

(51) International Patent Classification:
H04N 7/32 (2006.01)(21) International Application Number:
PCT/CN2012/076643(22) International Filing Date:
8 June 2012 (08.06.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/497,438 15 June 2011 (15.06.2011) US(71) Applicant (for all designated States except US): **MEDI-ATEK INC.** [CN/CN]; No.1, Dusing Rd. 1st, Science-Based Industrial Park, Hsin-Chu, Taiwan (CN).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LIN, Jian-Liang** [CN/CN]; No.133, Neipi Rd., Nanjian Vil. Su'ao Township, Yilan County, Taiwan (CN). **CHEN, Yi-Wen** [CN/CN]; 8F., No.66, Ln. 211, Yongkang Rd., Fengyuan Dist., Taichung City, Taiwan (CN). **TSAL, Yu-Pao** [CN/CN]; No.62, Huaming St., Jhensi Village, Fengshan Dist., Kaohsiung City, Taiwan (CN). **HUANG, Yu-Wen** [CN/CN]; 8F., No. 23, Ln. 298, Longjiang Rd., Zhongshan Dist., Taipei City, Taiwan (CN). **LEI, Shaw-Min** [CN/CN]; 6F., No.32, Liujia 6th St., Zhubei City, Hsinchu County, Taiwan (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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

[Continued on next page]

(54) Title: METHOD AND APPARATUS OF MOTION AND DISPARITY VECTOR PREDICTION AND COMPENSATION FOR 3D VIDEO CODING

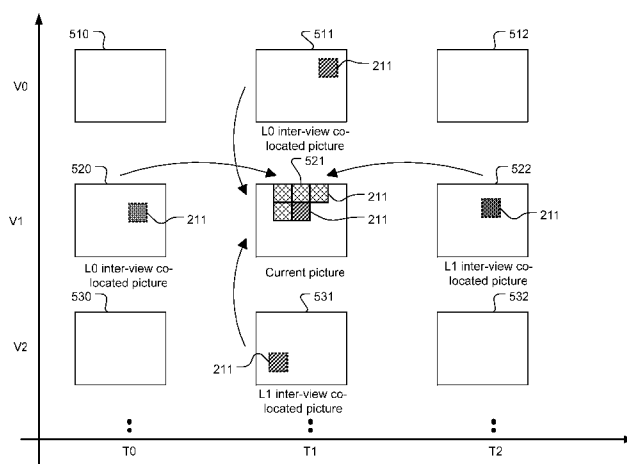


Fig. 3

(57) Abstract: A method and apparatus for deriving MV/MVP (motion vector or motion vector predictor) or DV/DVP (disparity vector or disparity vector predictor) associated Skip mode, Merge mode or Inter mode for a block of a current picture in three-dimensional (3D) video coding are disclosed. The 3D video coding may use temporal prediction and inter-view prediction to exploit temporal and inter-view correlation. MV/DV prediction is applied to reduce bitrate associated with MV/DV coding. The MV/MVP or DV/DVP for a block is derived from spatial candidates, temporal candidates and inter-view candidates. For the inter-view candidate, the position of the inter-view co-located block can be located using a global disparity vector (GDV) or warping the current block onto the co-located picture according to the depth information. The candidate can also be derived as the vector corresponding to warping the current block onto the co-located picture according to the depth information.

WO 2012/171442 A1



Published:

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

METHOD AND APPARATUS OF MOTION AND DISPARITY VECTOR PREDICTION AND COMPENSATION FOR 3D VIDEO CODING

BACKGROUND OF THE INVENTION

Cross Reference To Related Applications

[0001] The present invention claims priority to U.S. Provisional Patent Application, Serial No. 61/497,438, filed June 15, 2011, entitled "Method for motion vector prediction and disparity vector prediction in 3D video coding". The present invention is also related to U.S. Non-Provisional Patent Application, Serial No. 13/236,422, filed September 19, 2011, entitled "Method and Apparatus for Deriving Temporal Motion Vector Prediction". The U.S. Provisional Patent Application and U.S. Non-Provisional Patent Application are hereby incorporated by reference in their entireties.

Field of the Invention

[0002] The present invention relates to video coding. In particular, the present invention relates to motion/disparity vector prediction and information sharing of motion/disparity compensation in 3D video coding.

Description of the Related Art

[0003] Three-dimensional (3D) television has been a technology trend in recent years that is targeted to bring viewers sensational viewing experience. Various technologies have been developed to enable 3D. Among them, the multi-view video is a key technology for 3DTV application among others. The traditional video is a two-dimensional (2D) medium that only provides viewers a single view of a scene from the perspective of the camera. However, the multi-view video is capable of offering arbitrary viewpoints of dynamic scenes and provides viewers the sensation of realism.

[0004] The multi-view video is typically created by capturing a scene using multiple cameras simultaneously, where the multiple cameras are properly located so that each camera captures the scene from one viewpoint. Accordingly, the multiple cameras will capture multiple video sequences. In order to provide more views, more cameras have been used to generate multi-view video with a large number of video sequences associated with the views. Accordingly, the multi-view video will require a large storage space to store and/or a high bandwidth to transmit.

Therefore, multi-view video coding techniques have been developed in the field to reduce the required storage space of the transmission bandwidth. A straightforward approach may simply apply conventional video coding techniques to each single-view video sequence independently and disregard any correlation among different views. In order to improve multi-view video coding efficiency, typical multi-view video coding always exploits inter-view redundancy.

[0005] Fig. 1 illustrates an example of a prediction structure for 3D video coding. The vertical axis represents different views and the horizontal axis represents different time instances that the pictures are captured. In addition to a color image, a depth image is also captured at each view and each time instances. For example, for view V0, color images 110C, 111C, and 112C are captured corresponding to time instances T0, T1 and T2 respectively. Also, depth images 110D, 111D, and 112D are captured along with the color images corresponding to time instances T0, T1 and T2 respectively. Similarly, color images 120C, 121C, and 122C and associated depth images 120D, 121D, and 122D are captured corresponding to time instances T0, T1 and T2 respectively for view V1, and color images 130C, 131C, and 132C and associated depth images 130D, 131D, and 132D are captured corresponding to time instances T0, T1 and T2 respectively for view V2. Conventional video coding based on inter/intra-prediction can be applied to images in each video. For example, in view V1, images 120C and 122C are used for temporal prediction of image 121C. In addition, inter-view prediction serves as another dimension of prediction in addition to the temporal prediction. Accordingly, the term *prediction dimension* is used in this disclosure to refer to the prediction axis that video information along the axis is used for prediction. Therefore, the prediction dimension may refer to the inter-view prediction or the temporal prediction. For example, in time T1, image 111C from view V0 and image 131C from view V2 can be used to predict image 121C of view V1. Furthermore, the depth information associated with the scene is also included in the bit stream to provide support for interactive applications. The depth information can also be used for synthesizing virtual views from intermediate viewpoints.

[0006] In order to reduce the bit-rate for transmitting motion vectors (MVs) for coding the multi-view video, motion skip mode was disclosed to share the previously encoded motion information of adjacent views. As shown in Fig. 2, the motion skip mode includes two steps. In the first step, co-located block 212 of picture 222 in a neighboring view is identified for current block 210 of picture 220 in the current view. The co-located block 212 is identified by determining global disparity vector 230 between the current picture 220 in the current view and the co-located picture 222 in the neighboring view. In the second step, the motion information of the co-located block 212 in the co-located picture 222 is shared with the current block 210 in the current picture 220. For example, motion vectors 242 and 252 of the co-located block 212

can be shared by the current block 210. The motion vectors 240 and 250 for the current block 210 may be derived from motion vectors 242 and 252.

[0007] High Efficiency Video Coding (HEVC) is a new international video coding standard that is under development by the Joint Collaborative Team on Video Coding (JCT-VC). In the HEVC Working Draft Version 3.0 (WD-3.0) and the HEVC Test Model Version 3.0 (HM-3.0), a hybrid block-based motion-compensated DCT-like transform coding architecture, similar to previous coding standards such as MPEG-4 and AVC/H.264, is used. However, there are also new features and coding tools that are introduced. For example, the basic unit for compression, termed Coding Unit (CU), is a $2N \times 2N$ 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), where the PU is used as the block unit for prediction process. The PU sizes can be $2N \times 2N$, $2N \times N$, $N \times 2N$, and $N \times N$.

[0008] In order to increase the coding efficiency of motion vector coding in HEVC, the motion vector competition (MVC) based scheme is applied to select one motion vector predictor (MVP) among a given MVP candidate set, which includes spatial and temporal MVPs. There are three inter-prediction modes, i.e., Inter, Skip, and Merge included in HM-3.0. The Inter mode performs motion-compensated predictions based on transmitted motion vectors (MVs), while the Skip and Merge modes utilize motion inference methods to determine the motion information from spatially neighboring 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 as indicated in the slice header.

[0009] When a PU is coded in either Skip or Merge mode, no motion information is transmitted except for the index of the selected candidate. For a Skip-mode PU, the residual signal is not transmitted either. For the Inter in HM-3.0, 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. As for the Merge and Skip modes in HM-3.0, the Merge scheme is used to select a motion vector predictor among a Merge candidate set containing four spatial MVPs and one temporal MVP. Based on the rate-distortion optimization (RDO) decision, the encoder selects a final MVP from a given candidate set of MVPs for Inter, Skip, or Merge mode and transmits the index of the selected MVP to the decoder. The selected MVP may be linearly scaled according to temporal distances.

[0010] For the Inter mode, the reference picture index is explicitly transmitted to the decoder. The MVP is then selected among the candidate set for a given reference picture index. Fig. 3 illustrates the MVP candidate set for the Inter in HM-3.0, where two spatial MVPs and one temporal MVP are included:

1. Left predictor (the first available motion vector from A_0 or A_1)
2. Top predictor (the first available motion vector from B_0 , B_1 or B_{n+1})
3. Temporal predictor (the first available motion vector from T_{BR} or T_{CTR})

[0011] The temporal predictor is derived from a block (T_{BR} or T_{CTR}) located in a co-located picture where the co-located picture is the first reference picture in list 0 or list 1. The block where a temporal MVP is selected from may have two MVs: one from list 0 and the other from list 1. The temporal MVP is derived based on the MV from list 0 or list 1 according to the following rules:

1. The MV that crosses the current picture is chosen first.
2. If both MVs cross or both do not cross the current picture, the one with same reference list as the current list will be chosen.

[0012] A priority-based scheme is applied for deriving each spatial MVP. The spatial MVP can be derived from a different list and a different reference picture. The selection is based on a predefined order as follows:

1. The MV from the same reference list and the same reference picture;
2. The MV from the other reference list and the same reference picture;
3. The scaled MV from the same reference list and a different reference picture; and
4. The scaled MV from the other reference list and a different reference picture.

[0013] In HM-3.0, if a particular block is encoded as Merge or Skip modes, a MVP index is incorporated in the bitstream to indicate which MVP among the MVP candidate set is used for the block to be merged. To follow the essence of motion information sharing, each merged PU reuses the MV, prediction direction, and reference picture index of the selected candidate. The *prediction direction* refers to the temporal direction associated with reference picture, such as list 0 (L0)/list 1 (L1) or Bi-prediction. It is noted that if the selected MVP is a temporal MVP, the reference picture index is always set to the first reference picture. Fig. 4 illustrates the candidate set of MVPs for Merge and Skip modes in HM-3.0, where four spatial MVPs and one temporal MVP are included:

1. Left predictor (A_m)
2. Top predictor (B_n)
3. Temporal predictor (the first available motion vector from T_{BR} or T_{CTR})
4. Above right predictor (B_0)
5. Below left predictor (A_0)

[0014] As shown above, HEVC uses advanced MVP derivation to reduce the bitrate associated

with motion vectors. It is desirable to extend the advanced MVP technique to 3D video coding to improve the coding efficiency.

BRIEF SUMMARY OF THE INVENTION

[0015] A method and apparatus for deriving MV/MVP (motion vector or motion vector predictor) or DV/DVP (disparity vector or disparity vector predictor) associated Skip mode, Merge mode or Inter mode for a block of a current picture in three-dimensional video coding using spatial prediction, temporal prediction and inter-view prediction are disclosed. Embodiments according to the present invention select the MV/MVP or the DV/DVP from spatial candidates, temporal candidates and inter-view candidates. The spatial candidates are associated with neighboring blocks of the block in the current picture; the temporal candidates are associated with temporal co-located blocks of one or more temporal co-located pictures; and the inter-view candidates are associated with an inter-view co-located block associated with one or more inter-view co-located pictures corresponding to the block. The MVP or the DVP selected can be used as a candidate for the Inter mode in the three-dimensional video coding. The MV or the DV selected can be used as a candidate for the Merge or the Skip mode in the three-dimensional video coding.

[0016] One aspect of the present invention addresses derivation of the spatial candidates. The spatial candidates can be used to derive MV/MVP or DV/DVP. In this case, for a given prediction dimension and a target reference picture as indicated by a given reference picture index of a given reference list, the spatial candidate can be derived from the neighboring blocks associated with the target reference picture from the given reference list or other reference list. Alternatively, the spatial candidate can be derived from the neighboring blocks associated with other reference pictures from the given reference list or the other reference list.

[0017] Another aspect of the present invention addresses derivation of the temporal candidates. The temporal candidates can be used to derive MV/MVP or DV/DVP. In this case, for a given prediction dimension and a target reference picture as indicated by a given reference picture index of a given reference list, the temporal candidate can be derived from the temporal co-located blocks of temporal co-located pictures. The temporal co-located blocks are associated with the target reference picture in the given reference list or other reference list, or associated with other reference picture in the given reference list or the other reference list.

[0018] Yet another aspect of the present invention addresses derivation of the inter-view candidates. The inter-view candidates can be used to derive MV/MVP or DV/DVP. In this case, for a given prediction dimension and a target reference picture as indicated by a given reference

picture index of a given reference list, the inter-view candidate can be derived from the inter-view co-located blocks of inter-view co-located pictures. The inter-view co-located blocks are associated with the target reference picture in the given reference list or other reference list, or associated with other reference picture in the given reference list or the other reference list.

[0019] In another embodiment of the present invention, a depth candidate is derived from the DV associated with a corresponding co-located block by warping the block of the current picture onto the picture based on depth information.

BRIEF DESCRIPTION OF DRAWINGS

[0020] Fig. 1 illustrates an example of prediction structure for 3D video, where the prediction comprises temporal and inter-view predictions.

[0021] Fig. 2 illustrates an example of skip mode for 3D video, where the co-located block is determined using Global Disparity Vector (GDV).

[0022] Fig. 3 illustrates an example of Motion Vector Predictor (MVP) candidate set for Inter mode in HM-3.0.

[0023] Fig. 4 illustrates an example of Motion Vector Predictor (MVP) candidate set for Merge mode in HM-3.0.

[0024] Fig. 5 illustrates an example of Motion Vector (MV)/Disparity Vector (DV) candidate derivation for 3D video coding according to the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0025] In the present invention, various prediction schemes are applied to derive Motion Vector (MV)/Disparity Vector (DV) and Motion Vector Predictor (MVP)/ Disparity Vector Predictor (DVP) for Skip, Merge and Inter modes in 3D video coding.

[0026] Fig. 5 illustrates a scenario that the MV(P)/DV(P) candidates for a current block are derived from spatially neighboring blocks, temporally co-located blocks in the co-located pictures in list 0 (L0) or list 1(L1), and inter-view co-located blocks in the inter-view co-located picture. Pictures 510, 511 and 512 correspond to pictures from view V0 at time instances T0, T1 and T2 respectively. Similarly, pictures 520, 521 and 522 correspond to pictures from view V1 at time instances T0, T1 and T2 respectively and pictures 530, 531 and 532 correspond to pictures from view V2 at time instances T0, T1 and T2 respectively. The pictures shown in Fig. 5 can be the color images or the depth images. The derived candidates are termed as spatial candidate (spatial MVP), temporal candidate (temporal MVP) and inter-view candidate (inter-view MVP). In particular, for temporal and inter-view candidate derivation, the information to

indicate whether the co-located picture is in list 0 or list 1 can be implicitly derived or explicitly transmitted in different levels of syntax (e.g. sequence parameter set (SPS), picture parameter set (PPS), adaptive parameter set (APS), Slice header, CU level, largest CU level, leaf CU level, or PU level). The position of the inter-view co-located block can be determined by simply using the same position of the current block or using a Global Disparity Vector (GDV) or warping the current block onto the co-located picture according to the depth information.

[0027] The candidate can also be derived based on the vector corresponding to warping the current block onto the co-located picture according to the depth information. Accordingly, the candidate that is derived using the depth information is termed as *depth candidate*.

[0028] The motion vector competition (MVC) based scheme is then applied to select one Motion Vector Predictor (MVP)/Disparity Vector Predictor (DVP) among a candidate set of MVPs/DVPs which includes spatial, temporal, inter-view, and depth candidates. The index of the selected candidate is then transmitted to the decoder.

[0029] When a block is encoded as a Merge or Skip mode, the merge index is incorporated in the bitstream to indicate which MVP/DVP among the MVP/DVP candidate set is used for this block to be merged. The MVP/DVP candidate includes the spatial candidates (spatial MVPs/DVPs), temporal candidates (temporal MVPs/DVPs), inter-view candidates (inter-view MVPs/DVPs) and depth candidates. Bitrate associated with motion information is reduced by sharing the motion information with other coded blocks, where each merged PU reuses the MV/DV, prediction dimension, prediction direction, and reference picture index of the selected candidate. A merge index is transmitted to the decoder to indicate which candidate is selected for the Merge mode.

[0030] Various embodiments of the present invention to derive spatial candidate are disclosed herein. In one embodiment for spatial candidate derivation, the spatial candidate is derived from the MVs of the neighboring blocks if the spatial candidate is used to predict motion vectors. Similarly, the spatial candidate can also be derived from the DVs of the neighboring blocks if the spatial candidate is used to predict the disparity vector.

[0031] In another embodiment of the present invention for the spatial candidate derivation, the spatial candidate can be derived from the MVs and DVs of the neighboring blocks if the spatial candidate is used to predict motion vectors. Similarly, the spatial candidate can also be derived from the MVs and DVs of the neighboring blocks if the spatial candidate is used to predict the disparity vector.

[0032] In yet another embodiment of the present invention for the spatial candidate derivation,

the spatial candidate derived based on MV or MV/DV of neighboring blocks according to the above embodiments can be further used to derive the spatial candidate. When the target reference picture is identified as indicated by the given reference picture index of the given reference list, the spatial candidates can be derived from an MV/ DV pointing to the target reference picture either from the given reference list or the other reference list. For example, if all the neighboring blocks do not have the MV/DV pointing to the target reference in the given reference list, the candidate can be derived as the first available MV/DV pointing to the target reference picture in the other reference list from the neighboring blocks.

[0033] In an embodiment similar to the above embodiment, the spatial candidate derived based on MV or MV/DV of neighboring blocks according to the above embodiments can be further used to derive the spatial candidate. When the target reference picture is identified as indicated by the given reference picture index of the given reference list, the spatial candidates can be derived from an MV/ DV pointing to the target reference picture or from an MV/DV pointing to the reference picture other than target reference picture in the same given reference list. For example, if all the neighboring blocks do not have the MV/DV pointing to the target reference picture, the candidate can be derived as the scaled MV/DV based on the first available MV pointing to the other reference pictures from the neighboring blocks.

[0034] In another embodiment similar to the above embodiment, the spatial candidate derived based on MV or MV/DV of neighboring blocks according to the above embodiments can be further used to derive spatial candidate. When the target reference picture is identified as indicated by the given reference picture index of the given reference list, the spatial candidates can be derived from the other reference list or other reference picture index based on the following order:

- Search MV/DV pointing to the target reference picture within the given reference list;
- Search MV/DV pointing to the target reference picture within the other reference list;
- Search MV/DV pointing to the other reference pictures within the given reference list.

The derived MV/DV is then scaled according to the temporal distance/inter-view distance; and

- Search MV/DV pointing to the other reference pictures within the other reference list.

The derived MV/DV is then scaled according to the temporal distance/inter-view distance.

[0035] For the spatial candidate derivation for Merge and Skip mode, the prediction information of the spatial candidate includes the prediction dimension (Temporal or Inter-View), prediction direction (L0/L1 or Bi-prediction), reference picture index and MVs/DVs. The information of the spatial candidate directly reuses the prediction information of the selected neighboring block

used to derive the spatial candidate. The prediction information can be directly used by the current PU if that spatial candidate is selected.

[0036] Various embodiments of the present invention to derive temporal candidate are also disclosed herein. In one embodiment for temporal candidate derivation, the temporal candidate is derived from the MVs of the temporal co-located blocks if the temporal candidate is used to predict motion vectors. Similarly, the temporal candidate is derived from the DVs of the temporal co-located blocks if the temporal candidate is used to predict the disparity vector.

[0037] In another embodiment for temporal candidate derivation, the temporal candidate can be derived from the MVs and DVs of the temporal co-located blocks if the temporal candidate is used to predict motion vectors. Similarly, the temporal candidate can be derived from the MVs and DVs of the temporal co-located blocks if the temporal candidate is used to predict the disparity vector.

[0038] In yet another embodiment of the present invention for the temporal candidate derivation, the temporal candidate derived based on the MV or MV/DV of the temporal co-located blocks according to the above embodiments can be further used to derive the temporal candidate. For example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived by searching the MVs/DVs with the associated reference list same as the given reference list. The derived MV/DV is then scaled according to the temporal distance/inter-view distance. In another example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived by searching MV/DV crossing the current picture in the temporal/view dimension. The derived MV/DV is then scaled according to the temporal distance/inter-view distance. In yet another example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived according to the following order:

1. Search MV/DV crossing the current picture in the temporal/view dimension; and
2. If both MVs/DVs cross the current picture or both do not cross, the MV/DV with same reference list as the current list will be chosen.

The derived MV/DV is then scaled according to the temporal distance/inter-view distance.

[0039] In yet another embodiment of the present invention for the temporal candidate derivation, the temporal candidate derived based on MV or MV/DV of temporal co-located blocks according to the above embodiments can be further used to derive the temporal candidate. When the reference list is provided, the MV/DV candidate can be derived based on the MV/DV from list 0 or list 1 of the co-located block in the co-located picture in list 0 or list 1 according to a given priority order. The priority order is predefined, implicitly derived or explicitly transmitted to the decoder. The derived MV/DV is then scaled according to the

temporal distance/inter-view distance. An example of the priority order is shown as follows, where the current list is assumed to be list 0:

1. Scaled MV/DV from list 0 of the co-located block of the co-located picture in list 1;
2. Scaled MV/DV from list 1 of the co-located block of the co-located picture in list 0;
3. Scaled MV/DV from list 0 of the co-located block of the co-located picture in list 0; and
4. Scaled MV/DV from list 1 of the co-located block of the co-located picture in list 1.

[0040] For the temporal candidate derivation for Merge and Skip mode, if the prediction dimension of the temporal co-located block is inter-view dimension, the prediction information, such as the prediction dimension (Temporal or Inter-view), prediction direction (L0/L1 or Bi-prediction), reference picture index and DVs of the temporal co-located block can be directly used by the current PU if the temporal candidate is selected.

[0041] For the temporal candidate derivation for Merge and Skip mode, if the prediction dimension of the temporal co-located block is temporal dimension, the reference picture index can be transmitted explicitly or derived implicitly. The prediction information, such as the prediction dimension, prediction direction (L0/L1 or Bi-prediction) and MVs of the temporal co-located block can be directly used by the current PU if the temporal candidate is selected. The derived MV is then scaled according to the temporal distance. For the derivation of the reference picture index, it can be implicitly derived based on the median/mean or the majority of the reference picture indices from the neighboring blocks.

[0042] Various embodiments of the present invention to derive inter-view candidates are also disclosed herein. In one embodiment for inter-view candidate derivation, the inter-view candidate is derived from MVs of the inter-view co-located blocks if the inter-view candidate is used to predict a motion vector. Similarly, the inter-view candidate is derived from DVs of the inter-view co-located blocks if the inter-view candidate is used to predict a disparity vector. The position of the co-located block in inter-view dimension can be determined by using the same position of the current block in the inter-view co-located picture, using a Global Disparity Vector (GDV), or warping the current block onto the inter-view co-located picture according to the depth information.

[0043] In another embodiment for inter-view candidate derivation, the inter-view candidate can be derived from MVs and DVs of the inter-view co-located blocks if the inter-view candidate is used to predict the motion vector. Similarly, the inter-view candidate can be derived from the MVs and DVs of the inter-view co-located blocks if the inter-view candidate is used to predict the disparity vector. The position of the co-located block in inter-view dimension can be determined by using the same position of the current block in the inter-view co-located picture,

using a Global Disparity Vector (GDV), or warping the current block onto the inter-view co-located picture according to the depth information.

[0044] In yet another embodiment of the present invention for the inter-view candidate derivation, the inter-view candidate derived based on MV or MV/DV of the inter-view co-located blocks according to the above embodiments can be further used to derive the inter-view candidate. For example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived by searching the MVs/DVs with associated reference list same as the given reference list. The derived MV/DV is then scaled according to the temporal distance/inter-view distance. In another example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived by searching the MV/DV that crosses the current picture in the temporal/inter-view dimension. The derived MV/DV is then scaled according to the temporal distance/inter-view distance. In yet another example, when the reference list and the co-located picture are provided, the MV/DV candidate can be derived based on the following order:

1. Search the MV/DV that crosses the current picture in the temporal/inter-view dimension; and
2. If both MVs/DVs cross or both do not cross the current picture, the MV/DV with same reference list as the current list will be chosen.

The derived MV/DV is then scaled according to temporal distance/inter-view distance.

[0045] In yet another example, when the reference list is provided, the MV/DV candidate can be derived based on the MV/DV from list 0 or list 1 of the co-located block in the co-located picture in list 0 or list 1 according to a given priority order. The priority order can be pre-defined, implicitly derived, or explicitly transmitted to the decoder. The derived MV/DV is then scaled according to the temporal distance/inter-view distance. An example of the priority order is as follows, where the current list is assumed to be list 0:

1. Scaled MV/DV from list 0 of the co-located block of the co-located picture in list 1;
2. Scaled MV/DV from list 1 of the co-located block of the co-located picture in list 0;
3. Scaled MV/DV from list 0 of the co-located block of the co-located picture in list 0; and
4. Scaled MV/DV from list 1 of the co-located block of the co-located picture in list 1.

[0046] For the inter-view candidate derivation for Merge and Skip mode, if the prediction dimension of the inter-view co-located block is temporal dimension, the prediction information, such as prediction dimension, prediction direction (L0/L1 or Bi-prediction), reference picture index and MVs of the inter-view co-located block can be used directly by the current PU if the inter-view candidate is selected.

[0047] The position of the co-located block in inter-view dimension can be determined using the same position of the current block in the inter-view co-located picture, using a global disparity vector (GDV), or warping the current block onto the inter-view co-located picture according to the depth information.

[0048] For the inter-view candidate derivation for Merge and Skip mode, if the prediction dimension of the inter-view co-located block is inter-view dimension, the reference picture index could be transmitted explicitly or derived implicitly. The prediction information, such as prediction dimension, prediction direction (L0/L1 or Bi-prediction) and DVs of the inter-view co-located block can be used directly by the current PU if the inter-view candidate is selected. The derived DV is then scaled according to the inter-view distance. For the derivation of reference picture index, it can be implicitly derived based on the median/mean or the majority of the reference picture indices from the neighboring blocks.

[0049] The position of the co-located block in inter-view dimension can be determined using the same position of current block in the inter-view co-located picture or using a Global Disparity Vector (GDV) or warping the current block onto the inter-view co-located picture according to the depth information.

[0050] Embodiments of spatial candidate derivation, temporal candidate derivation or inter-view candidate derivation for 3D video coding according to 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 a circuit integrated into a video compression chip or program codes integrated into video compression software to perform the processing described herein. An embodiment of the present invention may also be program codes 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 codes 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.

[0051] 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.

1. A method of deriving MV/MVP (motion vector or motion vector predictor) or DV/DVP (disparity vector or disparity vector predictor) associated Skip mode, Merge mode or Inter mode for a block of a current picture in three-dimensional video coding using prediction dimension
5 consisting of temporal prediction and inter-view prediction, the method comprising:

determining one or more spatial candidates, one or more temporal candidates, or both said one or more spatial candidates and said one or more temporal candidates, wherein said one or more spatial candidates are associated with each of one or more neighboring blocks of the block; and wherein said one or more temporal candidates are associated with each of one or more
10 temporal co-located blocks of one or more temporal co-located pictures of the block;

determining one or more inter-view candidates associated with an inter-view co-located block associated with one or more inter-view co-located pictures corresponding to the block;

selecting the MV/MVP or DV/DVP from said one or more spatial candidates, said one or more temporal candidates and said one or more inter-view candidates; and

15 providing the selected MV/MVP or DV/DVP to the block.

2. The method of Claim 1, wherein the selected MVP or DVP is used for the Inter mode in the three-dimensional video coding.

20 3. The method of Claim 1, wherein the selected MV or DV is used for the Merge or the Skip mode in the three-dimensional video coding.

4. The method of Claim 1, wherein the spatial candidate is derived from the MV or a combination of the MV and the DV associated with the neighboring block if the spatial
25 candidate is used for deriving the MV/MVP; and wherein the spatial candidate is derived from the DV or a combination of the MV and the DV associated with the neighboring block if the spatial candidate is used for deriving the DV/DVP.

5. The method of Claim 4, wherein the spatial candidate is derived from said one or more
30 neighboring blocks for a given prediction dimension and a target reference picture as indicated by a given reference picture index of a given reference list, wherein said one or more neighboring blocks are associated with the target reference picture from the given reference list or other reference list, or associated with other reference picture from the given reference list or the other reference list.

6. The method of Claim 5, wherein the spatial candidate is derived based on a first available MV/DV in the given prediction dimension from said one or more neighboring blocks according to a search order, wherein the MV/DV of said one or more neighboring blocks pointing to the target reference picture in the given reference list is checked for availability before the MV/DV of said one or more neighboring blocks pointing to the other reference picture in the given reference list.

7. The method of Claim 5, wherein the spatial candidate is derived based on a first available MV/DV in the given prediction dimension from said one or more neighboring blocks according to a search order, wherein the MV/DV of said one or more neighboring blocks pointing to the target reference picture in the given reference list is checked for availability before the MV/DV of said one or more neighboring blocks pointing to the target reference picture in the other reference list.

8. The method of Claim 5, wherein the given prediction dimension, the reference picture index, or the given reference list is explicitly transmitted or implicitly derived.

9. The method of Claim 1, wherein the temporal candidate is derived from the MV or a combination of the MV and the DV associated with said one or more temporal co-located blocks of said one or more temporal co-located pictures if the temporal candidate is used for deriving the MV/MVP; and wherein the temporal candidate is derived from the DV or a combination of the MV and the DV associated with said one or more temporal co-located blocks of said one or more temporal co-located pictures if the temporal candidate is used for deriving the DV/DVP.

10. The method of Claim 9, wherein the temporal candidate is derived from said one or more temporal co-located blocks of said one or more temporal co-located pictures for a given prediction dimension and a target reference picture as indicated by a given reference picture index of a given reference list, wherein said one or more temporal co-located blocks of said one or more temporal co-located pictures are associated with the target reference picture from the given reference list or other reference list, or associated with other reference picture from the given reference list or the other reference list.

11. The method of Claim 10, wherein the temporal candidate is derived based on a first

available MV/DV in the given prediction dimension from said one or more temporal co-located blocks according to a search order, wherein the MV/DV of said one or more temporal co-located blocks crossing the current picture is checked for availability first.

5 12. The method of Claim 11, wherein if both the MV/DV of said one or more temporal co-located blocks corresponding to the given reference list and the MV/DV of said one or more temporal co-located blocks corresponding to the other reference list cross or do not cross the current picture, the MV/DV of said one or more temporal co-located blocks corresponding to the given reference list is checked for availability.

10 13. The method of Claim 10, wherein the temporal candidate is derived based on a first available MV/DV from said one or more temporal co-located blocks according to a search order, wherein the search order is related to reference list associated with pointing direction of the MV/DV or the reference list associated with said one or more temporal co-located pictures.

15 14. The method of Claim 10, wherein a flag is used to indicate which of said one or more temporal co-located pictures is used to determine said one or more temporal co-located blocks.

20 15. The method of Claim 14, wherein the flag is in a sequence level, a picture level or a slice level of a video bitstream.

25 16. The method of Claim 10, wherein the inter-view prediction or the temporal prediction used, the reference picture index, or the given reference list is explicitly transmitted or implicitly derived.

30 17. The method of Claim 1, wherein the inter-view candidate is derived from the MV or a combination of the MV and the DV associated with said one or more inter-view co-located blocks of said one or more inter-view co-located pictures if the inter-view candidate is used for deriving the MV/MVP; and wherein the inter-view candidate is derived from the DV or a combination of the MV and the DV associated with said one or more inter-view co-located blocks of said one or more inter-view co-located pictures if the inter-view candidate is used for deriving the DV/DVP.

18. The method of Claim 17, wherein the inter-view candidate is derived from said one or

more inter-view co-located blocks of said one or more inter-view co-located pictures for a given prediction dimension and a target reference picture as indicated by a given reference picture index of a given reference list, wherein said one or more inter-view co-located blocks of said one or more inter-view co-located pictures are associated with the target reference picture from the given reference list or other reference list, or associated with other reference picture from the given reference list or the other reference list.

19. The method of Claim 18, wherein a flag is used to indicate which of said one or more inter-view co-located pictures is used to determine said one or more inter-view co-located blocks.

20. The method of Claim 19, wherein the flag is in a sequence level, a picture level or a slice level of a video bitstream.

21. The method of Claim 18, wherein position of the inter-view co-located block is derived based on a global disparity vector between the current picture and the inter-view co-located picture corresponding to the inter-view co-located block.

22. The method of Claim 18, wherein position of the inter-view co-located block is derived by warping the block of the current picture according to depth information.

23. The method of Claim 18, wherein the inter-view candidate is derived based on a first available MV/DV from said one or more the inter-view co-located blocks according to a search order, wherein the MV/DV of said one or more the inter-view co-located blocks crossing the current picture in the given prediction dimension is checked for availability first.

24. The method of Claim 23, wherein if both the MV/DV of said one or more inter-view co-located blocks corresponding to the given reference list and the MV/DV of said one or more inter-view co-located blocks corresponding to the other reference list cross or do not cross the current picture in the given prediction dimension, the MV/DV of said one or more inter-view co-located blocks corresponding to the given reference list is checked for availability.

25. The method of Claim 18, wherein the inter-view candidate is derived based on a first available MV/DV from said one or more inter-view co-located blocks according to a search order, wherein the search order is related to reference list associated with pointing direction of the MV/DV or the reference list associated with said one or more inter-view co-located pictures.

5

26. The method of Claim 18, wherein the given prediction dimension, the reference picture index, or the given reference list is explicitly transmitted or implicitly derived.

27. The method of Claim 1, wherein, if the inter-view prediction is used, the inter-view candidate is derived as the DV by warping the block of the current picture onto a corresponding inter-view co-located block associated with said one or more inter-view co-located pictures based on depth information.

10

28. The method of Claim 1, wherein the prediction dimension is implicitly derived based on median, mean, or majority of the prediction dimension of said one or more neighboring blocks.

15

29. The method of Claim 1, wherein the MV points to a target reference picture indicated by a reference picture index of a given reference list, and the reference picture index is implicitly derived based on median, mean, or majority of reference picture indices of said one or more neighboring blocks.

20

30. The method of Claim 29, wherein the given reference list is implicitly derived based on median, mean, or majority of the reference lists of said one or more neighboring blocks.

25

31. The method of Claim 1, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the spatial candidate is derived from said one or more neighboring blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more neighboring blocks is directly used by the block of the current picture if the spatial candidate is selected.

30

32. The method of Claim 1, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the temporal candidate is derived from said one or more temporal co-located blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more temporal co-located blocks is directly used by the block of the current picture if the prediction dimension of the temporal co-located block is the inter-view prediction.

33. The method of Claim 1, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the temporal candidate is derived from said one or more temporal co-located blocks; wherein a reference picture index is explicitly transmitted or implicitly derived if the prediction dimension of the temporal co-located block is the temporal prediction; wherein, after the reference picture index is explicitly transmitted or implicitly derived, prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, and the MV/DV selected from one of said one or more temporal co-located blocks is directly used by the block of the current picture if the temporal candidate is selected; and wherein the MV or the DV selected is scaled according to a temporal distance.

34. The method of Claim 1, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the inter-view candidate is derived from said one or more inter-view co-located blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more inter-view co-located blocks is directly used by the block of the current picture if the prediction dimension of the inter-view co-located block is the temporal prediction.

35. The method of Claim 1, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the inter-view candidate is derived from said one or more inter-view co-located blocks; wherein a reference picture index is explicitly transmitted or implicitly

derived if the prediction dimension of the inter-view co-located block is the inter-view prediction; wherein, after the reference picture index is explicitly transmitted or implicitly derived, prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, and the MV/DV selected from one of said one or more inter-view co-located blocks is directly used by the block of the current picture if the inter-view candidate is selected; and wherein the MV or the DV selected is scaled according to an inter-view distance.

36. The method of Claim 1, further comprising determining one or more depth candidates derived based on a vector corresponding to warping the current block onto one or more inter-view co-located pictures according to depth information corresponding to the block, and selecting the MV/MVP or DV/DVP from said one or more spatial candidates, said one or more temporal candidates, said one or more inter-view candidates, and said one or more depth candidates.

37. An apparatus for deriving MV/MVP (motion vector or motion vector predictor) or DV/DVP (disparity vector or disparity vector predictor) associated Skip mode, Merge mode or Inter mode for a block of a current picture in three-dimensional video coding using prediction dimension consisting of temporal prediction and inter-view prediction, the apparatus comprising:

means for determining one or more spatial candidates, one or more temporal candidates, or both said one or more spatial candidates and said one or more temporal candidates, wherein said one or more spatial candidates are associated with each of one or more neighboring blocks corresponding to the block; and wherein said one or more temporal candidates are associated with each of one or more temporal co-located blocks of one or more temporal co-located pictures corresponding to the block;

means for determining one or more inter-view candidates associated with an inter-view co-located block associated with one or more inter-view co-located pictures corresponding to the block;

means for selecting the MV/MVP or the DV/DVP from said one or more spatial candidates, said one or more temporal candidates and said one or more inter-view candidates; and

means for providing the selected MV/MVP or DV/DVP to the block.

38. The apparatus of Claim 37, wherein the spatial candidate is derived from the MV or a combination the MV and the DV associated with the neighboring block if the spatial candidate is used for deriving the MV/MVP; and wherein the spatial candidate is derived from the DV or a combination of the MV and the DV associated with the neighboring block if the spatial candidate is used for deriving the DV/DVP.

39. The apparatus of Claim 37, wherein the temporal candidate is derived from the MV or a combination the MV and the DV associated with said one or more temporal co-located blocks of said one or more temporal co-located pictures if the temporal candidate is used for deriving the MV/MVP; and wherein the temporal candidate is derived from the DV or a combination of the MV and the DV associated with said one or more temporal co-located blocks of said one or more temporal co-located pictures if the temporal candidate is used for deriving the DV/DVP.

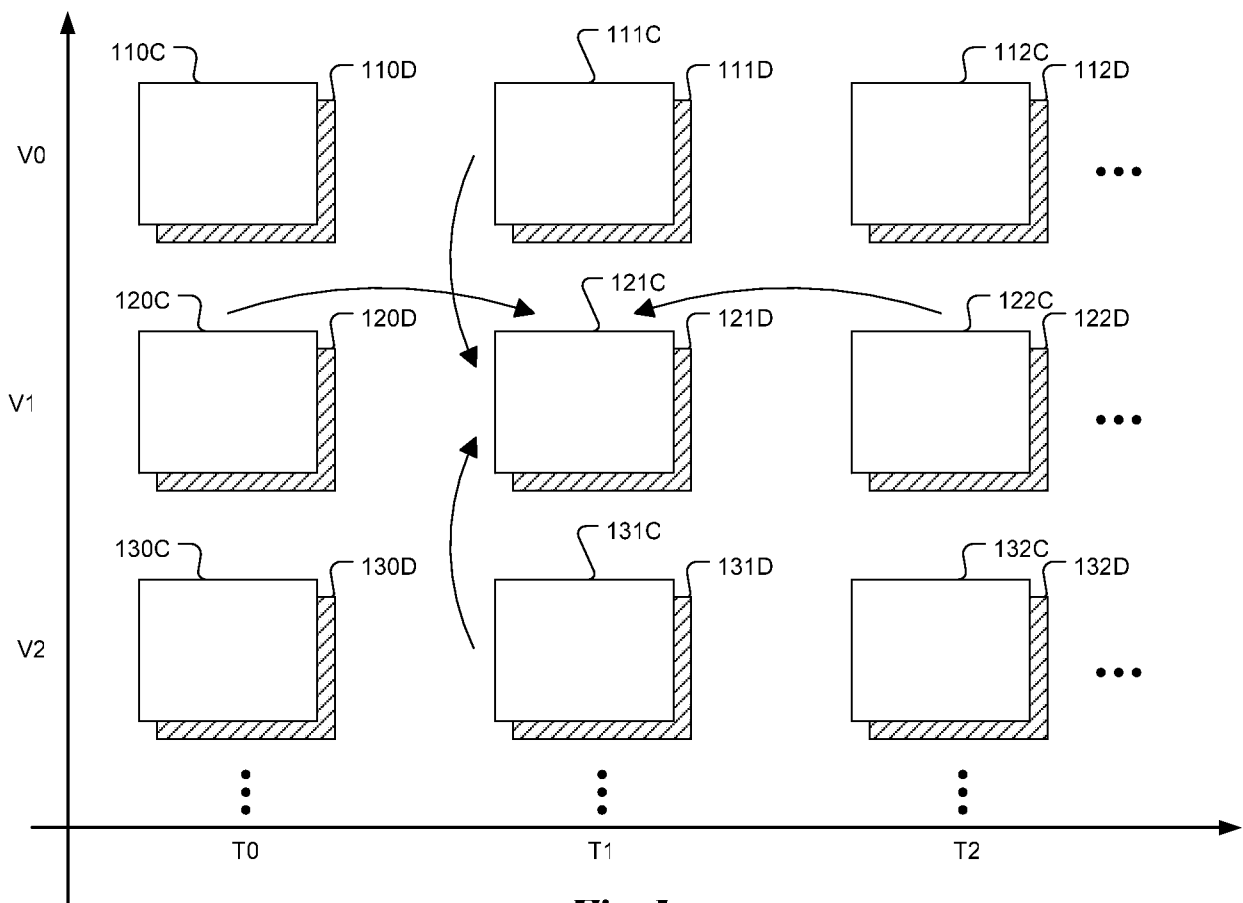
40. The apparatus of Claim 37, wherein the inter-view candidate is derived from the MV or a combination the MV and the DV associated with said one or more inter-view co-located blocks of said one or more inter-view co-located pictures if the inter-view candidate is used for deriving the MV/MVP; and wherein the inter-view candidate is derived from the DV or a combination of the MV and the DV associated with said one or more inter-view co-located blocks of said one or more inter-view co-located pictures if the inter-view candidate is used for deriving the DV/DVP.

41. The apparatus of Claim 37, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the spatial candidate is derived from said one or more neighboring blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more neighboring blocks is directly used by the block of the current picture if the spatial candidate is selected.

42. The apparatus of Claim 37, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the temporal candidate is derived from said one or more temporal co-located blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more temporal co-located blocks is directly used by the block of the current picture if the prediction dimension of

the temporal co-located block is the inter-view prediction.

- 5 43. The apparatus of Claim 37, wherein the MV or the DV is associated with the Merge mode or the Skip mode; wherein the inter-view candidate is derived from said one or more inter-view co-located blocks; and wherein prediction information including the prediction dimension, prediction direction consisting of reference list 0, reference list 1 and Bi-prediction, reference picture index, and the MV/DV selected from one of said one or more inter-view co-located blocks is directly used by the block of the current picture if the prediction dimension of the inter-view co-located block is the temporal prediction.

**Fig. 1**

V0

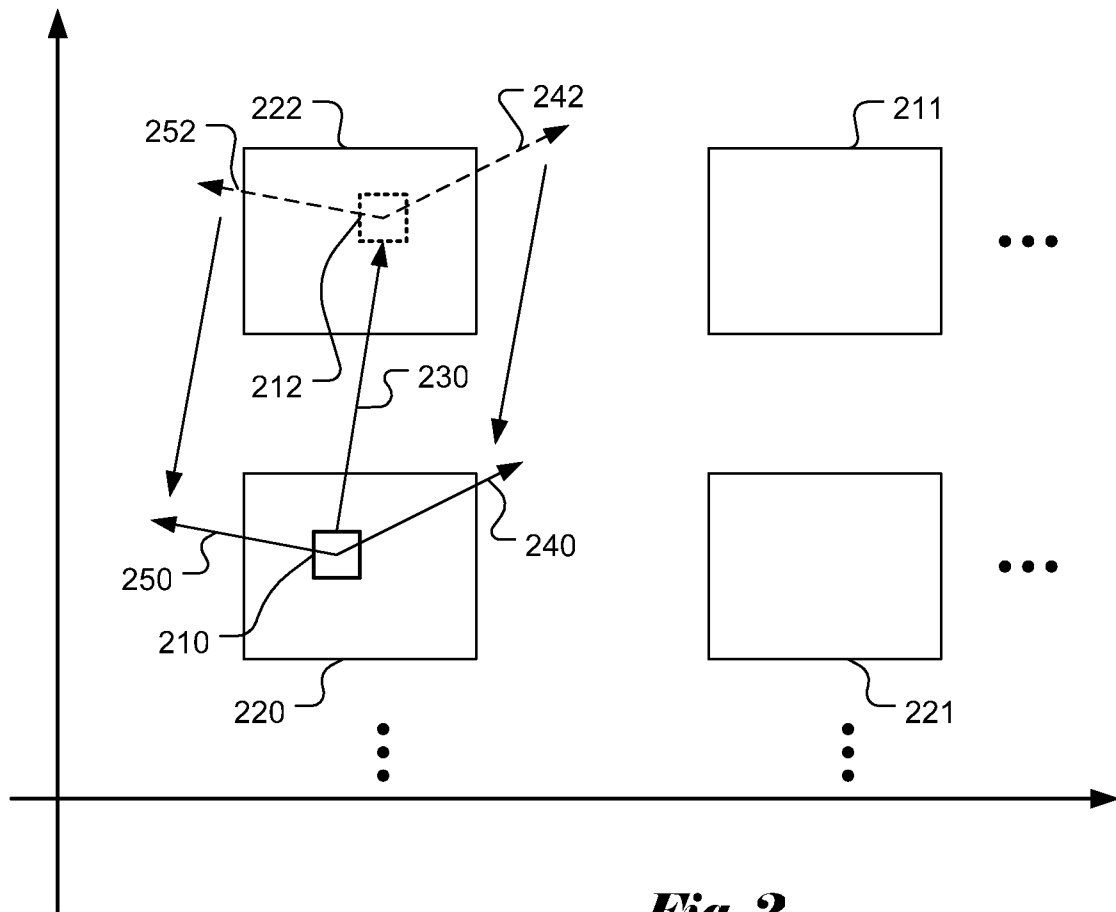


Fig. 2

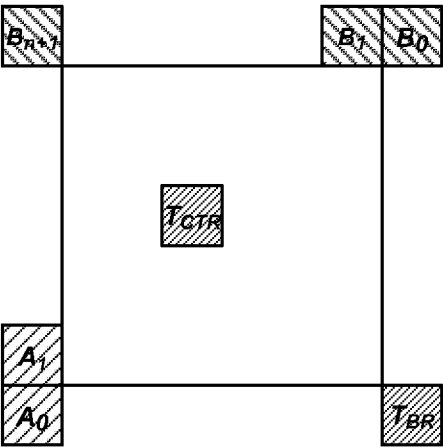


Fig. 3

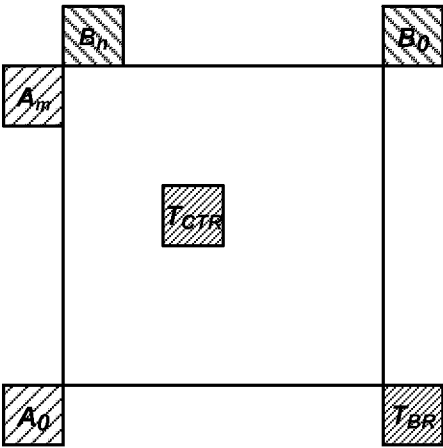
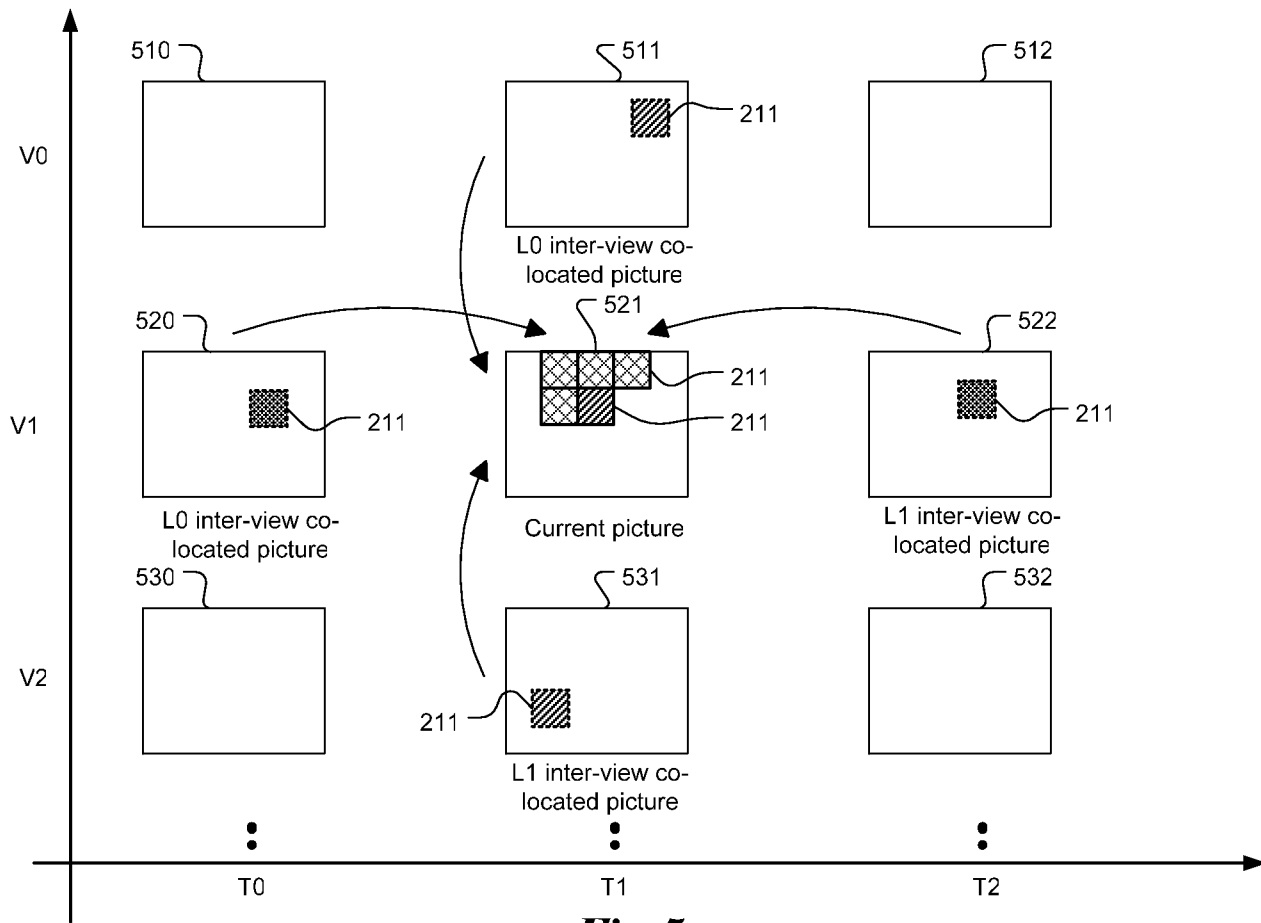


Fig. 4

**Fig. 5**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2012/076643

A. CLASSIFICATION OF SUBJECT MATTER

H04N 7/32 (2006.01) 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)

IPC: H04N7/-

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)

CNABS/CNKI/DWPI: motion, vector, disparity, predict+, derive+, mode?, (three w dimension+), video, view, spatial, temporal, block?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN101917619A (UNIV ZHEJIANG) 15 Dec. 2010(15.12.2010) the whole document	1-43
A	US2003202592A1 (Sohn et al.) 30 Oct. 2003(30.10.2003) the whole document	1-43

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search
23 Jul. 2012 (23.07.2012)Date of mailing of the international search report
23 Aug. 2012 (23.08.2012)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451Authorized officer
SHI, Yayun
Telephone No. (86-10)62411643

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2012/076643

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN101917619A	15.12.2010	none	
US2003202592A1	30.10.2003	US6999513B2	14.02.2006
		KR20030083285A	30.10.2003
		KR481732B	11.04.2005