

(51) International Patent Classification:
H04N 7/32 (2006.01)(21) International Application Number:
PCT/CN2012/084580(22) International Filing Date:
14 November 2012 (14.11.2012)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): **MEDI-ATEK SINGAPORE PTE. LTD.** [SG/SG]; No.1 Fusionopolis Walk, #03-01 Solaris, Singapore (SG).

(72) Inventors; and

(71) Applicants (for US only): **ZHANG, Kai** [CN/CN]; Institute of Computing Technology of the Chinese Academy of Sciences, No.6 Kexueyuan South Road Zhongguancun, Haidian District, Beijing 100080 (CN). **AN, Jicheng** [CN/CN]; No. 44 Zhanlanguan Road, Xicheng District, Beijing 100037 (CN). **LIN, Jian-Liang** [CN/CN]; No.133, Neipi Rd., Nanjian Vil. Su'ao Township, Yilan County, Taiwan 270 (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, 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, PA, 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).

Published:

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

(54) Title: METHODS FOR RESIDUAL PREDICTION WITH PSEUDO RESIDUES IN 3D VIDEO CODING

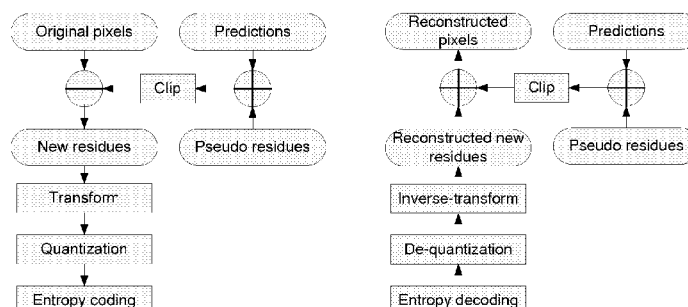


Fig.9

(57) **Abstract:** Methods of the pseudo residual prediction and Methods of DV/MV estimation for residual prediction for multi-view video coding and 3D video coding are disclosed. It is proposed that the pseudo residues are constructed using reconstructed samples and are used to predict the original residues in dependent views. The method to construct the pseudo residues for a block depends on the prediction method of this block. In addition, the estimated DV or MV of a block can be obtained by utilizing DVs or MVs of its spatial and temporal neighboring blocks, and exploring original or reconstructed samples in the current picture and reconstructed pictures. The proposed DV estimation method can also be used in other inter-view residual prediction method.



METHODS FOR RESIDUAL PREDICTION WITH PSEUDO RESIDUES IN 3D VIDEO CODING

FIELD OF INVENTION

5 The invention relates generally to Three-Dimensional (3D) video processing. In particular, the present invention relates to methods for the residual prediction with pseudo residues in the 3D video coding.

BACKGROUND OF THE INVENTION

10 3D video coding is developed for encoding or decoding video data of multiple views simultaneously captured by several cameras. Since all cameras capture the same scene from different viewpoints, multi-view video data contains a large amount of inter-view redundancy. To share the previously encoded residual information of reference views, an additional tool called inter-view residual prediction has been integrated to the current HEVC based 3D video
15 coding [1].

 The basic principle of the inter-view residual prediction (IVRP) is illustrated in Fig. 1. The inter-view residual prediction is based on a disparity vector derivation for the current block. A disparity vector is firstly derived for a current prediction unit (PU) and then the residual block in the reference view that is referenced by the disparity vector is used for predicting the residual
20 of the current PU.

 In the current 3D-HEVC, a DV derived from the spatial or temporal neighboring blocks is used as the estimated DV for the current block [2]. As a complement, DV-MCP [3] is adopted to supply an estimated DV when none of the neighboring blocks has a DV.

 A more detailed illustration of the concept for deriving a reference block location inside
25 the reference view is given in Fig. 2. A disparity vector is firstly determined for current PU. The disparity vector is added to the location of the top-left sample of the current PU yielding the location of the top-left sample of the reference block. Then, similar as for motion compensation, the block of residual samples in a reference view located at the derived reference location is subtracted from the current residual and only the resulting difference signal is transform coded.
30 If the disparity vector points to a sub-sample location, the residual prediction signal is obtained by interpolating the residual samples of the reference view using a bi-linear filter.

 The usage of the inter-view residual prediction can be adaptively controlled at the prediction unit (PU) level. An IVRP on/off control flag is signaled as part of the coding unit

(CU) syntax when all the following conditions are true:

1. The current CU is a texture CU in dependent view, and the current slice is not an I-slice.
 2. The current CU has at least one PU using motion-compensated prediction.
 3. One or more transform units (TUs) covered or partially covered by the reference block
- 5 in the reference view are non-intra coded and contain at least one nonzero coded block flag (CBF).

If the IVRP on/off control flag is signaled as 0 or not signaled for a CU, the IVRP is off for the CU, i.e., the residual of the CU is conventionally coded using the HEVC transform coding. Otherwise, if the IVRP on/off flag is signaled as 1 for a CU, then each PU in it should

10 determine whether to use the IVRP or not according to the reference picture type as follows:

1. If a PU only uses motion-compensated prediction, the IVRP is enabled in the PU.
 2. If a PU only uses disparity-compensated prediction, the IVRP is disabled in the PU.
 3. If the current PU is bi-prediction and one direction is motion compensated prediction and the other direction is disparity-compensated prediction, the IVRP is enabled, but the
- 15 reference residual used in IVRP is multiplied by 1/2.

An advanced residual prediction (ARP) method is proposed to improve IVRP [4]. It is proposed that motion of current view is applied to the corresponding block in a reference view to have a better prediction performance. Moreover, additional weighting factor is introduced to further compensate quality difference between different views.

20 Fig. 3 illustrates the prediction structure of ARP in multi-view video coding. As shown in Fig. 3, D_c represents the current block in the current view (view 1), B_c , and D_r denote the representation of D_c in the reference view (view 0) at time T_j and D_c 's temporal prediction from the same view (view 1) at time T_i . V_D denotes the motion from D_c to D_r . Since D_c and B_c are actually projections of the same object in two different views, these two blocks should share the

25 same motion information. Therefore, B_c 's temporal prediction B_r in view 0 at time T_i can be located from B_c by applying the motion information of V_D . The residual of B_c with motion information of V_D is then multiplied by a weighting factor and used as the residual predictor for current residual.

Main procedures of the proposed ARP at the decoder side can be described as follows:

- 30 1. Obtain an estimated disparity vector as specified in the current 3D-HEVC [1][2][3], pointing to a target reference view. Then, in the picture of the reference view within the same access unit, the corresponding block is located by the disparity vector.
2. Re-use the motion information of the current block to derive the motion information for the reference block. Apply motion compensation for the corresponding block based the same

motion vector of current block and derived reference picture in the reference view for the reference block, to derive a residue block. The relationship among current block, corresponding block and motion compensated block is shown in Fig. 4. The reference picture in the reference view (V_0) which has the same POC (Picture Order Count) value as the reference picture of current view (V_m) is selected as the reference picture of the corresponding block.

3. Apply the weighting factor to the residue block to get a weighted residue block and add the values of the weighted residue block to the predicted samples.

Three weighting factors are used in ARP, i.e., 0, 0.5 and 1. The one leading to minimal rate-distortion cost for the current CU is selected as the final weighting factor and the corresponding weighting index (0, 1 and 2 which correspond to weighting factor 0, 1, and 0.5, respectively) is transmitted in the bit-stream at the CU level. All PU predictions in one CU share the same weighting factor. When the weighting factor is equal to 0, ARP is not used for the current CU.

References

[1] H. Schwarz, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, D. Marpe, P. Merkle, K. Muller, H. Rhee, G. Tech, M. Winken, T. Wiegand, "Description of 3D Video Technology Proposal by Fraunhofer HHI (HEVC compatible; configuration B)", ISO/IEC JTC 1/SC 29/WG 11, m22571, Geneva, Switzerland, Nov., 2011.

[2] Li Zhang, Ying Chen, Marta Karczewicz, "3D-CE5.h related: Disparity vector derivation for multiview video and 3DV", Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, m24937, April 2012.

[3] Jaewon Sung, Moonmo Koo, Sehoon Yea, "3D-CE5.h: Simplification of disparity vector derivation for HEVC-based 3D video coding", Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A0126, July 2012, Stockholm.

[4] Li Zhang, Ying Chen, Xiang Li, Marta Karczewicz, "3D-CE5.h related: Advanced residual prediction for multiview coding", Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-B0051, October 2012.

SUMMARY OF THE INVENTION

A method of residual prediction with pseudo residues is proposed. In the proposed method, pseudo residues can be generated using already coded frames or already coded samples in the

current coding frame. Then the original residues in the current block in a dependent view can be predicted by the pseudo residues.

Pseudo residues can be generated in several ways. The way of generating pseudo residues can depend on the reference picture of the current block. Pseudo residues can be generated in a certain way, no matter the current block refers to one or more inter-time reference pictures, refers to one or more inter-view reference pictures, or refers to two or more reference pictures combined with inter-time and inter-view ones.

Estimated Disparity Vector (DV) or estimated Motion Vector (MV) is utilized to generate pseudo residues. These DVs or MVs can be estimated by the spatial or temporal neighboring blocks of the current block on the encoder, as well as on the decoder. Besides, these DVs or MVs can be estimated by searching a range of possible ones to find the optimal one on the encoder, and can be explicitly transmitted to the decoder. Moreover, these DVs or MVs can be estimated by searching a range of possible ones to find the optimal one on the decoder.

The method of pseudo residual prediction can be used in all blocks. Besides, the encoder can determine whether to use pseudo residual prediction on a block, and a flag for this block is transmitted to the decoder to indicate. Moreover, it can be determined whether to use pseudo residual prediction on a block according to the block coding modes of this block and its neighboring blocks.

Other aspects and features of the invention will become apparent to those with ordinary skill in the art upon review of the following descriptions of specific embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention can be more fully understood by reading the subsequent detailed description and examples with references made to the accompanying drawings, wherein:

Fig. 1 is a diagram illustrating the basic concept for inter-view residual prediction.

Fig. 2 is a diagram illustrating the derivation of the location of reference residual block.

Fig. 3 is a diagram illustrating the prediction structure of ARP in multi-view video coding.

Fig. 4 is a diagram illustrating the relationship among current block, corresponding block and motion compensated block in ARP.

Fig. 5 is a diagram illustrating the construction of pseudo residues for an inter block referring to an inter-time reference picture.

Fig. 6 is a diagram illustrating the construction of pseudo residues for an inter block referring to an inter-view reference picture.

Fig. 7 is a diagram illustrating how to use pseudo residues to predict the original residues

on the encoder.

Fig. 8 is a diagram illustrating how to use pseudo residues to predict the original residues on the decoder.

Fig. 9 is a diagram illustrating an implementation to use pseudo residues to predict the original residues on the encoder (left) and on the decoder (right).

Fig. 10 is a diagram illustrating how to use pseudo residues to predict the original residues if an inter block refers to an inter-time reference picture and an inter-view reference picture simultaneously.

Fig. 11 is a diagram illustrating the neighboring samples range for the current block and CR, denoted as A and B, for a decoder to derive the estimated DV. A and B has the same shape.

Fig. 12 is a diagram illustrating the neighboring samples range for the current block and CR, denoted as A and B, for a decoder to derive the estimated MV. A and B has the same shape.

DETAILED DESCRIPTION

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.

Pseudo residues can be generated for an inter block referring to an inter-time reference picture. Inter-time reference picture means that the reference picture and the current picture are in the same view but have different POCs. Fig. 5 illustrates this method. View 0 denotes the base view and view 1 denotes the dependent view. A block in a picture in view 1 is currently coding.

1. First, an estimated DV for the current block referring to an inter-view reference is derived. This inter-view reference denoted as Corresponding Picture (CP) is in the base view and has the same POC as that of the current picture in view 1. A Corresponding Region (CR) in CP for the current block in the current picture is located with the estimated DV. The reconstructed image in the CR is noted as S.

2. Second, a pseudo reference picture (PRP) in the base view which has the same POC as that of the reference picture for the current block is found. The same MV as that of the current block is used on CR to locate the Pseudo Reference Region (PRR) in PRP, whose relative displacement towards the current block is DV+MV. The reconstructed image in the PRR is noted as Q.

3. Third, the pseudo residues PR is calculated as $PR = S - Q$. The operation here is sample-

wised, *i.e.*, $PR[j,i]=S[j,i]-Q[j,i]$, where $PR[j,i]$ is a sample in PR, $S[j,i]$ is a sample in S, $Q[j,i]$ is a sample in Q, and $[j,i]$ is a relative position in the region. In the following descriptions, operations on a region are all sample-wised without specific explanation.

4. PR will be used as the pseudo residual prediction for the current block.

5 Pseudo residues can be generated for an inter block referring to an inter-view reference picture. Inter-view reference picture means that the reference picture and the current picture have the same POCs but are in different views. Fig. 6 illustrates this method. View 0 denotes the base view and view 1 denotes the dependent view. A block in a picture in view 1 is currently coding.

10 1. First, an estimated MV for the current block referring to an inter-time reference is derived. This inter-time reference denoted as Corresponding Picture (CP) is in view 1. A Corresponding Region (CR) in CP is located for the current block using the estimated MV. The reconstructed image in the CR is noted as S.

15 2. Second, a pseudo reference picture (PRP) in the base view which has the same POC as that of CP in view 1 is found. The same DV as that of the current block is used on CR to locate the Pseudo Reference Region (PRR) in PRP, whose relative displacement towards the current block is $MV+DV$. The reconstructed image in the PRR is noted as Q.

3. Third, the pseudo residue PR is obtained as $PR = S - Q$.

4. PR will be used as the pseudo residual prediction for the current block.

20 Some other forms of PR such as $f(PR)$ can replace the original PR to obtain NR or NP, where f can be any function. One example of $f(PR)$ is $a*PR + b$, where coefficients a and b can be transmitted explicitly from the encoder to decoder, or can be deduced implicitly in the same way both on the encoder and the decoder.

Pseudo residues can be generated for an inter block referring to more than one references.
25 Take a block with bi-prediction as an example. In this block, two predictions, P_0 and P_1 , refer to reference picture R_0 and R_1 from $List_0$ and $List_1$ respectively, and $(P_0 + P_1)/2$ is calculated as the bi-prediction. PR_x with $x = 0$ or 1 , can be obtained following the procedures described in paragraphs [0029]-[0031], depending on whether R_x is an inter-time reference picture or an inter-view reference picture. Then $PR = (PR_0 + PR_1)/2$ or $PR = (PR_0 + PR_1 + 1)/2$ is calculated
30 as the pseudo residues for this block.

Pseudo residues PR can be used to predict the original residues OR. The new residues NR are calculated as $NR = OR - PR$. On the encoder, NR replaces OR in the following coding processes such as transform, quantization and entropy coding, as depicted in Fig. 7. On the decoder, the reconstructed NR noted as NR' can be obtained after de-quantization and inverse-

transform. Then the reconstructed samples Y are calculated as $Y = NR' + PR + P$, where P represents the predictions, as depicted in Fig. 8.

There are some other implementations to generate NR . One of them is as follows depicted in Fig. 9. On the encoder, since $OR = X - P$, where X represents the original samples and P represents the predictions, then $NR = X - P - PR$. The new predictions (NP) can be calculated as $NP = P + PR$, therefore $NR = X - NP$. Before calculating NR , Values in NP can be clipped into the value range $[\text{min\_sample\_value}, \text{max\_sample\_value}]$, where min_sample_value and max_sample_value represent the minimal value and the maximal value of original samples respectively. On the decoder, NP is calculated and clipped as on the encoder. Then the reconstructed samples Y are calculated as $Y = NR' + NP$, where NR' is the reconstructed NR . In addition, some other forms of NP such as $f(NP)$ can replace the original NP to obtain NR and Y , where f can be any function. This replacement can happen before or after the clipping process. One example of $f(NP)$ is $a * NP + b$, where coefficients a and b can be transmitted explicitly from the encoder to decoder, or can be deduced implicitly in the same way both on the encoder and the decoder.

A special case arises if an inter block refers to an inter-time reference picture and an inter-view reference picture simultaneously as depicted in Fig. 10. In this case, the inter block has a current DV as well as a current MV. Therefore, the current DV referring to an inter-view reference and the current MV referring to an inter-time reference, instead of the estimated ones, can be used in the pseudo residue generation procedures as described in paragraphs [0029]-[0031]. For the inter-time prediction, PRP_0 and PRR_0 can be obtained with the current inter-view reference and DV replacing the estimated ones. On the other hand, for the inter-view prediction, PRP_1 and PRR_1 can be obtained with the current inter-time reference and MV replacing the estimated ones. It is obvious that PRP_0 and PRP_1 are the same picture, since they are both in the same base view as that of the current inter-view reference picture, and have the same POC as that of the current inter-time reference picture. Moreover, PRR_0 and PRR_1 are the same region, since their relative displacement towards the current block are $DV + MV$ and $MV + DV$ respectively, which point to the same region indeed. So PRR is used to denote PRR_0 and PRR_1 both.

In the above case, the new residues (NR) can be calculated in a special way. The original samples of the current block are noted as X . The reconstructed samples in the inter-view reference region are noted as P . The reconstructed samples in the inter-time reference region are noted as Q . The reconstructed samples in the PRR are noted as S . The original residues are $X - (P + Q)/2$. The pseudo residues are $((P - S) + (Q - S))/2$. Hence, NR are calculated as $X - (P + Q)/2 - ((P - S) + (Q - S))/2$.

$S)+(Q-S))/2$, which equals to $X-P-Q+S$. In a specific implement, the new predictions (NP) can be calculated as $NP=P+Q-S$, therefore $NR=X-NP$ on the encoder. Before calculating NR, values in NP can be clipped into the value range $[\text{min\_sample\_value}, \text{max\_sample\_value}]$, where min_sample_value and max_sample_value represent the minimal value and the maximal value of original samples respectively. On the decoder, NP is calculated and clipped as on the encoder. Then the reconstructed samples Y are calculated as $Y = NR' + NP$, where NR' is the reconstructed NR.

There are several methods to obtain an estimated DV referring to an inter-view reference picture.

1. Spatial neighboring blocks and temporal neighboring blocks of the current block are checked one by one to determine whether they have a DV referring to an inter-view reference picture. The first found DV from neighboring blocks is utilized as the estimated DV for the current block.

2. Spatial neighboring blocks and temporal neighboring blocks of the current block are checked one by one to determine whether they have a DV referring to an inter-view reference picture. All or some of DVs from neighboring blocks is utilized as candidates of the estimated DV for the current block. The encoder can transmit the index of the chosen candidate to the decoder explicitly, or the decoder can derive the chosen candidate implicitly in the same way as the encoder. The estimated index can be transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU).

3. The encoder can transmit the estimated DV to the decoder explicitly. The estimated DV can be transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU). The 'estimated DV' obtained in method 1 or method 2 above can be utilized as a prediction for the real estimated DV. Then only the difference between the estimated DV and the prediction is required to be transmitted.

4. The decoder can derive the estimated DV implicitly. The 'estimated DV' obtained in 1. or 2. above can be utilized as a hint for the derivation of the real estimated DV.

One possible way for an encoder to choose the estimated DV is as follows. For an estimated DV Candidate (EDVC) k , its CR is located as described in [0029]. Suppose S is the original samples in the current block and R is the reconstructed samples in CR, then an evaluating function $E(S, R)$ can be calculated. For each EDVC, its $E(S, R)$ is calculated, and the EDVC with the minimum or maximum $E(S, R)$ is chosen to be the estimated DV. For example, the evaluating function $E(S, R)$ can be a SAD function. $SAD(S, R) = \sum_j \sum_i |s_{i,j} - r_{i,j}|$, where $s_{i,j}$

and $r_{i,j}$ represent a sample on position (i, j) in S and R respectively. For each EDVC, its $SAD(S, R)$ is calculated, and the EDVC with the minimum $SAD(S, R)$ is chosen to be the estimated DV. For method 2 in [0037], EDVCs are obtained from neighboring blocks. For method 3 in [0037], EDVCs can involve any range of DVs.

5 One possible way for a decoder to derive the estimated DV is as follows. For an EDVC k , its CR is located as described in [0029]. As depicted in Fig. 11, the same neighboring samples region are located for the current block and CR, denoted as A and B respectively. Any neighboring reconstructed samples can be involved in this region, as long as A and B has the same shape relative to the current block and CR respectively. Suppose S is the reconstructed samples in A and R is the reconstructed samples in B, then an evaluating function $E(S, R)$ can be calculated. For each EDVC, its $E(S, R)$ is calculated, and the EDVC with the minimum or maximum $E(S, R)$ is chosen to be the estimated DV. For example, the evaluating function $E(S, R)$ can be a SAD function. $SAD(S, R) = \sum_j \sum_i |s_{i,j} - r_{i,j}|$, where $s_{i,j}$ and $r_{i,j}$ represent a sample on position (i, j) in S and R respectively. For each EDVC, its $SAD(S, R)$ is calculated, and the EDVC with the minimum $SAD(S, R)$ is chosen to be the estimated DV. For method 2 in [0037], EDVCs are obtained from neighboring blocks. For method 3 in [0037], EDVCs can involve any range of DVs. The 'estimated DV' obtained as in method 1- method 3 in [0037] can be utilized to determine the range of EDVCs.

20 The vertical component of the estimated DV can be always zero. In this case, the vertical component of the estimated DV need not to be transmitted in [0037] and vertical components of all EDVCs are zero in [0038] and [0039].

There are several methods to determine the inter-view reference picture which is referred by the estimated DV.

25 1. The inter-view reference picture referred can be determined by spatial neighboring blocks and temporal neighboring blocks of the current block implicitly. An EDVC can refer to the same inter-view reference picture as the neighboring block, where it comes from, refers to. Hence, if this EDVC is chosen as the estimated DV, the inter-view reference picture referred is also chosen.

30 2. The inter-view reference picture referred can be pre-specified. The encoder and the decoder can accept the same default picture, or the encoder can transmit a piece of information to the decoder about which inter-view reference is referred by the estimated DV. This information can be in Sequence Parameter Set (SPS), View Parameter Set (VPS), Picture Parameter Set (PPS), Slice header (SH), Macro-block (MB), Coding Unit (CU) or Prediction

Unit (PU). If the appointed inter-view reference picture is different from the inter-view reference picture referred by the neighboring block, the estimated DV or EDVC comes from this block must be scaled to the appointed inter-view reference picture.

There are several methods to obtain an estimated MV referring to an inter-time reference picture.

1. Spatial neighboring blocks and temporal neighboring blocks of the current block are checked one by one to determine whether they have a MV referring to an inter-time reference picture. The first found MV from neighboring blocks is utilized as the estimated MV for the current block.

2. Spatial neighboring blocks and temporal neighboring blocks of the current block are checked one by one to determine whether they have a MV referring to an inter-time reference picture. All or some of MVs from neighboring blocks is utilized as candidates of the estimated MV for the current block. The encoder can transmit the index of the chosen candidate to the decoder explicitly, or the decoder can derive the chosen candidate implicitly in the same way as the encoder. The estimated index can be transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU).

3. The encoder can transmit the estimated MV to the decoder explicitly. The estimated MV can be transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU). The 'estimated MV' obtained in method 1 or method 2 above can be utilized as a prediction for the real estimated MV. Then only the difference between the estimated MV and the prediction is required to be transmitted.

4. The decoder can derive the estimated MV implicitly. The 'estimated MV' obtained in 1. or 2. above can be utilized as a hint for the derivation of the real estimated MV.

One possible way for an encoder to choose the estimated MV is as follows. For an estimated MV Candidate (EMVC) k , its CR is located as described in [0030]. Suppose S is the original samples in the current block and R is the reconstructed samples in CR, then an evaluating function $E(S, R)$ can be calculated. For each EMVC, its $E(S, R)$ is calculated, and the EMVC with the minimum or maximum $E(S, R)$ is chosen to be the estimated MV. For example, the evaluating function $E(S, R)$ can be a SAD function. $SAD(S, R) = \sum_j \sum_i |s_{i,j} - r_{i,j}|$, where $s_{i,j}$ and $r_{i,j}$ represent a sample on position (i, j) in S and R respectively. For each EMVC, its $SAD(S, R)$ is calculated, and the EMVC with the minimum $SAD(S, R)$ is chosen to be the estimated MV. For method 2 in [0042], EMVCs are obtained from neighboring blocks. For method 3 in [0042], EMVCs can involve any range of MVs.

One possible way for a decoder to derive the estimated MV is as follows. For an EMVC k , its CR is located as described in [0030]. As depicted in Fig. 12, the same neighboring samples region are located for the current block and CR, denoted as A and B respectively. Any neighboring reconstructed samples can be involved in this region, as long as A and B has the same shape relative to the current block and CR respectively. Suppose S is the reconstructed samples in A and R is the reconstructed samples in B, then an evaluating function $E(S, R)$ can be calculated. For each EMVC, its $E(S, R)$ is calculated, and the EMVC with the minimum or maximum $E(S, R)$ is chosen to be the estimated MV. For example, the evaluating function $E(S, R)$ can be a SAD function. $SAD(S, R) = \sum_j \sum_i |s_{i,j} - r_{i,j}|$, where $s_{i,j}$ and $r_{i,j}$ represent a sample on position (i, j) in S and R respectively. For each EMVC, its $SAD(S, R)$ is calculated, and the EMVC with the minimum $SAD(S, R)$ is chosen to be the estimated MV. For method 2 in [0042], EMVCs are obtained from neighboring blocks. For method 3 in [0042], EMVCs can involve any range of MVs. The 'estimated MV' obtained as in method 1 -method 3 in [0042] can be utilized to determine the range of EMVCs.

There are several methods to determine the inter-time reference picture which is referred by the estimated MV.

1. The inter-time reference picture referred can be determined by spatial neighboring blocks and temporal neighboring blocks of the current block implicitly. An EMVC can refer to the same inter-time reference picture as the neighboring block, where it comes from, refers to. Hence, if this EMVC is chosen as the estimated MV, the inter-time reference picture referred is also chosen.

2. The inter-time reference picture referred can be pre-specified. The encoder and the decoder can accept the same default picture, or the encoder can transmit a piece of information to the decoder about which inter-time reference is referred by the estimated MV. This information can be in Sequence Parameter Set (SPS), View Parameter Set (VPS), Picture Parameter Set (PPS), Slice header (SH), Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU). If the appointed inter-time reference picture is different from the inter-time reference picture referred by the neighboring block, the estimated MV or EMVC comes from this block must be scaled to the appointed inter-time reference picture.

The DV or MV estimation methods proposed above can be used in the pseudo residual prediction procedure as well as in other residual prediction procedure. For example, the proposed DV estimation methods can replace the original DV estimation method in residual prediction procedure described in [1] or [4].

All the 'samples' discussed above can be sub-samples (at sub-pixels) as well as integer samples (at integer pixels). Sub-samples can be obtained by interpolation. In addition, 'samples' can be chroma samples as well as luma samples.

In the description, regions, such as CR and PRR in [0029] and [0030], A and B in [0039]
5 and [0044], can be located at integer samples or sub-samples. If a region is located at sub-samples, they can be rounded to the nearest integer-samples. Then these integer-samples can be treated as samples from this region and be used in the following process.

The pseudo residual prediction method can be used or not used in a slice, a Macro-block (MB), a Coding Unit (CU) or a Prediction Unit (PU). The encoder can transmit a piece of
10 information to the decoder to indicate whether the pseudo residual prediction method is used or not used in a slice, a Macro-block (MB), a Coding Unit (CU) or a Prediction Unit (PU). In another solution, the pseudo residual prediction method can be used only for Macro-blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) with some specific coding mode or coding type. For example, the pseudo residual prediction method must be used for Macro-
15 blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) with merge or skip mode and must not be used in others. In another example, the pseudo residual prediction method must be used for Macro-blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) which are coded in merge skip mode and use inter-view motion prediction, and must not be used in others.

The pseudo residual prediction and DV or MV estimation methods described above can be
20 used in a video encoder as well as in a video decoder. Embodiments of pseudo residual prediction methods 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
25 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-
30 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 format or style. The software code may also be compiled for different target platform. 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.

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. To the contrary, it is intended to cover various modifications and
5 similar arrangements (as would be apparent to those skilled in the art). Therefore, the scope of the appended claims should be accorded the broadest interpretation so as to encompass all such modifications and similar arrangements.

CLAIMS

1. A method of residual prediction for multi-view video coding or 3D video coding, wherein original residues in dependent views are predicted by pseudo residues.

5

2. A method as claimed in claim 1, wherein an estimated DV is utilized to derive the residues or pseudo residues used to predict the original residues.

3. The method as claimed in claim 1, wherein the pseudo residues are constructed using reconstructed samples, wherein the reconstructed samples are obtained after at least an in-loop or out-loop post-processing.

10

4. The method as claimed in claim 3, wherein the method to construct the pseudo residues for a block depends on the prediction method of this block.

15

5. The method as claimed in claim 4, wherein the pseudo residues are constructed as follows if a block on the dependent view V1 is predicted by an inter-time reference R with a motion vector MV. First, an estimated disparity vector DV referring to an inter-view reference picture CP on the base view V0 is obtained, and a region CR in CP is located, which has the same size as the current block and has a displacement DV from the position of the current block. Second, a reconstructed picture PRP on V0 which has the same POC as that of R is found. Third, a region PRR in PRP is located, which has the same size as the current block and has a displacement DV+MV from the position of the current block. Suppose S represents reconstructed samples in CR and Q represents reconstructed samples in PRR, then pseudo residues PR for the current block are calculated as $PR=S-Q$.

20

25

6. The method as claimed in claim 4, wherein the pseudo residues are constructed as follows if a block on the dependent view V1 is predicted by an inter-view reference R on the base view V0 with a disparity vector DV. First, an estimated motion vector MV referring to an inter-time reference picture CP on V1 is obtained, and a region CR in CP is located, which has the same size as the current block and has a displacement MV from the position of the current block. Second, a reconstructed picture PRP on V0 which has the same POC as that of CP is found. Third, a region PRR in PRP is located, which has the same size as the current block and has a displacement MV+DV from the position of the current block. Suppose S represents

30

reconstructed samples in CR and Q represents reconstructed samples in PRR, then pseudo residues PR for the current block are calculated as $PR=S-Q$.

7. The method as claimed in claim 1, wherein the pseudo residues PR are used to predict original residues OR in dependent views as follows. The new residues NR are calculated as $NR=OR-PR$. On the encoder, NR replaces OR in the following coding processes such as transform, quantization and entropy coding. On the decoder, the reconstructed NR noted as NR' can be obtained after decoding processes such as de-quantization and inverse-transform. Then the reconstructed samples Y are calculated as $Y = NR' + PR + P$, where P represents the predictions.

8. The method as claimed in claim 1, wherein the pseudo residues PR are used to predict original residues OR in dependent views as follows. On the encoder, the new predictions (NP) can be calculated as $NP=P+PR$, where P are the predictions for the current block. Before calculating new residues NR , values in NP can be clipped into the value range $[\text{min\_sample\_value}, \text{max\_sample\_value}]$, where min_sample_value and max_sample_value represent the minimal value and the maximal value of original samples respectively. Then $NR=X-NP$. NR replaces OR in the following coding processes such as transform, quantization and entropy coding. On the decoder, NP is calculated and clipped as on the encoder. Then the reconstructed samples Y are calculated as $Y = NR' + NP$, where NR' is the reconstructed NR after decoding processes such as de-quantization and inverse-transform.

9. The method as claimed in claim 2, wherein the estimated DV of a block is obtained by utilizing DVs of its spatial and temporal neighboring blocks, and exploring original or reconstructed samples in the current picture and pictures on the base-view.

10. The method as claimed in claim 9, wherein the estimated DV of a block is obtained by: checking spatial neighboring blocks and temporal neighboring blocks of the current block to determine whether they have a DV referring to an inter-view reference picture, utilizing at least some of DVs from neighboring blocks as candidates of the estimated DV for the current block; transmitting the index of the chosen candidate by an encoder to a decoder explicitly or the decoder derives the index implicitly; wherein the index is transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU).

11. The method as claimed in claim 10, the estimated DV obtained is utilized as a prediction for the real estimated DV; and the difference between the estimated DV and the prediction is transmitted to the decoder.

5 12. The method as claimed in claim 9, wherein a decoder derives the estimated DV implicitly, and the estimated DV obtained is utilized as a hint for the decoder to derive the real estimated DV.

10 13. The method as claimed in claim 2, wherein the estimated MV of a block can be obtained by utilizing MVs of its spatial and temporal neighboring blocks, and exploring original or reconstructed samples in the current picture and pictures on the current dependent view.

15 14. The method as claimed in claim 13, wherein the estimated MV of a block is obtained by: checking spatial neighboring blocks and temporal neighboring blocks of the current block to determine whether they have a MV referring to an inter-time reference picture; utilizing at least some of MVs from neighboring blocks as candidates of the estimated MV for the current block; the encoder transmits the index of the chosen candidate to the decoder explicitly; and the estimated index is transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU); the decoder can derive the chosen candidate implicitly.

20 15. The method as claimed in claim 14, wherein the encoder transmits the estimated MV to the decoder explicitly; the estimated MV is transmitted in Macro-block (MB), Coding Unit (CU) or Prediction Unit (PU); the estimated MV is utilized as a prediction for the real estimated MV; the difference between the estimated MV and the prediction is transmitted to the decoder.

25 16. The method as claimed in claim 14, wherein the decoder can derive the estimated MV implicitly; and the estimated MV is utilized as a hint for the decoder to derive the real estimated MV.

30 17. The method as claimed in claim 1, wherein the pseudo residual prediction method is applied in a slice, a Macro-block (MB), a Coding Unit (CU) or a Prediction Unit (PU); an encoder transmits a piece of information to a decoder to indicate whether the pseudo residual prediction method is used or not used in a slice, a Macro-block (MB), a Coding Unit (CU) or a Prediction Unit (PU).

18. The method as claimed in claim 1, the pseudo residual prediction method is used only for Macro-blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) with some specific coding mode or coding type, and there is no residual prediction on/off flag signaling in the bitstream. i.e., the on/off control for residual prediction is implicitly depending on the specific coding modes.

19. The method as claimed in claim 18, wherein the pseudo residual prediction method is used in Macro-blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) with merge/skip mode and uses inter-view merging candidate in temporal direction, and is not used in others.

10

20. The method as claimed in claim 18, wherein the pseudo residual prediction method is used in Macro-blocks (MBs), Coding Units (CUs) or Prediction Units (PUs) which are coded in merge/skip mode and use inter-view merging candidate in temporal direction and 2Nx2N partition, and is not used in other.

15

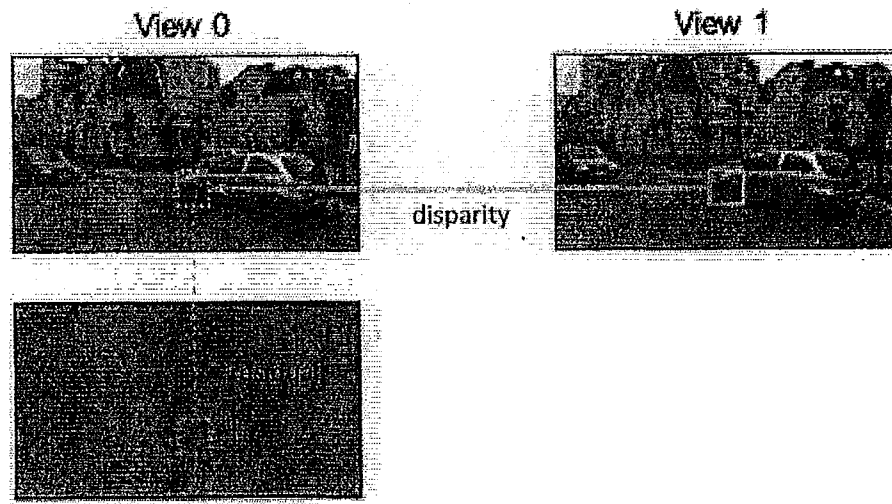


Fig.1

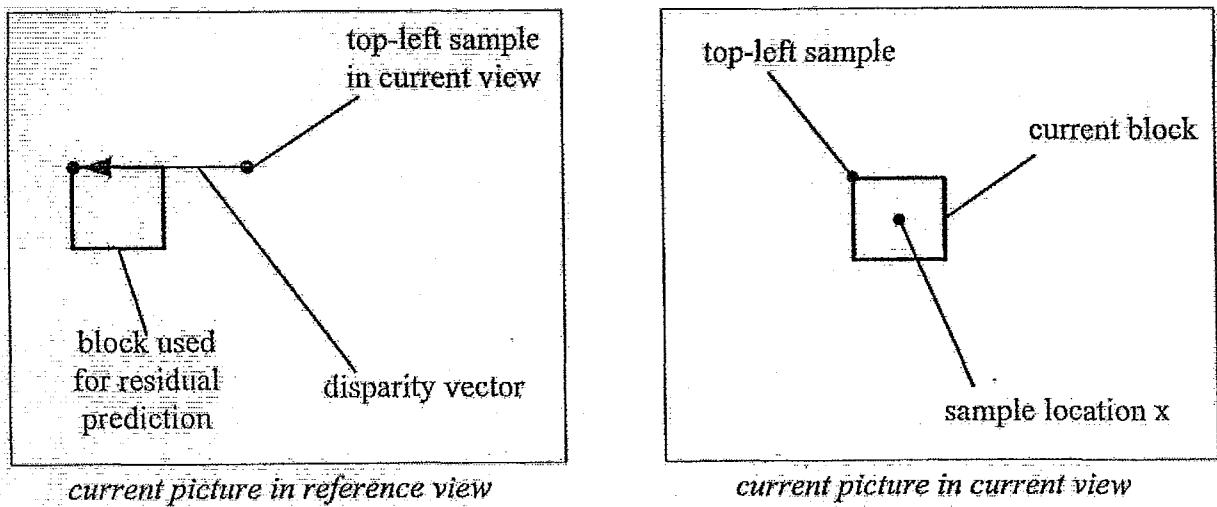


Fig.2

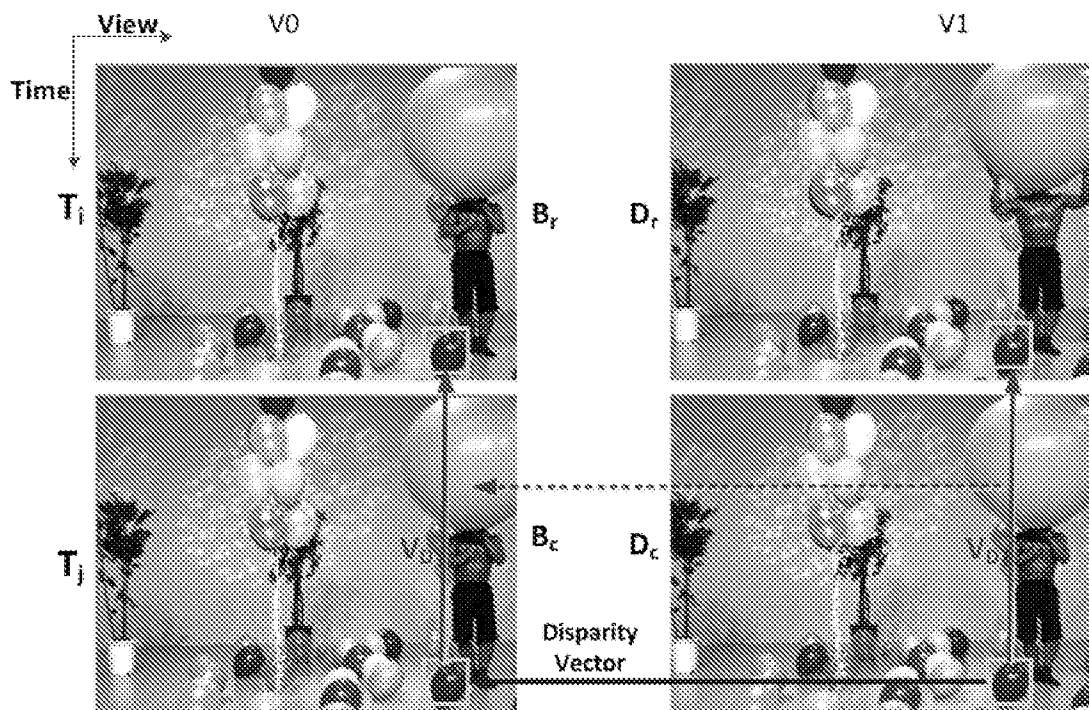


Fig.3

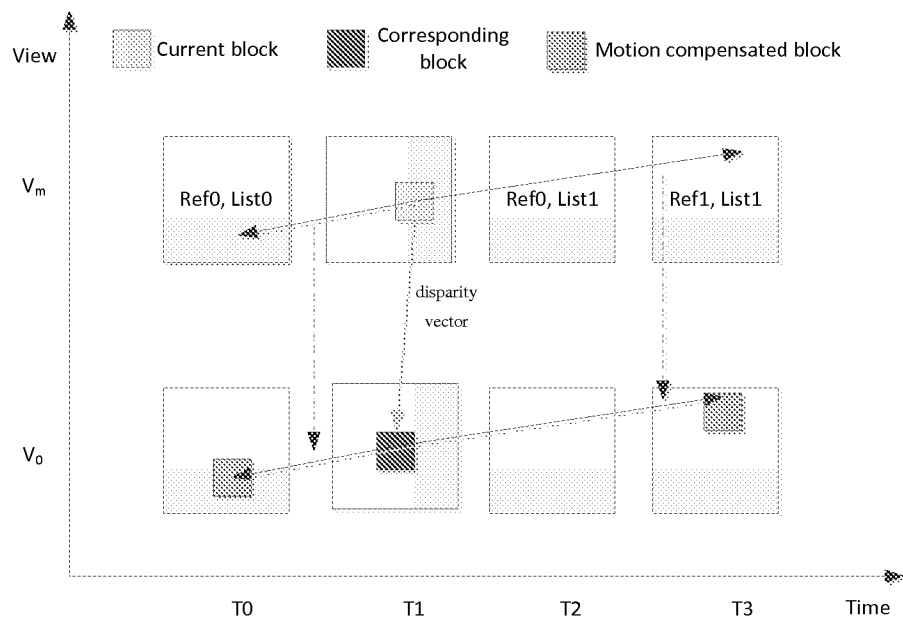


Fig.4

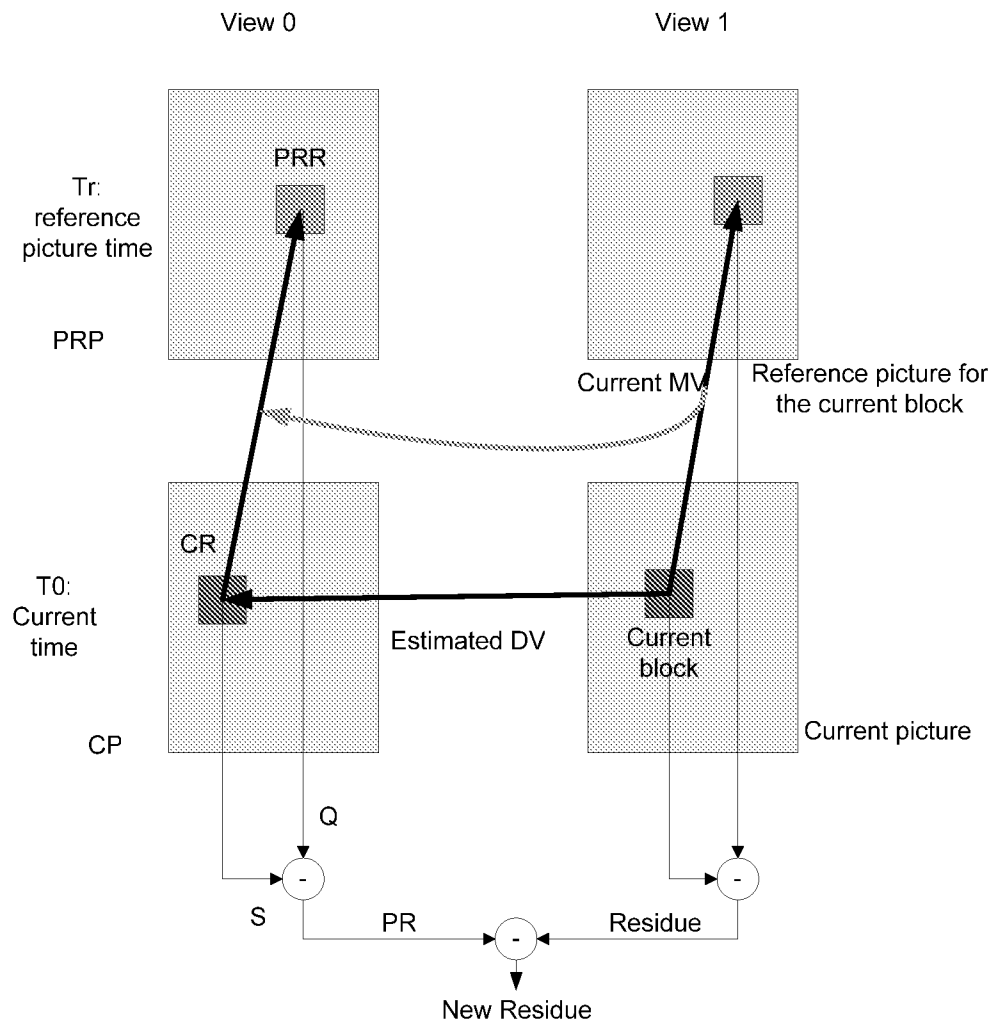


Fig.5

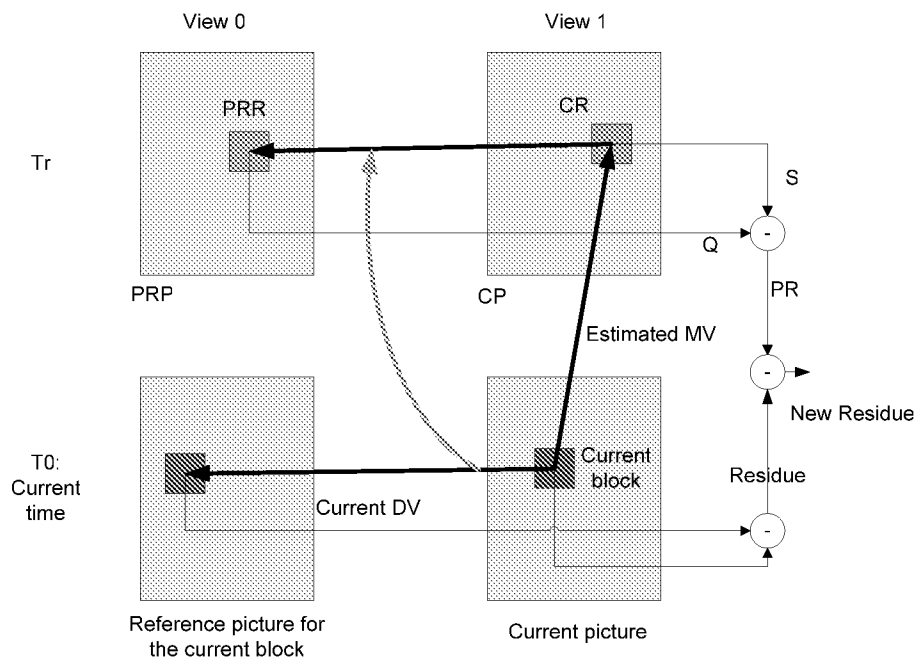


Fig.6

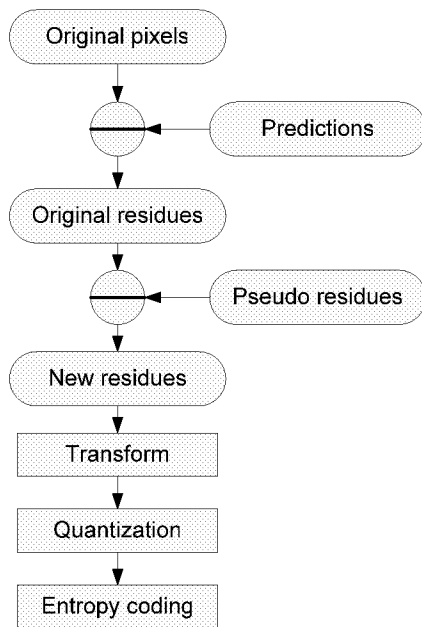


Fig.7

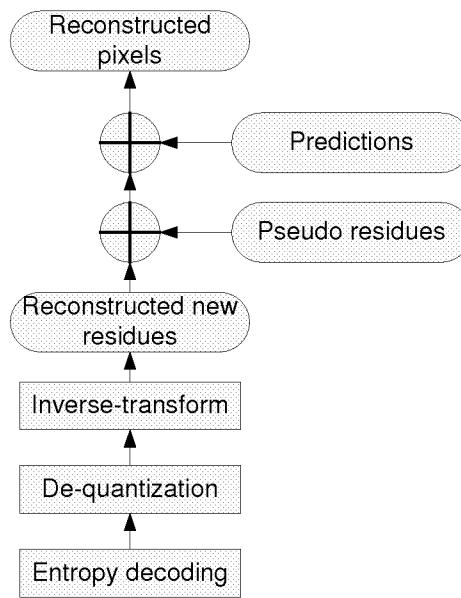


Fig.8

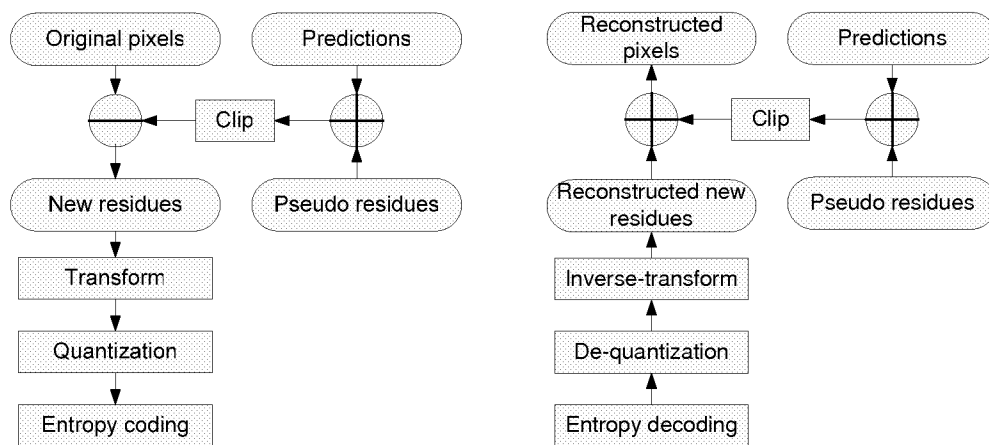


Fig.9

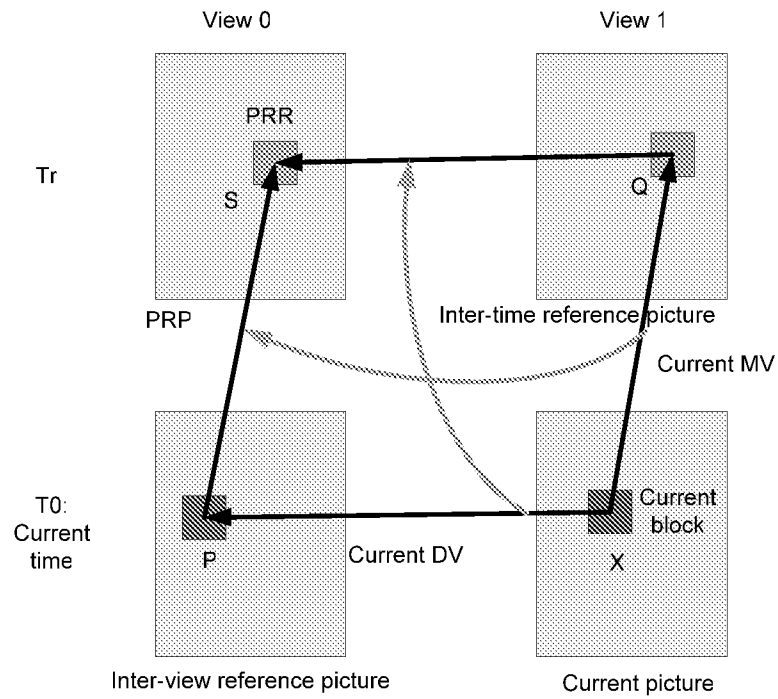


Fig.10

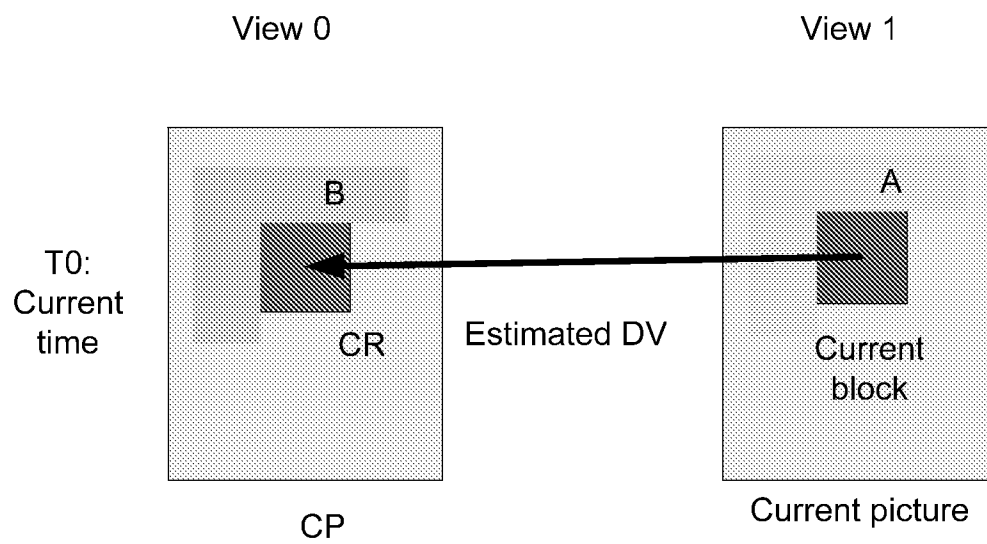


Fig.11

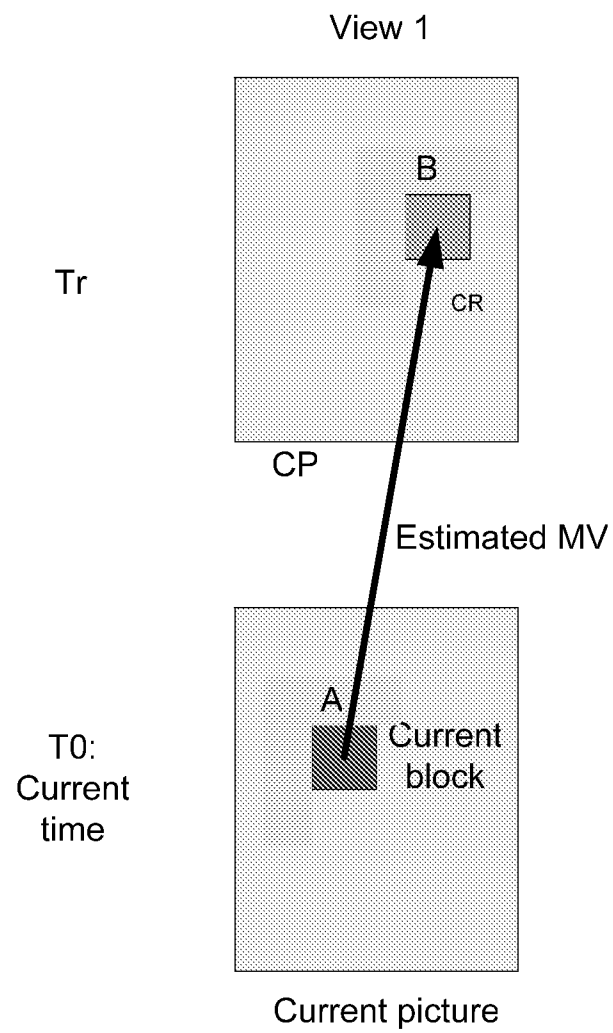


Fig.12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2012/084580

A. CLASSIFICATION OF SUBJECT MATTER

H04N7/32(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:H04N

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
VEN,CNABS,CNTEXT,IEEE: residual, residue, predict, 3D, three dimensional, disparity vector, DV, motion vector, MV, base view, dependent view, pseudo residue, inter-time, inter-view

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN102224731A(PANASONIC CORPORATION)19 Oct. 2011(19.10.2011) description paragraphs [0089]-[0096], [0108]-[0111]	1,3,17-20
Y		2,4
Y	CN101243692A(SAMSUNG ELECTRONICS CO LTD)13 Aug. 2008(13.08.2008) description page 6, lines 23-30	2
Y	CN101222627A(HUAWEI TECHNOLOGIES CO LTD)16 Jul. 2008(16.07.2008) description page 4, line 21 - page 5, line 17	4
A	CN101601304A(SAMSUNG ELECTRONICS CO LTD)09 Dec. 2009(09.12.2009) the whole document	1-20
A	WO2012096164A1(PANASONIC CORPORATION)19 Jul. 2012(19.07.2012) the whole document	1-20

☐ 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 10 Aug. 2013(10.08.2013)	Date of mailing of the international search report 22 Aug. 2013 (22.08.2013)
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. 86-10-62019451	Authorized officer WANG Bo Telephone No. (86-10)62411479

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2012/084580

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN102224731A	19.10.2011	WO2011037933A1	31.03.2011
		US8446958B2	21.05.2013
		JP2013505647W	14.02.2013
		US2011222605A1	15.09.2011
		KR20120089561A	13.08.2012
		EP2481214A1	01.08.2012
		IN201103439P4	07.09.2012
CN101243692A	13.08.2008	JP2009505604T	05.02.2009
		KR20070022568A	27.02.2007
		EP1917814A1	07.05.2008
		MX2008002391A	01.03.2008
		US2007041443A1	02.22.2007
		KR728009B1	13.06.2007
		MX279791B	07.10.2010
		CN101243692B	26.05.2010
		WO2007024072A1	01.03.2007
CN101222627A	16.07.2008	None	
CN101601304A	09.12.2009	EP2103144A1	23.09.2009
		KR20080066522A	16.07.2008
		US2008170618A1	17.07.2008
		JP2010516158T	13.05.2010
		WO2008084997A1	17.07.2008
WO2012096164A1	19.07.2012	None	