Claims

1. An apparatus for receiving a video data stream as an input bitstream, wherein the video data stream has a video encoded thereinto,

wherein, for a playback speed modification factor (Nx), the apparatus is to determine a decoding capability requirement depending on a decoding capability requirement limit,

wherein the playback speed modification factor (Nx) is a forward playback speed modification factor or is a backward playback speed modification factor.

2. An apparatus according to claim 1,

wherein the decoding capability requirement is a modified sample rate,

wherein the decoding capability requirement limit is a sample rate limit.

3. An apparatus according to claim 2,

wherein the video data stream comprises information on the sample rate limit.

4. An apparatus according to claim 2 or 3,

wherein the apparatus is to determine the modified sample rate depending on a picture size of a picture of the video.

5. An apparatus according to claim 4,

wherein the apparatus is to determine the modified sample rate further depending on a framerate and depending on the playback speed modification factor (Nx),

wherein the video data stream comprises information on the framerate.

6. An apparatus according to claim 5,

wherein the framerate depends on a sublayer of one or more sublayers,

wherein the video data stream comprises sub-layer specific framerate information for each of the one or more sublayers,

wherein the apparatus is to obtain the framerate from the sub-layer specific framerate information for one of the one or more sublayers.

7. An apparatus according to one of claims 4 to 6,

wherein the apparatus is to determine the modified sample rate depending on the picture size of a picture of the video, being a highest picture size among a plurality of pictures of the video.

8. An apparatus according to claim 7,

wherein the video data stream comprises an application factor,

wherein the apparatus is to compute the modified sample rate using the highest layer picture size and the application factor.

9. An apparatus according to claim 7 or 8,

wherein the highest picture size depends on a sublayer of one or more sublayers,

wherein the video data stream comprises sub-layer specific maximum picture size information for each of the one or more sublayers,

wherein the apparatus is to obtain the highest picture size from the sub-layer specific maximum picture size information for one of the one or more sublayers.

10. An apparatus according to one of claims 4 to 6,

wherein the video data stream comprises an application factor,

wherein the apparatus is to determine the modified sample rate using a sample rate of the video and using the application factor.

11. An apparatus according to claim 1,

wherein the decoding capability requirement is a modified bitrate,

wherein the decoding capability requirement limit is a bitrate limit.

12. An apparatus according to claim 11,

wherein the video data stream comprises information on the bitrate limit.

13. An apparatus according to claim 11 or 12,

wherein the apparatus is to determine the modified bitrate depending on an initial bitrate of the video data stream and depending on the playback speed modification factor (Nx).

14. An apparatus according to claim 13,

wherein the initial bitrate depends on a sublayer of one or more sublayers,

wherein the video data stream comprises sub-layer specific initial bitrate information for each of the one or more sublayers,

wherein the apparatus is to obtain the initial bitrate from the sub-layer specific framerate information for one of the one or more sublayers.

15. An apparatus according to one of claims 1 to 14,

wherein the video data stream comprises an IRAP-only flag which indicates whether or not decoding of only IRAP pictures may take place when an initial playback speed is increased,

wherein, depending on the IRAP-only flag, the apparatus is to read a decoding capability requirement from the video data stream.

16. An apparatus according to one of claims 1 to 15,

wherein the video data stream comprises an reference-pictures-only flag which indicates whether or not decoding of only reference pictures may take place when an initial playback speed is increased,

wherein, depending on the reference-pictures-only flag, the apparatus is to read a decoding capability requirement from the video data stream.

17. A video data stream having a video encoded thereinto,

wherein the video data stream comprises information on a sample rate limit, and/or

wherein the video data stream comprises information on a bitrate limit.

18. A video data stream according to claim 17,

wherein the video data stream comprises information on a framerate.

19. A video data stream according to claim 17 or 18,

wherein the video data stream comprises sub-layer specific framerate information for each of one or more sublayers.

20. A video data stream according to one of claims 17 to 19,

wherein the video data stream comprises an application factor for determining a modified sample rate.

21. A video data stream according to one of claims 17 to 20,

wherein the video data stream comprises sub-layer specific maximum picture size information for each of one or more sublayers.

22. A video data stream according to one of claims 17 to 21,

wherein the video data stream comprises sub-layer specific initial bitrate information for each of one or more sublayers.

23. A video data stream according to one of claims 17 to 22,

wherein the video data stream comprises an IRAP-only flag which indicates whether or not decoding of only IRAP pictures may take place when an initial playback speed is increased.

24. A video data stream according to one of claims 17 to 23,

wherein the video data stream comprises an reference-pictures-only flag which indicates whether or not decoding of only reference pictures may take place when an initial playback speed is increased.

25. A video data stream according to claim 23 or 24,

wherein, depending on the IRAP-only flag, the video data stream comprises level information, and/or

wherein, depending on the reference-pictures-only flag, the video data stream comprises level information.

26. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises information on a sample rate limit, and/or

wherein the video encoder is to generate the video data stream such that the video data stream comprises information on a bitrate limit.

27. A video encoder according to claim 26,

wherein the video encoder is to generate the video data stream such that the video data stream comprises information on a framerate.

28. A video encoder according to claim 26 or 27,

wherein the video encoder is to generate the video data stream such that the video data stream comprises sub-layer specific framerate information for each of one or more sublayers.

29. A video encoder according to one of claims 26 to 28,

wherein the video encoder is to generate the video data stream such that the video data stream comprises an application factor for determining a modified sample rate.

30. A video encoder according to one of claims 26 to 29,

wherein the video encoder is to generate the video data stream such that the video data stream comprises sub-layer specific maximum picture size information for each of one or more sublayers.

31. A video encoder according to one of claims 26 to 30,

wherein the video encoder is to generate the video data stream such that the video data stream comprises sub-layer specific initial bitrate information for each of one or more sublayers.

32. A video encoder according to one of claims 26 to 31,

wherein the video encoder is to generate the video data stream such that the video data stream comprises an IRAP-only flag which indicates whether or not decoding of only IRAP pictures may take place when an initial playback speed is increased.

33. A video encoder according to one of claims 26 to 32,

wherein the video encoder is to generate the video data stream such that the video data stream comprises an reference-pictures-only flag which indicates whether or decoding of only reference pictures may take place when an initial playback speed is increased.

34. A video encoder according to claim 32 or 33,

wherein the video encoder is to generate the video data stream such that, depending on the IRAP-only flag, the video data stream comprises level information, and/or

wherein the video encoder is to generate the video data stream such that, depending on the reference-pictures-only flag, the video data stream comprises level information.

35. An apparatus according to one of claims 1 to 16,

wherein the apparatus is to decode the input bitstream to decode the video.

36. A system for encoding a video into a video data stream, and for decoding the video, wherein the system comprises:

a video encoder according to one of claims 26 to 34, and

an apparatus according to claim 35,

wherein the video encoder is to encode the video into the video data stream, such that the video data stream has the video encoded thereinto,

wherein the apparatus is to receive the video data stream as an input bitstream,

wherein the apparatus is to decode the input bitstream to decode the video.

37. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein, for a playback speed modification factor (Nx), the method comprises determining a decoding capability requirement depending on a decoding capability requirement limit,

wherein the playback speed modification factor (Nx) is a forward playback speed modification factor or is a backward playback speed modification factor.

38. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises information on a sample rate limit, and/or

wherein the method comprises generating the video data stream such that the video data stream comprises information on a bitrate limit.

39. A computer program for implementing the method of claim 37 or 38 when being executed on a computer or signal processor.

40. A video data stream having a video encoded thereinto,

wherein the video data stream comprises a plurality of access units, wherein a group of access units of the video data stream comprises the plurality of access units,

wherein the video data stream comprises an indication for each access unit of a subset of access units indicating that a seamless splicing point exists after said access unit,

wherein the subset of access units is a proper subset of the group of access units of the video data stream.

41. A video data stream according to claim 40,

wherein the video data stream comprises a plurality of temporal sublayers,

wherein the indication that a seamless splicing point exists after an access unit, only exists for access units of a first group of temporal sublayers of the plurality of temporal sublayers,

but wherein the indication that a seamless splicing point exists after an access unit, does not exist for access units of a second group of temporal sublayers of the plurality of temporal sublayers.

42. A video data stream according to claim 40 or 41,

wherein the video data stream comprises a buffering period supplemental enhancement information,

wherein the buffering period supplemental enhancement information indicates a subset of access units of the plurality of access units for which the seamless splicing point exists after said access unit.

43. A video data stream according to one of claims 40 to 42,

wherein the video data stream comprises a divisor,

wherein a delay value is assigned to each access unit of the plurality of access units of the video data stream,

wherein the indication that indicates a seamless splicing point exists an said access unit, only exists for those access units of the plurality of access units of the video data stream which have a delay value assigned, which, when being divided by the divisor, is equal to a remainder value.

44. A video data stream according to claim 43,

wherein the remainder value is 0.

45. A video data stream according to claim 43,

wherein the video data stream comprises the remainder value.

46. A video data stream according to one of claims 40 to 45,

wherein the indication for each access unit of the subset of access units indicating that a seamless splicing point exists after said access unit, is a flag.

47. A video data stream according to claim 46,

wherein the flag is a delay_for_concatenation_ensured_flag.

48. A video data stream according to claim 46 or 47,

wherein the flag exhibits a first value, if a constant framerate for output pictures which result from decoding the video data stream is not ensured, and

wherein the flag exhibits a second value being different from the first value, if a constant framerate for output pictures which result from decoding the video data stream is ensured.

49. A video data stream according to claim 48,

wherein the flag exhibits the second value, if the output pictures which result from decoding the video data stream have an equidistant output time.

50. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a plurality of access units, wherein a group of access units of the video data stream comprises the plurality of access units,

wherein the video encoder is to generate the video data stream such that the video data stream comprises an indication for each access unit of a subset of access units indicating that a seamless splicing point exists after said access unit,

wherein the video encoder is to generate the video data stream such that the subset of access units is a proper subset of the group of access units of the video data stream.

51. A video encoder according to claim 50,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a plurality of temporal sublayers,

wherein the video encoder is to generate the video data stream such that the indication that a seamless splicing point exists after an access unit, only exists for access units of a first group of temporal sublayers of the plurality of temporal sublayers,

but wherein the video encoder is to generate the video data stream such that the indication that a seamless splicing point exists after an access unit, does not exist for access units of a second group of temporal sublayers of the plurality of temporal sublayers.

52. A video encoder according to claim 50 or 51,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a buffering period supplemental enhancement information,

wherein the video encoder is to generate the video data stream such that the buffering period supplemental enhancement information indicates a subset of access units of the plurality of access units for which the seamless splicing point exists after said access unit.

53. A video encoder according to one of claims 50 to 52,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a divisor,

wherein the video encoder is to generate the video data stream such that a delay value is assigned to each access unit of the plurality of access units of the video data stream,

wherein the video encoder is to generate the video data stream such that the indication that indicates a seamless splicing point exists an said access unit, only exists for those access units of the plurality of access units of the video data stream, which have a delay value assigned, which, when being divided by the divisor, is equal to a remainder value.

54. A video encoder according to claim 53,

wherein the video encoder is to generate the video data stream such that the remainder value is 0.

55. A video encoder according to claim 53,

wherein the video encoder is to generate the video data stream such that the video data stream comprises the remainder value.

56. A video encoder according to one of claims 50 to 55,

wherein the video encoder is to generate the video data stream such that the indication for each access unit of the subset of access units indicating that a seamless splicing point exists after said access unit, is a flag.

57. A video encoder according to claim 56,

wherein the video encoder is to generate the video data stream such that the flag is a delay_for_concatenation_ensured_flag.

58. A video encoder according to claim 56 or 57,

wherein the video encoder is to generate the video data stream such that the flag exhibits a first value, if a constant framerate for output pictures which result from decoding the video data stream is not ensured, and

wherein the video encoder is to generate the video data stream such that the flag exhibits a second value being different from the first value, if a constant framerate for output pictures which result from decoding the video data stream is ensured.

59. A video encoder according to claim 58,

wherein the video encoder is to generate the video data stream such that the flag exhibits the second value, if the output pictures which result from decoding the video data stream have an equidistant output time.

60. An apparatus for receiving a video data stream as an input bitstream, wherein the video data stream has a video encoded thereinto,

wherein the apparatus is to process the video data stream,

wherein the video data stream comprises a plurality of access units, wherein a group of access units of the video data stream comprises the plurality of access units,

wherein the video data stream comprises an indication for each access unit of a subset of access units indicating that a seamless splicing point exists after said access unit,

wherein the subset of access units is a proper subset of the group of access units of the video data stream,

wherein the apparatus is to process the indication.

61. An apparatus according to claim 60,

wherein the video data stream comprises a plurality of temporal sublayers,

wherein the indication that a seamless splicing point exists after an access unit, only exists for access units of a first group of temporal sublayers of the plurality of temporal sublayers,

but wherein the indication that a seamless splicing point exists after an access unit, does not exist for access units of a second group of temporal sublayers of the plurality of temporal sublayers.

62. An apparatus according to claim 60 or 61,

wherein the video data stream comprises a buffering period supplemental enhancement information,

wherein the buffering period supplemental enhancement information indicates a subset of access units of the plurality of access units for which the seamless splicing point exists after said access unit,

wherein the apparatus is to process the buffering period supplemental enhancement information.

63. An apparatus according to one of claims 60 to 62,

wherein the video data stream comprises a divisor,

wherein the apparatus is to receive the divisor,

wherein a delay value is assigned to each access unit of the plurality of access units of the video data stream,

wherein the indication that indicates a seamless splicing point exists an said access unit, only exists for those access units of the plurality of access units of the video data stream which have a delay value assigned, which, when being divided by the divisor, is equal to a remainder value.

64. An apparatus according to claim 63,

wherein the remainder value is 0.

65. An apparatus according to claim 63,

wherein the video data stream comprises the remainder value,

wherein the apparatus is to receive the remainder value.

66. An apparatus according to one of claims 60 to 65,

wherein the indication for each access unit of the subset of access units indicating that a seamless splicing point exists after said access unit, is a flag,

wherein the apparatus is to process the flag.

67. An apparatus according to claim 66,

wherein the flag is a delay_for_concatenation_ensured_flag.

68. An apparatus according to claim 66 or 67,

wherein the flag exhibits a first value, if a constant framerate for output pictures which result from decoding the video data stream is not ensured, and

wherein the flag exhibits a second value being different from the first value, if a constant framerate for output pictures which result from decoding the video data stream is ensured.

69. An apparatus according to claim 68,

wherein the flag exhibits the second value, if the output pictures which result from decoding the video data stream have an equidistant output time.

70. An apparatus according to one of claims 60 to 69,

wherein the apparatus is to decode the input bitstream to decode the video.

71. A system for encoding a video into a video data stream, and for decoding the video, wherein the system comprises:

a video encoder according to one of claims 50 to 59, and

an apparatus according to claim 70,

wherein the video encoder is to encode the video into the video data stream, such that the video data stream has the video encoded thereinto,

wherein the apparatus is to receive the video data stream as an input bitstream,

wherein the apparatus is to decode the input bitstream to decode the video.

72. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises a plurality of access units, wherein a group of access units of the video data stream comprises the plurality of access units,

wherein the method comprises generating the video data stream such that the video data stream comprises an indication for each access unit of a subset of access units indicating that a seamless splicing point exists after said access unit,

wherein the method comprises generating the video data stream such that the subset of access units is a proper subset of the group of access units of the video data stream.

73. A method for receiving a video data stream as an input bitstream, wherein the video data stream has a video encoded thereinto,

wherein the method comprises processing the video data stream,

wherein the video data stream comprises a plurality of access units, wherein a group of access units of the video data stream comprises the plurality of access units,

wherein the video data stream comprises an indication for each access unit of a subset of access units indicating that a seamless splicing point exists after said access unit,

wherein the subset of access units is a proper subset of the group of access units of the video data stream,

wherein the method comprises processing the indication.

74. A computer program for implementing the method of claim 72 or 73 when being executed on a computer or signal processor.

75. An apparatus for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the apparatus is to process the input bitstream depending on an output layer set to obtain a sub-bitstream,

wherein, if the output layer set does not comprise a predefined layer (vps_layer_id[0]) of a plurality of layers that are present in the video data stream, the apparatus is to remove at least one non-scalable-nested supplemental enhancement information message which is assigned with the predefined layer (vps_layer_id[0]), and

wherein, if the output layer set comprises the predefined layer (vps_layer_id[0]), the apparatus is configured to not remove any non-scalable-nested supplemental enhancement information message which is assigned with the predefined layer (vps_layer_id[0]).

76. An apparatus according to claim 75,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in a current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with at least one of a buffering period payload and a picture timing payload and a decoding unit information payload.

77. An apparatus according to claim 76,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with at least one of the buffering period payload and the picture timing payload and the decoding unit information payload and a subpicture level information payload.

78. An apparatus according to claim 77,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with a payload type being equal to a first value indicating the buffering period payload, or being equal to a second value indicating the picture timing payload, or being equal to a third value indicating the decoding unit information payload, or being equal to a fourth value indicating the subpicture level information payload.

79. An apparatus according to claim 78,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to the first value, or being equal to the second value, or being equal to the third value, or being equal to the fourth value, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

80. An apparatus according to claim 78 or 79,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with the payload type being equal to 0 indicating the buffering period payload, or being equal to 1 indicating the picture timing payload, or being equal to 130 indicating the decoding unit payload, or being equal to 203 indicating the subpicture level information payload.

81. An apparatus according to claim 80,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to 0 or being equal to 1 or being equal to 130 or being equal to 203, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

82. An apparatus according to one of claims 75 to 81,

wherein, if the output layer set does not comprise the predefined layer (vps_layer_id[0]), the apparatus is to remove all nested supplemental enhancement information messages which are assigned with the predefined layer (vps_layer_id[0]).

83. A video data stream having a video encoded thereinto,

wherein the video data stream comprises a video parameter set,

wherein the video parameter set comprises a plurality of output layer sets,

wherein, if none of the plurality of output layer sets comprises all layers of a plurality of layers that are present in the video data stream, the video data stream does not comprise any non-scalable nested supplemental enhancement information message for a hypothetical reference decoder.

84. A video data stream according to claim 83,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in a current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with at least one of a buffering period payload and a picture timing payload and a decoding unit information payload.

85. A video data stream according to claim 84,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no

non-scalable-nested supplemental enhancement information message with at least one of the buffering period payload and the picture timing payload and the decoding unit information payload and a subpicture level information payload.

86. A video data stream according to claim 85,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with a payload type being equal to a first value indicating the buffering period payload, or being equal to a second value indicating the picture timing payload, or being equal to a third value indicating the decoding unit information payload, or being equal to a fourth value indicating the subpicture level information payload.

87. A video data stream according to claim 86,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to the first value, or being equal to the second value, or being equal to the third value, or being equal to the fourth value, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

88. A video data stream according to claim 86 or 87,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with the payload type being equal to 0 indicating the buffering period payload, or being equal to 1 indicating the picture timing payload, or being equal to 130 indicating the decoding unit payload, or being equal to 203 indicating the subpicture level information payload.

89. A video data stream according to claim 88,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to 0 or being equal to 1 or being equal-

to 130 or being equal to 203, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

90. A video data stream according to one of claims 83 to 89,

wherein, if the video data stream comprises at least one non-scalable-nested supplemental enhancement information message for the hypothetical reference decoder, at least one of the plurality of output layer sets comprises all layers of the plurality of layers that are present in the video data stream.

91. A video data stream according to one of claims 83 to 90,

wherein all supplemental enhancement information message for the hypothetical reference decoder, which are not associated with an output layer set of the plurality of output layer sets which comprises all layers of the plurality of layers that are present in the video data stream, are scalable-nesting supplemental enhancement information messages.

92. A video data stream according to one of claims 83 to 91,

wherein all non-scalable-nested supplemental enhancement information message for the hypothetical reference decoder are associated with an output layer set of the plurality of output layer sets which comprises all layers of the plurality of layers that are present in the video data stream.

93. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a video parameter set,

wherein the video encoder is to generate the video data stream such that the video parameter set comprises a plurality of output layer sets,

wherein, if none of the plurality of output layer sets comprises all layers of a plurality of layers that are present in the video data stream, the video encoder is to generate

the video data stream such that the video data stream does not comprise any non-scalable nested supplemental enhancement information message for a hypothetical reference decoder.

94. A video encoder according to claim 93,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in a current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with at least one of a buffering period payload and a picture timing payload and a decoding unit information payload.

95. A video encoder according to claim 94,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with at least one of the buffering period payload and the picture timing payload and the decoding unit information payload and a subpicture level information payload.

96. A video encoder according to claim 95,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with a payload type being equal to a first value indicating the buffering period payload, or being equal to a second value indicating the picture timing payload, or being equal to a third value indicating the decoding unit information payload, or being equal to a fourth value indicating the subpicture level information payload.

97. A video encoder according to claim 96,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to the first value, or being equal to the second value, or being equal to the third value, or being equal to the fourth value, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

98. A video encoder according to claim 96 or 97,

wherein, if there is no output layer set that consists of all layers of the plurality of layers in the current coded video sequence of the video data stream, there is no non-scalable-nested supplemental enhancement information message with the payload type being equal to 0 indicating the buffering period payload, or being equal to 1 indicating the picture timing payload, or being equal to 130 indicating the decoding unit payload, or being equal to 203 indicating the subpicture level information payload.

99. A video encoder according to claim 98,

wherein, for a non-scalable-nested supplemental enhancement information message having the payload type being equal to 0 or being equal to 1 or being equal to 130 or being equal to 203, the non-scalable-nested supplemental enhancement information message applies to all the output layer sets, when present, that consist of all layers in the current coded video sequence in the entire bitstream.

100. A video encoder according to one of claims 93 to 99,

wherein, if the video data stream comprises at least one non-scalable-nested supplemental enhancement information message for the hypothetical reference decoder, the video encoder is to generate the video data stream such that at least one of the plurality of output layer sets comprises all layers of the plurality of layers that are present in the video data stream.

101. A video encoder according to one of claims 93 to 100,

wherein the video encoder is to generate the video data stream such that all supplemental enhancement information message for the hypothetical reference decoder, which are not associated with an output layer set of the plurality of output layer sets which comprises all layers of the plurality of layers that are present in the video data stream, are scalable-nesting supplemental enhancement information messages.

102. A video encoder according to one of claims 93 to 101,

wherein the video encoder is to generate the video data stream such that all non-scalable nested supplemental enhancement information message for the hypothetical reference decoder are associated with an output layer set of the plurality of output layer sets which comprises all layers of the plurality of layers that are present in the video data stream.

103. An apparatus according to one of claims 75 to 82,

wherein the apparatus is to decode the input bitstream to decode the video.

104. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the method comprises processing the input bitstream depending on an output layer set to obtain a sub-bitstream,

wherein, if the output layer set does not comprise a predefined layer (vps_layer_id[0]) of a plurality of layers that are present in the video data stream, the method comprises removing at least one non-scalable nested supplemental enhancement information message which is assigned with the predefined layer (vps_layer_id[0]), and

wherein, if the output layer set comprises the predefined layer (vps_layer_id[0]), the method comprises not removing any non-scalable nested supplemental enhancement information message which is assigned with the predefined layer (vps_layer_id[0]).

105. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises a video parameter set,

wherein the method comprises generating the video data stream such that the video parameter set comprises a plurality of output layer sets,

wherein, if none of the plurality of output layer sets comprises all layers of a plurality of layers that are present in the video data stream, the method comprises generating the video data stream such that the video data stream does not comprise any non-scalable nested supplemental enhancement information message for a hypothetical reference decoder.

106. A computer program for implementing the method of claim 104 or 105 when being executed on a computer or signal processor.

107. An apparatus for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises initial profile tier level information,

wherein the video data stream comprises a first group of two or more layers,

wherein the apparatus is to process the input bitstream depending on an output layer set of a plurality of output layer sets to generate an output bitstream by removing at least one layer from the first group of two or more layers to obtain a second group of one or more layers, such that the output bitstream comprises the second group of one or more layers, but not the at least one layer that has been removed from the first group,

wherein the apparatus is to remove those portions of the initial profile tier level information from the initial profile tier level information that is not associated with at least one layer of the second group of one or more layers.

108. An apparatus according to claim 107,

wherein the video data stream comprises initial decoding parameter set profile tier level information as the initial profile tier level information,

wherein the apparatus is to remove those portions of the initial decoding parameter set profile tier level information from the initial decoding parameter set profile tier level information that is not associated with the at least one layer of the second group of one or more layers.

109. An apparatus for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets,

wherein the apparatus is to process the input bitstream depending on an output layer set of the plurality of output layer sets to obtain an output bitstream, such that the output bitstream comprises the decoding parameter set information for said output layer set, and such that the output bitstream does not comprise the decoding parameter set information for any other output layer set of the plurality of output layer sets.

110. An apparatus according to claim 109,

wherein the video data stream comprises raw byte sequence payload,

wherein the raw byte sequence payload comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

111. An apparatus according to claim 109 or 110,

wherein the video data stream comprises a scalable nesting supplemental enhancement message which comprises the scalable nesting supplemental enhancement information that comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

112. An apparatus according to claim 111,

wherein the video data stream comprises a nesting type flag (nesting_type) comprising one or more bits.

113. An apparatus according to claim 112,

wherein, if the nesting type flag (nesting_type) exhibits a first value, this indicates that the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set

information, and that no further supplemental enhancement messages are nested within the scalable nesting supplemental enhancement message.

114. An apparatus according to claim 112 or 113,

wherein, if the nesting type flag (nesting_type) exhibits a second value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message apply to an output layer set.

115. An apparatus according to one of claims 112 to 114,

wherein, if the nesting type flag (nesting_type) exhibits a third value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message apply to a subset of an output layer set.

116. An apparatus according to one of claims 112 to 115,

wherein, if the nesting type flag (nesting_type) exhibits a fourth value, this indicates the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set information.

117. A video data stream having a video encoded thereinto,

wherein the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets.

118. A video data stream according to claim 117,

wherein the video data stream comprises raw byte sequence payload,

wherein the raw byte sequence payload comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

119. A video data stream according to claim 117 or 118,

wherein the video data stream comprises a scalable nesting supplemental enhancement message which comprises the scalable nesting supplemental enhancement information that comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

120. A video data stream according to claim 119,

wherein the video data stream comprises a nesting type flag (nesting_type) comprising one or more bits.

121. A video data stream according to claim 120,

wherein, if the nesting type flag (nesting_type) exhibits a first value, this indicates that the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set information, and that no further supplemental enhancement messages are nested within the scalable nesting supplemental enhancement message.

122. A video data stream according to claim 120 or 121,

wherein, if the nesting type flag (nesting_type) exhibits a second value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message apply to an output layer set.

123. A video data stream according to one of claims 120 to 122,

wherein, if the nesting type flag (nesting_type) exhibits a third value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message apply to a subset of an output layer set, e.g., wherein one or more subpictures are assigned to the subset.

124. A video data stream according to one of claims 120 to 123,

wherein, if the nesting type flag (nesting_type) exhibits a fourth value, this indicates the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set information.

125. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets.

126. A video encoder according to claim 125,

wherein the video encoder is to generate the video data stream such that the video data stream comprises raw byte sequence payload,

wherein the video encoder is to generate the video data stream such that the raw byte sequence payload comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

127. A video encoder according to claim 125 or 126,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a scalable nesting supplemental enhancement message which comprises the scalable nesting supplemental enhancement information that comprises the decoding parameter set information for each output layer set of the plurality of output layer sets.

128. A video encoder according to claim 127,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a nesting type flag (nesting_type) comprising one or more bits.

129. A video encoder according to claim 128,

wherein the video encoder is to generate the video data stream such that, if the nesting type flag (nesting_type) exhibits a first value, this indicates that the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set information, and that no

further supplemental enhancement messages are nested within the scalable nesting supplemental enhancement message.

130. A video encoder according to claim 128 or 129,

wherein the video encoder is to generate the video data stream such that, if the nesting type flag (nesting_type) exhibits a second value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message apply to an output layer set.

131. A video encoder according to one of claims 128 to 130,

wherein the video encoder is to generate the video data stream such that, if the nesting type flag (nesting_type) exhibits a third value, this indicates that one or more supplemental enhancement messages being nested within the scalable nesting supplemental enhancement message to a subset of an output layer set, e.g., wherein one or more subpictures are assigned to the subset.

132. A video encoder according to one of claims 128 to 131,

wherein the video encoder is to generate the video data stream such that, if the nesting type flag (nesting_type) exhibits a fourth value, this indicates the scalable nesting supplemental enhancement message comprises one or more decoding parameter sets which comprise the decoding parameter set information.

133. A video data stream having a video encoded thereinto,

wherein the video data stream comprises an indication of a plurality of operation points,

wherein the video data stream comprises a first mapping for each operation point of the plurality of operation points which assigns one or more profile tier levels of a plurality of profile tier levels to said operation point,

wherein the video data stream comprises a plurality of second mappings, wherein each of the plurality of mappings assigns one of the plurality of operation points to an output level set of a plurality of output level sets.

134. A video data stream according to claim 133,

wherein the video data stream comprises a decoding parameter set,

wherein the decoding parameter set comprises the indication of the plurality of operation points.

135. A video data stream according to claim 133 or 134,

wherein the video data stream comprises a number of layers indication indicating a number of layers of the video data stream, and/or

wherein the video data stream comprises a number of sublayers indication indicating a number of sublayers of the video data stream.

136. A video data stream according to claim 133 or 134,

wherein a number of layers of the video data stream is constant, and/or

wherein a number of sublayers of the video data stream is constant.

137. A video data stream according to one of claims 133 to 136,

wherein the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets.

138. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises an indication of a plurality of operation points,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a first mapping for each operation point of the plurality of operation points which assigns one or more profile tier levels of a plurality of profile tier levels to said operation point,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a plurality of second mappings, wherein each of the plurality of mappings assigns one of the plurality of operation points to an output level set of a plurality of output level sets.

139. A video encoder according to claim 138,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a decoding parameter set,

wherein the video encoder is to generate the video data stream such that the decoding parameter set comprises the indication of the plurality of operation points.

140. A video encoder according to claim 138 or 139,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a number of layers indication indicating a number of layers of the video data stream, and/or

wherein the video encoder is to generate the video data stream such that the video data stream comprises a number of sublayers indication indicating a number of sublayers of the video data stream.

141. A video encoder according to claim 138 or 139,

wherein the video encoder is to generate the video data stream such that a number of layers of the video data stream is constant, and/or

wherein the video encoder is to generate the video data stream such that a number of sublayers of the video data stream is constant.

142. A video encoder according to one of claims 138 to 141,

wherein the video encoder is to generate the video data stream such that the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets.

143. An apparatus for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises an indication of a plurality of operation points,

wherein the video data stream comprises a first mapping for each operation point of the plurality of operation points which maps one or more profile tier levels of a plurality of profile tier levels to said operation point,

wherein the video data stream comprises a plurality of second mappings, wherein each of the plurality of mappings maps one of the plurality of operation points to an output level set of a plurality of output level sets,

wherein the apparatus is to process the input bitstream depending on an output layer set of a plurality of output layer sets and depending on at least one of the second mappings, which maps one of the plurality of operation points said an output level set, to generate an output bitstream.

144. An apparatus according to claim 143,

wherein the video data stream comprises a decoding parameter set,

wherein the decoding parameter set comprises the indication of the plurality of operation points,

wherein the apparatus is to process the decoding parameter set.

145. An apparatus according to claim 143 or 144,

wherein the video data stream comprises a number of layers indication indicating a number of layers of the video data stream, and/or

wherein the video data stream comprises a number of sublayers indication indicating a number of sublayers of the video data stream.

146. An apparatus according to claim 143 or 144,

wherein a number of layers of the video data stream is constant, and/or

wherein a number of sublayers of the video data stream is constant.

147. An apparatus according to one of claims 143 to 146,

wherein the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets,

wherein the apparatus is to process the scalable nesting supplemental enhancement information.

148. An apparatus according to one of claims 107 to 116 or according to one of claims 143 to 147,

wherein the apparatus is to decode the input bitstream to decode the video.

149. A system for encoding a video into a video data stream, and for decoding the video, wherein the system comprises:

a video encoder according to one of claims 125 to 132 or according to one of claims 138 to 142, and

an apparatus according to claim 148,

wherein the video encoder is to encode the video into the video data stream, such that the video data stream has the video encoded thereinto,

wherein the apparatus is to receive the video data stream as an input bitstream,

wherein the apparatus is to decode the input bitstream to decode the video.

150. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises initial profile tier level information,

wherein the video data stream comprises a first group of two or more layers;

wherein the method comprises processing the input bitstream depending on an output layer set of a plurality of output layer sets to generate an output bitstream by removing at least one layer from the first group of two or more layers to obtain a second group of one or more layers, such that the output bitstream comprises the second group of one or more layers, but not the at least one layer that has been removed from the first group,

wherein the method comprises removing those portions of the initial profile tier level information from the initial profile tier level information that is not associated with at least one layer of the second group of one or more layers.

151. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets,

wherein the method comprises processing the input bitstream depending on an output layer set of the plurality of output layer sets to obtain an output bitstream, such that the output bitstream comprises the decoding parameter set information for said output layer set, and such that the output bitstream does not comprise the decoding parameter set information for any other output layer set of the plurality of output layer sets.

152. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises scalable nesting supplemental enhancement information comprising decoding parameter set information for each output layer set of a plurality of output layer sets.

153. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises an indication of a plurality of operation points,

wherein the method comprises generating the video data stream such that the video data stream comprises a first mapping for each operation point of the plurality of operation points which assigns one or more profile tier levels of a plurality of profile tier levels to said operation point,

wherein the method comprises generating the video data stream such that the video data stream comprises a plurality of second mappings, wherein each of the plurality of mappings assigns one of the plurality of operation points to an output level set of a plurality of output level sets.

154. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises an indication of a plurality of operation points,

wherein the video data stream comprises a first mapping for each operation point of the plurality of operation points which maps one or more profile tier levels of a plurality of profile tier levels to said operation point,

wherein the video data stream comprises a plurality of second mappings, wherein each of the plurality of mappings maps one of the plurality of operation points to an output level set of a plurality of output level sets,

wherein the method comprises processing the input bitstream depending on an output layer set of a plurality of output layer sets and depending on at least one of the second mappings, which maps one of the plurality of operation points said an output level set, to generate an output bitstream.

155. A computer program for implementing the method of one of claims 150 to 154 when being executed on a computer or signal processor.

156. A video data stream having a video encoded thereinto,

wherein the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the first indication value is smaller than or equal to the second indication value.

157. A video data stream having a video encoded thereinto,

wherein the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the second indication value of the decoding parameter set is an upper bound having precedence over the first indication value of the sequence parameter set.

158. A video data stream according to claim 156 or 157,

wherein the constant value is 1.

159. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the video encoder is to generate the video data stream such that the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the video encoder is to generate the video data stream such that the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the video encoder is to generate the video data stream such that the first indication value is smaller than or equal to the second indication value.

160. A video encoder for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the video encoder is to generate the video data stream such that the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the video encoder is to generate the video data stream such that the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the video encoder is to generate the video data stream such that the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the second indication value of the decoding parameter set is an upper bound having precedence over the first indication value of the sequence parameter set.

161. A video encoder according to claim 159 or 160,

wherein the constant value is 1,

wherein the video encoder is to generate the video data stream such that the sequence parameter set comprises the first indication value that indicates, according to the information stored in the sequence parameter set, the maximum number of sublayers or the maximum number of sublayers minus 1,

wherein the video encoder is to generate the video data stream such that the decoding parameter set comprises the second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus 1.

162. An apparatus for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein if the first indication value is greater than the second indication value, the apparatus is to process the input bitstream to generate an output bitstream such that the output bitstream comprises the second indication value as an indication on a maximum number of sublayers or on a maximum number of sublayers minus the constant value.

163. An apparatus according to claim 162,

wherein the constant value is 1,

wherein the sequence parameter set comprises the first indication value that indicates, according to the information stored in the sequence parameter set, the maximum number of sublayers or the maximum number of sublayers minus 1,

wherein the decoding parameter set comprises the second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus 1, and

wherein if the first indication value is greater than the second indication value, the apparatus is to process the input bitstream to generate an output bitstream such that the output bitstream comprises the second indication value as an indication on a maximum number of sublayers minus 1.

164. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the method comprises generating the video data stream such that the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the method comprises generating the video data stream such that the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the method comprises generating the video data stream such that the first indication value is smaller than or equal to the second indication value.

165. A method for encoding a video into a video data stream, such that the video data stream has the video encoded thereinto,

wherein the method comprises generating the video data stream such that the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the method comprises generating the video data stream such that the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the method comprises generating the video data stream such that the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein the second indication value of the decoding parameter set is an upper bound having precedence over the first indication value of the sequence parameter set.

166. A method for receiving a video data stream as an input bitstream, wherein video data stream has a video encoded thereinto,

wherein the video data stream comprises a sequence parameter set and a decoding parameter set,

wherein the sequence parameter set comprises a first indication value that indicates, according to the information stored in the sequence parameter set, a maximum number of sublayers or the maximum number of sublayers minus a constant value,

wherein the decoding parameter set comprises a second indication value that indicates, according to the information stored in the decoding parameter set, the maximum number of sublayers or the maximum number of sublayers minus the constant value,

wherein if the first indication value is greater than the second indication value, the method comprises processing the input bitstream to generate an output bitstream such that the output bitstream comprises the second indication value as an indication on a maximum number of sublayers or on a maximum number of sublayers minus the constant value.

~~167. A computer program for implementing the method of one of claims 164 to 166 when being executed on a computer or signal processor.~~

CLAIMS

1. A video encoder configured to:

encode, into a video data stream, a plurality of access units, the plurality of access units comprising:

a first subset of access units having a framerate, and

a second subset of access units having the framerate; and

encode, into the video data stream, a plurality of delay_for_concatenation_ensured_flags such that each access unit of the first subset of access units has an associated delay_for_concatenation_ensured_flag, wherein a delay_for_concatenation_ensured_flag having a value of 1 indicates that an access unit, to which the delay_for_concatenation_ensured_flag is associated, is at a position in the video bitstream where splicing will:

retain the framerate, and

be seamless with an access unit of the second subset of access units when the access of the second subset of access units has a BP SEI message with:

a concatenation_flag equal to 1, and

a selected InitCpbRemovalDelay less than or equal to the value of max_initial_removal_delay_for_concatination.

2. The video encoder of claim 1, wherein seamless means that a CPB removal time distance between the access unit associated with the delay_for_concatenation_ensured_flag having a value of 1 and the access unit of the second subset of access units is the same as consecutive access units in the first subset of access units.

3. The video encoder of claim 1, further configured to:

encode, into the video bitstream, a BP SEI message that designates a temporal sublayer, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units associated with the temporal sublayer.

4. The video encoder of claim 1, further configured to:

encode, into the video bitstream, a BP SEI message that designates the first subset of access units, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units in the first subset of access units.

5. The video encoder of claim 4, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder of zero, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder of zero.

6. The video encoder of claim 4, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor and a signaled remainder, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder.

7. A video decoder configured to:

decode, from a video data stream, a plurality of access units, the plurality of access units comprising:

a first subset of access units having a framerate, and

a second subset of access units having the framerate; and

decode, from the video data stream, a plurality of delay_for_concatenation_ensured_flags such that each access unit of the first subset of access units has an associated delay_for_concatenation_ensured_flag, wherein a delay_for_concatenation_ensured_flag having a value of 1 indicates that an access unit, to which the delay_for_concatenation_ensured_flag is associated, is at a position in the video bitstream where splicing will:

retain the framerate, and

be seamless with an access unit of the second subset of access units when the access of the second subset of access units has a BP SEI message with:

a concatenation_flag equal to 1, and

a selected InitCpbRemovalDelay less than or equal to the value of max_initial_removal_delay_for_concatination.

8. The video decoder of claim 7, wherein seamless means that a CPB removal time distance between the access unit associated with the delay_for_concatenation_ensured_flag having a value of 1 and the access unit of the second subset of access units is the same as consecutive access units in the first subset of access units.

9. The video decoder of claim 7, further configured to:

decode, from the video bitstream, a BP SEI message that designates a temporal sublayer, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units associated with the temporal sublayer.

10. The video decoder of claim 7, further configured to:

decode, from the video bitstream, a BP SEI message that designates the first subset of access units, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units in the first subset of access units.

11. The video decoder of claim 10, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder of zero, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder of zero.

12. The video decoder of claim 10, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor and a signaled remainder, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder.

13. A video decoding method comprising:

decoding, from a video data stream, a plurality of access units, the plurality of access units comprising:

a first subset of access units having a framerate, and

a second subset of access units having the framerate; and

decoding, from the video data stream, a plurality of delay_for_concatenation_ensured_flags such that each access unit of the first subset of access units has an associated delay_for_concatenation_ensured_flag, wherein a delay_for_concatenation_ensured_flag having a value of 1 indicates that an access unit, to which the delay_for_concatenation_ensured_flag is associated, is at a position in the video bitstream where splicing will:

retain the framerate, and

be seamless with an access unit of the second subset of access units when the access of the second subset of access units has a BP SEI message with:

a concatenation_flag equal to 1, and

a selected InitCpbRemovalDelay less than or equal to the value of max_initial_removal_delay_for_concatination.

14. The video decoder of claim 13, wherein seamless means that a CPB removal time distance between the access unit associated with the delay_for_concatenation_ensured_flag having a value of 1 and the access unit of the second subset of access units is the same as consecutive access units in the first subset of access units.

15. The video decoding method of claim 13, further comprising:

decoding, from the video bitstream, a BP SEI message that designates a temporal sublayer, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units associated with the temporal sublayer.

16. The video decoding method of claim 13, further comprising:

decoding, from the video bitstream, a BP SEI message that designates the first subset of access units, wherein the encoding of the plurality of delay_for_concatenation_ensured_flags is only performed for access units in the first subset of access units.

17. The video decoding method of claim 16, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder of zero, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder of zero.

18. The video decoding method of claim 16, wherein the BP SEI message that designates the first subset of access units includes a signaled divisor and a signaled remainder, and wherein the first subset of access units includes all of the access units that:

have a cpb_removal_delay_minus1 that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder, or

have a dpb_output_delay that, when divided by the signaled divisor, has a remainder that is equal to the signaled remainder.

19. A method of encoding a video, the method comprising the steps of:

encoding, into a video data stream, a plurality of access units, the plurality of access units comprising:

a first subset of access units having a framerate, and

a second subset of access units having the framerate; and

encoding, into the video data stream, a plurality of delay_for_concatenation_ensured_flags such that each access unit of the first subset of access units has an associated delay_for_concatenation_ensured_flag, wherein a delay_for_concatenation_ensured_flag having a value of 1 indicates that an access unit, to which the delay_for_concatenation_ensured_flag is associated, is at a position in the video bitstream where splicing will:

retain the framerate, and

be seamless with an access unit of the second subset of access units when the access of the second subset of access units has a BP SEI message with:

a concatenation_flag equal to 1, and

a selected InitCpbRemovalDelay less than or equal to the value of max_initial_removal_delay_for_concatination.