

DESCRIPTION OF EMBODIMENTS

[0106] The following clearly describes the technical solutions in the embodiments of this application with reference to the accompanying drawings in the embodiments of this application.

5 **[0107]** FIG. 1 is a schematic block diagram of a video coding system according to an embodiment of this application. In the system, a video encoder 20 and a video decoder 30 are configured to predict a predicted value of a pixel value of a picture block based on various picture prediction method examples provided in this application, and refine motion information, such as a motion vector, of a current coded or decoded
10 picture block, to further improve coding performance. As shown in FIG. 1, the system includes a source apparatus 12 and a destination apparatus 14. The source apparatus 12 generates coded video data that is to be decoded by the destination apparatus 14. The source apparatus 12 and the destination apparatus 14 may include any one of a wide range of apparatuses, including a desktop computer, a notebook computer, a tablet
15 computer, a set-top box, a telephone handset such as a "smart" phone, a "smart" touchpad, a television, a camera, a display apparatus, a digital media player, a video game console, a video streaming transmission apparatus, or the like.

[0108] The destination apparatus 14 may receive the to-be-decoded coded video data by using a link 16. The link 16 may include any type of medium or apparatus
20 capable of moving the coded video data from the source apparatus 12 to the destination apparatus 14. In a feasible implementation, the link 16 may include a communications medium capable of enabling the source apparatus 12 to directly transmit the coded video data to the destination apparatus 14 in real time. A communications standard (for example, a wireless communications protocol) may be followed to modulate the coded
25 video data and transmit modulated video data to the destination apparatus 14. The communications medium may include any wireless or wired communications medium, for example, a radio spectrum or one or more physical transmission lines. The communications medium may form a part of a packet-based network (for example, a local area network, a wide area network, or a global network of the internet). The

communications medium may include a router, a switch, a base station, or any other devices that may be configured to facilitate communication from the source apparatus 12 to the destination apparatus 14.

[0109] Alternatively, the coded data may be output from an output interface 22 to a

5 storage apparatus 24. Similarly, the coded data may be accessed from the storage apparatus 24 through an input interface. The storage apparatus 24 may include any of a plurality of scattered or local data storage media, for example, a hard disk drive, a Blu-ray disc, a DVD, a CD-ROM, a flash memory, a volatile or non-volatile memory, or any other appropriate digital storage medium used for storing the coded video data. In
10 another feasible implementation, the storage apparatus 24 may correspond to a file server or another intermediate storage apparatus capable of storing the coded video data generated by the source apparatus 12. The destination apparatus 14 may access the stored video data from the storage apparatus 24 through streaming transmission or downloading. The file server may be any type of server capable of storing the coded
15 video data and transmitting the coded video data to the destination apparatus 14. In a feasible implementation, the file server includes a website server, a file transfer protocol server, a network-attached storage apparatus, or a local disk drive. The destination apparatus 14 may access the coded video data through any standard data connection including an internet connection. The data connection may include a wireless channel
20 (for example, a Wi-Fi connection), a wired connection (for example, a cable modem), or a combination thereof that is applicable for accessing the coded video data stored in the file server. Transmission of the coded video data from the storage apparatus 24 may be streaming transmission, downloading transmission, or a combination thereof.

[0110] Techniques in this application are not necessarily limited to wireless

25 applications or settings. The techniques can be applied to video decoding, to support any one of a plurality of multimedia applications, for example, over-the-air television broadcasting, cable television transmission, satellite television transmission, streaming video transmission (for example, through the internet), digital video coding for storage on a data storage medium, decoding of a digital video stored on a data storage medium,
30 or another application. In some feasible implementations, the system may be configured

to support unidirectional or bidirectional video transmission, to support applications such as streaming video transmission, video playing, video broadcasting, and/or videotelephony.

[0111] In a feasible implementation of FIG. 1, the source apparatus 12 includes a video source 18, the video encoder 20, and the output interface 22. In some applications, the output interface 22 may include a modulator/demodulator (a modem) and/or a transmitter. In the source apparatus 12, the video source 18 may include, for example, the following sources: a video capturing apparatus (for example, a video camera), a video archive including a previously captured video, a video feed-in interface for receiving a video from a video content provider, and/or a computer graphics system for generating computer graphics data as a source video, or a combination thereof. In a feasible implementation, if the video source 18 is a video camera, the source apparatus 12 and the destination apparatus 14 can constitute a camera phone or a video phone. For example, the techniques described in this application may be applied to video decoding, and may be applied to wireless and/or wired applications.

[0112] The video encoder 20 may code a video that is captured, pre-captured, or generated by a computer. The coded video data may be directly transmitted to the destination apparatus 14 through the output interface 22 of the source apparatus 12. The coded video data may also (or alternatively) be stored on the storage apparatus 24 for subsequent access by the destination apparatus 14 or another apparatus for decoding and/or playing.

[0113] The destination apparatus 14 includes an input interface 28, the video decoder 30, and a display apparatus 32. In some applications, the input interface 28 may include a receiver and/or a modem. The input interface 28 of the destination apparatus 14 receives the coded video data by using the link 16. The coded video data transmitted or provided to the storage apparatus 24 by using the link 16 may include a plurality of syntax elements that are generated by the video encoder 20 and that are used by the video decoder 30 to decode the video data. These syntax elements may be included in the coded video data transmitted on the communications medium, and stored in the storage medium or stored in the file server.

mode. The PU obtained through partitioning may be in a non-square shape. For example, the syntactic data associated with the CU may also describe partitioning of one CU into one or more TUs based on the quadtree. The TU may be in a square or non-square shape.

[0123] The HEVC standard allows TU-based transformation, and TUs may be different for different CUs. A TU size is typically set based on a size of a PU within a given CU defined for a partitioned LCU. However, this may not always be the case. The TU size is generally the same as or less than the PU size. In some feasible implementations, a quadtree structure referred to as a "residual quadtree" (residual quadtree, RQT) may be used to divide a residual sample corresponding to the CU into smaller units. A leaf node of the RQT may be referred to as a TU. A pixel difference associated with the TU may be transformed to generate a transform coefficient, and the transform coefficient may be quantized.

[0124] Generally, transform and quantization processes are used for the TU. A given CU with one or more PUs may also include one or more TUs. After prediction, the video encoder 20 may calculate a residual value corresponding to the PU. The residual value includes a pixel difference, and the pixel difference may be transformed into a transform coefficient, and the transform coefficient is quantized and undergoes TU scanning to generate a serialized transform coefficient for entropy decoding. In this application, the term "picture block" is generally used to represent the decoding node of the CU. In some specific applications, in this application, the term "picture block" may also be used to represent the tree block including the decoding node, the PU, and the TU, for example, the LCU or the CU. In this embodiment of this application, various method examples described in an adaptive inverse quantization method in video encoding or decoding are described in detail below to perform an inverse quantization process of a transform coefficient corresponding to a current picture block (namely, a current transform block), to improve coding performance.

[0125] A video sequence generally includes a series of video frames or pictures. For example, a group of pictures (group of picture, GOP) includes a series of video pictures, one video picture, or a plurality of video pictures. The GOP may include syntactic data in header information of the GOP, in header information of one or more

the current picture block) is used by the current picture block. Optionally, the inter prediction data may further include a switch flag `block_based_enable_flag`, to indicate whether to perform picture prediction on the current picture block by using FIG. 3 or FIG. 14 (in other words, to indicate whether to perform inter prediction on the current picture block by using an MVD mirror constraint proposed in this application), or whether to perform picture prediction on the current picture block by using FIG. 12 or FIG. 16 (in other words, to indicate whether to perform inter prediction on the current picture block by using a proportional relationship that is based on a time domain distance and that is proposed in this application).

10 **[0153]** When the video slice is decoded into an intra-decoded (I) slice, the intra prediction unit 84 of the prediction module 81 may generate a prediction block of the picture block of the current video slice based on an intra prediction mode notified by sending a signal and data of a previously decoded block that is from a current frame or picture. When the video slice is decoded into an inter-decoded (namely, B or P) slice, 15 the motion compensation unit 82 of the prediction module 81 may determine, based on the syntax element received from the entropy decoding module 80, an inter prediction mode used to decode the current picture block of the current video slice, and decode (for example, perform inter prediction on) the current picture block based on the determined inter prediction mode. Specifically, the motion compensation unit 82 may 20 determine which picture prediction method is used to predict the current picture block of the current video slice, for example, the syntax element indicates that a picture prediction method based on the MVD mirror constraint is to be used to predict the current picture block. The motion information of the current picture block of the current video slice is predicted or refined, to obtain or generate, by using a motion 25 compensation process, the prediction block of the current picture block by using the predicted motion information of the current picture block. The motion information herein may include reference picture information and a motion vector. The reference picture information may include but is not limited to unidirectional/bidirectional prediction information, a reference picture list number, and a reference picture index 30 corresponding to a reference picture list. For inter prediction, the prediction block may

be generated from one of reference pictures in one of reference picture lists. The video decoder 30 may construct the reference picture lists, namely a list 0 and a list 1, based on the reference pictures stored in the reference picture memory 92. A reference frame index of the current picture may be included in one of or both a reference frame list 0 and a reference frame list 1. In some instances, the video encoder 20 may send a signal to indicate which new picture prediction method is used.

[0154] In this embodiment, the prediction module 81 is configured to generate a prediction block of a current coding picture block. Specifically, when the video slice is decoded into the intra-decoded (I) slice, the intra prediction unit 84 of the prediction module 81 may generate the prediction block of the picture block of the current video slice based on the signaled intra prediction mode sent and the data of the previously decoded picture block that is from the current frame or picture. When a video picture is decoded into an inter-decoded (for example, B, P, or GPB) slice, the motion compensation unit 82 of the prediction module 81 generates a prediction block of a picture block of the current video picture based on the motion vector and other syntax elements that are received from the entropy decoding unit 80.

[0155] The inverse quantization module 86 performs inverse quantization on, namely, dequantizes, a quantized transform coefficient provided in the bitstream obtained by the entropy decoding module 80 through decoding. The inverse quantization process may include: determining a to-be-applied quantization degree by using a quantization parameter calculated by the video encoder 20 for each picture block in the video slice, and similarly determining a to-be-applied inverse quantization degree. The inverse transform module 88 performs inverse transform, for example, inverse DCT, inverse integer transform, or a conceptually similar inverse transform process, to the transform coefficient, to generate a pixel-domain residual block.

[0156] After the motion compensation unit 82 generates the prediction block for the current picture block, the video decoder 30 sums the residual block from the inverse transform module 88 and the corresponding prediction block generated by the motion compensation unit 82, to obtain a reconstructed block, namely, a decoded picture block. A summator 90 represents a component that performs the summation operation. When

[0166] Manner 2:

[0167] In a non-merge mode of inter prediction, a motion vector predicted value list is constructed based on motion information of a neighboring block of the current picture block, and a motion vector is selected from the motion vector predicted value list as a motion vector predicted value of the current picture block. In this case, a motion vector of the current picture block may be a motion vector value of the neighboring block or a sum of differences between motion vectors of selected neighboring blocks and the motion vector of the current picture block. The motion vector difference is a difference between the motion vector obtained by performing motion estimation on the current picture block and the motion vector of the selected neighboring block. For example, motion vectors, corresponding to indexes 1 and 2, in the motion vector predicted value list are selected as a forward motion vector and a backward motion vector of the current picture block.

[0168] It should be understood that the foregoing manner 1 and manner 2 are merely two specific manners of obtaining the initial motion information of the picture block. In this application, a manner of obtaining the initial motion information of the picture block is not limited, and any manner in which the initial motion information of the picture block can be obtained shall fall within the protection scope of this application.

[0169] 302: Determine positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, where the N forward reference blocks are located in a forward reference picture, the N backward reference blocks are located in a backward reference picture, and N is an integer greater than 1.

[0170] Referring to FIG. 6, a current picture to which the current picture block in this embodiment of this application belongs has two reference pictures: the forward reference picture and the backward reference picture.

[0171] In an example, the initial motion information includes a first motion vector and a first reference picture index in a forward prediction direction, and a second motion vector and a second reference picture index in a backward prediction direction.

[0172] Correspondingly, step 302 may include:

$$\text{predSamples}'[x][y]=(\text{predSamplesL0}'[x][y]+\text{predSamplesL1}'[x][y]+1)>>1$$

(2)

[0202] In the foregoing formula, predSamplesL0' is the optimal forward reference block, predSamplesL1' is the optimal backward reference block, predSamples' is a prediction block of the current picture block, predSamplesL0'[x][y] is a pixel value of the optimal forward reference block at a sample (x, y), predSamplesL1'[x][y] is a pixel value of the optimal backward reference block at the sample (x, y), and predSamples'[x][y] is a final pixel value of the prediction block at the sample (x, y).

[0203] It should be noted that, in this embodiment of this application, a search method to be used is not limited, and any search method may be used. For each forward candidate block obtained through search, a difference between the forward candidate block and a corresponding backward candidate block is calculated, and a forward candidate block and a backward candidate block that have a minimum SAD, a forward motion vector corresponding to the forward candidate block, and a backward motion vector corresponding to the backward candidate block are respectively selected as the optimal forward reference block, the optimal backward reference block, the optimal forward motion vector corresponding to the optimal forward reference block, and the optimal backward motion vector corresponding to the optimal backward reference block. Alternatively, for each backward candidate block obtained through search, a difference between the backward candidate block and a corresponding forward candidate block in step 4 is calculated, and a backward candidate block and a forward candidate block that have a minimum SAD, a backward motion vector corresponding to the backward candidate block, and a forward motion vector corresponding to the forward candidate block are respectively selected as the optimal backward reference block, the optimal forward reference block, the optimal backward motion vector corresponding to the optimal backward reference block, and the optimal forward motion vector corresponding to the optimal forward reference block.

[0204] It should be noted that only an example of the search method based on the integer pixel step is provided in step 1004. Actually, in addition to search by using the integer pixel step, search by using a fractional pixel step can also be used. For example,

in step 1004, after the search by using the integer pixel step, the search by using the fractional pixel step is performed. Alternatively, the search by using the fractional pixel step is directly performed. A specific search method is not limited herein.

[0205] It should be noted that, in this embodiment of this application, a method for calculating the matching cost is not limited. For example, an SAD criterion, an MR-SAD criterion, or another criterion may be used. In addition, the matching cost may be calculated by using only a luminance component or by using both a luminance component and a chrominance component.

[0206] It should be noted that, in a search process, if the matching cost is 0 or reaches a preset threshold, a traversal operation or a search operation may be terminated in advance. An early termination condition of the search method is not limited herein.

[0207] It should be understood that a sequence of step 1005 and step 1006 is not limited, and they may be performed at the same time or may be performed in sequence.

[0208] It can be learned that, in an existing method, a template matching block needs to be calculated first, and forward search and backward search are separately performed by using the template matching block; but in this embodiment of this application, in a process of searching for a matching block, the matching cost is directly calculated by using the candidate block in the forward reference picture and the candidate block in the backward reference picture, to determine two blocks with the minimum matching cost. This simplifies a picture prediction process, improves picture prediction accuracy, and reduces complexity.

[0209] FIG. 11 is a schematic flowchart of a picture prediction method according to an embodiment of this application. The method shown in FIG. 11 may be performed by a video coding apparatus, a video codec, a video coding system, or another device having a video coding function. The method shown in FIG. 11 includes step 1101 to step 1105. For step 1101 to step 1103 and step 1105, refer to descriptions of step 1001 to step 1003 and step 1007 in FIG. 10. Details are not described herein again.

[0210] A difference between this embodiment of this application and the embodiment shown in FIG. 10 lies in that pixel values of current optimal forward and backward reference blocks are retained and updated in a search process. After the search

is completed, the predicted value of the pixel value of the current picture block can be calculated by using the pixel values of the current optimal forward and backward reference blocks.

[0211] For example, positions of N pairs of reference blocks need to be traversed.

5 $Cost_i$ is an i^{th} matching cost, and $MinCost$ indicates a current minimum matching cost. Bf_i is a pixel value of a forward reference block, and Bb_i is a pixel value of a backward reference block, where the pixel values are obtained for an i^{th} time. $BestBf$ is a value of the current optimal forward reference block and $BestBb$ is a value of the current optimal backward reference block. $CalCost(M, N)$ represents a matching cost of a block M and
10 a block N.

[0212] When search starts ($i=1$), $MinCost=Cost_0=CalCost(Bf_0, Bb_0)$, $BestBf=Bf_0$, and $BestBb=Bb_0$.

[0213] When other pairs of reference blocks are subsequently traversed, $BestBf$ and $BestBb$ are updated in real time. For example, when an i^{th} ($i>1$) round of search is
15 performed, if $Cost_i < MinCost$, $BestBf=Bf_i$ and $BestBb=Bb_i$; otherwise, no update is performed.

[0214] When the search ends, $BestBf$ and $BestBb$ are used to obtain the predicted value of the pixel value of the current block.

[0215] FIG. 12 is a schematic flowchart of a picture prediction method according
20 to an embodiment of this application. The method shown in FIG. 12 may be performed by a video coding apparatus, a video codec, a video coding system, or another device having a video coding function. The method shown in FIG. 12 may be used in a coding process or a decoding process. More specifically, the method shown in FIG. 12 may be used in an inter prediction process during encoding or decoding. A process 1200 may
25 be performed by a video encoder 20 or a video decoder 30, and specifically may be performed by a motion compensation unit of the video encoder 20 or the video decoder 30. It is assumed that for a video data stream with a plurality of video frames, the video encoder or the video decoder is being used to perform the process 1200 including the following steps, to obtain a predicted value of a pixel value of a current picture block of a current video frame.

30 **[0216]** The method shown in FIG. 12 includes step 1201 to step 1204. For step 1201,

block 904 in a forward reference picture Ref0 relative to a position of an $(i-1)^{\text{th}}$ -round target forward reference block 902 (namely, a forward search base point) is MVD0 (delta0x , delta0y). An offset of a position of a candidate backward reference block 905 in a backward reference picture Ref1 relative to a position of an $(i-1)^{\text{th}}$ -round target backward reference block 903 (namely, a backward search base point) is MVD1 (delta1x , delta1y).

[0247] MVD0= $-\text{MVD1}$, to be specific:

$\text{delta0x} = -\text{delta1x}$; and

$\text{delta0y} = -\text{delta1y}$.

10 **[0248]** In a different example, step 1403 may include:

determining, from the positions of the M pairs of reference blocks (one forward reference block and one backward reference block), that positions of a pair of reference blocks with a minimum matching error are the position of the i^{th} -round target forward reference block of the current picture block and the position of the i^{th} -round target backward reference block of the current picture block; or determining, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a matching error less than or equal to a matching error threshold are the position of the i^{th} -round target forward reference block of the current picture block and the position of the i^{th} -round target backward reference block of the current picture block, where M is less than or equal to N. In addition, a difference between a pixel value of a forward reference block and a pixel value of a backward reference block may be measured by using a sum of absolute differences (Sum of absolute differences, SAD), a sum of absolute transformation differences (Sum of absolute transformation differences, SATD), a sum of absolute square differences, or the like.

25 **[0249]** 1404: Obtain a predicted value of a pixel value of the current picture block based on a pixel value of the target forward reference block and a pixel value of the target backward reference block.

[0250] In an example, in step 1404, weighted processing is performed on the pixel value of the target forward reference block and the pixel value of the target backward reference block, to obtain the predicted value of the pixel value of the current picture

forward reference block may be used as a search base point, and positions of (N-1) candidate forward reference blocks are determined in the forward reference picture (also referred to as a forward reference region). A position of an initial backward reference block is used as a search base point, and positions of (N-1) candidate backward reference blocks are determined in the backward reference picture (also referred to as a backward reference region). For one or more pairs of reference block positions in positions of N pairs of reference blocks, a matching cost of two corresponding reference blocks is calculated, for example, a matching cost of the initial forward reference block and the initial backward reference block is calculated, and a matching cost of a candidate forward reference block and a candidate backward reference block that meet the MVD mirror constraint is calculated. In this way, a position of a first-round target forward reference block and a position of a first-round target backward reference block in the first round of search are obtained, and updated motion information is further obtained. The updated motion information includes a forward motion vector indicating that a position of the current picture block points to the position of the first-round target forward reference block, and a backward motion vector indicating that the position of the current picture block points to the position of the first-round target backward reference block. It should be understood that the updated motion information and the initial motion information include a same reference frame index and the like. Next, a second round of search is performed. The position of the first-round target forward reference block is used as a search base point, and positions of (N-1) candidate forward reference blocks are determined in the forward reference picture (also referred to as the forward reference region). The position of the first-round target backward reference block is used as a search base point, and positions of (N-1) candidate backward reference blocks are determined in the backward reference picture (also referred to as the backward reference region). For one or more pairs of reference block positions in positions of N pairs of reference blocks, a matching cost of two corresponding reference blocks is calculated, for example, a matching cost of the first-round target forward reference block and the first-round target backward reference block is calculated, and a matching cost of a candidate forward reference block and a

[0270] In addition, in step 1504, a fractional pixel step may also be directly used to search for the position of the at least one forward reference block and the position of the at least one backward reference block, or both the search by using the integer pixel step and the search by using the fractional pixel step are performed. A search method is not limited in this application.

[0271] In step 1504, for positions of each pair of reference blocks, when a difference between a pixel value of a forward reference block and a pixel value of a corresponding backward reference block is calculated, a difference between a pixel value of each forward reference block and a pixel value of a corresponding backward reference block may be measured by using an SAD, an SATD, a sum of absolute square differences, or the like. However, this application is not limited thereto.

[0272] When the predicted value of the pixel value of the current picture block is determined based on the optimal forward predicted block and the optimal backward predicted block, weighted processing may be performed on the pixel value of the optimal forward reference block and the pixel value of the optimal backward reference block that are obtained in step 1506 and step 1507, and a pixel value obtained after the weighted processing is used as the predicted value of the pixel value of the current picture block.

[0273] Specifically, the predicted value of the pixel value of the current picture block may be obtained based on a formula (8):

$$\text{predSamples}'[x][y] = (\text{predSamplesL0}'[x][y] + \text{predSamplesL1}'[x][y] + 1) >> 1$$

(8)

[0274] In the foregoing formula, $\text{predSamplesL0}'[x][y]$ is a pixel value of the optimal forward reference block at a sample (x, y), $\text{predSamplesL1}'[x][y]$ is a pixel value of the optimal backward reference block at the sample (x, y), and $\text{predSamples}'[x][y]$ is a pixel predicted value of the current picture block at the sample (x, y).

[0275] Referring to FIG. 11, pixel values of a current optimal forward reference block and a current optimal backward reference block may further be retained and updated in an iterative search process in this embodiment of this application. After the

example, $T=2$, or $T=3$. A specific value of T is not limited herein.

[0309] For another example, it is assumed that a ratio of a larger value of $|POC_c-POC_0|$ and $|POC_c-POC_1|$ to a smaller value of $|POC_c-POC_0|$ and $|POC_c-POC_1|$ is greater than a threshold R , in other words,

5 $(\text{Max}(|POC_c-POC_0|, |POC_c-POC_1|) / \text{Min}(|POC_c-POC_0|, |POC_c-POC_1|)) > R$.

[0310] $\text{Max}(A, B)$ indicates a larger value of A and B , and $\text{Min}(A, B)$ indicates a smaller value of A and B .

[0311] In this case, an interval needs to be considered for the mirror relationship
10 that is used. If the ratio of the larger value of $|POC_c-POC_0|$ and $|POC_c-POC_1|$ to the smaller value of $|POC_c-POC_0|$ and $|POC_c-POC_1|$ is not greater than the threshold R , no time interval is considered for the mirror relationship that is used. R herein is a preset threshold. For example, $R=2$, or $R=3$. A specific value of R is not limited herein.

[0312] It should be understood that the picture prediction method in this
15 embodiment of this application may be specifically performed by a motion compensation module in an encoder (for example, an encoder 20) or a decoder (for example, a decoder 30). In addition, the picture prediction method in this embodiment of this application may be performed in any electronic device or apparatus that needs to code and/or decode a video picture.

20 **[0313]** The following describes in detail a picture prediction apparatus in the embodiments of this application with reference to FIG. 18 to FIG. 21.

[0314] FIG. 18 is a schematic block diagram of a picture prediction apparatus 1800 according to an embodiment of this application. It should be noted that the prediction apparatus 1800 is applicable to both inter prediction for decoding a video picture and
25 inter prediction for coding a video picture. It should be understood that the prediction apparatus 1800 herein may correspond to the motion compensation unit 44 in FIG. 2A, or may correspond to the motion compensation unit 82 in FIG. 2B. The prediction apparatus 1800 may include:

a first obtaining unit 1801, configured to obtain initial motion information
30 of a current picture block;

CLAIMS

What is claimed is:

1. A picture prediction method, comprising:
 - obtaining initial motion information of a current picture block;
 - 5 determining positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, wherein the N forward reference blocks are located in a forward reference picture, the N backward reference blocks are located in a backward reference picture, and N is an integer greater than 1;
 - 10 determining, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward reference block of the current picture block, wherein positions of each pair of reference blocks comprise a position of a forward reference block and a position of a backward reference block;
 - 15 for the positions of each pair of reference blocks, a first position offset and a second position offset are in a mirror relationship, the first position offset represents an offset of the position of the forward reference block relative to a position of an initial forward reference block, and the second position offset represents an offset of the position of the backward reference block relative to a position of an initial backward reference block; and M is an integer greater than or equal to 1, and M is less than or equal to N; and
 - 20 obtaining a predicted value of a pixel value of the current picture block based on a pixel value of the target forward reference block and a pixel value of the target backward reference block.
- 25 2. The method according to claim 1, wherein a first position offset and a second position offset are in a mirror relationship comprises:
 - a direction of the first position offset is opposite to a direction of the second position offset, and an amplitude value of the first position offset is the same as an amplitude value of the second position offset.

initial motion information of a current picture block, the method further comprises: obtaining indication information from a bitstream of the current picture block, wherein the indication information is used to indicate the initial motion information of the current picture block.

5 8. The method according to any one of claims 1 to 7, wherein a first position offset is represented by $(\text{delta0x}, \text{delta0y})$ and the second position offset is represented by $(\text{delta1x}, \text{delta1y})$, wherein $\text{delta0x} = -\text{delta1x}$, and $\text{delta0y} = -\text{delta1y}$.

 9. A picture prediction method, comprising:

 obtaining initial motion information of a current picture block;

10 determining positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, wherein the N forward reference blocks are located in a forward reference picture, the N backward reference blocks are located in a backward reference picture, and N is an integer greater than 1;

15 determining, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward reference block of the current picture block, wherein positions of each pair of reference blocks comprise a position of a forward reference block and a position of a backward reference block; for the positions of each pair of reference blocks, a first position offset and a second position offset are in a proportional relationship based on a time domain distance, the first position offset represents an offset of the position of the forward reference block relative to a position of an initial forward reference block, and the second position offset represents an offset of the position of the backward reference block relative to a position of an initial backward reference block; and M is an integer greater than or equal to 1, and M is less than or equal to N; and

 obtaining a predicted value of a pixel value of the current picture block based on a pixel value of the target forward reference block and a pixel value of the target backward reference block.

30 10. The method according to claim 9, wherein that a first position offset and a

second position offset are in a proportional relationship based on a time domain distance comprises:

if a first time domain distance is the same as a second time domain distance, a direction of the first position offset is opposite to a direction of the second position offset, and an amplitude value of the first position offset is the same as an amplitude value of the second position offset; or

if a first time domain distance is different from a second time domain distance, a direction of the first position offset is opposite to a direction of the second position offset, and a proportional relationship between an amplitude value of the first position offset and an amplitude value of the second position offset is based on a proportional relationship between the first time domain distance and the second time domain distance; wherein

the first time domain distance represents a time domain distance between a current picture to which the current picture block belongs and the forward reference picture, and the second time domain distance represents a time domain distance between the current picture and the backward reference picture.

11. The method according to claim 9 or 10, wherein the method further comprises:

obtaining updated motion information of the current picture block, wherein the updated motion information comprises an updated forward motion vector and an updated backward motion vector, the updated forward motion vector points to the position of the target forward reference block, and the updated backward motion vector points to the position of the target backward reference block.

12. The method according to any one of claims 9 to 11, wherein the positions of the N forward reference blocks comprise a position of one initial forward reference block and positions of (N-1) candidate forward reference blocks, and an offset of a position of each candidate forward reference block relative to the position of the initial forward reference block is an integer pixel distance or a fractional pixel distance; or

the positions of the N backward reference blocks comprise a position of one initial backward reference block and positions of (N-1) candidate backward reference blocks, and an offset of a position of each candidate backward reference block relative to the

position of the initial backward reference block is an integer pixel distance or a fractional pixel distance.

13. The method according to any one of claims 9 to 12, wherein the initial motion information comprises a first motion vector and a first reference picture index in a forward prediction direction, and a second motion vector and a second reference picture index in a backward prediction direction; and

the determining positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block comprises:

10 determining, based on the first motion vector and the position of the current picture block, the position of the initial forward reference block of the current picture block in the forward reference picture corresponding to the first reference picture index, using the position of the initial forward reference block as a first search start point, and determining the positions of the (N-1) candidate forward reference blocks in the forward reference picture, wherein the positions of the N forward reference blocks comprise the position of the initial forward reference block and the positions of the (N-1) candidate forward reference blocks; and

determining, based on the second motion vector and the position of the current picture block, the position of the initial backward reference block of the current picture block in the backward reference picture corresponding to the second reference picture index, using the position of the initial backward reference block as a second search start point, and determining the positions of the (N-1) candidate backward reference blocks in the backward reference picture, wherein the positions of the N backward reference blocks comprise the position of the initial backward reference block and the positions of the (N-1) candidate backward reference blocks.

14. The method according to any one of claims 9 to 13, wherein the determining, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward reference block of the current picture block comprises:

determining, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a minimum matching error are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block; or

5 determining, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a matching error less than or equal to a matching error threshold are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block, wherein M is less than or equal to N.

10 15. The method according to any one of claims 9 to 14, wherein the method is used to code the current picture block; and the obtaining initial motion information of a current picture block comprises: obtaining the initial motion information from a candidate motion information list of the current picture block; or

the method is used to decode the current picture block; and before the obtaining
15 initial motion information of a current picture block, the method further comprises: obtaining indication information from a bitstream of the current picture block, wherein the indication information is used to indicate the initial motion information of the current picture block.

16. A picture prediction apparatus, comprising:

20 a first obtaining unit, configured to obtain initial motion information of a current picture block;

a first search unit, configured to: determine positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, wherein the N forward reference
25 blocks are located in a forward reference picture, the N backward reference blocks are located in a backward reference picture, and N is an integer greater than 1; and determine, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward
30 reference block of the current picture block, wherein positions of each pair of reference

blocks comprise a position of a forward reference block and a position of a backward reference block; for the positions of each pair of reference blocks, a first position offset and a second position offset are in a mirror relationship, the first position offset represents an offset of the position of the forward reference block relative to a position
5 of an initial forward reference block, and the second position offset represents an offset of the position of the backward reference block relative to a position of an initial backward reference block; and M is an integer greater than or equal to 1, and M is less than or equal to N ; and

a first prediction unit, configured to obtain a predicted value of a pixel value of the
10 current picture block based on a pixel value of the target forward reference block and a pixel value of the target backward reference block.

17. The apparatus according to claim 16, wherein that a first position offset and a second position offset are in a mirror relationship comprises: a direction of the first position offset is opposite to a direction of the second position offset, and an amplitude
15 value of the first position offset is the same as an amplitude value of the second position offset.

18. The apparatus according to claim 16 or 17, wherein the first prediction unit is further configured to obtain updated motion information of the current picture block, wherein the updated motion information comprises an updated forward motion vector
20 and an updated backward motion vector, the updated forward motion vector points to the position of the target forward reference block, and the updated backward motion vector points to the position of the target backward reference block.

19. The apparatus according to any one of claims 16 to 18, wherein the positions of the N forward reference blocks comprise a position of one initial forward reference
25 block and positions of $(N-1)$ candidate forward reference blocks, and an offset of a position of each candidate forward reference block relative to the position of the initial forward reference block is an integer pixel distance or a fractional pixel distance; or

the positions of the N backward reference blocks comprise a position of one initial backward reference block and positions of $(N-1)$ candidate backward reference blocks,
30 and an offset of a position of each candidate backward reference block relative to the

position of the initial backward reference block is an integer pixel distance or a fractional pixel distance.

20. The apparatus according to any one of claims 16 to 19, wherein the initial motion information comprises a first motion vector and a first reference picture index
5 in a forward prediction direction, and a second motion vector and a second reference picture index in a backward prediction direction; and

in the aspect of determining positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, the first search unit is specifically configured to:

10 determine, based on the first motion vector and the position of the current picture block, the position of the initial forward reference block of the current picture block in the forward reference picture corresponding to the first reference picture index, use the position of the initial forward reference block as a first search start point, and determine the positions of the (N-1) candidate forward reference blocks in the forward reference
15 picture, wherein the positions of the N forward reference blocks comprise the position of the initial forward reference block and the positions of the (N-1) candidate forward reference blocks; and

determine, based on the second motion vector and the position of the current picture block, the position of the initial backward reference block of the current picture
20 block in the backward reference picture corresponding to the second reference picture index, use the position of the initial backward reference block as a second search start point, and determine the positions of the (N-1) candidate backward reference blocks in the backward reference picture, wherein the positions of the N backward reference blocks comprise the position of the initial backward reference block and the positions
25 of the (N-1) candidate backward reference blocks.

21. The apparatus according to any one of claims 16 to 20, wherein in the aspect of determining, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward
30 reference block of the current picture block, the first search unit is specifically

configured to:

determine, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a minimum matching error are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block; or

determine, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a matching error less than or equal to a matching error threshold are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block, wherein M is less than or equal to N.

22. A picture prediction apparatus, comprising:

a second obtaining unit, configured to obtain initial motion information of a current picture block;

a second search unit, configured to: determine positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, wherein the N forward reference blocks are located in a forward reference picture, the N backward reference blocks are located in a backward reference picture, and N is an integer greater than 1; and determine, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward reference block of the current picture block, wherein positions of each pair of reference blocks comprise a position of a forward reference block and a position of a backward reference block; for the positions of each pair of reference blocks, a first position offset and a second position offset are in a proportional relationship based on a time domain distance, the first position offset represents an offset of the position of the forward reference block relative to a position of an initial forward reference block, and the second position offset represents an offset of the position of the backward reference block relative to a position of an initial backward reference block; and M is an integer greater than or equal to 1, and M is less than or equal to N; and

a second prediction unit, configured to obtain a predicted value of a pixel value of the current picture block based on a pixel value of the target forward reference block and a pixel value of the target backward reference block.

23. The apparatus according to claim 22, wherein that a first position offset and a
5 second position offset are in a proportional relationship based on a time domain distance comprises:

if a first time domain distance is the same as a second time domain distance, a direction of the first position offset is opposite to a direction of the second position offset, and an amplitude value of the first position offset is the same as an amplitude
10 value of the second position offset; or

if a first time domain distance is different from a second time domain distance, a direction of the first position offset is opposite to a direction of the second position offset, and a proportional relationship between an amplitude value of the first position offset and an amplitude value of the second position offset is based on a proportional
15 relationship between the first time domain distance and the second time domain distance; wherein

the first time domain distance represents a time domain distance between a current picture to which the current picture block belongs and the forward reference picture, and the second time domain distance represents a time domain distance between the
20 current picture and the backward reference picture.

24. The apparatus according to claim 22 or 23, wherein the second prediction unit is further configured to obtain updated motion information of the current picture block, wherein the updated motion information comprises an updated forward motion vector and an updated backward motion vector, the updated forward motion vector points to
25 the position of the target forward reference block, and the updated backward motion vector points to the position of the target backward reference block.

25. The apparatus according to any one of claims 22 to 24, wherein the positions of the N forward reference blocks comprise a position of one initial forward reference block and positions of (N-1) candidate forward reference blocks, and an offset of a
30 position of each candidate forward reference block relative to the position of the initial

forward reference block is an integer pixel distance or a fractional pixel distance; or

the positions of the N backward reference blocks comprise a position of one initial backward reference block and positions of (N-1) candidate backward reference blocks, and an offset of a position of each candidate backward reference block relative to the position of the initial backward reference block is an integer pixel distance or a fractional pixel distance.

26. The apparatus according to any one of claims 22 to 25, wherein the initial motion information comprises a first motion vector and a first reference picture index in a forward prediction direction, and a second motion vector and a second reference picture index in a backward prediction direction; and

in the aspect of determining positions of N forward reference blocks and positions of N backward reference blocks based on the initial motion information and a position of the current picture block, the second search unit is specifically configured to:

determine, based on the first motion vector and the position of the current picture block, the position of the initial forward reference block of the current picture block in the forward reference picture corresponding to the first reference picture index, use the position of the initial forward reference block as a first search start point, and determine the positions of the (N-1) candidate forward reference blocks in the forward reference picture, wherein the positions of the N forward reference blocks comprise the position of the initial forward reference block and the positions of the (N-1) candidate forward reference blocks; and

determine, based on the second motion vector and the position of the current picture block, the position of the initial backward reference block of the current picture block in the backward reference picture corresponding to the second reference picture index, use the position of the initial backward reference block as a second search start point, and determine the positions of the (N-1) candidate backward reference blocks in the backward reference picture, wherein the positions of the N backward reference blocks comprise the position of the initial backward reference block and the positions of the (N-1) candidate backward reference blocks.

27. The apparatus according to any one of claims 22 to 26, wherein in the aspect

of determining, from positions of M pairs of reference blocks based on a matching cost criterion, that positions of a pair of reference blocks are a position of a target forward reference block of the current picture block and a position of a target backward reference block of the current picture block, the second search unit is specifically
5 configured to:

determine, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a minimum matching error are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block; or

10 determine, from the positions of the M pairs of reference blocks, that positions of a pair of reference blocks with a matching error less than or equal to a matching error threshold are the position of the target forward reference block of the current picture block and the position of the target backward reference block of the current picture block, wherein M is less than or equal to N.

15 28. A video encoder, wherein the video encoder is configured to code a picture block, and comprises:

an inter prediction module, comprising the picture prediction apparatus according to any one of claims 16 to 27, wherein the inter prediction module is configured to obtain a predicted value of a pixel value of the picture block through prediction;

20 an entropy coding module, configured to code indication information into a bitstream, wherein the indication information is used to indicate initial motion information of the picture block; and

a reconstruction module, configured to reconstruct the picture block based on the predicted value of the pixel value of the picture block.

25 29. A video decoder, wherein the video decoder is configured to decode a bitstream to obtain a picture block, and comprises:

an entropy decoding module, configured to decode the bitstream to obtain indication information, wherein the indication information is used to indicate initial motion information of a current picture block obtained through decoding;

30 an inter prediction module, comprising the picture prediction apparatus according

to any one of claims 16 to 27, wherein the inter prediction module is configured to obtain a predicted value of a pixel value of the picture block through prediction; and

a reconstruction module, configured to reconstruct the picture block based on the predicted value of the pixel value of the picture block.

- 5 30. A non-transitory computer-readable medium carrying a program code which, when executed by a computer device, causes the computer device to perform the method of any one of the preceding claims.

31. A non-transitory storage medium comprising a bitstream encoded by the method of any one of the preceding claims.

10