CLAIMS

1. A method of decoding video data, the method comprising: determining a block of video data is encoded using a bi-directional inter prediction mode; determining a first motion vector (MV) for the block, wherein the first MV points to a first reference picture; determining a second MV for the block, wherein the second MV points to a second reference picture, the first reference picture being different than the second reference picture; using the first MV, locating a first predictive block in the first reference picture; using the second MV, locating a second predictive block in the second reference picture; for a first 4x4 sub-block of the first predictive block, determining a first amount of bi-directional optical flow (BIO) motion based on gradient values of samples in the first 4x4 sub-block; determining a first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, a first sub-block of the second predictive block, and the first amount of BIO motion, wherein determining the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion comprises, for each sample of the first final predictive sub-block, determining a sample value of the first final predictive sub-block based on co-located samples of the first 4x4 sub-block of the first predictive block and the first sub-block of the second predictive block using the first amount of BIO motion; for a second 4x4 sub-block of the first predictive block, determining a second amount of BIO motion based on gradient values of samples in the second 4x4 sub-block; determining a second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, a second sub-block of the second predictive block, and the second amount of BIO motion, wherein determining the second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, the second sub-block of the second predictive block, and the second amount of BIO motion comprises, for each sample of the second final predictive sub-block, determining a sample value of the second final predictive sub-block based on co-located samples of the second 4x4 sub-block of the first predictive block and the second sub-block of the second predictive block using the second amount of BIO motion; based on the first final predictive sub-block and the second final predictive sub-block, determining a final predictive block for the block of video data; and outputting a picture of video data comprising a decoded version of the block of video data.

2. The method of claim 1, wherein determining the first amount of BIO motion comprises determining the first amount of BIO motion based on samples in the first 4x4 sub-block and samples outside the first 4x4 sub-block.

3. The method of claim 1, wherein determining the first amount of BIO motion comprises determining the first amount of BIO motion based only on samples in the first 4x4 sub-block.

4. The method of claim 1, wherein determining the second amount of BIO motion comprises determining the second amount of BIO motion based on samples in the 4x4 second sub-block

and samples outside the second 4x4 sub-block.

5. The method of claim 1, wherein determining the second amount of BIO motion comprises determining the second amount of BIO motion based only on samples in the second 4x4 sub-block.

6. The method of claim 1, wherein the first amount of BIO motion comprises a motion vector field comprising a horizontal component and a vertical component.

7. The method of claim 1, wherein the first 4x4 sub-block is different than a coding unit, a prediction unit, and a transform unit for the block.

8. The method of claim 1, further comprising: adding residual data to the final predictive block to determine a reconstructed block for the block of video data.

9. The method of claim 1, wherein determining the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion comprises determining the first final predictive sub-block according to the equation:

$$pred_{BIO} = 1/2 \cdot \left(I^{(0)} + I^{(1)} + v_x \cdot \left(\tau_1 \partial I^{(1)}/\partial x - \tau_0 \partial I^{(0)}/\partial x\right) + v_y \cdot \left(\tau_1 \partial I^{(1)}/\partial y - \tau_0 \partial I^{(0)}/\partial y\right)\right)$$

wherein

$pred_{BIO}$ comprises a sample value of the first final predictive sub-block;

$I^{(0)}$ comprises a sample value of the first 4x4 sub-block of the first predictive block;

$I^{(1)}$ comprises a sample value of the first sub-block of the second predictive block;

v_x comprises a horizontal component of the first amount of BIO motion;

v_y comprises a vertical component of the first amount of BIO motion;

τ_0 comprise a distance to the first reference picture; and

τ_1 comprises a distance to the second reference picture.

10. The method of claim 1, wherein the method of decoding is performed as part of a decoding loop of a video encoding process, and wherein outputting the picture of video data comprising the decoded version of the block of video data comprises storing the picture of video data comprising the decoded version of the block of video data in a reference picture memory, the method further comprising: using the picture of video data comprising the decoded version of the block of video data as a reference picture in encoding another picture of the video data.

11. The method of claim 1, wherein outputting the picture of video data comprising the decoded version of the block of video data comprises outputting the picture of video data

comprising the decoded version of the block of video data to a display device.

12. A device for decoding video data, the device comprising: a memory configured to store video data; and one or more processors configured to determine a block of video data is encoded using a bi-directional inter prediction mode; determine a first motion vector (MV) for the block, wherein the first MV points to a first reference picture; determine a second MV for the block, wherein the second MV points to a second reference picture, the first reference picture being different than the second reference picture; using the first MV, locate a first predictive block in the first reference picture; using the second MV, locate a second predictive block in the second reference picture; for a first 4x4 sub-block of the first predictive block, determining a first amount of bi-directional optical flow (BIO) motion based on gradient values of samples in the first 4x4 sub-block; determine a first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, a first sub-block of the second predictive block, and the first amount of BIO motion, wherein to determine the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion, the one or more processors are further configured to, for each sample of the first final predictive sub-block, determine a sample value of the first final predictive sub-block based on co-located samples of the first 4x4 sub-block of the first predictive block and the first sub-block of the second predictive block using the first amount of BIO motion; for a second 4x4 sub-block of the first predictive block, determine a second amount of BIO motion based on gradient values of samples in the second 4x4 sub-block; determine a second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, a second sub-block of the second predictive block, and the second amount of BIO motion, wherein to determine the second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, the second sub-block of the second predictive block, and the second amount of BIO motion, the one or more processors are further configured to, for each sample of the second final predictive sub-block, determining a sample of the second final predictive sub-block value based on co-located samples of the second 4x4 sub-block of the first predictive block and the second sub-block of the second predictive block using the second amount of BIO motion; based on the first final predictive sub-block and the second final predictive sub-block, determine a final predictive block for the block of video data; and output a picture of video data comprising a decoded version of the block of video data.

13. The device of claim 12, wherein to determine the first amount of BIO motion, the one or more processors are configured to determine the first amount of BIO motion based on samples in the first 4x4 sub-block and samples outside the first 4x4 sub-block.

14. The device of claim 12, wherein to determine the first amount of BIO motion, the one or more processors are configured to determine the first amount of BIO motion based only on samples in the first 4x4 sub-block.

15. The device of claim 12, wherein to determine the second amount of BIO motion, the one

or more processors are configured to determine the second amount of BIO motion based on samples in the second 4x4 sub-block and samples outside the second 4x4 sub-block.

16. The device of claim 12, wherein to determine the second amount of BIO motion, the one or more processors are configured to determine the second amount of BIO motion based only on samples in the second 4x4 sub-block.

17. The device of claim 12, wherein the first amount of BIO motion comprises a motion vector field comprising a horizontal component and a vertical component.

18. The device of claim 12, wherein the first 4x4 sub-block is different than a coding unit, a prediction unit, and a transform unit for the block.

19. The device of claim 12, wherein the one or more processors are configured to: add residual data to the final predictive block to determine a reconstructed block for the block of video data.

20. The device of claim 12, wherein to determine the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion, the one or more processors are configured to determine the first final predictive sub-block according to the equation:

$$pred_{BIO} = 1/2 \cdot \left(I^{(0)} + I^{(1)} + v_x \cdot \left(\tau_1 \partial I^{(1)}/\partial x - \tau_0 \partial I^{(0)}/\partial x\right) + v_y \cdot \left(\tau_1 \partial I^{(1)}/\partial y - \tau_0 \partial I^{(0)}/\partial y\right)\right)$$

wherein

$pred_{BIO}$ comprises a sample value of the first final predictive sub-block;

$I^{(0)}$ comprises a sample value of the first 4x4 sub-block of the first predictive block;

$I^{(1)}$ comprises a sample value of the first sub-block of the second predictive block;

v_x comprises a horizontal component of the first amount of BIO motion;

v_y comprises a vertical component of the first amount of BIO motion;

τ_0 comprise a distance to the first reference picture; and

τ_1 comprises a distance to the second reference picture.

21. The device of claim 12, wherein the one or more processors decode the video data as part of a decoding loop of a video encoding process, and wherein to output the picture of video data comprising the decoded version of the block of video data, the one or more processors are configured to store the picture of video data comprising the decoded version of the block of video data in a reference picture memory, wherein the one or more processors are further configured to: use the picture of video data comprising the decoded version of the block of video data as a reference picture in encoding another picture of the video data.

22. The device of claim 12, wherein to output the picture of video data comprising the decoded version of the block of video data, the one or more processors are configured to output the picture of video data comprising the decoded version of the block of video data to a display device.

23. The device of claim 12, wherein the device comprises a wireless communication device, further comprising a receiver configured to receive encoded video data.

24. The device of claim 23, wherein the wireless communication device comprises a telephone handset and wherein the receiver is configured to demodulate, according to a wireless communication standard, a signal comprising the encoded video data.

25. The device of claim 12, wherein the device comprises a wireless communication device, further comprising a transmitter configured to transmit encoded video data.

26. The device of claim 25, wherein the wireless communication device comprises a telephone handset and wherein the transmitter is configured to modulate, according to a wireless communication standard, a signal comprising the encoded video data.

27. A non-transitory computer readable storage medium storing instructions that when executed by one or more processors cause the one or more processors to: determine a block of video data is encoded using a bi-directional inter prediction mode; determine a first motion vector (MV) for the block, wherein the first MV points to a first reference picture; determine a second MV for the block, wherein the second MV points to a second reference picture, the first reference picture being different than the second reference picture; using the first MV, locate a first predictive block in the first reference picture; using the second MV, locate a second predictive block in the second reference picture; for a first 4x4 sub-block of the first predictive block, determining a first amount of bi-directional optical flow (BIO) motion based on gradient values of samples in the first 4x4 sub-block; determine a first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, a first sub-block of the second predictive block, and the first amount of BIO motion, wherein to determine the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion, the instructions case the one or more processors to determine, for each sample of the first final predictive sub-block, a sample of the first final predictive sub-block value based on co-located samples of the first 4x4 sub-block of the first predictive block and the first sub-block of the second predictive block using the first amount of BIO motion; for a second 4x4 sub-block of the first predictive block, determine a second amount of BIO motion based on gradient values of samples in the second 4x4 sub-block; determine a second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, a second sub-block of the second predictive block, and the second amount of BIO motion, wherein to determine the second final predictive sub-block for the block of video data based on the second 4x4 sub-

block of the first predictive block, the second sub-block of the second predictive block, and the second amount of BIO motion, the instructions cause the one or more processors to determine, for each sample of the second final predictive sub-block, determine a sample of the second final predictive sub-block value based on co-located samples of the second 4x4 sub-block of the first predictive block and the second sub-block of the second predictive block using the second amount of BIO motion; based on the first final predictive sub-block and the second final predictive sub-block, determine a final predictive block for the block of video data; and output a picture of video data comprising a decoded version of the block of video data.

28. The non-transitory computer readable storage medium of claim 27, wherein the first amount of BIO motion comprises a motion vector field comprising a horizontal component and a vertical component.

29. The non-transitory computer readable storage medium of claim 27, wherein the first 4x4 sub-block is different than a coding unit, a prediction unit, and a transform unit for the block.

30. An apparatus for decoding video data, the apparatus comprising: means for determining a block of video data is encoded using a bi-directional inter prediction mode; means for determining a first motion vector (MV) for the block, wherein the first MV points to a first reference picture; means for determining a second MV for the block, wherein the second MV points to a second reference picture, the first reference picture being different than the second reference picture; means for locating a first predictive block in the first reference picture using the first MV; means for locating a second predictive block in the second reference picture using the second MV; means for determining a first amount of bi-directional optical flow (BIO) motion for a first 4x4 sub-block of the first predictive block based on gradient values of samples in the first 4x4 sub-block, wherein the means for determining the first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, the first sub-block of the second predictive block, and the first amount of BIO motion comprises, means for determining, for each sample of the first final predictive sub-block, a sample of the first final predictive sub-block value based on co-located samples of the first 4x4 sub-block of the first predictive block and the first sub-block of the second predictive block using the first amount of BIO motion; means for determining a first final predictive sub-block for the block of video data based on the first 4x4 sub-block of the first predictive block, a first sub-block of the second predictive block, and the first amount of BIO motion; means for determining a second amount of BIO motion for a second 4x4 sub-block of the first predictive block based on gradient values of samples in the second 4x4 sub-block; means for determining a second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, a second sub-block of the second predictive block, and the second amount of BIO motion, wherein the means for determining the second final predictive sub-block for the block of video data based on the second 4x4 sub-block of the first predictive block, the second sub-block of the second predictive block, and the second amount of BIO motion comprises means for determining, for each sample of the second final predictive sub-block, a sample value of the second final predictive sub-block

based on co-located samples of the second 4x4 sub-block of the first predictive block and the second sub-block of the second predictive block using the second amount of BIO motion; means for determining a final predictive block for the block of video data based on the first final predictive sub-block and the second final predictive sub-block; and means for outputting a picture of video data comprising a decoded version of the block of video data.

31. The method of claim 1, wherein determining the first amount of BIO motion based on the gradient values of the samples in the first 4x4 sub-block comprises, for each respective sample in the first 4x4 sub-block, applying a weighting function to a 5x5 window of samples around the respective sample for the first 4x4 sub-block, and wherein determining the second amount of BIO motion based on the gradient values of the samples in the second 4x4 sub-block comprises, for each respective sample in the second 4x4 sub-block, applying the weighting function to a 5x5 window of samples around the respective sample for the second 4x4 sub-block.

32. The method of claim 31, wherein the weighting function applies higher weights to samples of the 5x5 window of samples that are in the first 4x4 sub-block than to samples of the 5x5 window of samples that are outside the first 4x4 sub-block, and wherein the weighting function applies higher weights to samples of the 5x5 window of samples that are in the second 4x4 sub-block than to samples of the 5x5 window of samples that are outside the second 4x4 sub-block.

33. The device of claim 12, wherein to determine the first amount of BIO motion based on the gradient values of the samples in the first 4x4 sub-block, the one or more processors are further configured to, for each respective sample in the first 4x4 sub-block, apply a weighting function to a 5x5 window of samples around the respective sample for the first 4x4 sub-block, and wherein to determine the second amount of BIO motion based on the gradient values of the samples in the second 4x4 sub-block, the one or more processors are further configured to, for each respective sample in the second 4x4 sub-block, apply the weighting function to a 5x5 window of samples around the respective sample for the second 4x4 sub-block.

34. The device of claim 33, wherein the weighting function applies higher weights to samples of the 5x5 window of samples that are in the first 4x4 sub-block than to samples of the 5x5 window of samples that are outside the first 4x4 sub-block, and wherein the weighting function applies higher weights to samples of the 5x5 window of samples that are in the second 4x4 sub-block than to samples of the 5x5 window of samples that are outside the second 4x4 sub-block.