Claims

1. A method for decoding a coded video sequence by at least one processor, characterized by, the method comprising:

obtaining the coded video sequence; and

decoding the coded video sequence,

wherein a value of a temporal sublayer identifier of a video coding layer (VCL) network abstraction layer (NAL) unit in the coded video sequence is constrained to be less than or equal to a value of vps_max_sublayers_minus1, that specifies a maximum number of temporal sublayers capable of being present in a layer in each coded video sequence referring to the video parameter set (VPS), in the VPS referred to by the VCL NAL unit;

wherein a value of max_tid_il_ref_pics_plus1[i] being zero specifies that inter-layer prediction is not used by non-intra random access point (IRAP) pictures of the i-th layer.

2. The method of claim 1, wherein a value of max_tid_il_ref_pics_plus1[i] being greater than zero specifies that, for decoding pictures of the i-th layer, no picture with a temporal sublayer identifier greater than max_tid_il_ref_pics_plus1[i] − 1 is used as an inter-layer reference picture (ILRP).

3. The method of claim 1, wherein when not present, a value of max_tid_il_ref_pics_plus1[i] is inferred to be equal to vps_max_sublayers_minus1 + 1.

4. The method of claim 1, wherein max_tid_il_ref_pics_plus1[i] is constrained to be less than or equal to vps_max_sublayers_minus1 + 1.

5. The method of claim 1, wherein a sublayer-wise output layer set is constrained not be derived for independent layers.

6. The method of claim 5, wherein

a first variable NumSubLayersInLayerInOLS[i][j] specifies the number of sublayers in the j-th layer in the i-th OLS, wherein a second variable OutputLayerIdInOls[i][j] specifies the nuh_layer_id value of the j-th output layer in the i-th OLS, wherein a third variable LayerUsedAsOutputLayerFlag[k] specifies whether the k-th layer is used as an output layer in at least one OLS, and wherein the first variable, the second variable, and the third variable are derived from a value of variable max_tid_il_ref_pics_plus1[i].

7. The method of claim 5, wherein max_tid_il_ref_pics_plus1 and layerIncludedInOlsFlag are constrained not be derived for independent layers.

8. A device for decoding a coded video sequence, characterized by, the device comprising:

at least one memory configured to store program code; and

at least one processor configured to read the program code and operate as instructed by the program code, the program code including:

obtaining code configured to cause the at least one processor to obtain the coded video sequence; and

decoding code configured to cause the at least one processor to decode the coded video sequence,

wherein a value of a temporal sublayer identifier of a video coding layer (VCL) network abstraction layer (NAL) unit in the coded video sequence is constrained to be less than or equal to a value of vps_max_sublayers_minus1, that specifies a maximum number of temporal sublayers capable of being present in a layer in each coded video sequence referring to the video parameter set (VPS), in the VPS referred to by the VCL NAL unit;

wherein a value of max_tid_il_ref_pics_plus1[i] being zero specifies that inter-layer prediction is not used by non-intra random access point (IRAP) pictures of the i-th layer.

9. The device of claim 8, wherein a value of max_tid_il_ref_pics_plus1[i] being greater than zero specifies that, for decoding pictures of the i-th layer, no picture with a temporal sublayer identifier greater than max_tid_il_ref_pics_plus1[i] − 1 is used as an inter-layer reference picture (ILRP).

10. The device of claim 8, wherein when not present, a value of max_tid_il_ref_pics_plus1[i] is inferred to be equal to vps_max_sublayers_minus1 + 1.

11. The device of claim 8, wherein max_tid_il_ref_pics_plus1[i] is constrained to be less than or equal to vps_max_sublayers_minus1 + 1.

12. The device of claim 8, wherein a sublayer-wise output layer set is constrained not be derived for independent layers.

13. The device of claim 12, wherein max_tid_il_ref_pics_plus1 and layerIncludedInOlsFlag are constrained not be derived for independent layers.

14. A method for encoding a video sequence by an encoder, characterized by, the method comprising:

receiving video samples from a video source;

encoding the video samples to generate a coded video sequence, wherein a value of a temporal sublayer identifier of a video coding layer, VCL, network abstraction layer, NAL, unit in the coded video sequence is constrained to be less than or equal to a value of vps_max_sublayers_minus1, that specifies a maximum number of temporal sublayers capable of being present in a layer in each coded video sequence referring to the video parameter set, VPS, in the VPS referred to by the VCL NAL unit;

wherein a value of max_tid_il_ref_pics_plus1[i] being zero specifies that inter-layer prediction is not used by non-intra random access point (IRAP) pictures of the i-th layer.

15. A device for encoding a video sequence, characterized by, the device comprising:

at least one memory configured to store program code; and

at least one processor configured to read the program code and operate as instructed by the program code, the program code including:

receiving video samples from a video source;

encoding the video samples to generate a coded video sequence, wherein a value of a temporal sublayer identifier of a video coding layer, VCL, network abstraction layer, NAL, unit in the coded video sequence is constrained to be less than or equal to a value of vps_max_sublayers_minus1, that specifies a maximum number of temporal sublayers capable of being present in a layer in each coded video sequence referring to the video parameter set, VPS, in the VPS referred to by the VCL NAL unit;

wherein a value of max_tid_il_ref_pics_plus1[i] being zero specifies that inter-layer prediction is not used by non-intra random access point (IRAP) pictures of the i-th layer.

16. A non-transitory computer-readable medium storing instructions, characterized by, the instructions comprising: one or more instructions that, when executed by one or more processors of a device for decoding/encoding a video sequence, cause the one or more processors to perform the method according to any one of claims 1 to 13, and 14.