

WO/2014/011595

PCT/US2013/049662

What is claimed is:

1. A video coding method comprising:
receiving a video signal comprising one or more layers;
processing a base layer (BL) picture into an inter-layer reference (ILR) picture using picture level inter-layer prediction process;
selecting one or both of the processed ILR picture or an enhancement layer (EL) reference picture; and
predicting a current EL picture using the selected one or both of the ILR picture or the EL reference picture.
2. The method of claim 1, wherein the video signal comprises a BL and one or more ELs.
3. The method of claim 1 further comprising storing the processed ILR picture in an EL decoded picture buffer (DPB).
4. The method of claim 1, further comprising extracting a BL information from the BL, wherein the BL information comprises one or more of a BL mode information or a BL motion information.
5. The method of claim 4, further comprising processing the BL picture based on information comprising the extracted BL information.
6. The method of claim 1, further comprising:
converting a BL information from a first format to a second format; and
using the converted BL information in EL coding.
7. The method of claim 6, wherein the first format relates to a BL codec and the second format relates to an EL codec, and wherein the first format is different from the second format.
8. The method of claim 1, wherein the selected reference picture comprises one of the EL reference picture, or the ILR picture.
9. The method of claim 1, wherein the processed ILR picture is a non-collocated ILR picture.
10. The method of claim 1 wherein a plurality of the selected reference pictures correspond to a common time instance.
11. The method of claim 1, wherein processing the BL picture into an ILR picture comprises formatting the BL picture into a format that is suitable for prediction of the current EL picture.

REPLACEMENT SHEET 35

AMENDED SHEET - IPEA/US

WO/2014/011595

PCT/US2013/049662

12. The method of claim 11, wherein the format is dependent on a type of scalability between a BL codec and an EL codec.
13. The method of claim 12, wherein the type of scalability between the BL codec and the EL codec comprises one or more of spatial scalability or chroma format scalability.
14. The method of claim 13, wherein formatting the BL picture comprises upsampling the BL pictures.
15. The method of claim 12, wherein the type of scalability between the BL codec and the EL codec comprises one or more of quality scalability, spatial scalability, view scalability, standard scalability, or aspect ratio scalability.
16. The method of claim 15, wherein formatting a BL picture comprises one or more of denoising, restoration, or retargeting the BL pictures.
17. The method of claim 12, wherein the type of scalability between the BL codec and the EL codec comprises view scalability.
18. The method of claim 17, wherein formatting a BL picture comprises applying disparity compensation to the BL pictures.
19. The method of claim 12, wherein the type of scalability between the BL codec and the EL codec comprises bit-depth scalability.
20. The method of claim 19, wherein formatting a BL picture comprises inverse tone mapping the BL pictures.
21. The method of claim 1, further comprising packetizing inter-layer prediction (ILP) information.
22. The method of claim 21, wherein the ILP information comprises one or more of upsampling filter information, one or more coefficients for upsampling, one or more coefficients for denoising, one or more of disparity compensation parameters, or one or more of inverse tone mapping parameters.
23. The method of claim 21, further comprising sending the packetized ILP processing information via a network abstraction layer (NAL) unit.
24. A video coding system comprising:
 - a processor programmed to:
 - receive a video signal comprising one or more layers;
 - process a base layer (BL) picture into an inter-layer reference (ILR) picture using picture level inter-layer prediction process;

REPLACEMENT SHEET 36

AMENDED SHEET - IPEA/US

WO/2014/011595

PCT/US2013/049662

select one or both of the processed ILR picture or an enhancement layer (EL) reference picture; and

predict a current EL picture using the one or both of the selected ILR picture or the EL reference picture.

25. The system of claim 24, wherein the video signal comprises a BL and one or more ELs.
26. The system of claim 24, wherein the processor is further programmed to store the processed ILR picture in an EL decoded picture buffer (DPB).
27. The system of claim 24, wherein the processor is further programmed to extract a BL information from the BL, wherein the BL information comprises one or more of a BL mode information or a BL motion information.
28. The system of claim 27, wherein the processor is further programmed to process the BL picture based on information comprising the extracted BL information.
29. The system of claim 24, wherein the processor is further programmed to: convert a BL information from a first format to a second format; and use the converted BL information in EL coding.
30. The system of claim 29, wherein the first format relates to a BL codec and the second format relates to an EL codec, and wherein the first format is different from the second format.
31. The system of claim 24, wherein the selected reference picture comprises one of the EL reference picture, or the ILR picture.
32. The system of claim 24, wherein the processed ILR picture is a non-collocated ILR picture.
33. The system of claim 24 wherein a plurality of the selected reference pictures correspond to a common time instance.
34. The system of claim 24, wherein processing the BL picture into an ILR picture comprises formatting the BL picture into a format that is suitable for prediction of the current EL picture.
35. The system of claim 34, wherein the format is dependent on a type of scalability between a BL codec and an EL codec.
36. The system of claim 35, wherein the type of scalability between the BL codec and the EL codec comprises one or more of spatial scalability or chroma format scalability.
37. The system of claim 36, wherein formatting the BL picture comprises upsampling the BL pictures.

REPLACEMENT SHEET 37

AMENDED SHEET - IPEA/US

WO/2014/011595

PCT/US2013/049662

38. The system of claim 35, wherein the type of scalability between the BL codec and the EL codec comprises one or more of quality scalability, spatial scalability, view scalability, standard scalability, or aspect ratio scalability.
39. The system of claim 38, wherein formatting a BL picture comprises one or more of denoising, restoration, or retargeting the BL pictures.
40. The system of claim 35, wherein the type of scalability between the BL codec and the EL codec comprises view scalability.
41. The system of claim 40, wherein formatting a BL picture comprises applying disparity compensation to the BL pictures.
42. The system of claim 35, wherein the type of scalability between the BL codec and the EL codec comprises bit-depth scalability.
43. The system of claim 42, wherein formatting a BL picture comprises inverse tone mapping the BL pictures.
44. The system of claim 24, wherein the processor is further programmed to packetize an inter-layer prediction (ILP) information.
45. The system of claim 44, wherein the ILP information comprises one or more of upsampling filter information, one or more coefficients for upsampling, one or more coefficients for denoising, one or more of disparity compensation parameters, or one or more of inverse tone mapping parameters.
46. The system of claim 44, wherein the processor is further programmed to send the packetized ILP information via a network abstraction layer (NAL) unit.

4813-0615-9643, v. 1

REPLACEMENT SHEET 38

AMENDED SHEET - IPEA/US

of the predefined upsampling filters to upsample the pictures in the BL DPB and may send the one or more corresponding filter indices in the bitstream. For example, the decoder may perform the same upsampling process. The ILP processing and management unit may derive (e.g., adaptively derive) the one or more upsampling filters that are advantageous to use (for example, in the sense that the upsampled references may be used to predict the EL video more effectively and produce more favorable rate distortion characteristics). When adaptive upsampling filters are used, the filter coefficients and filter tap sizes may be included in the bitstream.

[0046] In another example of SNR or quality scalability, the BL video and the EL video may have the same resolution. The BL video may be coded with coarser quantization (for example, lower bit rate with lower quality), whereas the EL video may be coded with finer quantization (for example, higher bit rate with higher quality). The ILP processing and management unit 7012 may perform de-noising and/or image restoration type of operations on the reference pictures in the BL DPB 7010. Such de-noising or restoration operations may include, but are not limited to adaptive filters (e.g., based on Least Squared or LS techniques), sample adaptive offsets (SAO) (e.g., as may be supported by HEVC), and/or other types of de-noising filters targeted at, for example, reducing quantization noise. The corresponding parameters that may be used in the de-noising or restoration operations may be signaled to the decoder. Such de-noising or restoration techniques may be combined with up-sampling process to improve the prediction effectiveness for the spatial scalability case. As illustrated in FIG. 7, the coded EL bitstream, the coded BL bitstream, and the ILP information may be multiplexed (e.g., the MUX 7014) into scalable HEVC bitstream.

[0046A] In another example of spatial scalability or aspect ratio scalability, retargeting of the BL video may be performed. Retargeting may be used to establish the spatial correspondence between the two layers. For example, retargeting may correspond a spatial region associated with a picture (but not necessarily the entire picture) of the BL video to a spatial region associated with the picture (but not necessarily the entire picture) of the EL video. For example, when the EL video is higher resolution to be displayed on a large screen TV, and the BL video is lower resolution to be displayed on a mobile device with limited screen resolution, the BL video may represent a part of the EL video (e.g., a region of interest of the EL picture). Varying amount of horizontal and/or vertical scaling may be applied to the spatial region in the BL, e.g., depending on where the BL sample is located in the BL region. Retargeting may be utilized for aspect ratio scalability. For example, when the BL is in 4:3 aspect ratio and the EL is in 16:9 aspect ratio. The main object in the BL region (e.g., a news anchor) may be preserved in 4:3 aspect ratio by applying even (e.g., equal) scaling in horizontal and vertical directions, whereas the surrounding areas may be stretched horizontally to match the 16:9 aspect ratio in the EL."

[0047] As an example of view scalability, the BL video and the EL video may represent two views of a stereoscopic 3D video. The BL video may represent one view and the EL video may represent the other complementary view. The ILP processing and management unit may apply disparity compensation to the reference pictures in the BL DPB, such that they become effective prediction signal for the EL video. For a decoder to perform the same disparity compensation processing, parameters used during disparity compensation (for example, the affine transformation parameters if affine transformation is used, and/or the warping parameters if warping is used) may be signaled in the bitstream.

What is claimed is:

1. A video coding method comprising:
receiving a video signal comprising one or more layers;
processing a base layer (BL) picture into an inter-layer reference (ILR) picture using picture level inter-layer prediction process; and
on a condition that the processed ILR picture is used as a reference picture, predicting at least a portion of a current enhancement layer (EL) picture using the processed ILR picture as the reference picture.
2. The method of claim 1, further comprising predicting at least a portion of the current EL picture using an EL reference picture.
3. The method of claim 1 further comprising storing the processed ILR picture in an EL decoded picture buffer (DPB).
4. The method of claim 1, further comprising:
converting a BL information from a first format to a second format, wherein the first format relates to a BL codec and the second format relates to an EL codec, and wherein the first format is different from the second format; and
using the converted BL information in EL coding.
5. The method of claim 1, wherein the processed ILR picture is a non-collocated ILR picture.
6. The method of claim 1, wherein processing the BL picture into an ILR picture comprises formatting the BL picture into a format that is suitable for prediction of the current EL picture based on a type of scalability between the BL and the EL.
7. The method of claim 6, wherein the type of scalability between the BL and the EL comprises one or more of spatial scalability or chroma format scalability, and formatting the BL picture comprises upsampling at least a portion of the BL picture.
8. The method of claim 6, wherein the type of scalability between the BL and the EL comprises one or more of quality scalability, spatial scalability, view scalability, standard

- scalability, or aspect ratio scalability, and formatting the BL picture comprises one or more of denoising, restoration, or retargeting the BL picture.
9. The method of claim 6, wherein the type of scalability between the BL and the EL comprises view scalability, and wherein formatting the BL picture comprises applying disparity compensation to the BL picture.
 10. The method of claim 6, wherein the type of scalability between the BL and the EL comprises bit-depth scalability, wherein formatting the BL picture comprises inverse tone mapping the BL picture.
 11. The method of claim 1, further comprising receiving packetized inter-layer prediction (ILP) information.
 12. The method of claim 11, wherein the ILP information comprises one or more of upsampling filter information, one or more upsampling coefficients, one or more denoising coefficients, one or more of disparity compensation parameters, or one or more of inverse tone mapping parameters.
 13. The method of claim 11, wherein the packetized ILP processing information is received via a network abstraction layer (NAL) unit.
 14. A video coding device (VCD) comprising:
a processor configured to:
 receive a video signal comprising one or more layers;
 process a base layer (BL) picture into an inter-layer reference (ILR) picture using picture level inter-layer prediction process; and
 on a condition that the processed ILR picture is used as a reference picture,
 predict at least a portion of a current enhancement layer (EL) picture using the processed ILR picture as the reference picture.
 15. The VCD of claim 14, wherein the processor is further configured to predict at least a portion of the current EL picture using an EL reference picture.
 16. The VCD of claim 14 wherein the processor is further configured to store the processed ILR picture in an EL decoded picture buffer (DPB).

17. The VCD of claim 14, wherein the processor is further configured to:
convert a BL information from a first format to a second format, wherein the first format relates to a BL codec and the second format relates to an EL codec, and wherein the first format is different from the second format; and
use the converted BL information in EL coding.
18. The VCD of claim 14, wherein the processed ILR picture is a non-collocated ILR picture.
19. The VCD of claim 14, wherein processing the BL picture into an ILR picture comprises formatting the BL picture into a format that is suitable for prediction of the current EL picture based on a type of scalability between the BL and the EL.
20. The VCD of claim 19, wherein the type of scalability between the BL and the EL comprises one or more of spatial scalability or chroma format scalability, and formatting the BL picture comprises upsampling at least a portion of the BL picture.
21. The VCD of claim 19, wherein the type of scalability between the BL and the EL comprises one or more of quality scalability, spatial scalability, view scalability, standard scalability, or aspect ratio scalability, and formatting the BL picture comprises one or more of denoising, restoration, or retargeting the BL picture.
22. The VCD of claim 19, wherein the type of scalability between the BL and the EL comprises view scalability, and wherein formatting the BL picture comprises applying disparity compensation to the BL picture.
23. The VCD of claim 19, wherein the type of scalability between the BL and the EL comprises bit-depth scalability, wherein formatting the BL picture comprises inverse tone mapping the BL picture.
24. The VCD of claim 14, further comprising receiving packetized inter-layer prediction (ILP) information.
25. The VCD of claim 24, wherein the ILP information comprises one or more of upsampling filter information, one or more upsampling coefficients, one or more denoising coefficients, one or more of disparity compensation parameters, or one or more of inverse tone mapping parameters.

26. The VCD of claim 24, wherein the packetized ILP processing information is received via a network abstraction layer (NAL) unit.