METHOD AND APPARATUS OF MOTION COMPENSATION FOR VIDEO CODING BASED ON BI PREDICTION OPTICAL FLOW TECHNIQUES

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present invention claims priority to U.S. Provisional Patent Application, Serial No. 62/213,249, filed on September 2, 2015. The U.S. Provisional Patent Application is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] The present invention relates to motion compensation for video coding using bi-directional optical flow (BIO) techniques. In particular, the present invention relates to extending the BIO to more general cases, or applying BIO adaptively to improve performance or reducing complexity.

BACKGROUND

[0003] Bi-directional optical flow (BIO) is motion estimation/compensation technique disclosed in JCTVC-C204 (E. Alshina, et al., *Bi-directional optical flow*, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 3rd Meeting: Guangzhou, CN, 7-15 October, 2010, Document: JCTVC- C204) and VCEG-AZ05 (E. Alshina, et al., *Known tools performance investigation for next generation video coding*, ITU-T SG 16 Question 6, Video Coding Experts Group (VCEG), 52nd Meeting: 19–26 June 2015, Warsaw, Poland, Document: VCEG-AZ05). BIO derived the sample-level motion refinement based on the assumptions of optical flow and steady motion. It is applied only for truly bi-directional predicted blocks, which is predicted from two reference frames corresponding to the previous frame and the latter frame. In VCEG-AZ05, BIO utilizes a 5x5 window to derive the motion refinement of each sample. Therefore, for an NxN

block, the motion compensated results and corresponding gradient information of an (N+4)x(N+4) block are required to derive the sample-based motion refinement for the NxN block. According to VCEG-AZ05, a 6-Tap gradient filter and a 6-Tap interpolation filter are used to generate the gradient information for BIO. Therefore, the computation complexity of BIO is much higher than that of traditional bi-directional prediction. In order to further improve the performance of BIO, the following methods are proposed.

[0004] In a conventional bi-prediction(bi prediction) in HEVC, the predictor is generated using equation (1), where $P^{(0)}$ and $P^{(1)}$ are the list0 and list1 predictor, respectively.

$$P_{Conventional}[i,j]=\left(P^{(0)}[i,j]+P^{(1)}[i,j]+1\right)>>1 \quad (1)$$

[0005] In JCTVC-C204 and VECG-AZ05, the BIO predictor is generated using equation (2).

$$P_{OpticalFlow}=(P^{(0)}[i,j]+P^{(1)}[i,j]+v_x[i,j](I_x^{(0)}-I_x^{(1)}[i,j])+ v_y[i,j](I_y^{(0)}-I_y^{(1)}[i,j])+1)>>1 \quad (2)$$

[0006] In equation (2), $I_x^{(0)}$ and $I_x^{(1)}$ represent the x-directional gradient in list0 and list1 predictor, respectively; $I_y^{(0)}$ and $I_y^{(1)}$ represents the y-directional gradient in list0 and list1 predictor, respectively; v_x and v_y represents the offsets in x- and y-direction, respectively. The above equations are derived using differential techniques to compute velocity from spatiotemporal derivatives of image intensity as shown in eq. (3a) and eq. (3b), where *I(x, y, t)* represents image intensity in the spatiotemporal coordinates:

$$I(x,y,t)=I(x+MV0_x+v_{x,}y+MV0_y+v_{y,}t-\Delta t) \quad (3a)$$

$$=I(x+MV1_x-v_{x,}y+MV1_y-v_{y,}t+\Delta t) \quad (3b)$$

[0007] Eq. (3a) can be further derived as follows:

$$I(x+MV0_x+v_{x,}y+MV0_y+v_{y,}t-\Delta t)$$

$$=P^0(x,y)+v_x\frac{\partial P^0(x,y)}{\partial x}+v_y\frac{\partial P^0(x,y)}{\partial y} \quad (4a)$$

[0008] Similarly, eq. (3b) can be further derived as follows:

$$I(x+MV1_x-v_{x,}y+MV1_y-v_{y,}t+\Delta t)$$

$$= P^1(x,y) - v_x \frac{\partial P^1(x,y)}{\partial x} - v_y \frac{\partial P^1(x,y)}{\partial y} \quad (4b)$$

[0009] Accordingly, the bi-directional optical flow is derived as follows, which is equivalent to eq. (2) with $I_x^{(0)} = \partial P^0(x,y)/\partial x$, $I_x^{(1)} = \partial P^1(x,y)/\partial x$, $I_y^{(0)} = \partial P^0(x,y)/\partial y$ and $I_y^{(1)} = \partial P^1(x,y)/\partial y$:

$$P_{OpticalFlow} = (P^0(x,y) + P^1(x,y) + v_x(\frac{\partial P^0(x,y)}{\partial x} - \frac{\partial P^1(x,y)}{\partial x}) + v_y(\frac{\partial P^0(x,y)}{\partial y} - \frac{\partial P^1(x,y)}{\partial y}) + 1) >> 1 \quad (5)$$

[0010] The difference $\Delta[i,j]$ between values in two points can be derived according to:

$$\Delta[i,j] = P^{(0)}[i,j] - P^{(1)}[i,j] + v_x[i,j]\left(I_x^{(0)}[i,j] + I_x^{(1)}[i,j]\right) + v_y[i,j]\left(I_y^{(0)}[i,j] + I_y^{(1)}[i,j]\right)$$
$$= P^{(0)}[i,j] + v_x[i,j]I_x^{(0)}[i,j] + v_y[i,j]I_y^{(0)}[i,j] - \left(P^{(1)}[i,j] - v_x[i,j]I_x^{(1)}[i,j] - v_y[i,j]I_y^{(1)}[i,j]\right) \quad (6)$$

[0011] The difference $\Delta[i,j]$ between values in two points is referred as *flow difference* at two points in this disclosure. In eq. (6), $v_x[i,j]$ and $v_y[i,j]$ are pixel-wise motion vector refinement components, where only fine motion is considered and the major motion is compensated by MC. Also $\left(I_x^{(0)}[i,j], I_y^{(0)}[i,j]\right)$ and $\left(I_x^{(1)}[i,j], I_y^{(1)}[i,j]\right)$ are gradients of luminance I in the position $[i\ ,j]$ of list0 and list1 reference frames correspondently. The motion vector refinement components, $v_x[i,j]$ and $v_y[i,j]$ are also referred as the *x-offset value* and the *y-offset value* in this disclosure.

[0012] In order to solve $v_x[i,j]$ and $v_y[i,j]$, a window consisting the pixel being processed and (2M+1)×(2M+1) neighbours is used. The pixel set Ω represents pixels in the window, i.e., $[i', j'] \in \Omega$ if and only if $i-M \le i' \le i+M$ *and* $j-M \le j' \le j+M$. The $v_x[i,j]$ and $v_y[i,j]$ are selected based on the values that minimizes:

$$\sum_{[i',j]\in\Omega} \Delta^2[i',j].$$

[0013] The gradient calculation for integer pixel resolution is shown as follows:

$$I_x^{(k)}[i,j] = (P^{(k)}[i+1,j] - P^{(k)}[i,j])/2\,, \quad (7a)$$

$$I_y^{(k)}[i,j] = (P^{(k)}[i,j+1] - P^{(k)}[i,j])/2\,. \quad (7b)$$

[0014] For fractional pixel resolution, interpolation will be performed first and the

gradient is calculated as follows:

$$P^{(k)}[i] = \sum_{n=-M+1}^{M} F_n\left(\alpha_x^{(k)}\right) R^{(k)}[i+n],$$

$$I_x^{(k)}[i] = \sum_{n=-M+1}^{M} dF_n\left(\alpha_x^{(k)}\right) R^{(k)}[i+n], \quad k = 0, 1,$$

$$dF_n\left(\alpha_x^{(k)}\right) = \left(F_n\left(\alpha_x^{(k)} + h\right) - F_n\left(\alpha_x^{(k)} - h\right)\right) / 2h$$

$$I_y^{(k)}[i, j] = \sum_{n=-M+1}^{M} dF_n\left(\alpha_y^{(k)}\right) R^{(k)}[i, j+n], \quad k = 0, 1.$$

[0015] In the above equations, α is block motion vector, $R^{(k)}[i,j]$ is reference picture value in integer position $[i,j]$ for references k=0 or 1, $F_n(\alpha)$ is filter directly providing derivatives.

[0016] For x-directional gradient, if the y-location is an integer, the luma gradient filter is applied. If the y-location is fractional, interpolation in the y direction is performed and luma gradient filter is applied in the x direction. For y-directional gradient, if the x-location is an integer, the luma gradient filter is applied. If the x-location is fractional, luma gradient filter is applied in the y direction and interpolation in the x direction is performed .

[0017] In the existing BIO implementation, the window size for $v_x[i,j]$ and $v_y[i,j]$ are 5x5 and BIO is only applied to the luma component with truly bi-predicted 2N×2N coding units (CUs) only. For gradient calculation at fractional pixel resolution, an additional 6-tap interpolation/gradient filter is used. Furthermore, the vertical process is performed first followed by the horizontal process.

SUMMARY

[0018] A method and apparatus of motion compensation using the bi-directional optical flow (BIO) techniques are disclosed. According to one method of the present invention, the use of BIO is extended to general bi-prediction motion compensation by including the case that two reference pictures correspond to two previously coded pictures. In one embodiment, the two x-offset values and two y-offset values for two corresponding positions in two reference blocks have same values, but opposite sign. In another embodiment, the two x-offset values and two y-offset values for two

corresponding positions in two reference blocks have same values as well as the sign. In yet another embodiment, the two x-offset values and two y-offset values for two corresponding positions in two reference blocks are proportional to two relative temporal distances between the first reference picture and the current picture and between the second reference picture and the current picture.

[0019] According to another method of the present invention, the use of BIO is adaptively applied depending on the linearity of the two motion vectors associated with the two reference blocks or depending on block size of the current block. For example, the current block is encoded or decoded using the bi-directional optical-flow prediction if the linearity of the first motion vector and the second motion vector satisfies a linearity threshold or if the block size of the current block is larger than a threshold block size.

[0020] According to yet another method of the present invention, the refined motion vectors by compensating the original motion vectors with the respective x-offset values and y-offset values are stored in a motion-vector buffer for motion vector prediction of one or more following blocks. If the bi-directional optical-flow prediction is applied to the current block on block-level basis for sub-blocks of the current block, the refined motion vectors associated with the sub-blocks are stored in the motion-vector buffer.

BRIEF DESCRIPTION OF DRAWINGS

[0021] Fig. 1 illustrates an example of motion compensation using bi-directional optical flow technique.

[0022] Fig. 2 illustrates an exemplary flowchart of a video coding system incorporating an embodiment of the present invention, where the use of BIO is extended to general bi-prediction motion compensation by including the case that two reference pictures correspond to two previously coded pictures.

[0023] Fig. 3 illustrates an exemplary flowchart of a video coding system incorporating another embodiment of the present invention, where the use of BIO is adaptively applied depending on the linearity of the two motion vectors associated with the two reference blocks or depending on block size of the current block.

[0024] Fig. 4 illustrates an exemplary flowchart of a video coding system incorporating another embodiment of the present invention, where the refined motion

vectors by compensating the original motion vectors with the respective x-offset values and y-offset values are stored in a motion-vector buffer for motion vector prediction of one or more following blocks.

DETAILED DESCRIPTION

[0025] 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.

[0026] In VCEG-AZ05, the Bi-directional Optical flow (BIO) is implemented as an additional process to the process as specified in the HEVC reference software. The motion compensated prediction according to the conventional HEVC is generated as shown in eq. (1). On the other hand, the motion compensated prediction according to BIO is shown in eq. (2), where additional parameters are determined to modify the conventional motion compensated prediction. The BIO is always applied to those blocks that are predicted with true bi-directions. In order to avoid increasing the memory bandwidth in the worst case, a method of the present invention only applies BIO to larger blocks. For example, an 8-tap interpolation filter for the luma component and a 4-tap interpolation filter for the chroma component are used to perform fractional motion compensation in HEVC. In the case of using a 5×5 window for each to-be-processed pixel as specified in BIO, the worst-case bandwidth is increased from 3.52 (i.e., (8+7)×(8+7) / (8x8)) to 5.64 (i.e., (8+7+4)×(8+7+4) / (8×8)) samples accessed per to-be-processed sample per reference frame. If only blocks with size larger than 8x8 are allowed for the BIO process, the worst case memory requirement for each pixel in BIO is reduced from 5.64 to 2.84 (i.e., (16+7+4)×(16+7+4) / (16×16)), which is even smaller than the original worst-case bandwidth (i.e., 3.52 samples accessed per to-be-processed sample per reference frame). Therefore, the worst-case memory bandwidth will not be increased by restricting the BIO process to block sizes larger than a threshold block size (e.g. 8x8) according to the present invention.

[0027] A method is disclosed to reduce the complexity and/or cost associated with the BIO process. According to this method, the gradient filter and the interpolation filter in BIO are unified with the interpolation filter for fractional motion compensation.

Currently, the gradient filter and the interpolation filter in BIO are additional processes to the conventional HEVC. These filters are different from the interpolation filter used for motion compensation. The BIO related filters cause additional cost to the BIO process. However, the purpose of the interpolation filter in BIO and the purpose of the interpolation filter in motion compensation are similar since both are intended for approximating the fractional-pel motion. Furthermore, these filters will derive the related information such as interpolated pixel values and gradient values. The gradient filter in BIO can be derived directly from the interpolation filter in BIO. The method will further unify the interpolation filter in BIO with the interpolation filter in fractional-pel motion compensation, and derives the gradient filter from the interpolation filter.

[0028] According to the method of unifying interpolation filters as disclosed above, there is no need for an additional interpolation filter. Therefore, the computation becomes unified and simplified. An 8-tap interpolation filter or 4-tap interpolation filter can be used instead of 6-tap interpolation filter as specified in BIO. When 8-tap interpolation filter is used, the gradient filter is also changed and derived directly from the difference between filter coefficients with different fractional positions. For example, for the fractional position equal to 1/2-pel, the gradient filter coefficients can be derived from the differences between the interpolation filter coefficients for the fractional position equal to 3/4-pel and the interpolation filter coefficients for the fractional position equal to 1/4-pel divided by $2\times(1/4)$. The coding performance of BIO is improved because of the same interpolation filter is used for BIO and motion compensation. However, the computational complexity is increased also. If a 4-tap interpolation filter is used, no additional filter is required and the computation complexity can be further reduced.

[0029] Another method to improve the performance of BIO is to apply BIO for all bi-directional predicted blocks regardless of whether the blocks are “true bi-prediction” or not. According to the assumption of optical flow and steady motion, the corresponding equations and solutions for bi-directional predicted blocks can be used, where both reference frames are previously coded frames by using a similar approach. For example, the x-offset values and the y-offset values for the two corresponding positions (i.e., position A and B in Fig. 1) have the same value, but opposite sign. Accordingly, the x-offset values and the y-offset values for two corresponding positions in two reference blocks of two previously coded frames may have the same value, but opposite sign. In the assumption of steady motion, the temporal distances between current block and two

references blocks can be taken into account in the equations. For example, POC (picture order count) is often used for temporal distance. If the temporal distances between current block and two references blocks are m and n, the x-offset values and the y-offset values for two corresponding positions in two reference blocks of two previously coded frames can be proportional to m and n, where m and n are integers. In another embodiment, only the temporal direction should be considered in the corresponding equation for simplicity. In this case, the x-offset values and the y-offset values for two corresponding positions in two reference blocks of two previously coded frames may have the same value and the same sign.

[0030] In VCEG-AZ05 the BIO is applied in pixel-level basis. In an embodiment of the present invention, the process of the BIO is applied in the block-level basis. The block size can be N×M, where N and M are integers. All the pixels in an N×M block can share the same motion refinement. If N and M are equal to or greater than 4, the refined motion vector can be stored back to the MV buffers.

[0031] The BIO can be applied to sub-PUs (prediction units). For example, if a PU block is allowed for sub-PU partition and each sub-PU can have different motion information or modes, the BIO can be applied to each sub-PU. The initial MV for BIO can be different for each sub-PU.

[0032] In yet another embodiment, the BIO and the methods disclosed above can also be extended to the blocks (pixels) of multiple-hypothesis prediction such as Inter-prediction with more than two reference blocks (pixels).

[0033] In still yet another embodiment, the BIO operations can be adaptively applied according to the gradient calculations on $P^{(0)}$ and $P^{(1)}$ or the hybrid predictor ($P^{(0)}$ + $P^{(1)}$). For example, when the difference between the list0 gradient and list1 gradient is larger than a predefined threshold, the BIO is not applied.

[0034] In still yet another embodiment, the BIO operations can be adaptively applied according to the linearity of motion vectors that generates $P^{(0)}$ and $P^{(1)}$. In other words, if the motion vectors that generates $P^{(0)}$ and $P^{(1)}$ do not follow linear motion assumption, the refined pixel motions, v_x and v_y, are not reliable. Therefore, the decoder can check the linearity to adaptively apply BIO according to an embodiment of the present invention. For example, the BIO operations can be applied only if the linearity of motion vectors meets a required condition. For example, the current block can be encoded or decoded using the bi-directional optical-flow prediction only if the linearity

of the first motion vector and the second motion vector satisfies a linearity threshold.

[0035] In still yet another embodiment, if the motion vectors that generate $P^{(0)}$ and $P^{(1)}$ do not follow linear motion assumption, the decoder can calculate BIO according to the direction of the motion vectors that generates $P^{(0)}$ and $P^{(1)}$. For example, the decoder can derive pixel motion vectors in proportion to the motion vectors that generate $P^{(0)}$ and $P^{(1)}$.

[0036] In still yet another embodiment, the offsets calculated in the BIO process can be viewed as an offset to refine the motion vectors for all pixels in current block. The refined MVs can be stored in the MV buffer and used for the MV prediction of the following blocks. Note that, if the BIO is performed in a block level (e.g. 4×4 block), the refined MVs are also stored in the block level.

[0037] Fig. 2 illustrates an exemplary flowchart of a video coding system incorporating an embodiment of the present invention, where the use of BIO is extended to general bi-prediction motion compensation by including the case that two reference pictures correspond to two previously coded pictures. According to this method, input data associated with a current block in a current picture is received in step 210. A first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector are determined in step 220, where the first reference picture and the second reference picture are two previously coded pictures. The x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block is determined in step 230. The y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block is determined in step 240. An x-offset value and a y-offset value are determined according to an optical flow model in step 250, where the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block. Bi-directional optical-flow prediction corresponding to the given position is derived based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset

value, and the y-direction gradient difference weighted by the y-offset value as shown in step 260. Pixel data at the given position of the current block is encoded or decoded using the bi-directional optical-flow prediction corresponding to the given position as shown in step 270.

[0038] Fig. 3 illustrates an exemplary flowchart of a video coding system incorporating another embodiment of the present invention, where the use of BIO is adaptively applied depending on the linearity of the two motion vectors associated with the two reference blocks or depending on block size of the current block. According to this method, input data associated with a current block in a current picture is received in step 310. A first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector are determined in step 320. The x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block is determined in step 330. The y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block is determined in step 340. An x-offset value and a y-offset value are determined according to an optical flow model in step 350, where the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block. Bi-directional optical-flow prediction corresponding to the given position is derived based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value as shown in step 360. Pixel data at the given position of the current block is encoded or decoded using the bi-directional optical-flow prediction or not depending on linearity of the first motion vector and the second motion vector or depending on block size of the current block as shown in step 370.

[0039] Fig. 4 illustrates an exemplary flowchart of a video coding system incorporating another embodiment of the present invention, where the refined motion vectors by compensating the original motion vectors with the respective x-offset values and y-offset values are stored in a motion-vector buffer for motion vector prediction of

one or more following blocks. According to this method, input data associated with a current block in a current picture is received in step 410. A first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector are determined in step 420. The x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block is determined in step 430. The y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block is determined in step 440. An x-offset value and a y-offset value are determined according to an optical flow model in step 450, where the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block. Bi-directional optical-flow prediction corresponding to the given position is derived based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value as shown in step 460. Pixel data at the given position of the current block is encoded or decoded using the bi-directional optical-flow prediction corresponding to the given position as shown in step 470. The refined motion vectors for bi-directional optical-flow predicted pixels of the current block are stored in a motion-vector buffer for motion vector prediction of one or more following blocks in step 480, where the refined motion vectors are determined based on the first motion vector or the second motion vector modified by the x-offset value and the y-offset value.

[0040] The flowcharts shown are intended to illustrate an example of video coding according to the present invention. A person skilled in the art may modify each step, re-arranges the steps, split a step, or combine steps to practice the present invention without departing from the spirit of the present invention. In the disclosure, specific syntax and semantics have been used to illustrate examples to implement embodiments of the present invention. A skilled person may practice the present invention by substituting the syntax and semantics with equivalent syntax and semantics without departing from the spirit of the present invention.

[0041] The above description is presented to enable a person of ordinary skill in the

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

[0042] Embodiment of the present invention as described above may be implemented in various hardware, software codes, or a combination of both. For example, an embodiment of the present invention can be one or more circuits integrated into a video compression chip or program code integrated into video compression software to perform the processing described herein. An embodiment of the present invention may also be program code to be executed on a Digital Signal Processor (DSP) to perform the processing described herein. The invention may also involve a number of functions to be performed by a computer processor, a digital signal processor, a microprocessor, or field programmable gate array (FPGA). These processors can be configured to perform particular tasks according to the invention, by executing machine-readable software code or firmware code that defines the particular methods embodied by the invention. The software code or firmware code may be developed in different programming languages and different formats or styles. The software code may also be compiled for different target platforms. However, different code formats, styles and languages of software codes and other means of configuring code to perform the tasks in accordance with the invention will not depart from the spirit and scope of the invention.

[0043] The invention may be embodied in other specific forms without departing from its spirit or essential characteristics. The described examples are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.

CLAIMS

1. A method of motion compensation for video data, the method comprising:

receiving input data associated with a current block in a current picture;

determining a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector, wherein the first reference picture and the second reference picture are two previously coded pictures;

deriving x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

deriving y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determining an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

deriving bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value; and

encoding or decoding pixel data at the given position of the current block using the bi-directional optical-flow prediction corresponding to the given position, wherein two x-offset values for the first position and second position have a same x-offset value with a same sign and two y-offset values for the first position and the second position have a same y-offset value with the same sign.

2. The method of Claim 1, wherein two x-offset values for the first position and the second position are proportional to two relative temporal distances between the first reference picture and the current picture and between the second reference picture and the current picture and two y-offset values for the first position and the second position

are proportional to the two relative temporal distances between the first reference picture and the current picture and between the second reference picture and the current picture.

3. An apparatus for motion compensation of video data performed by a video coding system, the apparatus comprising one or more electronic circuits or processors configured to:

receive input data associated with a current block in a current picture;

determine a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector, wherein the first reference picture and the second reference picture are two previously coded pictures;

derive x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

derive y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determine an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

derive bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value, wherein two x-offset values for the first position and the second position have a same x-offset value with a same sign and two y-offset values for the first position and the second position have a same y-offset value with the same sign; and

encode or decode pixel data at the given position of the current block using the bi-directional optical-flow prediction corresponding to the given position.

4. The apparatus of Claim 3, wherein two x-offset values for the first position and the second position are proportional to two relative temporal distances between the

first reference picture and the current picture and between the second reference picture and the current picture and two y-offset values for the first position and the second position are proportional to the two relative temporal distances between the first reference picture and the current picture and between the second reference picture and the current picture.

5. A method of motion compensation for video data, the method comprising:

receiving input data associated with a current block in a current picture;

determining a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector;

deriving x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

deriving y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determining an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

deriving bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value; and

encoding or decoding pixel data at the given position of the current block using the bi-directional optical-flow prediction or not depending on linearity of the first motion vector and the second motion vector or depending on block size of the current block, wherein the current block is encoded or decoded using the bi-directional optical-flow prediction if the linearity of the first motion vector and the second motion vector satisfies a linearity threshold.

6. The method of Claim 5, wherein the current block is encoded or decoded using the bi-directional optical-flow prediction if the block size of the current block is larger

than a threshold block size.

7. The method of Claim 6, wherein the threshold block size is 8x8.

8. An apparatus for motion compensation of video data performed by a video coding system, the apparatus comprising one or more electronic circuits or processors configured to:

receive input data associated with a current block in a current picture;

determine a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector;

derive x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

derive y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determine an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

derive bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value; and

encode or decode pixel data at the given position of the current block using the bi-directional optical-flow prediction or not depending on linearity of the first motion vector and the second motion vector or depending on block size of the current block, wherein the current block is encoded or decoded using the bi-directional optical-flow prediction if the linearity of the first motion vector and the second motion vector satisfies a linearity threshold.

9. The apparatus of Claim 8, wherein the current block is encoded or decoded using the bi-directional optical-flow prediction if the block size of the current block is larger than a threshold block size.

10. The apparatus of Claim 9, wherein the threshold block size is 8x8.

11. A method of motion compensation for video data, the method comprising:

receiving input data associated with a current block in a current picture;

determining a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector;

deriving x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

deriving y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determining an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

deriving bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value;

encoding or decoding pixel data at the given position of the current block using the bi-directional optical-flow prediction corresponding to the given position; and

storing refined motion vectors for bi-directional optical-flow predicted pixels of the current block in a motion-vector buffer for motion vector prediction of one or more following blocks, wherein the refined motion vectors are determined based on the first motion vector or the second motion vector modified by the x-offset value and the y-offset value, wherein if the bi-directional optical-flow prediction is applied to the current block on block-level basis for sub-blocks of the current block, the refined motion vectors associated with the sub-blocks are stored in the motion-vector buffer.

12. An apparatus for motion compensation of video data performed by a video coding system, the apparatus comprising one or more electronic circuits or processors configured to:

receive input data associated with a current block in a current picture;

determine a first reference block in a first reference picture based on a first motion vector and a second reference block in a second reference picture based on a second motion vector;

derive x-direction gradient difference corresponding to a given position of the current block between first x-direction gradient of the first reference block and second x-direction gradient of the second reference block;

derive y-direction gradient difference corresponding to the given position of the current block between first y-direction gradient of the first reference block and second y-direction gradient of the second reference block;

determine an x-offset value and a y-offset value according to an optical flow model, wherein the x-offset value and the y-offset value are selected to obtain a reduced or minimum flow difference between a first position and a second position, and the first position and the second position are two positions in the first reference block and the second reference block respectively corresponding to the given position of the current block;

derive bi-directional optical-flow prediction corresponding to the given position based on the first reference block, the second reference block, the x-direction gradient difference weighted by the x-offset value, and the y-direction gradient difference weighted by the y-offset value;

encode or decode pixel data at the given position of the current block using the bi-directional optical-flow prediction corresponding to the given position; and

store refined motion vectors for bi-directional optical-flow predicted pixels of the current block in a motion-vector buffer for motion vector prediction of one or more following blocks, wherein the refined motion vectors are determined based on the first motion vector or the second motion vector modified by the x-offset value and the y-offset value, wherein if the bi-directional optical-flow prediction is applied to the current block on block-level basis for sub-blocks of the current block, the refined motion vectors associated with the sub-blocks are stored in the motion-vector buffer.