

PICTURE PREDICTION METHOD AND PICTURE PREDICTION APPARATUS

TECHNICAL FIELD

[0001] The present invention relates to the field of video encoding and video decoding,
5 and specifically, to a picture prediction method and a related device.

BACKGROUND

[0002] With development of an optoelectronic collection technology and an increasing
requirement for a high-definition digital video, a video data volume becomes increasingly
large. Limited heterogeneous transmission bandwidth and diversified video applications
10 impose a higher requirement on video coding efficiency. In this case, the High Efficiency
Video Coding (HEVC) standard starts to be formulated as required.

[0003] A basic principle of video coding compression is using correlation between a
space domain, a time domain, and a codeword to eliminate redundancy as much as possible.
At present, a common manner is using a block-based hybrid video coding framework to
15 implement video coding compression by means of steps such as prediction (including
intra-frame prediction and inter-frame prediction), transformation, quantization, and entropy
coding. This coding framework is powerful, and the block-based hybrid video coding
framework is also used for HEVC.

[0004] In various video encoding/decoding schemes, motion estimation/motion
20 compensation is a key technology affecting encoding/decoding performance. In many existing
video encoding/decoding schemes, it is generally assumed that a motion of an object meets a
requirement of a translational motion model, and various parts of the entire object are in a
same motion. An existing motion estimation/motion compensation algorithm is basically a
block-based motion compensation algorithm based on a translational motion model (English:
25 translational motion model). Existing inter-frame prediction is mainly block-based motion

with the phase of Q that is corresponding to the corresponding pixel unit.

[0012] With reference to the third possible implementation of the first aspect, in a fourth possible implementation of the first aspect, the obtaining, by means of calculation, a phase of each of the P pixel units by using the motion vector of each of the P pixel units includes:
5 obtaining, by means of calculation, the phase of each of the P pixel units according to the following formula by using the motion vector of each of the P pixel units:

$$\begin{aligned} X' &= \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1); \\ Y' &= \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1); \end{aligned} \quad \text{where}$$

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a
10 vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0013] With reference to the third possible implementation of the first aspect or the fourth possible implementation of the first aspect, in a fifth possible implementation of the first aspect, the phase includes a horizontal phase and a vertical phase; and the determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit includes: determining, based on the horizontal phase of each
20 pixel unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is
25 corresponding to the vertical phase.

[0014] With reference to the fifth possible implementation of the first aspect, in a sixth possible implementation of the first aspect, the performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using

the determined interpolation filter with the phase of Q that is corresponding to the pixel unit includes:

performing horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, where the pixel unit j is any one of the P pixel units.

[0015] With reference to any one of the first aspect, or the first to the sixth possible implementations of the first aspect, in a seventh possible implementation of the first aspect, the motion model is a translational motion model, an affine motion model, a rotational motion model, a parabolic motion model, a shearing motion model, a zooming motion model, a perspective motion model, or a bilinear motion model.

[0016] With reference to any one of the first aspect, or the first to the seventh possible implementations of the first aspect, in an eighth possible implementation of the first aspect, the motion model is represented as follows when W is equal to 2:

$$\begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{L} x - \frac{(v_{1y} - v_{0y}) \times N}{L} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{L} x + \frac{(v_{1x} - v_{0x}) \times N}{L} y + Nv_{0y} \right) / n \end{cases}; \text{ or}$$

pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0025] With reference to the third possible implementation of the second aspect or the fourth possible implementation of the second aspect, in a fifth possible implementation of the second aspect, the phase includes a horizontal phase and a vertical phase; and in an aspect of determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the corresponding pixel unit, the interpolation filtering unit is specifically configured to: determine, based on the horizontal phase of each pixel unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determine, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

[0026] With reference to the fifth possible implementation of the second aspect, in a sixth possible implementation of the second aspect, in an aspect of performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the corresponding pixel unit, the interpolation filtering unit is specifically configured to: perform horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and perform vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or the interpolation filtering unit is specifically configured to: perform vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and perform horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal

embodiment of the present invention; and

[0044] FIG. 6 is a schematic diagram of another picture prediction apparatus according to an embodiment of the present invention.

DESCRIPTION OF EMBODIMENTS

5 [0045] Embodiments of the present invention provide a picture prediction method and a related device, so as to reduce a quantity of intermediate caches and memory operations that are required for interpolation filtering during a picture prediction process, and reduce calculation complexity during the picture prediction process.

[0046] In the specification, claims, and accompanying drawings of the present invention,
10 the terms "first", "second", "third", and so on are intended to distinguish between different objects but do not indicate a particular order. In addition, the terms "including", "comprising", and any other variant thereof are intended to cover a non-exclusive inclusion. For example, a process, a method, a system, a product, or a device that includes a series of steps or units is not limited to the listed steps or units, but optionally further includes an unlisted step or unit,
15 or optionally further includes another inherent step or unit of the process, the method, the product, or the device.

[0047] The following first briefly describes some concepts that may be related to the embodiments of the present invention.

[0048] In most coding frameworks, a video sequence includes a series of pictures, a
20 picture is further divided into slices, and a slice is further divided into blocks. Video coding is based on a unit of a block, and coding processing may start to be performed at a location of an upper left corner of a picture and then performed line by line from left to right and from top to bottom. In some new video coding standards, a concept of the block is further extended. In the H.264 standard, a macroblock (MB) is described, and the MB may be further divided into
25 multiple prediction partitions that may be used for predictive coding. In the HEVC standard, basic concepts of a coding unit (CU), a prediction unit (PU), a transform unit

(TU), and the like are used. Multiple types of units are obtained by means of function division, and are described by using a new tree-based structure. For example, the CU may be divided into smaller CUs according to a quad tree, and a smaller CU may further continue to be divided to form a quad tree structure. Tree structures of the PU and the TU are similar to that of the CU. The CU, the PU, and the TU all belong to the concept of the block in essence. Similar to a macroblock MB or a coding block, the CU is a basic unit for dividing and encoding a coding picture. The PU is a basic unit for predictive coding, and may be corresponding to a prediction partition. According to a division manner, the CU is further divided into multiple PUs. The TU is a basic unit for transforming a predicted residual, and may be corresponding to a transform block. In the High Efficiency Video Coding (English: high efficiency video coding, HEVC for short) standard, the CU, the PU, and the TU may be collectively referred to as a coding tree block (English: coding tree block, CTB for short), and the like.

[0049] In the HEVC standard, the coding unit may include four levels in size: 64 x 64, 32 x 32, 16 x 16, and 8 x 8. Each level of coding unit may be divided into prediction units of different sizes according to intra-frame prediction and inter-frame prediction. For example, as shown in FIG. 1-a and FIG. 1-b, FIG. 1-a shows a prediction unit division manner corresponding to intra-frame prediction, and FIG. 1-b shows several prediction unit division manners corresponding to inter-frame prediction.

[0050] During a development and evolution process of a video coding technology, experts in video coding figure out various methods to utilize spatial-temporal correlation between adjacent encoded/decoded blocks to improve coding efficiency. In the H.264/Advanced Video Coding (English: advanced video coding, AVC for short) standard, a skip mode (skip mode) and a direct mode (direct mode) become effective means to improve coding efficiency. In a case of a low bit rate, a quantity of blocks using the two coding modes accounts for more than a half of blocks in an entire coding sequence. When the skip mode is used, a motion vector of a current picture block may be obtained by means of derivation by using a surrounding motion vector provided that a skip mode tag is transferred in a bitstream, and a value of a reference block is directly used as a reconstruction value of the current picture block according to the motion vector. Alternatively, when the direct mode is used, an encoder may

obtain, by means of derivation, a motion vector of a current picture block by using a surrounding motion vector, directly use a value of a reference block as a predicted value of the current picture block according to the motion vector, and perform predictive coding on the current picture block on an encoder side by using the predicted value. At present, some new coding means are used in the latest High Efficiency Video Coding (English: high efficiency video coding, HEVC for short) standard, to further improve video coding performance. A merge coding mode and an advanced motion vector prediction (AMVP) mode are two important inter-frame prediction means. In the merge coding mode, motion information (including a motion vector (MV), a prediction direction, a reference-frame index, and the like) of a coded block that surrounds a current coding block is used to form a set of candidate motion information. Candidate motion information with highest coding efficiency may be selected, by means of comparison, as motion information of the current coding block. Predictive coding is performed on the current coding block by using a predicted value, found in a reference frame, of the current coding block. In addition, an index value indexing a specific surrounding coded block from which motion information is selected may be written into a bitstream. When the advanced motion vector prediction mode is used, a motion vector of a surrounding coded block is used as a predicted value of a motion vector of a current coding block, a motion vector with highest coding efficiency may be selected to predict the motion vector of the current coding block, and an index value indicating selection of a specific surrounding motion vector may be written into a video bitstream.

[0051] The following continues to describe the technical solutions of the embodiments of the present invention.

[0052] The following first describes the picture prediction method provided in the embodiments of the present invention. The picture prediction method provided in the embodiments of the present invention is executed by a video encoding apparatus or a video decoding apparatus. The video encoding apparatus or the video decoding apparatus may be any apparatus that needs to output or store a video, for example, a notebook computer, a tablet computer, a personal computer, a mobile phone, a video server, or another device.

[0053] In an embodiment of the picture prediction method provided in the present

preset fixed value, and Q may be less than or equal to N. When the value of N may be a preset fixed value, it indicates that the precision of the motion vectors, obtained by means of calculation by using the motion model and the motion vectors of the W control points, of the P pixel units of the current picture block is unnecessarily related to a size of the current picture block. In other words, for example, based on the solution of this embodiment, the predicted pixel value with preset fixed precision can be obtained without an intermediate process for obtaining a lower-precision predicted pixel value by performing lower-precision interpolation filtering.

[0071] For example, W may be equal to 2, 3, 4, 5, 6, 8, or another value.

[0072] For example, P may be equal to 1, 2, 3, 4, 5, 6, 8, 10, 15, 16, 21, 32, 64, or another value.

[0073] For example, Q may be equal to 128, 9, 18, 24, 256, 8, 10, 15, 16, 21, 32, 64, or another value.

[0074] For example, N may be equal to 128, 9, 18, 24, 256, 8, 10, 15, 16, 21, 32, 64, or another value.

[0075] For example, n may be equal to 8, 2, 4, or another value.

[0076] In some possible implementations of the present invention, N may be an integral power of 2, or certainly, N may be another positive integer.

[0077] A pixel unit in the embodiments of the present invention may include one or more pixels. For example, the pixel unit may be a 2 x 2 pixel block, a 2 x 1 pixel block, a 2 x 1 pixel block, a 4 x 4 pixel block, or a 4 x 2 pixel block.

[0078] A control point in the embodiments of the present invention may include one or more pixels. For example, the control point may be a 2 x 2 pixel block, a 2 x 1 pixel block, a 2 x 1 pixel block, a 4 x 4 pixel block, or a 4 x 2 pixel block.

[0079] It can be learnt from the foregoing that, in the picture prediction method provided in this embodiment, the motion vector, whose precision is $1/N$ of the pixel precision, of each pixel unit of the current picture block is obtained by means of calculation by using the motion model and the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points, where N is greater than n. In other words, the precision of the motion vector that is obtained by means of calculation and that is of each pixel unit of the current picture block is

(v_{1x}, v_{1y}) represent motion vectors, whose precision is $1/n$ of the pixel precision, of two control points, v_{Nx} represents a horizontal component of a motion vector, whose precision is $1/N$ of the pixel precision, of a pixel unit with coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is $1/N$ of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0082] For another example, in some possible implementations of the present invention, the motion model may be represented as follows when W is equal to 3:

$$\begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{w} x + \frac{(v_{2x} - v_{0x}) \times N}{h} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{w} x + \frac{(v_{2y} - v_{0y}) \times N}{h} y + Nv_{0y} \right) / n \end{cases}; \text{ or}$$

$$\begin{cases} v_{Nx} = \frac{(v_{1x} - v_{0x}) \times (N/n)}{w} x + \frac{(v_{2x} - v_{0x}) \times (N/n)}{h} y + (N/n)v_{0x} \\ v_{Ny} = \frac{(v_{1y} - v_{0y}) \times (N/n)}{w} x + \frac{(v_{2y} - v_{0y}) \times (N/n)}{h} y + (N/n)v_{0y} \end{cases}, \text{ where}$$

v_{Nx} represents a horizontal component of a motion vector, whose precision is $1/N$ of the pixel precision, of a pixel unit with coordinates of (x, y) in the current picture block, v_{Ny} represents a vertical component of the motion vector, whose precision is $1/N$ of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, (v_{0x}, v_{0y}) , (v_{1x}, v_{1y}) , and (v_{2x}, v_{2y}) represent motion vectors, whose precision is $1/n$ of the pixel precision, of three control points, w represents a width of the current picture block, and h represents a height of the current picture block.

[0083] The foregoing examples are described by using the affine motion model. When a translational motion model, a rotational motion model, a shearing motion model, a zooming motion model, a perspective motion model, a parabolic motion model, a bilinear motion model, or the like is used, reference may be made to the foregoing examples. Details are not described herein again.

with the phase of Q that is corresponding to the corresponding pixel unit.

5 [0089] Specifically, for example, a phase of the pixel unit i may be obtained by means of calculation by using the motion vector of the pixel unit i of the P pixel units; an interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on the phase of the pixel unit i, where a filter coefficient used by the interpolation filter is corresponding to the phase; and interpolation filtering may be performed on a pixel of the
10 corresponding reference pixel unit, in the reference picture, of the pixel unit i by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit i. The pixel unit i may be any one of the P pixel units. For example, interpolation filtering may be performed on each of the P pixel units according to a manner similar to that of performing interpolation filtering on the pixel unit i.

15 [0090] In some possible implementations of the present invention, the obtaining, by means of calculation, a phase of each of the P pixel units by using the motion vector of each of the P pixel units includes: obtaining, by means of calculation, the phase of each of the P pixel units according to the following formula by using the motion vector of each of the P pixel units:

$$\begin{aligned} X' &= \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1). \\ Y' &= \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1). \end{aligned} \quad \text{where}$$

20

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the
25 pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0091] Optionally, in some possible implementations of the present invention, the phase

includes a horizontal phase and a vertical phase; and the determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit includes: determining, based on the horizontal phase of each pixel

unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

[0092] Specifically, for example, a horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on a horizontal phase of the pixel unit i, and a vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on a vertical phase of the pixel unit i, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase of the pixel unit i, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase of the pixel unit i. The pixel unit i may be any one of the P pixel units. For example, the interpolation filter with the phase of Q that is corresponding to each of the P pixel units may be determined according to a manner similar to that of determining the interpolation filter with the phase of Q that is corresponding to the pixel unit i.

[0093] Optionally, in some possible implementations of the present invention, the performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit includes:

performing horizontal interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of the pixel unit i by using the determined horizontal interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using the determined vertical interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j , to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j , to obtain a predicted pixel value of the pixel unit j , where the pixel unit j is any one of the P pixel units.

[0094] In addition, if a pixel unit has only one phase (that is, has only a horizontal phase or a vertical phase), interpolation filtering needs to be performed on a pixel of a reference pixel unit of the pixel unit only once, to obtain a predicted pixel value of the pixel unit.

[0095] Optionally, in some possible implementations of the present invention, the determining motion vectors of W control points in a current picture block includes:

determining the W control points in the current picture block, and determining a candidate motion information unit set corresponding to each of the W control points, where the candidate motion information unit set corresponding to each control point includes at least one candidate motion information unit;

determining a combined motion information unit set e including W motion information units, where each motion information unit in the combined motion information unit set e is selected from at least some motion information units in the candidate motion information unit set corresponding to each of the W control points, and each combined motion information unit in the combined motion information unit set e includes a motion vector; and

performing prediction or motion estimation on W motion vectors in the combined motion information unit set e to obtain the motion vectors of the W control points, or using W motion vectors included in the combined motion information unit set e as the motion vectors of the W control points.

[0096] In some possible implementations of the present invention, precision of each motion vector in the candidate motion information unit set may be $1/n$ of the pixel precision.

[0097] The picture prediction method may be applied to a video encoding process, or the picture prediction method may be applied to a video decoding process.

[0098] To better understand the technical solutions of the embodiments of the present

invention, the following provides descriptions from perspectives of an encoder side and a decoder side by using examples.

[0099] The following first describes a solution from a perspective of an encoder side.

[0100] Referring to FIG. 3-a, FIG. 3-a is a schematic flowchart of another picture prediction method according to another embodiment of the present invention. In an example shown in FIG. 3-a, the another picture prediction method provided in the another embodiment of the present invention may include the following steps.

[0101] 301. A video encoding apparatus determines W control points in a current picture block.

[0102] 302. The video encoding apparatus determines motion vectors of the W control points.

[0103] The determining, by the video encoding apparatus, motion vectors of the W control points may include: determining a candidate motion information unit set corresponding to each of the W control points, where the candidate motion information unit corresponding to each control point set includes at least one candidate motion information unit; determining a combined motion information unit set e including W motion information units, where each motion information unit in the combined motion information unit set e is selected from at least some motion information units in the candidate motion information unit set corresponding to each of the W control points, and each combined motion information unit in the combined motion information unit set e includes a motion vector; and performing prediction or motion estimation on W motion vectors in the combined motion information unit set e to obtain the motion vectors of the W control points, or using W motion vectors included in the combined motion information unit set e as the motion vectors of the W control points.

[0104] In some possible implementations of the present invention, precision of each motion vector in the candidate motion information unit set may be $1/n$ of pixel precision.

[0105] 303. The video encoding apparatus obtains, by means of calculation, motion vectors of P pixel units of the current picture block by using a motion model and the motion vectors of the W control points.

[0106] P is a positive integer, and W and N are integers greater than 1.

[0107] The P pixel units are some or all of pixel units of the current picture block.

[0108] A motion vector of each of the P pixel units is used to determine a corresponding reference pixel unit, in a reference picture, of a corresponding pixel unit. Therefore, the motion vector of each of the P pixel units may be used to determine the corresponding reference pixel unit, in the reference picture, of the corresponding pixel unit.

[0109] Precision of the determined motion vectors of the W control points is $1/n$ of the pixel precision.

[0110] Precision of the motion vector that is obtained by means of calculation and that is of each of the P pixel units is $1/N$ of the pixel precision.

[0111] W, n, and N are integers greater than 1.

[0112] N is greater than n. P is a positive integer.

[0113] Because N is greater than n, the precision of the motion vector that is obtained by means of calculation and that is of each of the P pixel units is higher than the precision of the determined motion vectors of the W control points. That is, the higher-precision motion vectors of the P pixel units of the current picture block are obtained.

[0114] In some possible implementations of the present invention, the motion vectors of the W control points are predicted based on a motion vector, whose precision is $1/n$ of the pixel precision, of an encoded picture block or a decoded picture block that surrounds the current picture block.

[0115] In some possible implementations of the present invention, when the motion vectors of the W control points are based on predicted values, precision of the predicted values is also $1/n$ of the pixel precision, and differences between the motion vectors, whose precision is $1/n$ of the pixel precision, of the control points and the corresponding predicted values may be written into a bitstream. Correspondingly, the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points may be obtained, by means of prediction, on a decoder side by using the predicted difference between the W motion vectors and the predicted motion vector and by using the motion vector, whose precision is $1/n$ of the pixel precision, of the decoded picture block that surrounds the current picture block, where the predicted difference is obtained from the bitstream.

[0116] In some possible implementations of the present invention, the motion model may

v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of a pixel unit with coordinates of (x, y) in the current picture block, v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, (v_{0x}, v_{0y}) , (v_{1x}, v_{1y}) , and (v_{2x}, v_{2y}) represent motion vectors, whose precision is 1/n of the pixel precision, of three control points, w represents a width of the current picture block, and h represents a height of the current picture block.

[0119] The foregoing examples are described by using the affine motion model. When a translational motion model, a rotational motion model, a shearing motion model, a zooming motion model, a perspective motion model, a parabolic motion model, a bilinear motion model, or the like is used, reference may be made to the foregoing examples. Details are not described herein again.

[0120] 304. The video encoding apparatus determines a corresponding reference pixel unit, in a reference picture, of each of the P pixel units by using a motion vector of each of the P pixel units.

[0121] Optionally, in some possible implementations of the present invention, the determining a corresponding reference pixel unit, in a reference picture, of each of the P pixel units by using a motion vector of each of the P pixel units may include: obtaining, by means of calculation, an integer-pixel location of each of the P pixel units by using the motion vector of each of the P pixel units; and searching, by using the motion vector of each of the P pixel units, the reference picture for a reference pixel unit corresponding to the integer-pixel location of each of the P pixel units, where the reference pixel unit that is corresponding to the integer-pixel location of each of the P pixel units and that is found in the reference picture is the reference pixel unit, in the reference picture, of each of the P pixel units.

[0122] Specifically, for example, an integer-pixel location of a pixel unit i may be obtained by means of calculation by using a motion vector of the pixel unit i of the P pixel units, and the reference picture may be searched, by using the motion vector of the pixel unit i, for a reference pixel unit corresponding to the integer-pixel location of the pixel unit i, where the reference pixel unit that is corresponding to the integer-pixel location of the pixel unit i

the phase of the pixel unit i, where a filter coefficient used by the interpolation filter is corresponding to the phase; and interpolation filtering may be performed on a pixel of the corresponding reference pixel unit, in the reference picture, of the pixel unit i by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit i.

5 The pixel unit i may be any one of the P pixel units. For example, interpolation filtering may be performed on each of the P pixel units according to a manner similar to that of performing interpolation filtering on the pixel unit i.

[0136] In some possible implementations of the present invention, the obtaining, by means of calculation, a phase of each of the P pixel units by using the motion vector of each
10 of the P pixel units includes: obtaining, by means of calculation, the phase of each of the P pixel units according to the following formula by using the motion vector of each of the P pixel units:

$$\begin{aligned} X' &= \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1); \\ Y' &= \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1); \end{aligned} \quad \text{where}$$

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal
15 phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the
20 pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0137] Optionally, in some possible implementations of the present invention, the phase includes a horizontal phase and a vertical phase; and the determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit includes: determining, based on the horizontal phase of each pixel unit, a horizontal
25 interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter

coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

[0138] Specifically, for example, a horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on a horizontal phase of the pixel unit i, and a vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on a vertical phase of the pixel unit i, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase of the pixel unit i, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase of the pixel unit i. The pixel unit i may be any one of the P pixel units. For example, the interpolation filter with the phase of Q that is corresponding to each of the P pixel units may be determined according to a manner similar to that of determining the interpolation filter with the phase of Q that is corresponding to the pixel unit i.

[0139] Optionally, in some possible implementations of the present invention, the performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit includes:

performing horizontal interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of the pixel unit i by using the determined horizontal interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using the determined vertical interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, where the pixel unit j is any one of the P pixel units.

[0140] In addition, if a pixel unit has only one phase (that is, has only a horizontal phase or a vertical phase), interpolation filtering needs to be performed on a pixel of a reference pixel unit of the pixel unit only once, to obtain a predicted pixel value of the pixel unit.

[0141] 306. The video encoding apparatus may obtain a predicted residual of the current picture block by using original pixel values of the P pixel units and predicted pixel values of the P pixel units of the current picture block, and the video encoding apparatus may write the predicted residual of the current picture block into a video bitstream.

[0142] It can be learnt from the foregoing that, in the solution of this embodiment, the video encoding apparatus obtains, by means of calculation, the motion vector, whose precision is $1/N$ of the pixel precision, of each pixel unit of the current picture block by using the motion model and the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points, where N is greater than n. In other words, the precision of the motion vector that is obtained by means of calculation and that is of each pixel unit of the current picture block is higher than the precision of the determined motion vectors of the W control points.

The higher-precision motion vector is obtained first. Therefore, the higher-precision motion vector of each pixel unit of the current picture block is used to determine the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block, and interpolation filtering is performed on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block by using the interpolation filter with the phase of Q (Q is greater than n), to obtain the predicted pixel value of each pixel unit of the current picture block. It can be learnt that, the foregoing manner helps reduce a quantity of times of interpolation filtering required for obtaining, by means of prediction, the higher-precision predicted pixel value of the current picture block (for example, an intermediate process for obtaining the lower-precision predicted pixel value by performing lower-precision interpolation filtering may not be required), so as to reduce a quantity of intermediate caches and memory operations that are required for interpolation filtering during a picture prediction process, and reduce calculation complexity during the picture prediction process.

[0143] The following describes a solution from a perspective of a decoder side.

[0144] Referring to FIG. 4, FIG. 4 is a schematic flowchart of another picture prediction

method according to another embodiment of the present invention. In an example shown in FIG. 4, the another picture prediction method provided in the another embodiment of the present invention may include the following steps.

[0145] 401. A video decoding apparatus determines W control points in a current picture block.

[0146] 402. The video decoding apparatus determines motion vectors of the W control points.

[0147] The determining, by the video decoding apparatus, motion vectors of the W control points may include: determining a candidate motion information unit set corresponding to each of the W control points, where the candidate motion information unit set corresponding to each control point includes at least one candidate motion information unit; determining a combined motion information unit set e including W motion information units, where each motion information unit in the combined motion information unit set e is selected from at least some motion information units in the candidate motion information unit set corresponding to each of the W control points, and each combined motion information unit in the combined motion information unit set e includes a motion vector; and performing prediction or motion estimation on W motion vectors in the combined motion information unit set e to obtain the motion vectors of the W control points, or using W motion vectors included in the combined motion information unit set e as the motion vectors of the W control points.

[0148] In some possible implementations of the present invention, precision of each motion vector in the candidate motion information unit set may be $1/n$ of pixel precision.

[0149] 403. The video decoding apparatus obtains, by means of calculation, motion vectors of P pixel units of the current picture block by using a motion model and the motion vectors of the W control points.

[0150] P is a positive integer, and W and N are integers greater than 1.

[0151] The P pixel units are some or all of pixel units of the current picture block.

[0152] A motion vector of each of the P pixel units is used to determine a corresponding reference pixel unit, in a reference picture, of a corresponding pixel unit. Therefore, the motion vector of each of the P pixel units may be used to determine the corresponding

translational motion model, a rotational motion model, a shearing motion model, a zooming motion model, a perspective motion model, a parabolic motion model, a bilinear motion model, or the like is used, reference may be made to the foregoing examples. Details are not described herein again.

5 **[0164]** 404. The video decoding apparatus determines a corresponding reference pixel unit, in a reference picture, of each of the P pixel units by using a motion vector of each of the P pixel units.

10 **[0165]** Optionally, in some possible implementations of the present invention, the determining the corresponding reference pixel unit, in the reference picture, of each of the P pixel units by using the motion vector of each of the P pixel units may include: obtaining, by means of calculation, an integer-pixel location of each of the P pixel units by using the motion vector of each of the P pixel units; and searching, by using the motion vector of each of the P pixel units, the reference picture for a reference pixel unit corresponding to the integer-pixel location of each of the P pixel units, where the reference pixel unit that is corresponding to the
15 integer-pixel location of each of the P pixel units and that is found in the reference picture is the reference pixel unit, in the reference picture, of each of the P pixel units.

20 **[0166]** Specifically, for example, an integer-pixel location of a pixel unit i may be obtained by means of calculation by using a motion vector of the pixel unit i of the P pixel units, and the reference picture may be searched, by using the motion vector of the pixel unit i, for a reference pixel unit corresponding to the integer-pixel location of the pixel unit i, where the reference pixel unit that is corresponding to the integer-pixel location of the pixel unit i and that is found in the reference picture is a reference pixel unit, in the reference picture, of the pixel unit i. The pixel unit i may be any one of the P pixel units. For example, the corresponding reference pixel unit, in the reference picture, of each of the P pixel units may
25 be determined according to a manner similar to that of determining the reference pixel unit, in the reference picture, of the pixel unit i.

[0167] In some possible implementations of the present invention, the obtaining, by means of calculation, an integer-pixel location of each of the P pixel units by using the motion vector of each of the P pixel units includes:

30 obtaining, by means of calculation, the integer-pixel location of each of the P pixel

pixel units:

$$X' = \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1). \\ Y' = \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1). \quad \text{where}$$

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

10 **[0180]** Optionally, in some possible implementations of the present invention, the phase includes a horizontal phase and a vertical phase; and the determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit includes: determining, based on the horizontal phase of each pixel unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and
15 determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

[0181] Specifically, for example, a horizontal interpolation filter with a phase of Q that is
20 corresponding to the pixel unit i may be determined based on a horizontal phase of the pixel unit i, and a vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i may be determined based on a vertical phase of the pixel unit i, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase of the pixel unit i, and a filter coefficient used by the vertical interpolation filter is corresponding to the
25 vertical phase of the pixel unit i. The pixel unit i may be any one of the P pixel units. For example, the interpolation filter with the phase of Q that is corresponding to each of the P pixel units may be determined according to a manner similar to that of determining the

interpolation filter with the phase of Q that is corresponding to the pixel unit i.

[0182] Optionally, in some possible implementations of the present invention, the performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit includes:

performing horizontal interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of the pixel unit i by using the determined horizontal interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using the determined vertical interpolation filter with the phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, where the pixel unit j is any one of the P pixel units.

[0183] In addition, if a pixel unit has only one phase (that is, has only a horizontal phase or a vertical phase), interpolation filtering needs to be performed on a pixel of a reference pixel unit of the pixel unit only once, to obtain a predicted pixel value of the pixel unit.

[0184] 406. The video decoding apparatus reconstructs the current picture block by using a predicted pixel value of the current picture block and a predicted residual, in a video bitstream, of the current picture block.

[0185] It can be learnt from the foregoing that, in the solution of this embodiment, the video decoding apparatus obtains, by means of calculation, the motion vector, whose precision is $1/N$ of the pixel precision, of each pixel unit of the current picture block by using the motion model and the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points, where N is greater than n. In other words, the precision of the motion vector

that is obtained by means of calculation and that is of each pixel unit of the current picture block is higher than the precision of the determined motion vectors of the W control points. The higher-precision motion vector is obtained first. Therefore, the higher-precision motion vector of each pixel unit of the current picture block is used to determine the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block, and interpolation filtering is performed on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block by using the interpolation filter with the phase of Q (Q is greater than n), to obtain the predicted pixel value of each pixel unit of the current picture block. It can be learnt that, the foregoing manner helps reduce a quantity of times of interpolation filtering required for obtaining, by means of prediction, the higher-precision predicted pixel value of the current picture block (for example, an intermediate process for obtaining the lower-precision predicted pixel value by performing lower-precision interpolation filtering may not be required), so as to reduce a quantity of intermediate caches and memory operations that are required for interpolation filtering during a picture prediction process, and reduce calculation complexity during the picture prediction process.

[0186] The following describes, by using examples, some possible specific implementations of determining a combined motion information unit set e including W motion information units in the foregoing embodiments.

[0187] The determining a combined motion information unit set e including W motion information units may include: determining, from A candidate combined motion information unit sets, the combined motion information unit set e including the W motion information units, where each motion information unit included in each of the A candidate combined motion information unit sets is selected from at least some motion information units, complying with a constraint condition, in a candidate motion information unit set corresponding to each of the W control points, A is a positive integer, the A candidate combined motion information unit sets are different from each other, and each of the A candidate combined motion information unit sets includes W motion information units.

[0188] Optionally, in some possible implementations of the present invention, the A candidate combined motion information unit sets satisfy at least one of a first condition, a

second condition, a third condition, a fourth condition, or a fifth condition.

[0189] The first condition includes: A motion manner, indicated by a motion information unit in any one of the A candidate combined motion information unit sets, of the current picture block is a non-translational motion.

5 **[0190]** The second condition includes: Prediction directions corresponding to two motion information units in any one of the A candidate combined motion information unit sets are the same.

[0191] The third condition includes: Reference-frame indexes corresponding to two motion information units in any one of the A candidate combined motion information unit sets
10 are the same.

[0192] The fourth condition includes: An absolute value of a difference between horizontal components of motion vectors of two motion information units in any one of the A candidate combined motion information unit sets is less than or equal to a horizontal component threshold, or an absolute value of a difference between horizontal components of a
15 motion vector of a control point Z and a motion vector of one motion information unit in any one of the A candidate combined motion information unit sets is less than or equal to a horizontal component threshold, where the control point Z of the current picture block is different from any one of the W control points.

[0193] The fifth condition includes: An absolute value of a difference between vertical
20 components of motion vectors of two motion information units in any one of the A candidate combined motion information unit sets is less than or equal to a vertical component threshold, or an absolute value of a difference between vertical components of a motion vector of a control point Z and a motion vector of any motion information unit in one of the A candidate combined motion information unit sets is less than or equal to a vertical component threshold,
25 where the control point Z of the current picture block is different from any one of the W control points.

[0194] Optionally, in some possible implementations of the present invention, the determining, from A candidate combined motion information unit sets by the video encoding/decoding apparatus, the combined motion information unit set e including the W
30 motion information units may include: determining, from the A candidate combined motion

(for example, the combined motion information unit set e) may be specifically, for example, a sum of squared differences (SSD), a sum of absolute differences (SAD), a sum of differences, or another distortion parametric value that can measure distortion, between the original pixel value of the picture block (for example, the current picture block) and the predicted pixel value that is of the picture block and that is obtained by means of pixel value prediction performed on the picture block by using the candidate combined motion information unit set (for example, the combined motion information unit set e).

[0200] The following describes, by using examples, some possible manners of determining W control points in a current picture block.

[0201] Optionally, in some possible implementations of the present invention, the W control points include W control points of an upper left control point, an upper right control point, a lower left control point, and a central control point a1 in the current picture block.

[0202] The upper left control point in the current picture block is an upper left vertex of the current picture block or a pixel block, in the current picture block, including an upper left vertex of the current picture block. The lower left control point in the current picture block is a lower left vertex of the current picture block or a pixel block, in the current picture block, including a lower left vertex of the current picture block. The upper right control point in the current picture block is an upper right vertex of the current picture block or a pixel block, in the current picture block, including an upper right vertex of the current picture block. The central control point a1 of the current picture block is a central pixel of the current picture block or a pixel block, in the current picture block, including a central pixel of the current picture block.

[0203] Optionally, in some possible implementations of the present invention, a candidate motion information unit set corresponding to the upper left control point in the current picture block includes a motion information unit of x1 pixel units. The x1 pixel units include at least one pixel unit (for example, picture blocks A, B, and C in FIG. 2-b) spatially adjacent to the upper left control point in the current picture block and/or at least one pixel unit temporally adjacent to the upper left control point in the current picture block, where x1 is a positive integer.

[0204] For example, the x1 pixel units may include at least one of the following: a pixel

unit, in a video frame temporally adjacent to a video frame to which the current picture block belongs, whose location is the same as that of the upper left control point in the current picture block; a pixel unit that is spatially adjacent to and that is on the left of the current picture block; a pixel unit that is spatially adjacent to and that is on the upper left of the current picture block; or a pixel unit that is spatially adjacent to and that is on an upper side of the current picture block.

[0205] Optionally, in some possible implementations of the present invention, a candidate motion information unit set corresponding to the upper right control point in the current picture block includes a motion information unit of x_2 pixel units. The x_2 pixel units include at least one pixel unit (for example, picture blocks E and D in FIG. 2-b) spatially adjacent to the upper right control point in the current picture block and/or at least one pixel unit temporally adjacent to the upper right control point in the current picture block, where x_2 is a positive integer.

[0206] For example, the x_2 pixel units may include at least one of the following: a pixel unit, in a video frame temporally adjacent to a video frame to which the current picture block belongs, whose location is the same as that of the upper right control point in the current picture block; a pixel unit that is spatially adjacent to and that is on the right of the current picture block; a pixel unit that is spatially adjacent to and that is on the upper right of the current picture block; or a pixel unit that is spatially adjacent to and that is on an upper side of the current picture block.

[0207] Optionally, in some possible implementations of the present invention, a candidate motion information unit set corresponding to the lower left control point in the current picture block includes a motion information unit of x_3 pixel units. The x_3 pixel units include at least one pixel unit spatially adjacent to the lower left control point in the current picture block and/or at least one pixel unit temporally adjacent to the lower left control point in the current picture block, where x_3 is a positive integer.

[0208] For example, the x_3 pixel units may include at least one of the following: a pixel unit, in a video frame temporally adjacent to a video frame to which the current picture block belongs, whose location is the same as that of the lower left control point in the current picture block; a pixel unit that is spatially adjacent to and that is on the left of the current

Phase	Filter coefficient
54	{-1, 1, -3, 11, 61, -7, 2, 0}
55	{-1, 1, -3, 10, 62, -6, 1, 0}
56	{0, 1, -3, 8, 62, -6, 2, 0}
57	{0, 1, -2, 7, 62, -5, 1, 0}
58	{-1, 1, -2, 6, 63, -4, 1, 0}
59	{0, 1, -2, 5, 63, -4, 1, 0}
60	{0, 0, -1, 4, 63, -3, 1, 0}
61	{0, 0, -2, 3, 64, -2, 1, 0}
62	{0, 0, -1, 3, 64, -2, 0, 0}
63	{0, 0, 0, 1, 64, -1, 0, 0}

[0214] For example, when a horizontal phase of a pixel unit is $x = 12$, if a 64-phase interpolation filter with a gain scale factor of 256 is used, a horizontal interpolation filtering result S may be represented as follows:

$$S = ((-2) \times x_0 + 10 \times x_1 + (-31) \times x_2 + 239 \times x_2 + 52 \times x_4 + (-17) \times x_5 + 5 \times x_6 + 0 \times x_7) + offset / 256$$

[0215] A predicted pixel value of the pixel unit may be obtained by using the horizontal interpolation filtering result S , where *offset* may be any integer.

[0216] The following uses some test data to present technical effects of the solutions of the embodiments of the present invention.

10 [0217] Table 3 and Table 4 provide contrast of some test performance between a conventional solution and the solutions of the embodiments of the present invention. Table 3 lists test performance of the conventional solution, and Table 4 lists test performance of solutions of some embodiments of the present invention.

Table 3

Class	Sequence	Y BD-Rate	U BD-Rate	V BD-Rate	EncTime	DecTime
Class Affine	Dolphin	-1.4%	0.0%	-0.2%	209.2%	209.6%
	City	-4.2%	-3.8%	-4.5%	218.5%	169.5%

of performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the corresponding pixel unit, the interpolation filtering unit is specifically configured to:

5 perform horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and perform vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with
10 a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

 perform vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a
15 vertical interpolation filtering result; and perform horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, where the pixel unit j is any one of the P pixel units.

[0234] Optionally, in some possible implementations of the present invention, the motion
20 model is a translational motion model, an affine motion model, a rotational motion model, a zooming motion model, a parabolic motion model, a shearing motion model, a perspective motion model, or a bilinear motion model.

[0235] Optionally, in some possible implementations of the present invention,
the motion model is represented as follows when W is equal to 2:

$$25 \quad \begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{L} x - \frac{(v_{1y} - v_{0y}) \times N}{L} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{L} x + \frac{(v_{1x} - v_{0x}) \times N}{L} y + Nv_{0y} \right) / n \end{cases}; \text{ or}$$

[0237] Optionally, in some possible implementations of the present invention, the motion vectors of the W control points are predicted based on a motion vector, whose precision is $1/n$ of the pixel precision, of an encoded picture block or a decoded picture block that surrounds the current picture block.

5 [0238] Optionally, in some possible implementations of the present invention, the picture prediction apparatus 500 is applied to a video encoding apparatus, or the picture prediction apparatus 500 is applied to a video decoding apparatus.

[0239] It can be understood that, functions of the functional modules of the picture prediction apparatus 500 in this embodiment may be specifically implemented according to
10 the methods in the foregoing method embodiments. For a specific implementation process of the picture prediction apparatus 500, reference may be made to related descriptions in the foregoing method embodiments. Details are not described herein again. The picture prediction apparatus 500 may be any apparatus that needs to output or play a video, for example, a notebook computer, a tablet computer, a personal computer, a mobile phone, or another
15 device.

[0240] It can be learnt that, in this embodiment, the picture prediction apparatus 500 obtains, by means of calculation, the motion vector, whose precision is $1/N$ of the pixel precision, of each pixel unit of the current picture block by using the motion model and the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points, where N
20 is greater than n. In other words, the precision of the motion vector that is obtained by means of calculation and that is of each pixel unit of the current picture block is higher than the precision of the determined motion vectors of the W control points. The higher-precision motion vector is obtained first. Therefore, the higher-precision motion vector of each pixel unit of the current picture block is used to determine the corresponding reference pixel unit, in
25 the reference picture, of each pixel unit of the current picture block, and interpolation filtering is performed on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block by using the interpolation filter with the phase of Q (Q is greater than n), to obtain the predicted pixel value of each pixel unit of the current picture block. It can be learnt that, the foregoing manner helps reduce a quantity of times of
30 interpolation filtering required for obtaining, by

means of prediction, the higher-precision predicted pixel value of the current picture block (for example, an intermediate process for obtaining a lower-precision predicted pixel value by performing lower-precision interpolation filtering may not be required), so as to reduce a quantity of intermediate caches and memory operations that are required for interpolation filtering during a picture prediction process, and reduce calculation complexity during the picture prediction process.

[0241] Referring to FIG. 6, FIG. 6 is a structural block diagram of a picture prediction apparatus 600 according to another embodiment of the present invention. The picture prediction apparatus 600 may include at least one processor 601, a memory 605, and at least one communications bus 602. The communications bus 602 is configured to implement a connection and communication between the components.

[0242] The picture prediction apparatus 600 may optionally include at least one network interface 604 and/or a user interface 603. The user interface 603 may include a display (for example, holographic imaging device, a cathode-ray tube, or a projector), a pointing device (for example, a mouse, a trackball, a touch panel, or a touchscreen), a camera and/or a sound pickup apparatus, or the like.

[0243] The memory 605 may include a read-only memory and a random access memory, and provide an instruction and data for the processor 601. A part of the memory 605 may further include a non-volatile random access memory.

[0244] In some implementations, the memory 605 stores the following elements: an executable module or a data structure, their subsets, or their extended sets; and the memory 605 includes:

an operating system 6051, including various system programs, and configured to implement various basic services and handle hardware-based tasks; and

an application program module 6052, including various application programs, and configured to implement various application services.

[0245] In this embodiment of the present invention, by invoking the program or the instruction that is stored in the memory 605, the processor 601 is configured to: determine motion vectors of W control points in a current picture block; obtain, by means of calculation, motion vectors of P pixel units of the current picture block by using a motion model and the

$$X' = \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1);$$

$$Y' = \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1); \text{ where}$$

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block,

5 v_{Nx} represents a horizontal component of a motion vector, whose precision is $1/N$ of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is $1/N$ of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

[0252] Optionally, in some possible implementations of the present invention, the phase includes a horizontal phase and a vertical phase; and the determining, by the processor and based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit includes: determining, based on the horizontal phase of each pixel unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, where a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

[0253] Optionally, in some possible implementations of the present invention, the performing, by the processor, interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit includes:

performing horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with

a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, where the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, where the pixel unit j is any one of the P pixel units.

10 **[0254]** Optionally, in some possible implementations of the present invention, the motion model is a translational motion model, an affine motion model, a rotational motion model, a parabolic motion model, a shearing motion model, a zooming motion model, a perspective motion model, or a bilinear motion model.

[0255] Optionally, in some possible implementations of the present invention,
15 the motion model is represented as follows when W is equal to 2:

$$\begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{L} x - \frac{(v_{1y} - v_{0y}) \times N}{L} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{L} x + \frac{(v_{1x} - v_{0x}) \times N}{L} y + Nv_{0y} \right) / n \end{cases}; \text{ or}$$

$$\begin{cases} v_{Nx} = \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} x - \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} y + (N/n)v_{0x} \\ v_{Ny} = \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} x + \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} y + (N/n)v_{0y} \end{cases}, \text{ where}$$

L represents a width or a height of the current picture block, (v_{0x}, v_{0y}) and (v_{1x}, v_{1y}) represent motion vectors, whose precision is 1/n of the pixel precision, of two control points, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of a pixel unit with coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

the picture prediction apparatus 600, reference may be made to related descriptions in the foregoing method embodiments. Details are not described herein again. The picture prediction apparatus 600 may be any apparatus that needs to output or play a video, for example, a notebook computer, a tablet computer, a personal computer, a mobile phone, or another device.

[0261] It can be learnt that, in the picture prediction method provided in this embodiment, the picture prediction apparatus 600 obtains, by means of calculation, the motion vector, whose precision is $1/N$ of the pixel precision, of each pixel unit of the current picture block by using the motion model and the motion vectors, whose precision is $1/n$ of the pixel precision, of the W control points, where N is greater than n . In other words, the precision of the motion vector that is obtained by means of calculation and that is of each pixel unit of the current picture block is higher than the precision of the determined motion vectors of the W control points. The higher-precision motion vector is obtained first. Therefore, the higher-precision motion vector of each pixel unit of the current picture block is used to determine the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block, and interpolation filtering is performed on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit of the current picture block by using the interpolation filter with the phase of Q (Q is greater than n), to obtain the predicted pixel value of each pixel unit of the current picture block. It can be learnt that, the foregoing manner helps reduce a quantity of times of interpolation filtering required for obtaining, by means of prediction, the higher-precision predicted pixel value of the current picture block (for example, an intermediate process for obtaining a lower-precision predicted pixel value by performing lower-precision interpolation filtering may not be required), so as to reduce a quantity of intermediate caches and memory operations that are required for interpolation filtering during a picture prediction process, and reduce calculation complexity during the picture prediction process.

[0262] In the foregoing embodiments, the description of each embodiment has respective focuses. For a part that is not described in detail in an embodiment, reference may be made to related descriptions in other embodiments.

[0263] In the several embodiments provided in this application, it should be understood

CLAIMS

1. A picture prediction method, comprising:

determining motion vectors of W control points in a current picture block;

obtaining, by means of calculation, motion vectors of P pixel units of the current picture

5 block by using a motion model and the motion vectors of the W control points, wherein precision of the determined motion vectors of the W control points is $1/n$ of pixel precision, precision of the motion vector that is obtained by means of calculation and that is of each of the P pixel units is $1/N$ of the pixel precision, the P pixel units are some or all of pixel units of the current picture block, the motion vector of each of the P pixel units is used to determine a
10 corresponding reference pixel unit, in a reference picture, of a corresponding pixel unit, W , n , and N are integers greater than 1, N is greater than n , and P is a positive integer; and

performing interpolation filtering on a pixel of the corresponding reference pixel unit, in the reference picture, of each of the P pixel units by using an interpolation filter with a phase of Q , to obtain a predicted pixel value of each of the P pixel units, wherein Q is an integer
15 greater than n .

2. The method according to claim 1, wherein a value of N is a preset fixed value, and Q is less than or equal to N .

3. The method according to claim 1 or 2, wherein a horizontal component or a vertical component of one of the motion vectors of the W control points is amplified N times in the
20 motion model by using the N , or a component difference between motion vectors of any two of the W control points is amplified N times in the motion model by using N .

4. The method according to any one of claims 1 to 3, wherein the performing interpolation filtering on a pixel of the corresponding reference pixel unit, in the reference picture, of each of the P pixel units by using an interpolation filter with a phase of Q
25 comprises:

obtaining, by means of calculation, a phase of each of the P pixel units by using the motion vector of each of the P pixel units; determining, based on the phase of each pixel unit,

the interpolation filter with the phase of Q that is corresponding to the pixel unit, wherein a filter coefficient used by the interpolation filter is corresponding to the phase; and performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the corresponding pixel unit.

5. The method according to claim 4, wherein the obtaining, by means of calculation, a phase of each of the P pixel units by using the motion vector of each of the P pixel units comprises: obtaining, by means of calculation, the phase of each of the P pixel units according to the following formula by using the motion vector of each of the P pixel units:

$$\begin{aligned} X' &= \text{abs}(v_{Nx}) \% N, \text{ or } X' = v_{Nx} \& ((1 \leq M) - 1); \\ Y' &= \text{abs}(v_{Ny}) \% N, \text{ or } Y' = v_{Ny} \& ((1 \leq M) - 1); \end{aligned} \quad \text{wherein}$$

M is equal to $\log_2 N$ when N is an integral power of 2, X' represents a horizontal phase of a pixel unit with coordinates of (x, y) in the current picture block, Y' represents a vertical phase of the pixel unit with the coordinates of (x, y) in the current picture block, v_{Nx} represents a horizontal component of a motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block, and v_{Ny} represents a vertical component of the motion vector, whose precision is 1/N of the pixel precision, of the pixel unit with the coordinates of (x, y) in the current picture block.

6. The method according to claim 4 or 5, wherein the phase comprises a horizontal phase and a vertical phase; and the determining, based on the phase of each pixel unit, the interpolation filter with the phase of Q that is corresponding to the pixel unit comprises: determining, based on the horizontal phase of each pixel unit, a horizontal interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit; and determining, based on the vertical phase of each pixel unit, a vertical interpolation filter with a phase of Q that is corresponding to the corresponding pixel unit, wherein a filter coefficient used by the horizontal interpolation filter is corresponding to the horizontal phase, and a filter coefficient used by the vertical interpolation filter is corresponding to the vertical phase.

7. The method according to claim 6, wherein the performing interpolation filtering on the pixel of the corresponding reference pixel unit, in the reference picture, of each pixel unit by

using the determined interpolation filter with the phase of Q that is corresponding to the pixel unit comprises:

performing horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and performing vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, wherein the pixel unit i is any one of the P pixel units; or

performing vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and performing horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, wherein the pixel unit j is any one of the P pixel units.

8. The method according to any one of claims 1 to 7, wherein the motion model is a translational motion model, an affine motion model, a rotational motion model, a parabolic motion model, a shearing motion model, a zooming motion model, a perspective motion model, or a bilinear motion model.

9. The method according to any one of claims 1 to 8, wherein the motion model is represented as follows when W is equal to 2:

$$\begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{L} x - \frac{(v_{1y} - v_{0y}) \times N}{L} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{L} x + \frac{(v_{1x} - v_{0x}) \times N}{L} y + Nv_{0y} \right) / n \end{cases} ; \text{ or}$$

$$\begin{cases} v_{Nx} = \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} x - \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} y + (N/n)v_{0x} \\ v_{Ny} = \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} x + \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} y + (N/n)v_{0y} \end{cases}, \text{ wherein}$$

L represents a width or a height of the current picture block, (v_{0x}, v_{0y}) and (v_{1x}, v_{1y})

reference picture, of each pixel unit by using the determined interpolation filter with the phase of Q that is corresponding to the corresponding pixel unit, the interpolation filtering unit is specifically configured to:

perform horizontal interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit i by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a horizontal interpolation filtering result; and perform vertical interpolation filtering on the horizontal interpolation filtering result by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit i, to obtain a predicted pixel value of the pixel unit i, wherein the pixel unit i is any one of the P pixel units; or

perform vertical interpolation filtering on a pixel of a corresponding reference pixel unit, in the reference picture, of a pixel unit j by using a determined vertical interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a vertical interpolation filtering result; and perform horizontal interpolation filtering on the vertical interpolation filtering result by using a determined horizontal interpolation filter with a phase of Q that is corresponding to the pixel unit j, to obtain a predicted pixel value of the pixel unit j, wherein the pixel unit j is any one of the P pixel units.

20. The picture prediction apparatus according to any one of claims 13 to 19, wherein the motion model is a translational motion model, an affine motion model, a rotational motion model, a zooming motion model, a parabolic motion model, a shearing motion model, a perspective motion model, or a bilinear motion model.

21. The picture prediction apparatus according to any one of claims 13 to 20, wherein the motion model is represented as follows when W is equal to 2:

$$\begin{cases} v_{Nx} = \left(\frac{(v_{1x} - v_{0x}) \times N}{L} x - \frac{(v_{1y} - v_{0y}) \times N}{L} y + Nv_{0x} \right) / n \\ v_{Ny} = \left(\frac{(v_{1y} - v_{0y}) \times N}{L} x + \frac{(v_{1x} - v_{0x}) \times N}{L} y + Nv_{0y} \right) / n \end{cases}; \text{ or}$$

$$\begin{cases} v_{Nx} = \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} x - \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} y + (N/n)v_{0x} \\ v_{Ny} = \frac{(v_{1y} - v_{0y}) \times (N/n)}{L} x + \frac{(v_{1x} - v_{0x}) \times (N/n)}{L} y + (N/n)v_{0y} \end{cases}, \text{ wherein}$$

picture prediction apparatus is applied to a video encoding apparatus, or the picture prediction apparatus is applied to a video decoding apparatus.

25. A picture prediction apparatus, comprising a memory and a processor and a program stored in the memory, when the program is running by the processor for implementing the
- 5 method of any of claim 1 to 12.