

DESCRIPTION**Interpolation filter for an inter prediction apparatus and method for video coding**5 TECHNICAL FIELD

Generally, the present invention relates to the field of video coding. More specifically, the present invention relates to an interpolation filter for an inter prediction apparatus and method for video coding as well as an encoding apparatus and a decoding apparatus
10 comprising such an inter prediction apparatus.

BACKGROUND

Digital video communication and storage applications are implemented by a wide range of
15 digital devices, e.g. digital cameras, cellular radio telephones, laptops, broadcasting systems, video teleconferencing systems, etc. One of the most important and challenging tasks of these applications is video compression. The task of video compression is complex and is constrained by two contradicting parameters: compression efficiency and computational complexity. Video coding standards, such as ITU-T H.264/AVC or ITU-T
20 H.265/HEVC, provide a good tradeoff between these parameters. For that reason support of video coding standards is a mandatory requirement for almost any video compression application.

The state-of-the-art video coding standards are based on partitioning of a source frame or
25 picture into frame or picture blocks. Processing of these blocks depend on their size, spatial position and a coding mode specified by an encoder. Coding modes can be classified into two groups according to the type of prediction: intra- and inter-prediction modes. Intra-prediction modes use pixels of the same frame (also referred to as picture or image) to generate reference samples to calculate the prediction values for the pixels of
30 the block being reconstructed. Intra-prediction is also referred to as spatial prediction. Inter-prediction modes are designed for temporal prediction and uses reference samples of previous or next frames to predict pixels of the block of the current frame. After a prediction stage, transform coding is performed for a prediction error that is the difference between an original signal and its prediction. Then, the transform coefficients and side
35 information are encoded using an entropy coder (e.g., CABAC for AVC/H.264 and HEVC/H.265). The recently adopted ITU-T H.265/HEVC standard (ISO/IEC 23008-

2:2013, "Information technology - High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding", November 2013) declares a set of state-of-the-art video coding tools that provide a reasonable tradeoff between coding efficiency and computational complexity. An overview on the ITU-T

5 H.265/HEVC standard has been given by Gary J. Sullivan, "Overview of the High Efficiency Video Coding (HEVC) Standard", in IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, No. 12, December 2012, the entire content of which is incorporated herein by reference.

10 Similarly to the ITU-T H.264/AVC video coding standard, the HEVC/H.265 video coding standard provides for a division of the source frame into frame blocks in the form of so-called coding units (CUs). Each of the CUs can be further split into either smaller CUs or prediction units (PUs). A PU can be intra- or inter-predicted according to the type of processing applied for the pixels of PU. In case of inter-prediction, a PU represents an

15 area of pixels that is processed by motion compensation using a motion vector specified for a PU. For intra prediction, the adjacent pixels of neighbor blocks are used as reference samples to predict a current block. A PU specifies a prediction mode that is selected from the set of intra-prediction modes for all the transform units (TUs) contained in this PU. A TU can have different sizes (e.g., 4x4, 8x8, 16x16 and 32x32 pixels) and can be

20 processed in different ways. For a TU, transform coding is performed, i.e. the prediction error is transformed with a discrete cosine transform or a discrete sine transform (in the HEVC/H.265 standard, it is applied to intra-coded blocks) and quantized. Hence, reconstructed pixels contain quantization noise (it can become apparent, for examples, as blockiness between units, ringing artifacts along with sharp edges, etc.) that in-loop filters

25 such as Deblocking Filter (DBF), Sample Adaptive Offset (SAO) and Adaptive Loop Filter (ALF) try to suppress.

Sub-integer pixel interpolation techniques have been developed in order to improve the prediction accuracy for a level of compression that can be achieved for inter prediction. In

30 this case, the predictive data generated during motion compensation, which is used to code a video block, may correspond to sub-integer pixels, whose values may be interpolated from values for full pixels of video blocks of a reference video frame or other coded unit to which a motion vector refers. The video encoder may calculate values for sub-integer pixel positions using interpolation techniques, e.g., by applying interpolation

35 filters to a set of support pixels, e.g., values for full integer pixels and/or previously coded values of other sub-integer pixel positions.

Today's standards H.264/AVC and H.265/HEVC are based on a $\frac{1}{4}$ pel displacement resolution. The Joint Video Exploration Team (JVET) group is exploring post-HEVC video compression technologies, including non-translational motion compensation models, such as affine transforms. In order to estimate and compensate the fractional-pel (or sub-integer) displacements, the full-integer pixels of the reference image has to be interpolated on the fractional-pel, i.e. sub-integer positions. To get the interpolated image on the fractional-pel positions interpolation filters are used. The problem of interpolation for non-translation motion models is the variable fractional-pel displacements.

The quality of the interpolated image strongly depends on the properties of the interpolation filter(s). Short-tap filters (e.g. bilinear) may suppress high frequencies and make interpolated picture blurred. On the other hand, long-tap filters (e.g. sinc-based) require more memory bandwidth and may preserve high frequencies but generate some ringing artifacts in the vicinity of sharp edges. Another consideration is that for motion compensation of non-translational models the complexity has been reduced by decreasing of accuracy of interpolation and prediction.

In the proposed JEM affine motion model there are two types of motion supported: zoom and rotation. Most of the fractional-pel positions are non-constant within a prediction unit. Pixels are replaced by sub-blocks to speed up the interpolation. Within one sub-block the displacement vector is constant and translational. The complexity was reduced gradually but accuracy was reduced as well. To improve the quality of prediction the precision of motion compensation was improved by increasing the precision of motion vector displacement for sub-blocks with increasing of amount of interpolation filters. Current accuracy of interpolation filtering for non-translational motion model still requires to be improved.

Thus, there is a need for an improved interpolation filter for an inter prediction apparatus and method for video coding providing an improved video coding efficiency.

SUMMARY

It is an object of the invention to provide an improved interpolation filter for an inter prediction apparatus and method for video coding providing an improved video coding efficiency.

The foregoing and other objects are achieved by the subject matter of the independent claims. Further implementation forms are apparent from the dependent claims, the description and the figures.

5

According to a first aspect the invention relates to an apparatus for inter prediction of a sample value of a current full-integer pixel of a plurality of pixels of a current block of a current frame of a video signal. The apparatus comprises a processing unit configured to:
10 determine a motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal and/or a motion compensation model; determine for the current full-integer pixel a corresponding sub-integer pixel in the reference frame on the basis of the motion vector of the current full-integer pixel; generate on the basis of a predefined set of filter support pixels in the current frame a set of corresponding filter support pixels in the reference frame, wherein the predefined set of filter support pixels in
15 the current frame comprises one or more neighboring full-integer pixels and/or sub-integer pixels of the current full-integer pixel; determine a respective sample value, in particular luminance value, of the corresponding sub-integer pixel of the current full-integer pixel and the corresponding filter support pixels in the reference frame; and determine the inter predicted sample value of the current pixel in the current frame by applying a spatial high-
20 pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels in the reference frame.

Thus, an improved inter prediction apparatus is provided allowing improving the efficiency
25 for video coding.

In a further possible implementation form of the first aspect, the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring half-integer pixels of the current full-integer pixel in the current frame.

30

In a further possible implementation form of the first aspect, the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring full-integer pixels of the current full-integer pixel in the current frame.

35 In a further possible implementation form of the first aspect, the spatial high-pass filter is a 5-tap filter or a 3-tap filter. In an implementation form, the 5-tap filter or the 3-tap filter is a

symmetric filter, i.e. a filter where the first and the fifth filter coefficients (or, in the case of the 3-tap filter, the first and the third filter coefficients) are identical and the second and the fourth filter coefficients are identical. In an implementation form, the first and the fifth filter coefficients are negative, while the other filter coefficients of the 5-tap filter are
5 positive.

In a further possible implementation form of the first aspect, the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and half-integer pixels of the current full-integer pixel and the 5-tap filter has the following
10 filter coefficients $(-6, 9, 26, 9, -6)$, which can be normalized by the factor $1/32$.

In a further possible implementation form of the first aspect, the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and half-integer pixels of the current full-integer pixel and the 5-tap filter has the following
15 filter coefficients $(-1, 0, 10, 0, -1)$, which can be normalized by the factor $1/8$. As will be appreciated, in a further possible implementation form without the intermediate half-integer support pixels, this filter becomes a 3-tap filter with the coefficients $(-1, 10, -1)$.

In a further possible implementation form of the first aspect, the processing unit of the
20 apparatus is configured to determine the respective sample values of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and of the corresponding filter support pixels in the reference frame on the basis of a bilinear interpolation of the respective neighboring full-integer pixels in the reference frame.

In a further possible implementation form of the first aspect, the processing unit of the
25 apparatus is further configured to determine a respective motion vector of each of the sub-integer pixels of the filter support pixels in the current frame on the basis of the motion vector of the current full-integer pixel and one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel.

30 In a further possible implementation form of the first aspect, the processing unit of the apparatus is configured to determine a respective motion vector of each of the sub-integer pixels of the filter support pixels in the current frame by determining a mean vector of the motion vector of the current full-integer pixel and one or more motion vectors of
35 neighboring full-integer pixels of the current full-integer pixel.

In a further possible implementation form of the first aspect, the processing unit of the apparatus is further configured to determine the one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel on the basis of the current frame and the reference frame of the video signal and/or a motion compensation model.

5

In a further possible implementation form of the first aspect, the processing unit of the apparatus is configured to determine the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels in the reference frame, in case the motion vector of the current full-integer pixel is determined on the basis of a non-translational motion compensation model, and to determine the inter predicted sample value of the current pixel in the current frame on the basis of a conventional scheme, such as the conventional scheme defined in H.264/AVC and H.265/HEVC, in case the motion vector of the current full-integer pixel is determined on the basis of a translational motion compensation model.

10

15

In a further possible implementation form of the first aspect, the non-translational motion compensation model is an affine, warping and/or panoramic motion compensation model.

20

According to a second aspect the invention relates to a corresponding method for inter prediction of a sample value of a current full-integer pixel of a plurality of pixels of a current block of a current frame of a video signal. The method comprises the steps of: determining a motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal and/or on the basis of a motion compensation model; determining for the current full-integer pixel a corresponding sub-integer pixel in the reference frame on the basis of the motion vector of the current full-integer pixel; generating on the basis of a predefined set of filter support pixels in the current frame a set of corresponding filter support pixels in the reference frame, wherein the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and/or sub-integer pixels of the current full-integer pixel; determining a respective sample value, in particular luminance value, of the corresponding sub-integer pixel of the current full-integer pixel and the corresponding filter support pixels in the reference frame; and determining the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the

25

30

35

reference frame and to the sample values of the corresponding filter support pixels in the reference frame.

5 Thus, an improved inter prediction method is provided allowing improving the efficiency for video coding.

The inter prediction method according to the second aspect of the invention can be performed by the inter prediction apparatus according to the first aspect of the invention. Further features of the inter prediction method according to the second aspect of the
10 invention result directly from the functionality of the inter prediction apparatus according to the first aspect of the invention and its different implementation forms described above and below.

According to a third aspect the invention relates to an encoding apparatus for encoding a
15 current frame of a video signal, wherein the encoding apparatus comprises an inter prediction apparatus according to the first aspect.

According to a fourth aspect the invention relates to a decoding apparatus for decoding a current reconstructed frame of a compressed video signal, wherein the decoding
20 apparatus comprises an inter prediction apparatus according to the first aspect.

According to a fifth aspect the invention relates to a computer program product comprising program code for performing the method according to the second aspect when executed on a computer or a processor.
25

Embodiments of the invention provide in particular for the following advantages. Embodiments of the invention allow performing interpolation with pixel-wise accuracy while keeping the complexity at a low level. Embodiments of the invention support any kind of non-translational movements. Embodiments of the invention allow removing
30 blocking artefacts across sub-blocks edges. Embodiments of the invention reduce memory bandwidth. Embodiments of the invention reduce memory requirements for storing a set of filter coefficients. Embodiments of the invention allow reusing well optimized in HW bilinear transform. Embodiments of the invention align orientation of filtering along transformation. Embodiments of the invention allow reducing ringing
35 artifacts caused by long motion interpolation filter while improving quality of interpolated edges. Embodiments of the invention allow increasing subjective quality of edges in

reconstructed pictures. Moreover, embodiments of the invention do not require any additional signaling and, thus, can seamlessly replace existing interpolation methods for non-translational movements. Splitting the interpolation into two steps allows separating the fractional offset compensation from the high-pass filtering.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Further embodiments of the invention will be described with respect to the following figures, wherein:

10

Fig. 1 shows a schematic diagram illustrating an encoding apparatus according to an embodiment comprising an inter prediction apparatus according to an embodiment;

15

Fig. 2 shows a schematic diagram illustrating a decoding apparatus according to an embodiment comprising an inter prediction apparatus according to an embodiment;

Fig. 3 shows a schematic diagram illustrating an aspect of an interpolation process implemented in an inter prediction apparatus according to an embodiment;

20

Fig. 4 shows a schematic diagram illustrating different aspects of an interpolation process implemented in an inter prediction apparatus according to an embodiment;

Fig. 5 shows a schematic diagram illustrating different aspects of an interpolation process implemented in an inter prediction apparatus according to an embodiment; and

25

Fig. 6 shows a flow diagram illustrating steps of an inter prediction method according to an embodiment.

Fig. 7 shows Table 1.

30

In the various figures, identical reference signs will be used for identical or functionally equivalent features.

DETAILED DESCRIPTION OF EMBODIMENTS

35

In the following description, reference is made to the accompanying drawings, which form part of the disclosure, and in which are shown, by way of illustration, specific aspects in which the present invention may be placed. It is understood that other aspects may be utilized and structural or logical changes may be made without departing from the scope of the present invention. The following detailed description, therefore, is not to be taken in a limiting sense, as the scope of the present invention is defined by the appended claims.

For instance, it is understood that a disclosure in connection with a described method may also hold true for a corresponding device or system configured to perform the method and vice versa. For example, if a specific method step is described, a corresponding device may include a unit to perform the described method step, even if such unit is not explicitly described or illustrated in the figures. Further, it is understood that the features of the various exemplary aspects described herein may be combined with each other, unless specifically noted otherwise.

Figure 1 shows an encoding apparatus 100 according to an embodiment comprising an inter prediction apparatus 144 according to an embodiment. The encoding apparatus 100 is configured to encode a block of a frame of a video signal comprising a plurality of frames (also referred to as pictures or images herein), wherein each frame is dividable into a plurality of blocks and each block comprises a plurality of pixels. In an embodiment, the blocks could be macro blocks, coding tree units, coding units, prediction units and/or prediction blocks.

In the exemplary embodiment shown in figure 1, the encoding apparatus 100 is implemented in the form of a hybrid video coding encoder. Usually, the first frame of a video signal is an intra frame, which is encoded using only intra prediction. To this end, the embodiment of the encoding apparatus 100 shown in figure 2 further comprises an intra prediction unit 154 for intra prediction. An intra frame can be decoded without information from other frames. The intra prediction unit 154 can perform the intra prediction of a block on the basis of information provided by the intra estimation unit 152.

The blocks of subsequent frames following the first intra frame can be coded using inter or intra prediction, as selected by a mode selection unit 160. Generally, the inter prediction unit 144 can be configured to perform motion compensation of a block based on motion estimation, as will be described in more detail further below. In an embodiment, the motion estimation can be performed by an inter estimation unit 142 of the encoding

apparatus. However, in other embodiments, the functionality of the inter estimation unit 142 can be implemented as part of the inter prediction unit 144 as well.

Furthermore, in the hybrid encoder embodiment shown in figure 1 a residual calculation unit 104 determines the difference between the original block and its prediction, i.e. the residual block defining the prediction error of the intra/inter picture prediction. This residual block is transformed by the transformation unit 106 (for instance using a DCT) and the transformation coefficients are quantized by the quantization unit 108. The output of the quantization unit 108 as well as the coding or side information provided, for instance, by the inter prediction unit 144 are further encoded by an entropy encoding unit 170.

A hybrid video encoder, such as the encoding apparatus 100 shown in figure 1, usually duplicates the decoder processing such that both will generate the same predictions. Thus, in the embodiment shown in figure 1 the inverse quantization unit 110 and the inverse transformation unit perform the inverse operations of the transformation unit 106 and the quantization unit 108 and duplicate the decoded approximation of the residual block. The decoded residual block data is then added to the results of the prediction, i.e. the prediction block, by the reconstruction unit 114. Then, the output of the reconstruction unit 114 can be provided to a line buffer 116 to be used for intra prediction and is further processed by an in-loop filter 120 for removing image artifacts. The final picture is stored in a decoded picture buffer 130 and can be used as a reference frame for the inter prediction of subsequent frames.

Figure 2 shows a decoding apparatus 200 according to an embodiment comprising an inter prediction apparatus 244 according to an embodiment. The decoding apparatus 200 is configured to decode a block of a frame of an encoded video signal. In the embodiment shown in figure 2 the decoding apparatus 200 is implemented as a hybrid decoder. An entropy decoding unit 204 performs entropy decoding of the encoded picture data, which generally can comprise prediction errors (i.e. residual blocks), motion data and other side information, which are needed, in particular, for the inter prediction apparatus 244 and an intra prediction unit 254 as well as other components of the decoding apparatus 200. In the embodiment shown in figure 2, the inter prediction apparatus 244 or the intra prediction unit 254 of the decoding apparatus 200 shown in figure 3 are selected by a mode selection unit 260 and function in the same way as the inter prediction apparatus 144 and the intra prediction unit 154 of the encoding apparatus 100 shown in figure 1, so that identical predictions can be generated by the encoding apparatus 100 and the

decoding apparatus 200. A reconstruction unit 214 of the decoding apparatus 200 is configured to reconstruct the block on the basis of the filtered predicted block and the residual block provided by the inverse quantization unit 210 and the inverse transformation unit 212. As in the case of the encoding apparatus 100, the reconstructed
5 block can be provided to a line buffer 216 used for intra prediction and the filtered block/frame can be provided to a decoded picture buffer 230 by the in-loop filter 220 for future inter predictions.

As already described above, the apparatus 144, 244 is configured to perform an inter
10 prediction of a sample value of a current full-integer pixel of a plurality of pixels of a current block of a current frame of a video signal. The apparatus 144, 244 comprises a processing unit, which can be implemented in software and/or hardware.

The processing unit of the apparatus 144, 244 is configured to determine a motion vector
15 of the current full-integer pixel. In an embodiment, the processing unit of the apparatus 144, 244 is configured to determine the motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal by determining the position of the pixel in the reference frame corresponding to the current full-integer pixel in the current frame. In an embodiment, the processing unit of the apparatus 144,
20 244 is configured to determine the motion vector of the current full-integer pixel on the basis of an adopted motion compensation model. For instance, in the case of a translational or affine motion compensation model the motion vector of the current full-integer pixel can be determined using one or more motion vectors of pixels of the same block the current full-integer pixel belongs to. As used herein, "motion compensation
25 model" is also referred to as motion transformational model, motion model description and the like.

The processing unit of the apparatus 144, 244 is further configured to determine for the current full-integer pixel the corresponding sub-integer pixel in the reference frame on the
30 basis of the motion vector of the current full-integer pixel.

The processing unit of the apparatus 144, 244 is further configured to generate on the basis of a predefined set of filter support pixels in the current frame a set of corresponding filter support pixels in the reference frame. The predefined set of filter support pixels in the
35 current frame comprises one or more neighboring full-integer pixels and/or sub-integer pixels of the current full-integer pixel.

In an embodiment, the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring half-integer pixels of the current full-integer pixel in the current frame. For instance, in an embodiment the

5 predefined set of filter support pixels in the current frame comprises the neighboring half-integer pixels above, to the left of, below and to the right of the current full-integer pixel.

In an embodiment, the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring full-integer pixels of the

10 current full-integer pixel in the current frame. For instance, in an embodiment the predefined set of filter support pixels in the current frame comprises the neighboring full-integer pixels above, to the left of, below and to the right of the current full-integer pixel. In an embodiment, the predefined set of filter support pixels in the current frame can comprise the neighboring half-integer and full-integer pixels above, to the left of, below

15 and to the right of the current full-integer pixel in the current frame.

The processing unit of the apparatus 144, 244 is further configured to determine a respective sample value, in particular a luminance value of the corresponding sub-integer pixel of the current full-integer pixel and the corresponding filter support pixels in the

20 reference frame.

In an embodiment, the processing unit of the apparatus 144, 244 is configured to determine the respective sample values of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and of the corresponding filter support

25 pixels in the reference frame on the basis of a bilinear interpolation of the respective neighboring full-integer pixels in the reference frame. Figure 3 illustrates an example of using bilinear interpolation for determining the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame. In figure 3, a reference block in the reference frame is enlarged and rotated relative to a current block comprising an exemplary current pixel of the current frame. Moreover, figure 3 illustrated the

30 increased resolution used for the filter support pixels.

As can be taken from the enlarged view in figure 3, in an embodiment the sample value L of the corresponding sub-integer pixel of the current full-integer pixel in the reference

35 frame can be determined by the processing unit as follows. The corresponding sub-integer pixel of the current full-integer pixel has the fractional position (fdX, fdY) in a

corresponding cell of the sample grid of the reference frame. L_0, L_1, L_2, L_3 are the known sample values of the neighboring full-integer pixels in the reference frame (i.e. the full-integer pixels located at the corners of the corresponding cell of the sample grid of the reference frame the corresponding sub-integer pixel of the current full-integer pixel is

5 located in). On the basis of the fractional position (fdX, fdY) the respective areas of the rectangles corresponding to s_0, s_1, s_2, s_3 can be calculated as follows: $s_0 = fdX * fdY, s_1 = (1 - fdX) * fdY, s_2 = fdX * (1 - fdY), s_3 = (1 - fdX) * (1 - fdY)$. Bilinear interpolation can be expressed using a 2-tap filter having the following horizontal coefficients ($1 - fdX, fdX$) and the following vertical coefficients ($1 - fdY, fdY$). On the basis of these weighting factors the

10 sample value L of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame can be determined on the basis of the following equation:

$L = L_0 * s_3 + L_1 * s_2 + L_2 * s_1 + L_3 * s_0$. As already mentioned above, the same bilinear interpolation can be used for determining the sample values for the corresponding filter support pixels in the reference frame.

15

The processing unit of the apparatus 144, 244 is further configured to determine an inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter

20 support pixels in the reference frame.

20

In an embodiment, the spatial high-pass filter is a 5-tap filter. In an embodiment, the 5-tap filter is a symmetric filter, i.e. a filter where the first and the fifth filter coefficients are identical and the second and the fourth filter coefficients are identical. In an embodiment,

25 the first and the fifth filter coefficients are negative, while the other filter coefficients of the 5-tap filter are positive. In an embodiment, the spatial high-pass filter can be applied separately in the vertical and the horizontal direction.

25

In an embodiment, the predefined set of filter support pixels in the current frame

30 comprises five neighboring full-integer pixels and half-integer pixels of the current full-integer pixel and the 5-tap filter has the following filter coefficients $(-6, 9, 26, 9, -6)$ within a given numerical precision, which can be normalized by the factor $1/32$.

30

In a further embodiment, the predefined set of filter support pixels in the current frame

35 comprises five neighboring full-integer pixels and half-integer pixels of the current full-integer pixel and the 5-tap filter has the following filter coefficients $(-1, 0, 10, 0, -1)$ within a

35

given numerical precision, which can be normalized by the factor $1/8$. As will be appreciated, in an embodiment without the intermediate half-integer support pixels, this filter becomes a 3-tap filter with three full-integer support pixels and with the coefficients $(-1, 10, -1)$.

5

In an embodiment, the processing unit of the apparatus 144, 244 is configured to determine the inter predicted sample value of the current pixel in the current frame on the basis of a two-dimensional separable convolution process on the basis of the filter support pixels. For instance, for the case of a 5-tap filter the processing unit of the apparatus 144, 244 is configured to apply the 5-tap filter to each horizontal and vertical line defined by the five horizontally and vertically neighbouring half-integer and full-integer pixels of the current pixel.

10

Figure 4 illustrates different stages of the processing unit performed by the processing unit of the apparatus 144, 244 using a 5-tap filter in the vertical and the horizontal direction for the example shown in figure 3. As in the example shown in figure 3, the reference block is enlarged and rotated (corresponding to an affine transformation) relative to the current block, the 5-tap filters, which are vertical and horizontal in the current frame, are rotated in the reference frame.

15

20

In the following further embodiments of the inter prediction apparatus 144, 244, the encoding apparatus 100 and the decoding apparatus 200 will be described. In this context it will be understood that embodiments of the inter prediction apparatus 144, 244 relate to embodiments of the inter prediction apparatus 133 as implemented in the encoding apparatus 100 as well as embodiments of the inter prediction apparatus 244 as implemented in the decoding apparatus 200.

25

In an embodiment, the processing unit of the apparatus 144, 244 is further configured to determine a respective motion vector of each of the sub-integer pixels of the filter support pixels in the current frame on the basis of the motion vector of the current full-integer pixel and one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel. To this end, in an embodiment, the processing unit of the apparatus 144, 244 is configured to determine a mean vector of the motion vector of the current full-integer pixel and one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel. For instance, for determining the motion vector of the half-integer pixel above the current full-integer pixel in the current frame the processing unit of the apparatus 144, 244

30

35

can compute the average, i.e. mean of the motion vector of the current full-integer pixel and the motion vector of the neighboring full-integer pixel above the current full-integer pixel.

- 5 Similar to the determination of the motion vector of the current pixel the processing unit of the apparatus 144, 244 can be configured to determine the one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel on the basis of the current frame and the reference frame of the video signal and/or a motion compensation model.
- 10 In an embodiment, the processing unit of the apparatus 144, 244 is configured to determine the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels in the reference frame, in case the motion vector of the
- 15 current full-integer pixel is determined on the basis of a non-translational motion compensation model, and to determine the inter predicted sample value of the current pixel in the current frame on the basis of a conventional scheme, such as the conventional scheme defined in H.264/AVC and H.265/HEVC, in case the motion vector of the current full-integer pixel is determined on the basis of a translational motion compensation model.
- 20 In an embodiment, the non-translational motion compensation model is an affine, warping and/or panoramic motion compensation model.

In an another embodiment for translational motion compensation model the processing unit of the apparatus 144, 244 is configured to construct the set of separable 4-tap filters

- 25 by convolving the proposed filters (i.e. the bilinear filter and the subsequent high-pass filter); for example, as follows: $[f, 1-f] * [c_0, c_1, c_2] = [(1-f) * (c_0), f * (c_0) + (1-f) * c_1, f * (c_1) + (1-f) * c_2, f * (c_2)]$, where f is a fractional offset of the motion vector in the horizontal or vertical direction, and $[c_0, c_1, c_2]$ are coefficients of the high-pass filter. The fractional position in this case may be discrete with a certain precision to obtain the finite table size. The final
- 30 values of the coefficients may be multiplied by a norm factor and rounded so that the sum of coefficients for each offset equals the norm factor. For some specific scenarios, a filter with zero fractional offset may be equal to 1 (701) (no filtering for zero offset). For instance, having norm factor equal to 64 and fractional offset precision 1/32, the set of filter coefficients (700) may be the one shown in Table 1. This technique may be used, for
- 35 example, for filtering chroma components. The technique may also be used for filtering of small blocks in order to keep worst case complexity and/or memory bandwidth low.

In any embodiments, the intermediate rounding or clipping at any step of implementation could be applied in order to keep the fixed number of bits as input range for computational operations. For instance, after any separable step of filtering (like horizontal filtering) intermediate values could be normalized and rounded to 16 bit of precision for using in the
5 following multiplication by filter coefficient of the next step (like vertical filtering).

Figure 5 summarizes several aspects of embodiments of the invention described above.

Figure 6 shows a flow diagram illustrating steps of a corresponding inter prediction
10 method 600 according to an embodiment. The method 600 comprises the following steps: determining 601 a motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal and/or on the basis of a motion compensation model; determining 603 for the current full-integer pixel a corresponding sub-integer pixel in the reference frame on the basis of the motion vector of the current
15 full-integer pixel; generating 605 on the basis of a predefined set of filter support pixels in the current frame a set of corresponding filter support pixels in the reference frame, wherein the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and/or sub-integer pixels of the current full-integer pixel; determining 607 a respective sample value of the corresponding sub-integer pixel of
20 the current full-integer pixel and the corresponding filter support pixels in the reference frame; and determining 609 the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels in the reference frame.

25 While a particular feature or aspect of the disclosure may have been disclosed with respect to only one of several implementations or embodiments, such feature or aspect may be combined with one or more other features or aspects of the other implementations or embodiments as may be desired and advantageous for any given or particular
30 application. Furthermore, to the extent that the terms "include", "have", "with", or other variants thereof are used in either the detailed description or the claims, such terms are intended to be inclusive in a manner similar to the term "comprise". Also, the terms "exemplary", "for example" and "e.g." are merely meant as an example, rather than the best or optimal. The terms "coupled" and "connected", along with derivatives may have
35 been used. It should be understood that these terms may have been used to indicate that

two elements cooperate or interact with each other regardless whether they are in direct physical or electrical contact, or they are not in direct contact with each other.

Although specific aspects have been illustrated and described herein, it will be

5 appreciated by those of ordinary skill in the art that a variety of alternate and/or equivalent implementations may be substituted for the specific aspects shown and described without departing from the scope of the present disclosure. This application is intended to cover any adaptations or variations of the specific aspects discussed herein.

10 Although the elements in the following claims are recited in a particular sequence with corresponding labeling, unless the claim recitations otherwise imply a particular sequence for implementing some or all of those elements, those elements are not necessarily intended to be limited to being implemented in that particular sequence.

15 Many alternatives, modifications, and variations will be apparent to those skilled in the art in light of the above teachings. Of course, those skilled in the art readily recognize that there are numerous applications of the invention beyond those described herein. While the present invention has been described with reference to one or more particular embodiments, those skilled in the art recognize that many changes may be made thereto
20 without departing from the scope of the present invention. It is therefore to be understood that within the scope of the appended claims and their equivalents, the invention may be practiced otherwise than as specifically described herein.

25

CLAIMS

1. An apparatus (144, 244) for inter prediction of a sample value of a current full-integer pixel of a plurality of pixels of a current block of a current frame of a video signal,
5 wherein the apparatus (144, 244) comprises a processing unit configured to:

determine a motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal or a non-translational motion compensation model;

10 determine for the current full-integer pixel a corresponding sub-integer pixel in the reference frame on the basis of the motion vector of the current full-integer pixel;

generate on the basis of a predefined set of filter support pixels in the current frame a set
15 of corresponding filter support pixels in the reference frame, wherein the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and/or sub-integer pixels of the current full-integer pixel;

determine a respective motion vector of each of the sub-integer pixels of the filter support
20 pixels in the current frame by determining a mean vector of the motion vector of the current full-integer pixel and one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel;

determine a respective sample value of the corresponding sub-integer pixel of the current
25 full-integer pixel and the corresponding filter support pixels in the reference frame; and

determine the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the
30 corresponding filter support pixels in the reference frame.

2. The apparatus (144, 244) of claim 1, wherein the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring half-integer pixels of the current full-integer pixel in the current frame.

3. The apparatus (144, 244) of claim 1 or 2, wherein the predefined set of filter support pixels in the current frame comprises one or more vertically and/or horizontally neighboring full-integer pixels of the current full-integer pixel in the current frame.
- 5 4. The apparatus (144, 244) of any one of the preceding claims, wherein the spatial high-pass filter is a 5-tap or a 3-tap filter.
5. The apparatus (144, 244) of claim 4, wherein the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and half-integer pixels of the current full-integer pixel and wherein the 5-tap filter has the following
10 filter coefficients (-6, 9, 26, 9, -6).
6. The apparatus (144, 244) of claim 4, wherein the predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer pixels and/or
15 half-integer pixels of the current full-integer pixel and wherein the 5-tap filter has the filter coefficients (-1, 0, 10, 0, -1) or the 3-tap filter has the filter coefficients (-1, 10, -1).
7. The apparatus (144, 244) of any one of the preceding claims, wherein the processing unit of the apparatus (144, 244) is configured to determine the respective
20 sample values of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and of the corresponding filter support pixels in the reference frame on the basis of a bilinear interpolation of the respective neighboring full-integer pixels in the reference frame.
- 25 8. The apparatus (144, 244) of claim 1, wherein the processing unit of the apparatus (144, 244) is further configured to determine the one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel on the basis of the current frame and the reference frame of the video signal and/or a motion compensation model.
- 30 9. The apparatus (144, 244) of any one of the preceding claims, wherein the processing unit of the apparatus (144, 244) is configured to determine the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels
35 in the reference frame, in case the motion vector of the current full-integer pixel is determined on the basis of a non-translational motion compensation model, and to

determine the inter predicted sample value of the current pixel in the current frame on the basis of a conventional scheme, in case the motion vector of the current full-integer pixel is determined on the basis of a translational motion compensation model.

5 10. The apparatus (144, 244) of claim 9, wherein the non-translational motion compensation model is an affine, warping and/or panoramic motion compensation model.

11. The apparatus (144, 244) of any one of the preceding claims, wherein the processing unit is configured to apply an intermediate rounding operation in the filtering
10 process.

12. A method (600) for inter prediction of a sample value of a current full-integer pixel of a plurality of pixels of a current block of a current frame of a video signal, wherein the method (600) comprises:

15 determining (601) a motion vector of the current full-integer pixel on the basis of the current frame and a reference frame of the video signal or on the basis of a non-translational motion compensation model;

20 determining (603) for the current full-integer pixel a corresponding sub-integer pixel in the reference frame on the basis of the motion vector of the current full-integer pixel;

generating (605) on the basis of a predefined set of filter support pixels in the current frame a set of corresponding filter support pixels in the reference frame, wherein the
25 predefined set of filter support pixels in the current frame comprises one or more neighboring full-integer and/or sub-integer pixels of the current full-integer pixel;

determining a respective motion vector of each of the sub-integer pixels of the filter support pixels in the current frame by determining a mean vector of the motion vector of
30 the current full-integer pixel and one or more motion vectors of neighboring full-integer pixels of the current full-integer pixel;

determining (607) a respective sample value of the corresponding sub-integer pixel of the current full-integer pixel and the corresponding filter support pixels in the reference frame;
35 and

determining (609) the inter predicted sample value of the current pixel in the current frame by applying a spatial high-pass filter to the sample value of the corresponding sub-integer pixel of the current full-integer pixel in the reference frame and to the sample values of the corresponding filter support pixels in the reference frame.

5

13. An encoding apparatus (100) for encoding a current frame of a video signal, wherein the encoding apparatus (100) comprises an inter prediction apparatus (144) according to any one of claims 1 to 11, wherein the processing unit is configured to determine the motion vector of the current full-integer pixel on the basis of the current frame and reference frame of the video signal or on the basis of the non-translational motion compensation model.

14. A decoding apparatus (200) for decoding a current reconstructed frame of a compressed video signal, wherein the decoding apparatus (200) comprises an inter prediction apparatus (244) according to any one of claims 1 to 11, wherein the processing unit is configured to determine the motion vector of the current full-integer pixel on the basis of the current frame and reference frame of the video signal.

15. A computer program product comprising program code for performing the method of claim 12 when executed on a computer or a processor.