

REFERENCE PICTURE RESAMPLING AND INTER-CODING TOOLS FOR VIDEO CODING

[0001] This application claims priority to U.S. Patent Application No. 16/874,091 filed May 14, 2020, which claims the benefit of U.S. Provisional Patent Application No. 62/849,754, filed May 17, 2019, the entire content of each of which is hereby incorporated by reference.

TECHNICAL FIELD

[0002] This disclosure relates to video encoding and video decoding.

BACKGROUND

[0003] Digital video capabilities can be incorporated into a wide range of devices, including digital televisions, digital direct broadcast systems, wireless broadcast systems, personal digital assistants (PDAs), laptop or desktop computers, tablet computers, e-book readers, digital cameras, digital recording devices, digital media players, video gaming devices, video game consoles, wireless phones (e.g., cellular or satellite radio telephones or so-called “smart phones”), video teleconferencing devices, video streaming devices, and the like. Digital video devices implement video coding techniques, such as those described in the standards defined by MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Part 10, Advanced Video Coding (AVC), ITU-T H.265/High Efficiency Video Coding (HEVC), and extensions of such standards. The video devices may transmit, receive, encode, decode, and/or store digital video information more efficiently by implementing such video coding techniques.

[0004] Video coding techniques include spatial (intra-picture) prediction and/or temporal (inter-picture) prediction to reduce or remove redundancy inherent in video sequences. For block-based video coding, a video slice (e.g., a video picture or a portion of a video picture) may be partitioned into video blocks, which may also be referred to as coding tree units (CTUs), coding units (CUs) and/or coding nodes. Video blocks in an intra-coded (I) slice of a picture are encoded using spatial prediction with respect to reference samples in neighboring blocks in the same picture. Video blocks in an inter-coded (P or B) slice of a picture may use spatial prediction with respect to reference samples in neighboring blocks in the same picture or temporal prediction with

respect to reference samples in other reference pictures. Pictures may be referred to as frames, and reference pictures may be referred to as reference frames.

SUMMARY

[0005] In general, this disclosure describes techniques for utilizing inter-coding tools for coding (e.g., encoding or decoding) a current block. The example techniques may be usable in video coding cases where sizes (e.g., resolution, dimensions, etc.) of the current picture and a reference picture are not the same. As one example, the example techniques may be usable where reference picture resampling (RPR) is enabled but the techniques should not be considered limited to when RPR is enabled.

[0006] When the current picture and the reference picture have different sizes, there may be limitations on which inter-coding tools are available. However, various inter-coding tools that may have been disabled when the reference picture and current picture have different sizes need not be disabled. This disclosure describes example ways in which inter-coding tools may be utilized even where a block of the current picture has a reference picture that is of a different size than the current picture. In this manner, a video coder (e.g., video encoder or video decoder) may leverage the coding efficiency gains that are available from various inter-coding tools even when a block of the current picture has a reference picture that is a different size than the current picture.

[0007] In one example, a method includes determining a first reference block in a first reference picture for a first block of a current picture; determining whether the first reference picture is a same size as the current picture; based on the first reference picture being a different size than the current picture, coding the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled; determining a second reference block in a second reference picture for a second block of the current picture; determining whether the second reference picture is a same size as the current picture; and based on the second reference picture being the same size as the current picture, coding the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0008] In one example, a device includes a memory configured to store video data and one or more processors implemented in circuitry and coupled to the memory, the one or more processors being configured to: determine a first reference block in a first

reference picture for a first block of a current picture; determine whether the first reference picture is a same size as the current picture; based on the first reference picture being a different size than the current picture, code the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled; determine a second reference block in a second reference picture for a second block of the current picture; determine whether the second reference picture is a same size as the current picture; and based on the second reference picture being the same size as the current picture, code the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0009] In one example, a device includes means for determining a first reference block in a first reference picture for a first block of a current picture; means for determining whether the first reference picture is a same size as the current picture; means for coding the first block with a first inter-coding tool of a first set of inter-coding tools based on the first reference picture being a different size than the current picture, wherein a first particular tool of a plurality of inter-coding tools is disabled; means for determining a second reference block in a second reference picture for a second block of the current picture; means for determining whether the second reference picture is a same size as the current picture; and means for coding the second block with a second inter-coding tool of a second set of inter-coding tools based on the second reference picture being the same size as the current picture, wherein the first particular tool is enabled.

[0010] In one example, a non-transitory computer-readable storage medium stores instructions that, when executed, cause one or more processors to determine a first reference block in a first reference picture for a first block of a current picture; determine whether the first reference picture is a same size as the current picture; based on the first reference picture being a different size than the current picture, code the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled; determine a second reference block in a second reference picture for a second block of the current picture; determine whether the second reference picture is a same size as the current picture; and based on the second reference picture being the same size as the current picture, code the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0011] The details of one or more examples are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description, drawings, and claims.

BRIEF DESCRIPTION OF DRAWINGS

[0012] FIG. 1 is a block diagram illustrating an example video encoding and decoding system that may perform the techniques of this disclosure.

[0013] FIGS. 2A and 2B are conceptual diagrams illustrating an example quadtree binary tree (QTBT) structure, and a corresponding coding tree unit (CTU).

[0014] FIG. 3 is a block diagram illustrating an example video encoder that may perform the techniques of this disclosure.

[0015] FIG. 4 is a block diagram illustrating an example video decoder that may perform the techniques of this disclosure.

[0016] FIG. 5 is a flowchart illustrating an example of selectively enabling use of coding tools according to the techniques of this disclosure.

[0017] FIG. 6 is a flowchart illustrating another example of selectively enabling use of coding tools according to the techniques of this disclosure.

[0018] FIG. 7 a flowchart illustrating another example of selectively enabling use of coding tools according to the techniques of this disclosure.

[0019] FIG. 8 is a flowchart illustrating an example of encoding video data.

[0020] FIG. 9 is a flowchart illustrating an example of decoding video data.

DETAILED DESCRIPTION

[0021] Reference picture resampling (RPR) may be used when a reference picture has a different size than a current picture. In RPR, motion vector scaling and interpolation may be used when performing inter prediction. In some cases, decoder side tools use pixel information around the reference block. If the size of the reference picture is different than the size of the current picture, this may cause problems, such as an increase in memory needed to pull samples outside of the reference block.

[0022] Some techniques disallow certain coding tools when using RPR. Those techniques disallow the certain coding tools when using RPR because leaving these coding tools available may result in poor video quality on the decoder side and, therefore, be a waste of computing resources to implement decoder side tools with different resolutions. There may also be an increase in memory bandwidth needed due

to the video coder pulling more information from the reference picture if these coding tools are working. However, by disabling these coding tools, a video encoder or video decoder may not take advantage of such coding tools when they may be beneficial (e.g., provide better quality decoded video, use less bandwidth, use less processing power, etc.).

[0023] With the techniques of this disclosure, the decoder side tools are selectively enabled when reference picture resampling is used. For example, the decoder side tools are enabled based on a comparison of resolution (e.g., height and width (size)) of the reference picture and the current picture. In this manner, the video decoder does not waste computation cycles or perform coding tools that result in poor quality blocks. In addition, memory bandwidth is saved by not implementing decoder side tools that requires additional information from the reference picture. Another advantage of selectively enabling these tools on a block by block basis is that blocks that would benefit from the tools can have the tools enabled and blocks that would not benefit from tools may have the tools disabled, rather than disabling the tools for all blocks of a current picture.

[0024] FIG. 1 is a block diagram illustrating an example video encoding and decoding system 100 that may perform the techniques of this disclosure. The techniques of this disclosure are generally directed to coding (encoding and/or decoding) video data. In general, video data includes any data for processing a video. Thus, video data may include raw, unencoded video, encoded video, decoded (e.g., reconstructed) video, and video metadata, such as signaling data.

[0025] As shown in FIG. 1, system 100 includes a source device 102 that provides encoded video data to be decoded and displayed by a destination device 116, in this example. In particular, source device 102 provides the video data to destination device 116 via a computer-readable medium 110. Source device 102 and destination device 116 may comprise any of a wide range of devices, including desktop computers, notebook (i.e., laptop) computers, tablet computers, set-top boxes, telephone handsets such smartphones, televisions, cameras, display devices, digital media players, video gaming consoles, video streaming device, or the like. In some cases, source device 102 and destination device 116 may be equipped for wireless communication, and thus may be referred to as wireless communication devices.

[0026] In the example of FIG. 1, source device 102 includes video source 104, memory 106, video encoder 200, and output interface 108. Destination device 116 includes

input interface 122, video decoder 300, memory 120, and display device 118. In accordance with this disclosure, video encoder 200 of source device 102 and video decoder 300 of destination device 116 may be configured to apply the techniques for utilizing inter-coding tools for coding (e.g., encoding or decoding) a current block such as in cases where a block of a current picture has a reference picture that is a different size (e.g., resolution, dimensions, etc.) than the current picture. Thus, source device 102 represents an example of a video encoding device, while destination device 116 represents an example of a video decoding device. In other examples, a source device and a destination device may include other components or arrangements. For example, source device 102 may receive video data from an external video source, such as an external camera. Likewise, destination device 116 may interface with an external display device, rather than including an integrated display device.

[0027] System 100 as shown in FIG. 1 is merely one example. In general, any digital video encoding and/or decoding device may perform techniques for utilizing inter-coding tools for coding (e.g., encoding or decoding) a current block such as in cases where a block of a current picture has a reference picture with a different size (e.g., resolution, dimension, etc.) than the current picture. Source device 102 and destination device 116 are merely examples of such coding devices in which source device 102 generates coded video data for transmission to destination device 116. This disclosure refers to a “coding” device as a device that performs coding (encoding and/or decoding) of data. Thus, video encoder 200 and video decoder 300 represent examples of coding devices, in particular, a video encoder and a video decoder, respectively. In some examples, source device 102 and destination device 116 may operate in a substantially symmetrical manner such that each of source device 102 and destination device 116 include video encoding and decoding components. Hence, system 100 may support one-way or two-way video transmission between source device 102 and destination device 116, e.g., for video streaming, video playback, video broadcasting, or video telephony.

[0028] In general, video source 104 represents a source of video data (i.e., raw, unencoded video data) and provides a sequential series of pictures (also referred to as “frames”) of the video data to video encoder 200, which encodes data for the pictures. Video source 104 of source device 102 may include a video capture device, such as a video camera, a video archive containing previously captured raw video, and/or a video feed interface to receive video from a video content provider. As a further alternative,

video source 104 may generate computer graphics-based data as the source video, or a combination of live video, archived video, and computer-generated video. In each case, video encoder 200 encodes the captured, pre-captured, or computer-generated video data. Video encoder 200 may rearrange the pictures from the received order (sometimes referred to as “display order”) into a coding order for coding. Video encoder 200 may generate a bitstream including encoded video data. Source device 102 may then output the encoded video data via output interface 108 onto computer-readable medium 110 for reception and/or retrieval by, e.g., input interface 122 of destination device 116.

[0029] Memory 106 of source device 102 and memory 120 of destination device 116 represent general purpose memories. In some example, memories 106, 120 may store raw video data, e.g., raw video from video source 104 and raw, decoded video data from video decoder 300. Additionally or alternatively, memories 106, 120 may store software instructions executable by, e.g., video encoder 200 and video decoder 300, respectively. Although shown separately from video encoder 200 and video decoder 300 in this example, it should be understood that video encoder 200 and video decoder 300 may also include internal memories for functionally similar or equivalent purposes. Furthermore, memories 106, 120 may store encoded video data, e.g., output from video encoder 200 and input to video decoder 300. In some examples, portions of memories 106, 120 may be allocated as one or more video buffers, e.g., to store raw, decoded, and/or encoded video data.

[0030] Computer-readable medium 110 may represent any type of medium or device capable of transporting the encoded video data from source device 102 to destination device 116. In one example, computer-readable medium 110 represents a communication medium to enable source device 102 to transmit encoded video data directly to destination device 116 in real-time, e.g., via a radio frequency network or computer-based network. Output interface 108 may modulate a transmission signal including the encoded video data, and input interface 122 may modulate the received transmission signal, according to a communication standard, such as a wireless communication protocol. The communication medium may comprise any wireless or wired communication medium, such as a radio frequency (RF) spectrum or one or more physical transmission lines. The communication medium may form part of a packet-based network, such as a local area network, a wide-area network, or a global network such as the Internet. The communication medium may include routers, switches, base

stations, or any other equipment that may be useful to facilitate communication from source device 102 to destination device 116.

[0031] In some examples, source device 102 may output encoded data from output interface 108 to storage device 112. Similarly, destination device 116 may access encoded data from storage device 112 via input interface 122. Storage device 112 may include any of a variety of distributed or locally accessed data storage media such as a hard drive, Blu-ray discs, DVDs, CD-ROMs, flash memory, volatile or non-volatile memory, or any other suitable digital storage media for storing encoded video data.

[0032] In some examples, source device 102 may output encoded video data to file server 114 or another intermediate storage device that may store the encoded video generated by source device 102. Destination device 116 may access stored video data from file server 114 via streaming or download. File server 114 may be any type of server device capable of storing encoded video data and transmitting that encoded video data to the destination device 116. File server 114 may represent a web server (e.g., for a website), a File Transfer Protocol (FTP) server, a content delivery network device, or a network attached storage (NAS) device. Destination device 116 may access encoded video data from file server 114 through any standard data connection, including an Internet connection. This may include a wireless channel (e.g., a Wi-Fi connection), a wired connection (e.g., DSL, cable modem, etc.), or a combination of both that is suitable for accessing encoded video data stored on file server 114. File server 114 and input interface 122 may be configured to operate according to a streaming transmission protocol, a download transmission protocol, or a combination thereof.

[0033] Output interface 108 and input interface 122 may represent wireless transmitters/receiver, modems, wired networking components (e.g., Ethernet cards), wireless communication components that operate according to any of a variety of IEEE 802.11 standards, or other physical components. In examples where output interface 108 and input interface 122 comprise wireless components, output interface 108 and input interface 122 may be configured to transfer data, such as encoded video data, according to a cellular communication standard, such as 4G, 4G-LTE (Long-Term Evolution), LTE Advanced, 5G, or the like. In some examples where output interface 108 comprises a wireless transmitter, output interface 108 and input interface 122 may be configured to transfer data, such as encoded video data, according to other wireless standards, such as an IEEE 802.11 specification, an IEEE 802.15 specification (e.g., ZigBee™), a Bluetooth™ standard, or the like. In some examples, source device 102

and/or destination device 116 may include respective system-on-a-chip (SoC) devices. For example, source device 102 may include an SoC device to perform the functionality attributed to video encoder 200 and/or output interface 108, and destination device 116 may include an SoC device to perform the functionality attributed to video decoder 300 and/or input interface 122.

[0034] The techniques of this disclosure may be applied to video coding in support of any of a variety of multimedia applications, such as over-the-air television broadcasts, cable television transmissions, satellite television transmissions, Internet streaming video transmissions, such as dynamic adaptive streaming over HTTP (DASH), digital video that is encoded onto a data storage medium, decoding of digital video stored on a data storage medium, or other applications.

[0035] Input interface 122 of destination device 116 receives an encoded video bitstream from computer-readable medium 110 (e.g., storage device 112, file server 114, or the like). The encoded video bitstream computer-readable medium 110 may include signaling information defined by video encoder 200, which is also used by video decoder 300, such as syntax elements having values that describe characteristics and/or processing of video blocks or other coded units (e.g., slices, pictures, groups of pictures, sequences, or the like). Display device 118 displays decoded pictures of the decoded video data to a user. Display device 118 may represent any of a variety of display devices such as a cathode ray tube (CRT), a liquid crystal display (LCD), a plasma display, an organic light emitting diode (OLED) display, or another type of display device.

[0036] Although not shown in FIG. 1, in some examples, video encoder 200 and video decoder 300 may each be integrated with an audio encoder and/or audio decoder, and may include appropriate MUX-DEMUX units, or other hardware and/or software, to handle multiplexed streams including both audio and video in a common data stream. If applicable, MUX-DEMUX units may conform to the ITU H.223 multiplexer protocol, or other protocols such as the user datagram protocol (UDP).

[0037] Video encoder 200 and video decoder 300 each may be implemented as any of a variety of suitable encoder and/or decoder circuitry, such as one or more microprocessors, digital signal processors (DSPs), application specific integrated circuits (ASICs), field programmable gate arrays (FPGAs), discrete logic, software, hardware, firmware or any combinations thereof. When the techniques are implemented partially in software, a device may store instructions for the software in a suitable, non-

transitory computer-readable medium and execute the instructions in hardware using one or more processors to perform the techniques of this disclosure. Each of video encoder 200 and video decoder 300 may be included in one or more encoders or decoders, either of which may be integrated as part of a combined encoder/decoder (CODEC) in a respective device. A device including video encoder 200 and/or video decoder 300 may comprise an integrated circuit, a microprocessor, and/or a wireless communication device, such as a cellular telephone.

[0038] Video encoder 200 and video decoder 300 may operate according to a video coding standard, such as ITU-T H.265, also referred to as High Efficiency Video Coding (HEVC) or extensions thereto, such as the multi-view and/or scalable video coding extensions. Alternatively, video encoder 200 and video decoder 300 may operate according to other proprietary or industry standards, such as ITU-T H.266, also referred to as Versatile Video Coding (VVC). A draft of the VVC standard is described in Bross, et al. “Versatile Video Coding (Draft 5),” Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 14th Meeting: Geneva, CH, 19-27 March 2019, JVET-N1001-v5 (hereinafter “VVC Draft 5”). A more recent draft is described in Bross, et al. “Versatile Video Coding (Draft 9),” Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 18th Meeting: by teleconference, 15-24 April 2020, JVET-R2001-v3. The techniques of this disclosure, however, are not limited to any particular coding standard.

[0039] In general, video encoder 200 and video decoder 300 may perform block-based coding of pictures. The term “block” generally refers to a structure including data to be processed (e.g., encoded, decoded, or otherwise used in the encoding and/or decoding process). For example, a block may include a two-dimensional matrix of samples of luminance and/or chrominance data. In general, video encoder 200 and video decoder 300 may code video data represented in a YUV (e.g., Y, Cb, Cr) format. That is, rather than coding red, green, and blue (RGB) data for samples of a picture, video encoder 200 and video decoder 300 may code luminance and chrominance components, where the chrominance components may include both red hue and blue hue chrominance components. In some examples, video encoder 200 converts received RGB formatted data to a YUV representation prior to encoding, and video decoder 300 converts the YUV representation to the RGB format. Alternatively, pre- and post-processing units (not shown) may perform these conversions.

[0040] This disclosure may generally refer to coding (e.g., encoding and decoding) of pictures to include the process of encoding or decoding data of the picture. Similarly, this disclosure may refer to coding of blocks of a picture to include the process of encoding or decoding data for the blocks, e.g., prediction and/or residual coding. An encoded video bitstream generally includes a series of values for syntax elements representative of coding decisions (e.g., coding modes) and partitioning of pictures into blocks. Thus, references to coding a picture or a block should generally be understood as coding values for syntax elements forming the picture or block.

[0041] HEVC defines various blocks, including coding units (CUs), prediction units (PUs), and transform units (TUs). According to HEVC, a video coder (such as video encoder 200) partitions a coding tree unit (CTU) into CUs according to a quadtree structure. That is, the video coder partitions CTUs and CUs into four equal, non-overlapping squares, and each node of the quadtree has either zero or four child nodes. Nodes without child nodes may be referred to as “leaf nodes,” and CUs of such leaf nodes may include one or more PUs and/or one or more TUs. The video coder may further partition PUs and TUs. For example, in HEVC, a residual quadtree (RQT) represents partitioning of TUs. In HEVC, PUs represent inter-prediction data, while TUs represent residual data. CUs that are intra-predicted include intra-prediction information, such as an intra-mode indication.

[0042] As another example, video encoder 200 and video decoder 300 may be configured to operate according to VVC. According to VVC, a video coder (such as video encoder 200) partitions a picture into a plurality of coding tree units (CTUs). Video encoder 200 may partition a CTU according to a tree structure, such as a quadtree-binary tree (QTBT) structure or Multi-Type Tree (MTT) structure. The QTBT structure removes the concepts of multiple partition types, such as the separation between CUs, PUs, and TUs of HEVC. A QTBT structure includes two levels: a first level partitioned according to quadtree partitioning, and a second level partitioned according to binary tree partitioning. A root node of the QTBT structure corresponds to a CTU. Leaf nodes of the binary trees correspond to coding units (CUs).

[0043] In an MTT partitioning structure, blocks may be partitioned using a quadtree (QT) partition, a binary tree (BT) partition, and one or more types of triple tree (TT) partitions. A triple tree partition is a partition where a block is split into three sub-blocks. In some examples, a triple tree partition divides a block into three sub-blocks

without dividing the original block through the center. The partitioning types in MTT (e.g., QT, BT, and TT), may be symmetrical or asymmetrical.

[0044] In some examples, video encoder 200 and video decoder 300 may use a single QTBT or MTT structure to represent each of the luminance and chrominance components, while in other examples, video encoder 200 and video decoder 300 may use two or more QTBT or MTT structures, such as one QTBT/MTT structure for the luminance component and another QTBT/MTT structure for both chrominance components (or two QTBT/MTT structures for respective chrominance components).

[0045] Video encoder 200 and video decoder 300 may be configured to use quadtree partitioning per HEVC, QTBT partitioning, MTT partitioning, or other partitioning structures. For purposes of explanation, the description of the techniques of this disclosure is presented with respect to QTBT partitioning. However, it should be understood that the techniques of this disclosure may also be applied to video coders configured to use quadtree partitioning, or other types of partitioning as well.

[0046] This disclosure may use “N×N” and “N by N” interchangeably to refer to the sample dimensions of a block (such as a CU or other video block) in terms of vertical and horizontal dimensions, e.g., 16×16 samples or 16 by 16 samples. In general, a 16×16 CU will have 16 samples in a vertical direction ($y = 16$) and 16 samples in a horizontal direction ($x = 16$). Likewise, an N×N CU generally has N samples in a vertical direction and N samples in a horizontal direction, where N represents a nonnegative integer value. The samples in a CU may be arranged in rows and columns. Moreover, CUs need not necessarily have the same number of samples in the horizontal direction as in the vertical direction. For example, CUs may comprise N×M samples, where M is not necessarily equal to N.

[0047] Video encoder 200 encodes video data for CUs representing prediction and/or residual information, and other information. The prediction information indicates how the CU is to be predicted in order to form a prediction block for the CU. The residual information generally represents sample-by-sample differences between samples of the CU prior to encoding and the prediction block.

[0048] To predict a CU, video encoder 200 may generally form a prediction block for the CU through inter-prediction or intra-prediction. Inter-prediction generally refers to predicting the CU from data of a previously coded picture, whereas intra-prediction generally refers to predicting the CU from previously coded data of the same picture. To perform inter-prediction, video encoder 200 may generate the prediction block using

one or more motion vectors. Video encoder 200 may generally perform a motion search to identify a reference block that closely matches the CU, e.g., in terms of differences between the CU and the reference block. Video encoder 200 may calculate a difference metric using a sum of absolute difference (SAD), sum of squared differences (SSD), mean absolute difference (MAD), mean squared differences (MSD), or other such difference calculations to determine whether a reference block closely matches the current CU. In some examples, video encoder 200 may predict the current CU using uni-directional prediction or bi-directional prediction.

[0049] Some examples of VVC also provide an affine motion compensation mode, which may be considered an inter-prediction mode. In affine motion compensation mode, video encoder 200 may determine two or more motion vectors that represent non-translational motion, such as zoom in or out, rotation, perspective motion, or other irregular motion types.

[0050] To perform intra-prediction, video encoder 200 may select an intra-prediction mode to generate the prediction block. Some examples of VVC provide sixty-seven intra-prediction modes, including various directional modes, as well as planar mode and DC mode. In general, video encoder 200 selects an intra-prediction mode that describes neighboring samples to a current block (e.g., a block of a CU) from which to predict samples of the current block. Such samples may generally be above, above and to the left, or to the left of the current block in the same picture as the current block, assuming video encoder 200 codes CTUs and CUs in raster scan order (left to right, top to bottom).

[0051] Video encoder 200 encodes data representing the prediction mode for a current block. For example, for inter-prediction modes, video encoder 200 may encode data representing which of the various available inter-prediction modes is used, as well as motion information for the corresponding mode. For uni-directional or bi-directional inter-prediction, for example, video encoder 200 may encode motion vectors using advanced motion vector prediction (AMVP) or merge mode. Video encoder 200 may use similar modes to encode motion vectors for affine motion compensation mode.

[0052] Following prediction, such as intra-prediction or inter-prediction of a block, video encoder 200 may calculate residual data for the block. The residual data, such as a residual block, represents sample by sample differences between the block and a prediction block for the block, formed using the corresponding prediction mode. Video encoder 200 may apply one or more transforms to the residual block, to produce

transformed data in a transform domain instead of the sample domain. For example, video encoder 200 may apply a discrete cosine transform (DCT), an integer transform, a wavelet transform, or a conceptually similar transform to residual video data.

Additionally, video encoder 200 may apply a secondary transform following the first transform, such as a mode-dependent non-separable secondary transform (MDNSST), a signal dependent transform, a Karhunen-Loeve transform (KLT), or the like. Video encoder 200 produces transform coefficients following application of the one or more transforms.

[0053] As noted above, following any transforms to produce transform coefficients, video encoder 200 may perform quantization of the transform coefficients.

Quantization generally refers to a process in which transform coefficients are quantized to possibly reduce the amount of data used to represent the coefficients, providing further compression. By performing the quantization process, video encoder 200 may reduce the bit depth associated with some or all of the coefficients. For example, video encoder 200 may round an n -bit value down to an m -bit value during quantization, where n is greater than m . In some examples, to perform quantization, video encoder 200 may perform a bitwise right-shift of the value to be quantized.

[0054] Following quantization, video encoder 200 may scan the transform coefficients, producing a one-dimensional vector from the two-dimensional matrix including the quantized transform coefficients. The scan may be designed to place higher energy (and therefore lower frequency) coefficients at the front of the vector and to place lower energy (and therefore higher frequency) transform coefficients at the back of the vector. In some examples, video encoder 200 may utilize a predefined scan order to scan the quantized transform coefficients to produce a serialized vector, and then entropy encode the quantized transform coefficients of the vector. In other examples, video encoder 200 may perform an adaptive scan. After scanning the quantized transform coefficients to form the one-dimensional vector, video encoder 200 may entropy encode the one-dimensional vector, e.g., according to context-adaptive binary arithmetic coding (CABAC). Video encoder 200 may also entropy encode values for syntax elements describing metadata associated with the encoded video data for use by video decoder 300 in decoding the video data.

[0055] To perform CABAC, video encoder 200 may assign a context within a context model to a symbol to be transmitted. The context may relate to, for example, whether

neighboring values of the symbol are zero-valued or not. The probability determination may be based on a context assigned to the symbol.

[0056] Video encoder 200 may further generate syntax data, such as block-based syntax data, picture-based syntax data, and sequence-based syntax data, to video decoder 300, e.g., in a picture header, a block header, a slice header, or other syntax data, such as a sequence parameter set (SPS), picture parameter set (PPS), or video parameter set (VPS). Video decoder 300 may likewise decode such syntax data to determine how to decode corresponding video data.

[0057] In this manner, video encoder 200 may generate a bitstream including encoded video data, e.g., syntax elements describing partitioning of a picture into blocks (e.g., CUs) and prediction and/or residual information for the blocks. Ultimately, video decoder 300 may receive the bitstream and decode the encoded video data.

[0058] In general, video decoder 300 performs a reciprocal process to that performed by video encoder 200 to decode the encoded video data of the bitstream. For example, video decoder 300 may decode values for syntax elements of the bitstream using CABAC in a manner substantially similar to, albeit reciprocal to, the CABAC encoding process of video encoder 200. The syntax elements may define partitioning information of a picture into CTUs, and partitioning of each CTU according to a corresponding partition structure, such as a QTBT structure, to define CUs of the CTU. The syntax elements may further define prediction and residual information for blocks (e.g., CUs) of video data.

[0059] The residual information may be represented by, for example, quantized transform coefficients. Video decoder 300 may inverse quantize and inverse transform the quantized transform coefficients of a block to reproduce a residual block for the block. Video decoder 300 uses a signaled prediction mode (intra- or inter-prediction) and related prediction information (e.g., motion information for inter-prediction) to form a prediction block for the block. Video decoder 300 may then combine the prediction block and the residual block (on a sample-by-sample basis) to reproduce the original block. Video decoder 300 may perform additional processing, such as performing a deblocking process to reduce visual artifacts along boundaries of the block.

[0060] In some cases, the size (e.g., resolution, actual dimensions, etc.) of a reference picture and a current picture may be different. For example, in Chen, et al. "AHG 19: Adaptive Resolution Change," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 14th Meeting: Geneva, CH, 19-27 March 2019,

JVET-N0279 (hereinafter “JVET-N0279”) reference picture resampling (RPR) is proposed. In JVET-N0279, a current picture and a reference picture may have different resolutions (e.g., height and width, or size), so during motion compensation, video encoder 200 or video decoder 300 may consider the difference in resolution during the interpolation filtering by adjusting the interpolation filter phase.

[0061] There may be some inter coding tools, especially the ones that have a derivation operation on the decoder side (e.g., video decoder 300), that use reference pictures. In JVET-N0279, the examples of such tools are temporal motion vector prediction (TMVP), decoder side motion vector refinement (DMVR), and bi-directional optical flow (BDOF). That is TMVP, DMVR, and BDOF are examples of inter-coding tools from a plurality of inter-coding tools that rely on the use of a reference picture. JVET-N0279 proposed to be disabled at least these tools at the sequence level if reference picture resampling (RPR) is allowed. The sequence level generally refers to parameters that apply to a coded video sequence (i.e., a plurality of pictures), and allows signaling of information one time for the entire sequence of pictures (e.g., the sequence parameter set (SPS)), rather than for a picture-by-picture basis. JVET-N0279 described disabling TMVP, DMVR, and BDOR coding tools, as a few examples, at the sequence level for a plurality of pictures in the sequence. As one example, JVET-N0279 may allow the disabling of these examples of a plurality of inter-coding tools at the sequence level when RPR is allowed.

[0062] In accordance with the techniques of this disclosure, video encoder 200 and video decoder 300 may be configured to utilize one or more of the plurality of inter-coding tools even where a block of the current picture has a reference picture that is of a different size than the current picture (e.g., where RPR is allowed). Accordingly, the example techniques described in this disclosure may promote video coding efficiency by allowing various inter-coding tools to be enabled for certain blocks of a current picture even where there is a block of the current picture that has a reference picture that is of a different size of the current picture. In some examples, such as with DMVR, enabling of an otherwise disabled tool may be possible by video decoder 300 without additional signaling by video encoder 200. Thus, decoded video quality may be improved without additional signaling.

[0063] This disclosure may generally refer to “signaling” certain information, such as syntax elements. The term “signaling” may generally refer to the communication of values syntax elements and/or other data used to decode encoded video data. That is,

video encoder 200 may signal values for syntax elements in the bitstream. In general, signaling refers to generating a value in the bitstream. As noted above, source device 102 may transport the bitstream to destination device 116 substantially in real time, or not in real time, such as might occur when storing syntax elements to storage device 112 for later retrieval by destination device 116.

[0064] FIGS. 2A and 2B are conceptual diagram illustrating an example quadtree binary tree (QTBT) structure 130, and a corresponding coding tree unit (CTU) 132. The solid lines represent quadtree splitting, and dotted lines indicate binary tree splitting. In each split (i.e., non-leaf) node of the binary tree, one flag is signaled to indicate which splitting type (i.e., horizontal or vertical) is used, where 0 indicates horizontal splitting and 1 indicates vertical splitting in this example. For the quadtree splitting, there is no need to indicate the splitting type, since quadtree nodes split a block horizontally and vertically into 4 sub-blocks with equal size. Accordingly, video encoder 200 may encode, and video decoder 300 may decode, syntax elements (such as splitting information) for a region tree level of QTBT structure 130 (i.e., the solid lines) and syntax elements (such as splitting information) for a prediction tree level of QTBT structure 130 (i.e., the dashed lines). Video encoder 200 may encode, and video decoder 300 may decode, video data, such as prediction and transform data, for CUs represented by terminal leaf nodes of QTBT structure 130.

[0065] In general, CTU 132 of FIG. 2B may be associated with parameters defining sizes of blocks corresponding to nodes of QTBT structure 130 at the first and second levels. These parameters may include a CTU size (representing a size of CTU 132 in samples), a minimum quadtree size (MinQTSIZE, representing a minimum allowed quadtree leaf node size), a maximum binary tree size (MaxBTSIZE, representing a maximum allowed binary tree root node size), a maximum binary tree depth (MaxBTDepth, representing a maximum allowed binary tree depth), and a minimum binary tree size (MinBTSIZE, representing the minimum allowed binary tree leaf node size).

[0066] The root node of a QTBT structure corresponding to a CTU may have four child nodes at the first level of the QTBT structure, each of which may be partitioned according to quadtree partitioning. That is, nodes of the first level are either leaf nodes (having no child nodes) or have four child nodes. The example of QTBT structure 130 represents such nodes as including the parent node and child nodes having solid lines for branches. If nodes of the first level are not larger than the maximum allowed binary

tree root node size (MaxBTSIZE), they can be further partitioned by respective binary trees. The binary tree splitting of one node can be iterated until the nodes resulting from the split reach the minimum allowed binary tree leaf node size (MinBTSIZE) or the maximum allowed binary tree depth (MaxBTDepth). The example of QTBT structure 130 represents such nodes as having dashed lines for branches. The binary tree leaf node is referred to as a coding unit (CU), which is used for prediction (e.g., intra-picture or inter-picture prediction) and transform, without any further partitioning. As discussed above, CUs may also be referred to as “video blocks” or “blocks.”

[0067] In one example of the QTBT partitioning structure, the CTU size is set as 128x128 (luma samples and two corresponding 64x64 chroma samples), the MinQTSIZE is set as 16x16, the MaxBTSIZE is set as 64x64, the MinBTSIZE (for both width and height) is set as 4, and the MaxBTDepth is set as 4. The quadtree partitioning is applied to the CTU first to generate quad-tree leaf nodes. The quadtree leaf nodes may have a size from 16x16 (i.e., the MinQTSIZE) to 128x128 (i.e., the CTU size). If the leaf quadtree node is 128x128, it will not be further split by the binary tree, since the size exceeds the MaxBTSIZE (i.e., 64x64, in this example). Otherwise, the leaf quadtree node will be further partitioned by the binary tree. Therefore, the quadtree leaf node is also the root node for the binary tree and has the binary tree depth as 0. When the binary tree depth reaches MaxBTDepth (4, in this example), no further splitting is permitted. When the binary tree node has width equal to MinBTSIZE (4, in this example), it implies no further horizontal splitting is permitted. Similarly, a binary tree node having a height equal to MinBTSIZE implies no further vertical splitting is permitted for that binary tree node. As noted above, leaf nodes of the binary tree are referred to as CUs, and are further processed according to prediction and transform without further partitioning.

[0068] In previous video standards, before the draft VVC standard, a reference picture was always the same size (e.g., dimensions or resolution) as the current picture. VVC introduced the concept of permitting reference pictures to have different sizes than the current picture. Reference picture resampling (RPR) is a technique used, e.g., by video decoder 300, to resample reference pictures stored in a decoded picture buffer that are of different sizes than the current picture. As described above, when the reference picture and the current picture have different resolutions, certain inter-coding tools have been disabled. Enabling certain inter tools, when the reference picture has a different size or resolution than the current picture, may lead to an additional implementation

complexity, such as increased bandwidth. However, when RPR is allowed, that does not mean that all blocks will use reference pictures with sizes that are different than the size of the current picture.

[0069] For example, a current picture may contain several blocks. Each of those blocks may have a corresponding reference block in a reference picture, but the reference pictures for various blocks of the current picture may be different. As such, there may be many reference pictures, some of which may have different sizes than the current picture and some of which have a same size as a current picture. However, one or more of the inter-coding tools (e.g., inter tools) described above are disabled for all cases when one block of a current picture has a reference block in a reference picture that is a different size than the current picture.

[0070] This disclosure describes one or more example technical solutions to the example technical problems. In this manner, the example techniques provide for a practical application that improves the operation of the video coding process. For instance, this disclosure describes examples of coding (e.g., encoding or decoding) one or more blocks of a current picture having one or more of a plurality of inter-coding tools selectively enabled based on a block of a current picture having a reference picture of the same size as the current picture. The example techniques described in this disclosure may promote video coding efficiency by allowing various inter-coding tools to be enabled for certain blocks even where there is a block of the current picture that has a reference picture that is of a different size of the current picture. In some examples, such as with DMVR, enabling of an otherwise disabled tool may be possible by video decoder 300 without additional signaling by video encoder 200. Thus, decoded video quality may be improved without additional signaling.

[0071] As one example, in one or more examples described in this disclosure, the TMVP, DMVR, BDOF tools and similar coding tools are disabled only for a current block when the current block uses reference picture(s) with sizes different from the current picture, and if the current block uses reference picture(s) that have the same size as the current picture, the tool usage is allowed. The reference picture to be used for inter-prediction is identified by the reference indices, which are associated with the inter coded block. In this case, some blocks in the current picture may have the tool usage disabled, but some blocks in the same current picture may use the tool. For example, video encoder 200 or video decoder 300 may identify whether the tool may be used for a given block by comparing the reference picture sizes of a given block of the current

picture to the current picture size. In some examples, video encoder 200 may utilize the techniques of this disclosure in a reconstruction loop, for example, when decoding encoded video data.

[0072] For example, the DMVR tool may be selectively enabled for the blocks of the current picture which have reference picture sizes that are the same as the current picture size and may be disabled for blocks of the current picture that use reference picture(s) with sizes different than the current picture sizes for inter prediction.

[0073] Accordingly, in one example, for coding one or more blocks, a video coder (e.g., video encoder 200 and video decoder 300) may determine a first reference block in a first reference picture for a first block of a current picture. The video coder may determine whether the first reference picture is a same size as the current picture. Based on the first reference picture being a different size than the current picture, the video coder may code (e.g., encode or decode) the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled. The video coder may also determine a second reference block in a second reference picture for a second block of the current picture. The video coder may determine whether the second reference picture is a same size as the current picture. Based on the second reference picture being the same size as the current picture, the video coder may code (e.g., encode or decode) the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0074] In some examples, the first particular tool includes one of TMVP, DMVR, and BDOF. In some examples, the techniques of this disclosure are performed based on a determination that RPR is enabled. In some examples, the size of the reference picture or the size of the current picture refers to a resolution (e.g., height and width) of the reference picture or a resolution (e.g., height and width) of the current picture.

[0075] In the case when several reference pictures are used, such as for bi-prediction, one reference picture may have the same size as the current picture and the other reference picture may have a different size. In this example, the inter-coding tool may only be disabled for the reference block of the reference picture that is a different size than the current picture. The inter-coding tool may be enabled for the reference block of the reference picture that is the same size as the current picture. For example, a video coder (e.g., video encoder 200 or video decoder 300) may determine a third reference block in a third reference picture for a third block of the current picture. The video

coder may determine a fourth reference block in a fourth reference picture for the third block of the current picture. The video coder may determine whether the third reference picture is a same size as the current picture. Based on the third reference picture being a different size than the current picture, the video coder may code (e.g., encode or decode) the third block with a third inter-coding tool of a third set of inter-coding tools, wherein a second particular tool of the plurality of inter-coding tools is disabled for the third reference block. The video coder may also determine whether the fourth reference picture is a same size as the current picture. Based on the fourth reference picture being the same size as the current picture, the video coder may code (e.g., encode or decode) the third block with a fourth inter-coding tool of a fourth set of inter-coding tools, wherein the second particular tool is enabled for the fourth reference block.

[0076] In yet another example, when using bi-prediction, the two reference pictures (e.g., one reference picture in list0 and another reference picture in list1) may have the same sizes, but have a different size than the current picture size. In this case, depending on the tool, the tool may be enabled. If some parameters are derived based on the reference pictures and those derived parameters are applied to the predicted block, which may be of different resolution, then the parameter derivation may not create a problem, since both reference pictures have the same resolution. For example, a video coder (e.g., video encoder 200 or video decoder 300) may determine a third reference block in a third reference picture for a third block. The video coder may also determine a fourth reference block in a fourth reference picture for the third block. The video coder may determine whether the third reference picture is a same size as the fourth reference picture. Based on third reference picture being the same size as the fourth reference picture, the video coder may code the third block with a third inter-coding tool of a third set of inter-coding tools, wherein the first particular tool is enabled.

[0077] Some tools, such as BDOF, may not use the reference pictures directly, but only use the result of the interpolation, which has already the same resolution ratio as a block in the current picture. In this case, there is no reason to disable such tools and, according to the techniques of this disclosure, the tools that do not use the reference pictures directly when RPR is allowed are enabled.

[0078] Accordingly, in one example, for coding one or more blocks, a video coder (e.g., video encoder 200 or video decoder 300) may determine whether the disabled first particular tool uses a result of interpolation of the first reference block and does not

directly use the first reference block. Based on the disabled first particular tool using a result of interpolation of the first reference block and not directly using the first reference block, the video coder may enable the disabled first particular tool and select the first particular tool as the first inter-coding tool and code (e.g., encode or decode) the first block with the first particular tool.

[0079] In the above examples, which may be performed separately or in any combination, one or more inter-coding tools include TMVP, DMVR, and BDOF. Also, any one or combination of the example techniques may be performed based on a determination RPR is enabled. A size of a reference picture or a size of a current picture may refer to a resolution of the reference picture or a resolution of the current picture.

[0080] In one or more examples, the decision to apply a tool on a block by block basis may be based on the following conditions solely or in any combination: 1) Whether any reference picture size is different than the current picture size; 2) Whether a reference picture of the plurality of the reference pictures has a size that is different than the current picture size, but other reference pictures have the same size as the current picture; 3) Whether reference pictures have the same size, but their size is different than the current picture size; and 4) Whether a tool does not use the reference pictures directly, as an example, where the tool may operate on the derived prediction signal. Accordingly, the example techniques described in this disclosure may promote video coding efficiency by allowing various inter-coding tools to be enabled for certain blocks of the current picture even where there is a block of the current picture that has a reference picture that is of a different size of the current picture. In some examples, such as with DMVR, enabling of an otherwise disabled tool may be possible by video decoder 300 without additional signaling by video encoder 200. Thus, decoded video quality may be improved without additional signaling.

[0081] FIG. 3 is a block diagram illustrating an example video encoder 200 that may perform the techniques of this disclosure. FIG. 3 is provided for purposes of explanation and should not be considered limiting of the techniques as broadly exemplified and described in this disclosure. For purposes of explanation, this disclosure describes video encoder 200 in the context of video coding standards such as the HEVC video coding standard and the H.266 video coding standard in development. However, the techniques of this disclosure are not limited to these video coding standards, and are applicable generally to video encoding and decoding.

[0082] In the example of FIG. 3, video encoder 200 includes video data memory 230, mode selection unit 202, residual generation unit 204, transform processing unit 206, quantization unit 208, inverse quantization unit 210, inverse transform processing unit 212, reconstruction unit 214, filter unit 216, decoded picture buffer (DPB) 218, and entropy encoding unit 220. Any or all of video data memory 230, mode selection unit 202, residual generation unit 204, transform processing unit 206, quantization unit 208, inverse quantization unit 210, inverse transform processing unit 212, reconstruction unit 214, filter unit 216, DPB 218, and entropy encoding unit 220 may be implemented in one or more processors or in processing circuitry. Moreover, video encoder 200 may include additional or alternative processors or processing circuitry to perform these and other functions.

[0083] Video data memory 230 may store video data to be encoded by the components of video encoder 200. Video encoder 200 may receive the video data stored in video data memory 230 from, for example, video source 104 (FIG. 1). DPB 218 may act as a reference picture memory that stores reference video data for use in prediction of subsequent video data by video encoder 200. Video data memory 230 and DPB 218 may be formed by any of a variety of memory devices, such as dynamic random access memory (DRAM), including synchronous DRAM (SDRAM), magnetoresistive RAM (MRAM), resistive RAM (RRAM), or other types of memory devices. Video data memory 230 and DPB 218 may be provided by the same memory device or separate memory devices. In various examples, video data memory 230 may be on-chip with other components of video encoder 200, as illustrated, or off-chip relative to those components.

[0084] In this disclosure, reference to video data memory 230 should not be interpreted as being limited to memory internal to video encoder 200, unless specifically described as such, or memory external to video encoder 200, unless specifically described as such. Rather, reference to video data memory 230 should be understood as reference memory that stores video data that video encoder 200 receives for encoding (e.g., video data for a current block that is to be encoded). Memory 106 of FIG. 1 may also provide temporary storage of outputs from the various units of video encoder 200.

[0085] The various units of FIG. 3 are illustrated to assist with understanding the operations performed by video encoder 200. The units may be implemented as fixed-function circuits, programmable circuits, or a combination thereof. Fixed-function circuits refer to circuits that provide particular functionality, and are preset on the

operations that can be performed. Programmable circuits refer to circuits that can be programmed to perform various tasks, and provide flexible functionality in the operations that can be performed. For instance, programmable circuits may execute software or firmware that cause the programmable circuits to operate in the manner defined by instructions of the software or firmware. Fixed-function circuits may execute software instructions (e.g., to receive parameters or output parameters), but the types of operations that the fixed-function circuits perform are generally immutable. In some examples, the one or more of the units may be distinct circuit blocks (fixed-function or programmable), and in some examples, the one or more units may be integrated circuits.

[0086] Video encoder 200 may include arithmetic logic units (ALUs), elementary function units (EFUs), digital circuits, analog circuits, and/or programmable cores, formed from programmable circuits. In examples where the operations of video encoder 200 are performed using software executed by the programmable circuits, memory 106 (FIG. 1) may store the object code of the software that video encoder 200 receives and executes, or another memory within video encoder 200 (not shown) may store such instructions.

[0087] Video data memory 230 is configured to store received video data. Video encoder 200 may retrieve a picture of the video data from video data memory 230 and provide the video data to residual generation unit 204 and mode selection unit 202. Video data in video data memory 230 may be raw video data that is to be encoded.

[0088] Mode selection unit 202 includes a motion estimation unit 222, motion compensation unit 224, and an intra-prediction unit 226. Mode selection unit 202 may include additional functional units to perform video prediction in accordance with other prediction modes. As examples, mode selection unit 202 may include a palette unit, an intra-block copy unit (which may be part of motion estimation unit 222 and/or motion compensation unit 224), an affine unit, a linear model (LM) unit, or the like.

[0089] Mode selection unit 202 generally coordinates multiple encoding passes to test combinations of encoding parameters and resulting rate-distortion values for such combinations. The encoding parameters may include partitioning of CTUs into CUs, prediction modes for the CUs, transform types for residual data of the CUs, quantization parameters for residual data of the CUs, and so on. Mode selection unit 202 may ultimately select the combination of encoding parameters having rate-distortion values that are better than the other tested combinations. In this manner, mode selection unit

202 may select a mode that is beneficial (e.g., provides better quality decoded video, uses less bandwidth, uses less processing power, etc.).

[0090] Video encoder 200 may partition a picture retrieved from video data memory 230 into a series of CTUs, and encapsulate one or more CTUs within a slice. Mode selection unit 202 may partition a CTU of the picture in accordance with a tree structure, such as the QTBT structure or the quad-tree structure of HEVC described above. As described above, video encoder 200 may form one or more CUs from partitioning a CTU according to the tree structure. Such a CU may also be referred to generally as a “video block” or “block.”

[0091] In general, mode selection unit 202 also controls the components thereof (e.g., motion estimation unit 222, motion compensation unit 224, and intra-prediction unit 226) to generate a prediction block for a current block (e.g., a current CU, or in HEVC, the overlapping portion of a PU and a TU). For inter-prediction of a current block, motion estimation unit 222 may perform a motion search to identify one or more closely matching reference blocks in one or more reference pictures (e.g., one or more previously coded pictures stored in DPB 218). In particular, motion estimation unit 222 may calculate a value representative of how similar a potential reference block is to the current block, e.g., according to sum of absolute difference (SAD), sum of squared differences (SSD), mean absolute difference (MAD), mean squared differences (MSD), or the like. Motion estimation unit 222 may generally perform these calculations using sample-by-sample differences between the current block and the reference block being considered. Motion estimation unit 222 may identify a reference block having a lowest value resulting from these calculations, indicating a reference block that most closely matches the current block.

[0092] Motion estimation unit 222 may form one or more motion vectors (MVs) that defines the positions of the reference blocks in the reference pictures relative to the position of the current block in a current picture. Motion estimation unit 222 may then provide the motion vectors to motion compensation unit 224. For example, for uni-directional inter-prediction, motion estimation unit 222 may provide a single motion vector, whereas for bi-directional inter-prediction, motion estimation unit 222 may provide two motion vectors. Motion compensation unit 224 may then generate a prediction block using the motion vectors.

[0093] In some examples, motion compensation unit may include size unit (SU) 223. Size unit 223 may compare the size of the reference picture pointed to by the motion

vector for the current block to the size of the current picture to determine whether the size of the reference picture for the current block is the same as the size of the current picture. In some examples, such as with a bi-predicted block, size unit 223 may compare the size of the two reference pictures for the bi-predicted block to determine whether the size of the two reference pictures are the same size. In some examples, size unit 223 may be part of motion estimation unit 222 or split between motion estimation unit 222 and motion compensation unit 224.

[0094] Motion compensation unit 224 may retrieve data of the reference block using the motion vector. As another example, if the motion vector has fractional sample precision, motion compensation unit 224 may interpolate values for the prediction block according to one or more interpolation filters. Moreover, for bi-directional inter-prediction, motion compensation unit 224 may retrieve data for two reference blocks identified by respective motion vectors and combine the retrieved data, e.g., through sample-by-sample averaging or weighted averaging.

[0095] As another example, for intra-prediction, or intra-prediction coding, intra-prediction unit 226 may generate the prediction block from samples neighboring the current block. For example, for directional modes, intra-prediction unit 226 may generally mathematically combine values of neighboring samples and populate these calculated values in the defined direction across the current block to produce the prediction block. As another example, for DC mode, intra-prediction unit 226 may calculate an average of the neighboring samples to the current block and generate the prediction block to include this resulting average for each sample of the prediction block.

[0096] Mode selection unit 202 provides the prediction block to residual generation unit 204. Residual generation unit 204 receives a raw, unencoded version of the current block from video data memory 230 and the prediction block from mode selection unit 202. Residual generation unit 204 calculates sample-by-sample differences between the current block and the prediction block. The resulting sample-by-sample differences define a residual block for the current block. In some examples, residual generation unit 204 may also determine differences between sample values in the residual block to generate a residual block using residual differential pulse code modulation (RDPCM). In some examples, residual generation unit 204 may be formed using one or more subtractor circuits that perform binary subtraction.

[0097] In examples where mode selection unit 202 partitions CUs into PUs, each PU may be associated with a luma prediction unit and corresponding chroma prediction units. Video encoder 200 and video decoder 300 may support PUs having various sizes. As indicated above, the size of a CU may refer to the size of the luma coding block of the CU and the size of a PU may refer to the size of a luma prediction unit of the PU. Assuming that the size of a particular CU is $2N \times 2N$, video encoder 200 may support PU sizes of $2N \times 2N$ or $N \times N$ for intra prediction, and symmetric PU sizes of $2N \times 2N$, $2N \times N$, $N \times 2N$, $N \times N$, or similar for inter prediction. Video encoder 200 and video decoder 300 may also support asymmetric partitioning for PU sizes of $2N \times nU$, $2N \times nD$, $nL \times 2N$, and $nR \times 2N$ for inter prediction.

[0098] In examples where mode selection unit does not further partition a CU into PUs, each CU may be associated with a luma coding block and corresponding chroma coding blocks. As above, the size of a CU may refer to the size of the luma coding block of the CU. Video encoder 200 and video decoder 300 may support CU sizes of $2N \times 2N$, $2N \times N$, or $N \times 2N$.

[0099] For other video coding techniques such as an intra-block copy mode coding, an affine-mode coding, and linear model (LM) mode coding, as a few examples, mode selection unit 202, via respective units associated with the coding techniques, generates a prediction block for the current block being encoded. In some examples, such as palette mode coding, mode selection unit 202 may not generate a prediction block, and instead generate syntax elements that indicate the manner in which to reconstruct the block based on a selected palette. In such modes, mode selection unit 202 may provide these syntax elements to entropy encoding unit 220 to be encoded.

[0100] As described above, residual generation unit 204 receives the video data for the current block and the corresponding prediction block. Residual generation unit 204 then generates a residual block for the current block. To generate the residual block, residual generation unit 204 calculates sample-by-sample differences between the prediction block and the current block.

[0101] Transform processing unit 206 applies one or more transforms to the residual block to generate a block of transform coefficients (referred to herein as a “transform coefficient block”). Transform processing unit 206 may apply various transforms to a residual block to form the transform coefficient block. For example, transform processing unit 206 may apply a discrete cosine transform (DCT), a directional transform, a Karhunen-Loeve transform (KLT), or a conceptually similar transform to a

residual block. In some examples, transform processing unit 206 may perform multiple transforms to a residual block, e.g., a primary transform and a secondary transform, such as a rotational transform. In some examples, transform processing unit 206 does not apply transforms to a residual block.

[0102] Quantization unit 208 may quantize the transform coefficients in a transform coefficient block, to produce a quantized transform coefficient block. Quantization unit 208 may quantize transform coefficients of a transform coefficient block according to a quantization parameter (QP) value associated with the current block. Video encoder 200 (e.g., via mode selection unit 202) may adjust the degree of quantization applied to the coefficient blocks associated with the current block by adjusting the QP value associated with the CU. Quantization may introduce loss of information, and thus, quantized transform coefficients may have lower precision than the original transform coefficients produced by transform processing unit 206.

[0103] Inverse quantization unit 210 and inverse transform processing unit 212 may apply inverse quantization and inverse transforms to a quantized transform coefficient block, respectively, to reconstruct a residual block from the transform coefficient block. Reconstruction unit 214 may produce a reconstructed block corresponding to the current block (albeit potentially with some degree of distortion) based on the reconstructed residual block and a prediction block generated by mode selection unit 202. For example, reconstruction unit 214 may add samples of the reconstructed residual block to corresponding samples from the prediction block generated by mode selection unit 202 to produce the reconstructed block.

[0104] Filter unit 216 may perform one or more filter operations on reconstructed blocks. For example, filter unit 216 may perform deblocking operations to reduce blockiness artifacts along edges of CUs. Operations of filter unit 216 may be skipped, in some examples.

[0105] Video encoder 200 stores reconstructed blocks in DPB 218. For instance, in examples where operations of filter unit 216 are not needed, reconstruction unit 214 may store reconstructed blocks to DPB 218. In examples where operations of filter unit 216 are needed, filter unit 216 may store the filtered reconstructed blocks to DPB 218. Motion estimation unit 222 and motion compensation unit 224 may retrieve a reference picture from DPB 218, formed from the reconstructed (and potentially filtered) blocks, to inter-predict blocks of subsequently encoded pictures. In addition, intra-prediction

unit 226 may use reconstructed blocks in DPB 218 of a current picture to intra-predict other blocks in the current picture.

[0106] In general, entropy encoding unit 220 may entropy encode syntax elements received from other functional components of video encoder 200. For example, entropy encoding unit 220 may entropy encode quantized transform coefficient blocks from quantization unit 208. As another example, entropy encoding unit 220 may entropy encode prediction syntax elements (e.g., motion information for inter-prediction or intra-mode information for intra-prediction) from mode selection unit 202. Entropy encoding unit 220 may perform one or more entropy encoding operations on the syntax elements, which are another example of video data, to generate entropy-encoded data. For example, entropy encoding unit 220 may perform a context-adaptive variable length coding (CAVLC) operation, a CABAC operation, a variable-to-variable (V2V) length coding operation, a syntax-based context-adaptive binary arithmetic coding (SBAC) operation, a Probability Interval Partitioning Entropy (PIPE) coding operation, an Exponential-Golomb encoding operation, or another type of entropy encoding operation on the data. In some examples, entropy encoding unit 220 may operate in bypass mode where syntax elements are not entropy encoded.

[0107] Video encoder 200 may output a bitstream that includes the entropy encoded syntax elements needed to reconstruct blocks of a slice or picture. In particular, entropy encoding unit 220 may output the bitstream.

[0108] The operations described above are described with respect to a block. Such description should be understood as being operations for a luma coding block and/or chroma coding blocks. As described above, in some examples, the luma coding block and chroma coding blocks are luma and chroma components of a CU. In some examples, the luma coding block and the chroma coding blocks are luma and chroma components of a PU.

[0109] In some examples, operations performed with respect to a luma coding block need not be repeated for the chroma coding blocks. As one example, operations to identify a motion vector (MV) and reference picture for a luma coding block need not be repeated for identifying a MV and reference picture for the chroma blocks. Rather, the MV for the luma coding block may be scaled to determine the MV for the chroma blocks, and the reference picture may be the same. As another example, the intra-prediction process may be the same for the luma coding blocks and the chroma coding blocks.

[0110] Video encoder 200 represents an example of a device configured to encode video data including a memory configured to store video data, and one or more processing units implemented in circuitry and configured to determine a first reference block in a first reference picture for a first block of a current picture; determine whether the first reference picture is a same size as the current picture; based on the first reference picture being a different size than the current picture, code the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled; determine a second reference block in a second reference picture for a second block of the current picture; determine whether the second reference picture is a same size as the current picture; and based on the second reference picture being the same size as the current picture, code the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0111] FIG. 4 is a block diagram illustrating an example video decoder 300 that may perform the techniques of this disclosure. FIG. 4 is provided for purposes of explanation and is not limiting on the techniques as broadly exemplified and described in this disclosure. For purposes of explanation, this disclosure describes video decoder 300 is described according to the techniques of VVC and HEVC. However, the techniques of this disclosure may be performed by video coding devices that are configured to other video coding standards.

[0112] In the example of FIG. 4, video decoder 300 includes coded picture buffer (CPB) memory 320, entropy decoding unit 302, prediction processing unit 304, inverse quantization unit 306, inverse transform processing unit 308, reconstruction unit 310, filter unit 312, and decoded picture buffer (DPB) 314. Any or all of CPB memory 320, entropy decoding unit 302, prediction processing unit 304, inverse quantization unit 306, inverse transform processing unit 308, reconstruction unit 310, filter unit 312, and DPB 314 may be implemented in one or more processors or in processing circuitry. Moreover, video decoder 300 may include additional or alternative processors or processing circuitry to perform these and other functions.

[0113] Prediction processing unit 304 includes motion compensation unit 316 and intra-prediction unit 318. Prediction processing unit 304 may include addition units to perform prediction in accordance with other prediction modes. As examples, prediction processing unit 304 may include a palette unit, an intra-block copy unit (which may form part of motion compensation unit 316), an affine unit, a linear model (LM) unit, or

the like. In other examples, video decoder 300 may include more, fewer, or different functional components.

[0114] CPB memory 320 may store video data, such as an encoded video bitstream, to be decoded by the components of video decoder 300. The video data stored in CPB memory 320 may be obtained, for example, from computer-readable medium 110 (FIG. 1). CPB memory 320 may include a CPB that stores encoded video data (e.g., syntax elements) from an encoded video bitstream. Also, CPB memory 320 may store video data other than syntax elements of a coded picture, such as temporary data representing outputs from the various units of video decoder 300. DPB 314 generally stores decoded pictures, which video decoder 300 may output and/or use as reference video data when decoding subsequent data or pictures of the encoded video bitstream. CPB memory 320 and DPB 314 may be formed by any of a variety of memory devices, such as dynamic random access memory (DRAM), including synchronous DRAM (SDRAM), magnetoresistive RAM (MRAM), resistive RAM (RRAM), or other types of memory devices. CPB memory 320 and DPB 314 may be provided by the same memory device or separate memory devices. In various examples, CPB memory 320 may be on-chip with other components of video decoder 300, or off-chip relative to those components. [0115] Additionally or alternatively, in some examples, video decoder 300 may retrieve coded video data from memory 120 (FIG. 1). That is, memory 120 may store data as discussed above with CPB memory 320. Likewise, memory 120 may store instructions to be executed by video decoder 300, when some or all of the functionality of video decoder 300 is implemented in software to be executed by processing circuitry of video decoder 300.

[0116] The various units shown in FIG. 4 are illustrated to assist with understanding the operations performed by video decoder 300. The units may be implemented as fixed-function circuits, programmable circuits, or a combination thereof. Similar to FIG. 3, fixed-function circuits refer to circuits that provide particular functionality, and are preset on the operations that can be performed. Programmable circuits refer to circuits that can be programmed to perform various tasks, and provide flexible functionality in the operations that can be performed. For instance, programmable circuits may execute software or firmware that cause the programmable circuits to operate in the manner defined by instructions of the software or firmware. Fixed-function circuits may execute software instructions (e.g., to receive parameters or output parameters), but the types of operations that the fixed-function circuits perform are generally immutable. In

some examples, the one or more of the units may be distinct circuit blocks (fixed-function or programmable), and in some examples, the one or more units may be integrated circuits.

[0117] Video decoder 300 may include ALUs, EFUs, digital circuits, analog circuits, and/or programmable cores formed from programmable circuits. In examples where the operations of video decoder 300 are performed by software executing on the programmable circuits, on-chip or off-chip memory may store instructions (e.g., object code) of the software that video decoder 300 receives and executes.

[0118] Entropy decoding unit 302 may receive encoded video data from the CPB and entropy decode the video data to reproduce syntax elements. Prediction processing unit 304, inverse quantization unit 306, inverse transform processing unit 308, reconstruction unit 310, and filter unit 312 may generate decoded video data based on the syntax elements extracted from the bitstream.

[0119] In general, video decoder 300 reconstructs a picture on a block-by-block basis. Video decoder 300 may perform a reconstruction operation on each block individually (where the block currently being reconstructed, i.e., decoded, may be referred to as a “current block”).

[0120] Entropy decoding unit 302 may entropy decode syntax elements defining quantized transform coefficients of a quantized transform coefficient block, as well as transform information, such as a quantization parameter (QP) and/or transform mode indication(s). Inverse quantization unit 306 may use the QP associated with the quantized transform coefficient block to determine a degree of quantization and, likewise, a degree of inverse quantization for inverse quantization unit 306 to apply. Inverse quantization unit 306 may, for example, perform a bitwise left-shift operation to inverse quantize the quantized transform coefficients. Inverse quantization unit 306 may thereby form a transform coefficient block including transform coefficients.

[0121] After inverse quantization unit 306 forms the transform coefficient block, inverse transform processing unit 308 may apply one or more inverse transforms to the transform coefficient block to generate a residual block associated with the current block. For example, inverse transform processing unit 308 may apply an inverse DCT, an inverse integer transform, an inverse Karhunen-Loeve transform (KLT), an inverse rotational transform, an inverse directional transform, or another inverse transform to the coefficient block.

[0122] Furthermore, prediction processing unit 304 generates a prediction block according to prediction information syntax elements that were entropy decoded by entropy decoding unit 302. For example, if the prediction information syntax elements indicate that the current block is inter-predicted, motion compensation unit 316 may generate the prediction block. In this case, the prediction information syntax elements may indicate a reference picture in DPB 314 from which to retrieve a reference block, as well as a motion vector identifying a location of the reference block in the reference picture relative to the location of the current block in the current picture. Motion compensation unit 316 may generally perform the inter-prediction process in a manner that is substantially similar to that described with respect to motion compensation unit 224 (FIG. 3).

[0123] For example, motion compensation unit 316 may include size unit (SU) 315. Size unit 315 may compare the size of the reference picture pointed to by the motion vector for the current block to the size of the current picture to determine whether the size of the reference picture for the current block is the same as the size of the current picture. In some examples, such as with a bi-predicted block, size unit 315 may compare the size of the two reference pictures for the bi-predicted block to determine whether the size of the two reference pictures are the same size.

[0124] As another example, if the prediction information syntax elements indicate that the current block is intra-predicted, intra-prediction unit 318 may generate the prediction block according to an intra-prediction mode indicated by the prediction information syntax elements. Again, intra-prediction unit 318 may generally perform the intra-prediction process in a manner that is substantially similar to that described with respect to intra-prediction unit 226 (FIG. 3). Intra-prediction unit 318 may retrieve data of neighboring samples to the current block from DPB 314.

[0125] Reconstruction unit 310 may reconstruct the current block using the prediction block and the residual block. For example, reconstruction unit 310 may add samples of the residual block to corresponding samples of the prediction block to reconstruct the current block.

[0126] Filter unit 312 may perform one or more filter operations on reconstructed blocks. For example, filter unit 312 may perform deblocking operations to reduce blockiness artifacts along edges of the reconstructed blocks. Operations of filter unit 312 are not necessarily performed in all examples.

[0127] Video decoder 300 may store the reconstructed blocks in DPB 314. As discussed above, DPB 314 may provide reference information, such as samples of a current picture for intra-prediction and previously decoded pictures for subsequent motion compensation, to prediction processing unit 304. Moreover, video decoder 300 may output decoded pictures from DPB for subsequent presentation on a display device, such as display device 118 of FIG. 1.

[0128] In this manner, video decoder 300 represents an example of a video decoding device including a memory configured to store video data, and one or more processing units implemented in circuitry and configured to determine a first reference block in a first reference picture for a first block of a current picture; determine whether the first reference picture is a same size as the current picture; based on the first reference picture being a different size than the current picture, code the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled; determine a second reference block in a second reference picture for a second block of the current picture; determine whether the second reference picture is a same size as the current picture; and based on the second reference picture being the same size as the current picture, code the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled.

[0129] FIG. 5 is a flowchart illustrating an example of selectively enabling use of coding tools according to the techniques of this disclosure. In the example of FIG. 5, motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may determine a first reference block in a first reference picture for a first block of a current picture (322). For example, motion estimation unit 222 may perform a motion search to identify one or more closely matching reference blocks to the first block in one or more reference pictures stored in DPB 218 as discussed with respect to FIG. 3. Motion compensation unit 316 may receive prediction information syntax elements that indicate a reference picture in DPB 314 as discussed with respect to FIG. 4.

[0130] Size unit 223 of video encoder 200 or size unit 315 of video decoder 300 may determine whether the first reference picture is a same size as the current picture (324). For example, size unit 223 or size unit 315 may compare the dimensions (e.g., height, width) or resolution of the first reference picture to the current picture. Based on the first reference picture being a different size than the current picture, video encoder 200

or video decoder 300 may code the first block with a first inter-coding tool of a first set of inter-coding tools, wherein a first particular tool of a plurality of inter-coding tools is disabled (326). For example, video encoder 200 or video decoder 300 may not use the first particular tool to code the first block as the first particular tool is disabled. In some examples, video encoder 200 (e.g., motion estimation unit 222) or video decoder 300 (e.g., motion compensation unit 316) may disable the first particular tool of the plurality of inter-coding tools based on the first reference picture being a different size than the current picture.

[0131] Motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may also determine a second reference block in a second reference picture for a second block of the current picture (328). For example, motion estimation unit 222 may perform a motion search to identify one or more closely matching reference blocks in one or more reference pictures stored in DPB 218 as discussed with respect to FIG. 3. Motion compensation unit 316 may receive prediction information syntax elements that indicate a reference picture in DPB 314 as discussed with respect to FIG. 4.

[0132] Size unit 223 of video encoder 200 or size unit 315 of video decoder 300 may determine whether the second reference picture is a same size as the current picture (330). For example, size unit 223 or size unit 315 may compare the dimensions (e.g., height, width) or resolution of the first reference picture to the current picture. Based on the second reference picture being the same size as the current picture, video encoder 200 or video decoder 300 may code the second block with a second inter-coding tool of a second set of inter-coding tools, wherein the first particular tool is enabled (332). For example, video encoder 200 or video decoder 300 may use the first particular tool to code the second block as the first particular tool is enabled. In some examples, video encoder 200 (e.g., motion estimation unit 222) or video decoder 300 (e.g., motion compensation unit 316) may enable the first particular tool of the plurality of inter-coding tools based on the second reference picture being the same size as the current picture.

[0133] In some examples, video encoder 200 may determine whether application of the first particular tool would be beneficial (e.g., provide better quality decoded video, use less bandwidth, use less processing power, etc.) for the second block. Based on the application of the first particular tool being beneficial for the second block, video

encoder 200 may select the first particular tool as the second inter-coding tool and thus code the second block using the first particular tool.

[0134] In some examples, the first set of inter-coding tools is the same as the second set of inter-coding tools. In some examples, the first set of inter-coding tools is different than the second set of inter coding tools. In some examples, the first set of inter-coding tools does not include the first particular tool. In some examples, the second set of inter-coding tools does include the first particular tool. In some examples, the first inter-coding tool is the same as the second inter-coding tool. In some examples, the first inter-coding tool is different than the second inter-coding tool.

[0135] FIG. 6 is a flowchart illustrating another example of selectively enabling use of coding tools according to the techniques of this disclosure. For example, the techniques of FIG. 6 may be used with the techniques of FIG. 5 for a bi-predicted block. In the example of FIG. 6, motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may determine a third reference block in a third reference picture for a third block (334). For example, motion estimation unit 222 may perform a motion search to identify one or more closely matching reference blocks to the third block in one or more reference pictures stored in DPB 218 as discussed with respect to FIG. 3. Motion compensation unit 316 may receive prediction information syntax elements that indicate a reference picture in DPB 314 as discussed with respect to FIG. 4. Motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may also determine a fourth reference block in a fourth reference picture for the third block (336).

[0136] Size unit 223 of video encoder 200 or size unit 317 of video decoder 300 may determine whether the third reference picture is a same size as the fourth reference picture (338). For example, size unit 223 or size unit 317 may compare the dimensions (e.g., height, width) or resolution of the third reference picture to the fourth reference picture.

[0137] Based on third reference picture being the same size as the fourth reference picture, video encoder 200 or video decoder 300 may code the third block with a third inter-coding tool of a third set of inter-coding tools, wherein the first particular tool is enabled (340). In some examples, video encoder 200 or video decoder 300 may use the first particular tool to code the third block as the first particular tool is enabled. In some examples, video encoder 200 (e.g., motion estimation unit 222) or video decoder 300 (e.g., motion compensation unit 316) may enable the first particular tool of the plurality

of inter-coding tools based on the third reference picture being the same size as the fourth picture.

[0138] In some examples, video encoder 200 may determine whether application of the first particular tool would be beneficial (e.g., provide better quality decoded video, use less bandwidth, use less processing power, etc.) for the third block. Based on the application of the first particular tool being beneficial for the third block, video encoder 200 may select the first particular tool as the third inter-coding tool and thus code the third block using the first particular tool.

[0139] FIG. 7 is a flowchart illustrating another example of selectively enabling use of coding tools according to the techniques of this disclosure. For example, the techniques of FIG. 7 may be used with a bi-predicted block and may be used with the techniques of any of FIGS. 5 and 6 or may be used on their own. In the example of FIG. 7, motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may determine a third reference block in a third reference picture for a third block of the current picture (382). For example, motion estimation unit 222 may perform a motion search to identify one or more closely matching reference blocks to the third block in one or more reference pictures stored in DPB 218 as discussed with respect to FIG. 3. Motion compensation unit 316 may receive prediction information syntax elements that indicate a reference picture in DPB 314 as discussed with respect to FIG. 4. Motion estimation unit 222 of video encoder 200 or motion compensation unit 316 of video decoder 300 may also determine a fourth reference block in a fourth reference picture for the third block of the current picture (384).

[0140] Size unit 223 of video encoder 200 or size unit 317 of video decoder 300 may determine whether the third reference picture is a same size as the current picture (386). For example, size unit 223 or size unit 317 may compare the dimensions (e.g., height, width) or resolution of the third reference picture to the current picture.

[0141] Based on the third reference picture being a different size than the current picture, video encoder 200 or video decoder 300 may code the third block with a third inter-coding tool of a third set of inter-coding tools, wherein a second particular tool of the plurality of inter-coding tools is disabled for the third reference block (388). For example, video encoder 200 or video decoder 300 may not use the second particular tool to code the third block using the third reference block as the second particular tool is disabled. In some examples, video encoder 200 (e.g., motion estimation unit 222) or video decoder 300 (e.g., motion compensation unit 316) may disable the second

particular tool of the plurality of inter-coding tools based on the third reference picture being a different size than the current picture.

[0142] Size unit 223 of video encoder 200 or size unit 317 of video decoder 300 may also determine whether the fourth reference picture is a same size as the current picture (390). For example, size unit 223 or size unit 317 may compare the dimensions (e.g., height, width) or resolution of the fourth reference picture to the current picture.

[0143] Based on the fourth reference picture being the same size as the current picture, video encoder 200 or video decoder 300 may code the third block with a fourth inter-coding tool of a fourth set of inter-coding tools, wherein the second particular tool is enabled for the fourth reference block (392). For example, video encoder 200 or video decoder 300 may use the second particular tool to code the third block using the fourth reference block as the first particular tool is enabled. In some examples, video encoder 200 (e.g., motion estimation unit 222) or video decoder 300 (e.g., motion compensation unit 316) may enable the second particular tool of the plurality of inter-coding tools based on the fourth reference picture being the same size as the current picture.

[0144] In some examples, video encoder 200 may determine whether application of the second particular tool would be beneficial (e.g., provide better quality decoded video, use less bandwidth, use less processing power, etc.) for the third block. Based on the application of the second particular tool being beneficial for the third block, video encoder 200 may select the second particular tool as the fourth inter-coding tool and thus code the third block using the second particular tool.

[0145] In some examples, the third set of inter-coding tools is the same as the fourth set of inter-coding tools. In some examples, the third set of inter-coding tools is different than the fourth set of inter-coding tools. In some examples, the third set of inter-coding tools does not include the second particular tool. In some examples, the fourth set of inter-coding tools does include the second particular tool. In some examples, the third inter-coding tool is the same as the fourth inter-coding tool. In some examples, the third inter-coding tool is different than the fourth inter-coding tool. It should also be understood that the sets of inter-coding tools, the inter-coding tools and the particular coding tool of FIG. 7 may be the same as or different than those of FIGS. 5-6.

[0146] FIG. 8 is a flowchart illustrating an example method for encoding a current block. The current block may comprise a current CU. Although described with respect to video encoder 200 (FIGS. 1 and 3), it should be understood that other devices may be configured to perform a method similar to that of FIG. 8.

[0147] In this example, video encoder 200 initially predicts the current block (350). For example, video encoder 200 may form a prediction block for the current block. While and/or after forming the prediction block for the current block, video encoder 200 may perform any of the techniques set forth in FIGS. 5-7. Video encoder 200 may then calculate a residual block for the current block (352). To calculate the residual block, video encoder 200 may calculate a difference between the original, unencoded block and the prediction block for the current block. Video encoder 200 may then transform and quantize coefficients of the residual block (354). Next, video encoder 200 may scan the quantized transform coefficients of the residual block (356). During the scan, or following the scan, video encoder 200 may entropy encode the coefficients (358). For example, video encoder 200 may encode the coefficients using CAVLC or CABAC. Video encoder 200 may then output the entropy coded data of the block (360).

[0148] FIG. 9 is a flowchart illustrating an example method for decoding a current block of video data. The current block may comprise a current CU. Although described with respect to video decoder 300 (FIGS. 1 and 4), it should be understood that other devices may be configured to perform a method similar to that of FIG. 9.

[0149] Video decoder 300 may receive entropy coded data for the current block, such as entropy coded prediction information and entropy coded data for coefficients of a residual block corresponding to the current block (370). Video decoder 300 may entropy decode the entropy coded data to determine prediction information for the current block and to reproduce coefficients of the residual block (372). Video decoder 300 may predict the current block (374), e.g., using an intra- or inter-prediction mode as indicated by the prediction information for the current block, to calculate a prediction block for the current block. During determining prediction information, during predicting the current block and/or thereafter, video decoder 300 may perform any of the techniques of FIGS. 5-7. Video decoder 300 may then inverse scan the reproduced coefficients (376), to create a block of quantized transform coefficients. Video decoder 300 may then inverse quantize and inverse transform the coefficients to produce a residual block (378). Video decoder 300 may ultimately decode the current block by combining the prediction block and the residual block (380).

[0150] The present disclosure includes the following examples.

[0151] Example 1. A method of coding video data, the method comprising: coding one or more blocks of a current picture having one or more of a plurality of inter-coding

tools selectively enabled based on a current picture and a reference picture having different sizes.

[0152] Example 2. The method of example 1, wherein coding the one or more blocks comprises: determining a first reference block in a first reference picture for a first block of a current picture; based on a size of the first reference picture being different than a size of the current picture, coding the first block with having a particular tool of one or more of a plurality of inter-coding tools disabled; determining a second reference block in a second reference picture for a second block of the same current picture; and based on a size of the second reference picture being same as the size of the current picture, coding the second block with having the particular tool of the one or more of the plurality of inter-coding tools enabled.

[0153] Example 3. The method of example 1, wherein coding the one or more blocks comprises: determining a first reference block in a first reference picture for a current block of a current picture; determining a second reference block in a second reference picture for a current block; and based on a size of the first reference picture being different than a size of the current picture and a size of the second reference picture being the same as the size of the current picture, coding the current block with having a particular tool of one or more of a plurality of inter-coding tools disabled for the first reference block and having the particular tool of the one or more of the plurality of inter-coding tools enabled for the second reference block.

[0154] Example 4. The method of example 1, wherein coding the one or more blocks comprises: determining a first reference block in a first reference picture for a current block of a current picture; determining a second reference block in a second reference picture for a current block; and based on sizes of the first reference picture and the second reference being the same as one another and being different than a size of the current picture, coding the current block with a particular tool of the one or more of a plurality of inter-coding tools being enabled, wherein the particular tool would have been disabled if first and second reference pictures had different sizes.

[0155] Example 5. The method of example 1, wherein coding the one or more blocks comprises: determining that a reference picture has a different size than a current picture; and based on an inter-coding tool using a result of interpolation of a reference block of the reference picture and not directly using the reference picture, coding a current block of a current picture with the inter-coding tool being enabled.

[0156] Example 6. The method of any of examples 1–5, wherein the particular tool includes one of temporal motion vector prediction (TMVP), decode side motion vector refinement (DMVR), and bi-directional optical flow (BDOF).

[0157] Example 7. The method of any of examples 1–6, wherein the method is performed based on a determination that reference picture resampling (RPR) is enabled.

[0158] Example 8. The method of any of examples 1–7, wherein a size of a reference picture or a size of a current picture refers to a resolution of the reference picture or a resolution of the current picture.

[0159] Example 9. A method comprising any one or combination of examples 1–8.

[0160] Example 10. The method of example 9, wherein coding comprises decoding.

[0161] Example 11. The method of example 9, wherein coding comprises encoding.

[0162] Example 12. A device for coding video data, the device comprising: a memory configured to store video data; and a video coder comprising at least one of programmable or fixed-function circuitry, wherein the video coder is configured to perform the method of any one or combination of examples 1–9.

[0163] Example 13. The device of example 12, wherein the video coder comprises a video decoder.

[0164] Example 14. The device of example 12, wherein the video coder comprises a video encoder.

[0165] Example 15. The device of any of examples 12–14, wherein the device further comprises one or more of a display to display the tiles and a camera to capture pictures that includes the tiles.

[0166] Example 16. The device of any of examples 12–15, wherein the device comprises one or more of a computer, a mobile device, a broadcast receiver device, or a set-top box.

[0167] Example 17. A computer-readable storage medium having stored thereon instructions that, when executed, cause one or more processors to perform the method of any of examples 1–11.

[0168] Example 18. A device for coding video data, the device comprising means for performing the method of any of examples 1–11.

[0169] It is to be recognized that depending on the example, certain acts or events of any of the techniques described herein can be performed in a different sequence, may be added, merged, or left out altogether (e.g., not all described acts or events are necessary for the practice of the techniques). Moreover, in certain examples, acts or events may

be performed concurrently, e.g., through multi-threaded processing, interrupt processing, or multiple processors, rather than sequentially.

[0170] In one or more examples, the functions described may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on or transmitted over as one or more instructions or code on a computer-readable medium and executed by a hardware-based processing unit. Computer-readable media may include computer-readable storage media, which corresponds to a tangible medium such as data storage media, or communication media including any medium that facilitates transfer of a computer program from one place to another, e.g., according to a communication protocol. In this manner, computer-readable media generally may correspond to (1) tangible computer-readable storage media which is non-transitory or (2) a communication medium such as a signal or carrier wave. Data storage media may be any available media that can be accessed by one or more computers or one or more processors to retrieve instructions, code and/or data structures for implementation of the techniques described in this disclosure. A computer program product may include a computer-readable medium.

[0171] By way of example, and not limitation, such computer-readable storage media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage, or other magnetic storage devices, flash memory, or any other medium that can be used to store desired program code in the form of instructions or data structures and that can be accessed by a computer. Also, any connection is properly termed a computer-readable medium. For example, if instructions are transmitted from a website, server, or other remote source using a coaxial cable, fiber optic cable, twisted pair, digital subscriber line (DSL), or wireless technologies such as infrared, radio, and microwave, then the coaxial cable, fiber optic cable, twisted pair, DSL, or wireless technologies such as infrared, radio, and microwave are included in the definition of medium. It should be understood, however, that computer-readable storage media and data storage media do not include connections, carrier waves, signals, or other transitory media, but are instead directed to non-transitory, tangible storage media. Disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and Blu-ray disc, where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Combinations of the above should also be included within the scope of computer-readable media.

[0172] Instructions may be executed by one or more processors, such as one or more digital signal processors (DSPs), general purpose microprocessors, application specific integrated circuits (ASICs), field programmable gate arrays (FPGAs), or other equivalent integrated or discrete logic circuitry. Accordingly, the terms “processor” and “processing circuitry,” as used herein may refer to any of the foregoing structures or any other structure suitable for implementation of the techniques described herein. In addition, in some aspects, the functionality described herein may be provided within dedicated hardware and/or software modules configured for encoding and decoding, or incorporated in a combined codec. Also, the techniques could be fully implemented in one or more circuits or logic elements.

[0173] The techniques of this disclosure may be implemented in a wide variety of devices or apparatuses, including a wireless handset, an integrated circuit (IC) or a set of ICs (e.g., a chip set). Various components, modules, or units are described in this disclosure to emphasize functional aspects of devices configured to perform the disclosed techniques, but do not necessarily require realization by different hardware units. Rather, as described above, various units may be combined in a codec hardware unit or provided by a collection of interoperative hardware units, including one or more processors as described above, in conjunction with suitable software and/or firmware.

[0174] Various examples have been described. These and other examples are within the scope of the following claims.

CLAIMS

1. A method of coding video data, the method comprising:
for a first block of a current picture, determining one or more reference blocks in one or more first directly referenced pictures;
determining whether the one or more first directly referenced pictures has the same size as the current picture;
based on one or more of the one or more first directly referenced pictures being a different size than the current picture, coding the first block with an inter-coding tool of a set of inter-coding tools, wherein temporal motion vector prediction (TMVP), decoder side motion vector refinement (DMVR), and bi-directional optical flow (BDOF) is disabled for coding the first block;
for a second block of the current picture, determining one or more reference blocks in one or more second directly referenced pictures;
determining whether the one or more second directly referenced pictures has the same size as the current picture; and
based on the one or more second directly referenced pictures being the same size as the current picture, coding the second block with an inter-coding tool of the set of inter-coding tools, wherein the TMVP, DMVR, and BDOF disabled for the coding of the first block is enabled for the coding of the second block.
2. The method of claim 1, wherein at least one of the first block and the second block are bi-directionally coded.
3. The method of claim 1, wherein the method is performed based on a determination that reference picture resampling (RPR) is enabled.
4. The method of claim 1, wherein size comprises a resolution.
5. A device for coding video data, the device comprising: a memory configured to store video data; and one or more processors implemented in circuitry and coupled to the memory, the one or more processors being configured to perform the method of any one of claims 1-4.

6. The device of claim 5, comprising a wireless phone.
7. A non-transitory computer-readable medium having stored thereon instructions that, when executed, cause one or more processors to perform the method of any one of claims 1-4.
8. A device for coding video data, the device comprising:
 - means for determining, for a first block of a current picture, one or more reference blocks in one or more first directly referenced pictures;
 - means for determining whether the one or more first directly referenced pictures has the same size as the current picture;
 - means for coding the first block with an inter-coding tool of a set of inter-coding tools, wherein based on at least one of the first directly referenced pictures being a different size than the current picture, at least one of temporal motion vector prediction (TMVP), decoder side motion vector refinement (DMVR), and bi-directional optical flow (BDOF) is disabled for coding the first block;
 - means for determining, for a second block of the current picture, one or more reference blocks in one or more second directly referenced pictures;
 - means for determining whether the one or more second directly referenced pictures has the same size as the current picture; and
 - means for coding the second block with an inter-coding tool of a set of interceding tools, wherein, based on the second directly referenced pictures being the same size as the current picture, the at least one of TMVP, DMVR, and BDOF disabled for the coding of the first block is enabled for coding the second block.
9. The device of claim 8, wherein at least one of the first block and the second block are bi-directionally coded.
10. The device of claim 8, further comprising means for determining that reference picture resampling (RPR) is enabled.
11. The device of claim 8, wherein size comprises a resolution.