

CLAIMS

1. A method of processing media data, the method comprising: processing a first sample of a first track of a media file, the first track having a first composition time; determining whether a second track of the media file, the second track being associated with the first track, has an available sample at the first composition time; when the second track has the available sample at the first composition time, selecting the available sample of the second track as a second sample having a common composition time with the first composition time of the first sample; when the second track does not have the available sample at the first composition time, selecting a sample of the second track having a composition time that is a closest available preceding composition time as the second sample having the common composition time with the first composition time of the first sample; and in response to selecting the second sample, processing the second sample.
2. The method of claim 1, wherein the first sample and the second sample have different decoding times.
3. The method of claim 1, further comprising, when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) or an AdaptationSet element for the first track and the second track, preventing processing of a region-wise packing (RWPK) descriptor of the AdaptationSet element.
4. The method of claim 1, further comprising determining that the first track either belongs to a sub-picture composition track group or is a tile track in response to determining that a sample entry corresponding to the first track contains a composition restriction box.
5. The method of claim 1, further comprising determining that the second track either belongs to a sub-picture composition track group or is a tile track in response to determining that a sample entry corresponding to the second track contains a composition restriction box.
6. The method of claim 1, further comprising, when at least one of the first track or the second

track is a tile track based on having a corresponding sample entry having a composition restriction box, preventing presentation of media data of the at least one of the first track or the second track.

7. The method of claim 1, further comprising processing a projected omnidirectional video box of media data, the omnidirectional video box comprising an extension of a FullBox data structure.

8. The method of claim 1, further comprising preventing projected regions of media data of at least one of the first track or the second track indicated by region-wise packing boxes from overlapping, wherein the projected regions comprise a first projected region comprising a rectangle specified by proj_reg_width[i], proj_reg_height[i], proj_reg_top[i], and proj_reg_left[i] and a second projected region comprising a rectangle specified by proj_reg_width[j], proj_reg_height[j], proj_reg_top[j], and proj_reg_left[j] for any value of j in the range of 0 to i-1, inclusive.

9. The method of claim 1, further comprising preventing processing a global coverage information box of media data of at least one of the first track or the second track when a region-wise packing box is not present in the media data of the at least one of the first track or the second track.

10. The method of claim 1, further comprising determining that a global coverage information box of media data of at least one of the first track or the second track indicates that coverage is an entire sphere when a region-wise packing box of the media data of the at least one of the first track or the second track is not present.

11. The method of claim 1, further comprising at least one of: when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) for the first track and the second track, preventing processing of a region-wise packing (RWPK) descriptor of the MPD; when the PF descriptor is not present directly in the MPD and not present directly in an AdaptationSet element, preventing processing of a RWPK descriptor of the AdaptationSet

element; or when the PF descriptor is not present directly in a representation element, is not present directly in the AdaptationSet element containing the representation element, and not present directly in the MPD, preventing processing of a RWPK descriptor of the representation element.

12. The method of claim 1, wherein processing the first sample comprises encapsulating the first sample, the first sample comprising first encoded video data, and wherein processing the second sample comprises encapsulating the second sample, the second sample comprising second encoded video data.

13. The method of claim 1, wherein processing the first sample comprises decapsulating the first sample, the first sample comprising first encoded video data, and sending the first encoded video data to a video decoder, and wherein processing the second sample comprises decapsulating the second sample, the second sample comprising second encoded video data, and sending the second encoded video data to the video decoder.

14. The method of claim 1, wherein the first track comprises a tile base track, and the second track comprises a tile track associated with the tile base track.

15. A device for processing media data, the device comprising: a memory configured to store media data; and a processor comprising circuitry configured to: process a first sample of a first track of a media file, the first track having a first composition time, the first sample including first media data; determine whether a second track of the media file, the second track being associated with the first track, has an available sample at the first composition time; when the second track has the available sample at the first composition time, select the available sample of the second track as a second sample having a common composition time with the first composition time of the first sample; when the second track does not have the available sample at the first composition time, select a sample of the second track having a composition time that is a closest available preceding composition time as the second sample having the common composition time with the first composition time of the first sample; and in response to selecting the second sample, process the second sample, the second sample including second media data.

16. The device of claim 15, wherein the first sample and the second sample have different decoding times.
17. The device of claim 15, wherein the processor is further configured to, when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) or an AdaptationSet element for the first track and the second track, prevent processing of a region-wise packing (RWPK) descriptor of the AdaptationSet element.
18. The device of claim 15, wherein the processor is further configured to determine that the first track either belongs to a sub-picture composition track group or is a tile track in response to determining that a sample entry corresponding to the first track contains a composition restriction box.
19. The device of claim 15, wherein the processor is further configured to determine that the second track either belongs to a sub-picture composition track group or is a tile track in response to determining that a sample entry corresponding to the second track contains a composition restriction box.
20. The device of claim 15, wherein the processor is further configured to, when at least one of the first track or the second track is a tile track based on having a corresponding sample entry having a composition restriction box, prevent presentation of media data of the at least one of the first track or the second track.
21. The device of claim 15, wherein the processor is further configured to process a projected omnidirectional video box of media data, the omnidirectional video box comprising an extension of a FullBox data structure.
22. The device of claim 15, wherein the processor is further configured to prevent projected regions of media data of at least one of the first track or the second track indicated by region-wise packing boxes from overlapping, wherein the projected regions comprise a first projected

region comprising a rectangle specified by proj_reg_width[i], proj_reg_height[i], proj_reg_top[i], and proj_reg_left[i] and a second projected region comprising a rectangle specified by proj_reg_width[j], proj_reg_height[j], proj_reg_top[j], and proj_reg_left[j] for any value of j in the range of 0 to i-1, inclusive.

23. The device of claim 15, wherein the processor is further configured to prevent processing a global coverage information box of media data of at least one of the first track or the second track when a region-wise packing box is not present in the media data of the at least one of the first track or the second track.

24. The device of claim 15, wherein the processor is further configured to determine that a global coverage information box of media data of at least one of the first track or the second track indicates that coverage is an entire sphere when a region-wise packing box of the media data of the at least one of the first track or the second track is not present.

25. The device of claim 15, wherein the processor is further configured to at least one of: when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) for the first track and the second track, prevent processing of a region-wise packing (RWPK) descriptor of the MPD; when the PF descriptor is not present directly in the MPD and not present directly in an AdaptationSet element, prevent processing of a RWPK descriptor of the AdaptationSet element; or when the PF descriptor is not present directly in a representation element, is not present directly in the AdaptationSet element containing the representation element, and not present directly in the MPD, prevent processing of a RWPK descriptor of the representation element.

26. The device of claim 15, further comprising a video encoder configured to encode first video data to form first encoded video data and to encode second video data to form second encoded video data, the first media data comprising the first encoded video data, and the second media data comprising the second encoded video data, wherein to process the first sample, the processor is configured to encapsulate the first sample, the first sample comprising the first encoded video data, and wherein to process the second sample, the processor is configured to

encapsulate the second sample, the second sample comprising the second encoded video data.

27. The device of claim 15, further comprising a video decoder configured to decode video data, wherein to process the first sample, the processor is configured to decapsulate the first sample, the first sample comprising first encoded video data, the first media data comprising the first encoded video data, and send the first encoded video data to the video decoder, and wherein to process the second sample, the processor is configured to decapsulate the second sample, the second sample comprising second encoded video data, the second media data comprising the second encoded video data, and send the second encoded video data to the video decoder.

28. The device of claim 15, wherein the device comprises at least one of: an integrated circuit; a microprocessor; or a wireless communication device.

29. A device for processing media data, the device comprising: means for processing a first sample of a first track of a media file, the first track having a first composition time; means for determining whether a second track of the media file, the second track being associated with the first track, has an available sample at the first composition time; means for selecting, when the second track has the available sample at the first composition time, the available sample of the second track as a second sample having a common composition time with the first composition time of the first sample; means for selecting, when the second track does not have the available sample at the first composition time, a sample of the second track having a composition time that is a closest available preceding composition time as the second sample having the common composition time with the first composition time of the first sample; and means for processing, in response to selecting the second sample, the second sample.

30. The device of claim 29, wherein the first sample and the second sample have different decoding times.

31. The device of claim 29, further comprising means for preventing, when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) or an AdaptationSet element for the first track and the second track, processing of a region-wise

packing (RWPK) descriptor of the AdaptationSet element.

32. The device of claim 29, further comprising a video encoder configured to encode first video data to form first encoded video data and to encode second video data to form second encoded video data, wherein the means for processing the first sample comprises means for encapsulating the first sample, the first sample comprising the first encoded video data, and wherein the means for processing the second sample comprises means for encapsulating the second sample, the second sample comprising the second encoded video data.

33. The device of claim 29, further comprising a video decoder, wherein the means for processing the first sample comprises means for decapsulating the first sample, the first sample comprising first encoded video data, and means for sending the first encoded video data to the video decoder, and wherein the means for processing the second sample comprises means for decapsulating the second sample, the second sample comprising second encoded video data, and means for sending the second encoded video data to the video decoder.

34. A non-transitory computer-readable storage medium having stored thereon instructions that, when executed, cause a processor to: process a first sample of a first track of a media file, the first track having a first composition time; determine whether a second track of the media file, the second track being associated with the first track, has an available sample at the first composition time; when the second track has the available sample at the first composition time, select the available sample of the second track as a second sample having a common composition time with the first composition time of the first sample; when the second track does not have the available sample at the first composition time, select a sample of the second track having a composition time that is a closest available preceding composition time as the second sample having the common composition time with the first composition time of the first sample; and in response to selecting the second sample, process the second sample.

35. The non-transitory computer-readable storage medium of claim 34, wherein the first sample and the second sample have different decoding times.

36. The non-transitory computer-readable storage medium of claim 34, further comprising instructions that cause the processor to, when a projection format (PF) descriptor is not present directly in a media presentation description (MPD) or an AdaptationSet element for the first track and the second track, prevent processing of a region-wise packing (RWPK) descriptor of the AdaptationSet element.

37. The non-transitory computer-readable storage medium of claim 34, wherein the instructions that cause the processor to process the first sample comprise instructions that cause the processor to encapsulate the first sample, the first sample comprising first encoded video data, and wherein the instructions that cause the processor to process the second sample comprise instructions that cause the processor to encapsulate the second sample, the second sample comprising second encoded video data.

38. The non-transitory computer-readable storage medium of claim 34, wherein the instructions that cause the processor to process the first sample comprise instructions that cause the processor to decapsulate the first sample, the first sample comprising first encoded video data, and send the first encoded video data to a video decoder, and wherein the instructions that cause the processor to process the second sample comprise instructions that cause the processor to decapsulate the second sample, the second sample comprising second encoded video data, and send the second encoded video data to the video decoder.