVP
absMvdVer
220
110
210
(-,+)
(+,+)
220
CU
214
210

Fig. 2

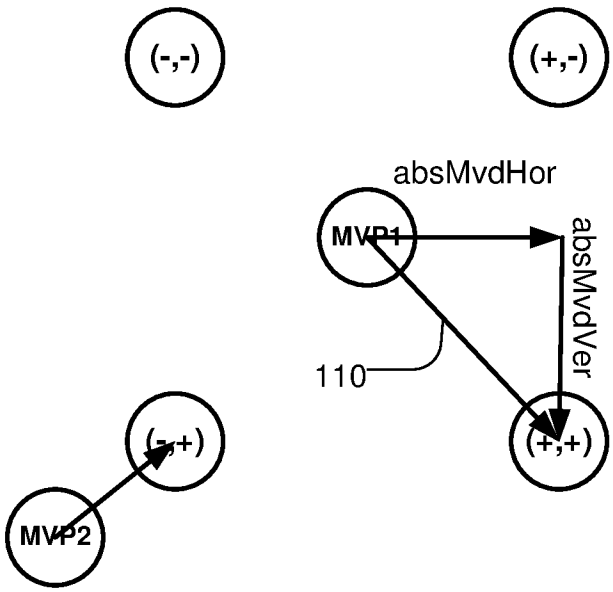


Fig. 3

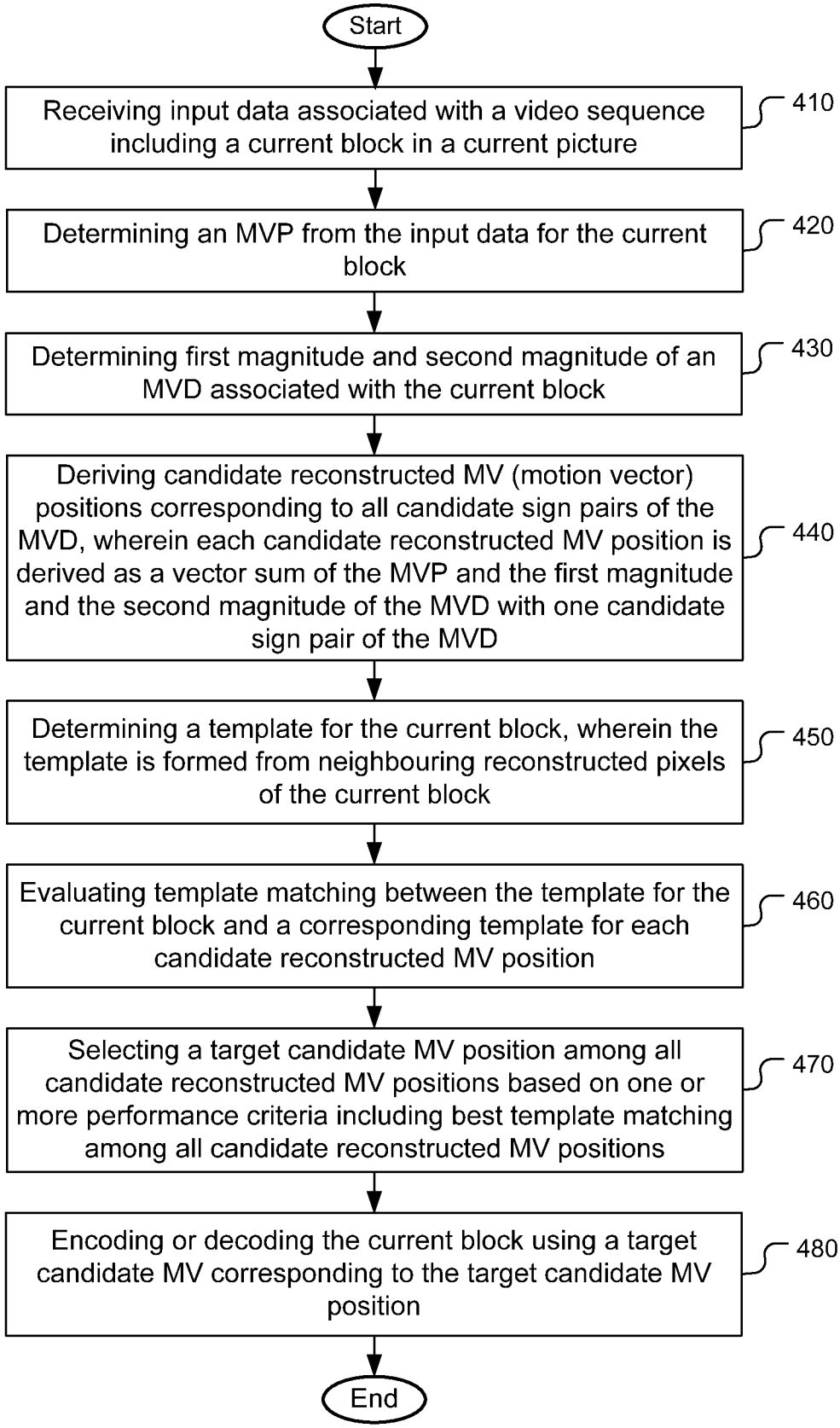

Start
Receiving input data associated with a video sequence including a current block in a current picture
410
Determining an MVP from the input data for the current block
420
Determining first magnitude and second magnitude of an MVD associated with the current block
430
Deriving candidate reconstructed MV (motion vector) positions corresponding to all candidate sign pairs of the MVD, wherein each candidate reconstructed MV position is derived as a vector sum of the MVP and the first magnitude and the second magnitude of the MVD with one candidate sign pair of the MVD
440
Determining a template for the current block, wherein the template is formed from neighbouring reconstructed pixels of the current block
450
Evaluating template matching between the template for the current block and a corresponding template for each candidate reconstructed MV position
460
Selecting a target candidate MV position among all candidate reconstructed MV positions based on one or more performance criteria including best template matching among all candidate reconstructed MV positions
470
Encoding or decoding the current block using a target candidate MV corresponding to the target candidate MV position
480
End

Fig. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/112140

A. CLASSIFICATION OF SUBJECT MATTER

H04N 19/139(2014.01)i

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,G06T

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 practicable, search terms used)

CNPAT,CNKI,EPODOC,WPI: HEVC,JCT VC,motion vector,reconstruct+,cod+,decoder,MVP,MVD,predict+,difference, magnitude,sign,template,candidate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2013003849 A1 (QUALCOMM INCORPORATED) 03 January 2013 (2013-01-03) description, paragraphs [0020], [0051]-[0053], [0138]	1-20
A	CN 104519363 A (THOMSON LICENSING) 15 April 2015 (2015-04-15) the whole document	1-20
A	STANKOWSKI, Jakub et al. "Analysis of Compressed Data Stream Content in HEVC Video Encoder" *INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS,* Vol. 61, No. 2, 30 June 2015 (2015-06-30), the whole document	1-20
A	SZE, Vivienne et al. "High Throughput CABAC Entropy Coding in HEVC" *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY,* Vol. 22, No. 12, 31 December 2012 (2012-12-31), the whole document	1-20

☐ 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	Date of mailing of the international search report
26 January 2018	**23 February 2018**

Name and mailing address of the ISA/CN	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088** **China**	**FENG,Wei**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)010-62413235**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/112140

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2013003849	A1	03 January 2013	ES	2546678	T3	25 September 2015
				CA	2840728	C	27 June 2017
				JP	2014523714	A	11 September 2014
				AU	2012279234	A1	09 January 2014
				CN	103703781	A	02 April 2014
				KR	20140043807	A	10 April 2014
				CN	103703781	B	05 April 2017
				EP	2727353	B1	15 July 2015
				DK	2727353	T3	19 October 2015
				MX	2013014936	A	11 February 2014
				CA	2840728	A1	10 January 2013
				BR	112013033809	A2	14 February 2017
				AU	2012279234	B2	07 May 2015
				SG	195279	A1	30 December 2013
				EP	2727353	A1	07 May 2014
				HU	E025216	T2	29 February 2016
				PT	2727353	E	23 October 2015
				KR	101660133	B1	26 September 2016
				WO	2013006483	A1	10 January 2013
				US	2014341297	A1	20 November 2014
CN	104519363	A	15 April 2015	US	9654793	B2	16 May 2017
				JP	2015070610	A	13 April 2015
				US	2015085934	A1	26 March 2015
				EP	2854405	A1	01 April 2015
				KR	20150034636	A	03 April 2015
				EP	2854406	A1	01 April 2015