CLAIMS

1. A method of processing video data, the method comprising: obtaining a media file associated with 360-degree video data, the 360-degree video data including a spherical representation of a scene, the media file including first signaling information and second signaling information of a viewport region corresponding to a region of interest (ROI) in the spherical representation, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; extracting pixels corresponding to the viewport region from data of the picture based on the first signaling information and second signaling information; and providing the pixels to render the viewport region for display, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles, that includes the viewport region, and wherein the minimum set is the least set, of the tiles of the plurality of tiles, that includes all of the pixels corresponding to the viewport region.

2. The method of claim 1, wherein the first signaling information includes a first angle and a second angle of a center of the viewport region with respect to a spherical center of the spherical representation of the scene, the first angle being formed on a first plane and the second angle being formed on a second plane, the first plane being perpendicular to the second plane.

3. The method of claim 1, wherein the first signaling information further includes a third angle associated with a width of the viewport region and a fourth angle associated with a height of the viewport region.

4. The method of claim 3, wherein the third angle is formed between a first edge and a second edge of the viewport region; and wherein the fourth angle is formed between a third edge and a fourth edge of the viewport region.

5. The method of claim 2, wherein the ROI is defined by at least four planes that intersect with

the spherical representation; and wherein each of the four planes also intersects with the spherical center.

6. The method of claim 5, further determining a shape of the viewport region based on the intersecting of the at least four planes with the spherical representation.

7. The method of claim 6, wherein the pixels corresponding to the viewport region are extracted based on the shape.

8. The method of claim 1, wherein the extracting comprises using the first signaling information to determine a boundary of the viewport within the identified minimum set.

9. The method of claim 1, wherein the second signaling information includes one or more coordinates associated with the one or more tiles in the picture.

10. The method of claim 1, wherein the one or more tiles form a tile group, and wherein the second signaling information includes a group identifier associated with the tile group.

11. The method of claim 1, wherein the plurality of tiles are motion-constrained tiles.

12. The method of claim 1, wherein the second signaling information includes pixel coordinates associated with a pre-determined location within a viewport region formed by projecting the ROI on a plane, a width of the viewport region, and a height of the viewport region.

13. The method of claim 1, wherein the media file is based on an International Standards Organization (ISO) base media file format (ISOBMFF).

14. The method of claim 13, wherein the media file identifies a sample group including a video sample corresponding to the spherical representation of the scene; and wherein the first signaling information and the second signaling information are included in one or more syntax elements of the sample group.

15. The method of claim 1, wherein: the media file is based on a media presentation description (MPD) format and includes one or more adaptation sets; each of the one or more adaptation sets includes one or more representations; and the first signaling information, the second signaling information, and a link to the picture are included in one or more elements associated with the ROI included in the one or more representations; and wherein the method comprises: obtaining the picture based on the link included in the media file.

16. The method of claim 15, wherein the one or more representations are tile-based representations, and wherein the second signaling information includes identifiers associated with tiles including the ROI included in the tile-based representations.

17. The method of claim 1, wherein the spherical representation of the scene is projected onto the plane using a rectilinear projection.

18. The method of claim 1, further comprising: extracting pixels of multiple ROIs from the picture based on the first signaling information and the second signaling information.

19. An apparatus for processing video data, comprising: a memory configured to store 360-degree video data; and a processor configured to: obtain a media file associated with the 360-degree video data, the 360-degree video data including a spherical representation of a scene, the media file including first signaling information and second signaling information of a viewport region corresponding to a region of interest (ROI) in the spherical representation, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; extract pixels corresponding to the viewport region from data of the picture based on the first signaling information and second signaling information; and provide the pixels to render the viewport region for display, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles, that includes the

viewport region, and wherein the minimum set is the least set, of the tiles of the plurality of tiles, that includes all of the pixels corresponding to the viewport region.

20. The apparatus of claim 19, wherein the processor is further configured to: determine, from the first signaling information, a first angle and a second angle of a center of the viewport region with respect to a spherical center of the spherical representation of the scene, the first angle being formed on a first plane and the second angle being formed on a second plane, the first plane being perpendicular to the second plane.

21. The apparatus of claim 19, wherein the processor is further configured to: determine, from the first signaling information, a third angle associated with a width of the viewport region and a fourth angle associated with a height of the viewport region.

22. The apparatus of claim 20, wherein the third angle is formed between a first edge and a second edge of the viewport region; and wherein the fourth angle is formed between a third edge and a fourth edge of the viewport region.

23. The apparatus of claim 19, wherein the ROI is defined by at least four planes that intersect with the spherical representation; and wherein each of the four planes also intersects with the spherical center.

24. The apparatus of claim 23, wherein the processor is further configured to determine a shape of the viewport region based on the intersecting of the at least four planes with the spherical representation.

25. The apparatus of claim 24, wherein the processor is configured to extract the pixels corresponding to the viewport region based on the shape.

26. The apparatus of claim 19, wherein the processor configured to extract the pixels is configured to use the first signaling information to determine a boundary of the viewport within the identified minimum set.

27. The apparatus of claim 19, wherein the processor is further configured to determine, from the second signaling information, one or more coordinates associated with the one or more tiles in the picture.

28. The apparatus of claim 19, wherein the one or more tiles form a tile group, and wherein the processor is further configured to determine, from the second signaling information, a group identifier associated with the tile group.

29. The apparatus of claim 19, wherein the plurality of tiles are motion-constrained tiles.

30. The apparatus of claim 19, wherein the processor is further configured to determine, from the second signaling information, pixel coordinates associated with a pre-determined location within a viewport region formed by projecting the ROI on a plane, a width of the viewport region, and a height of the viewport region.

31. The apparatus of claim 19, wherein the media file is based on an International Standards Organization (ISO) base media file format (ISOBMFF).

32. The apparatus of claim 31, wherein the media file identifies a sample group including a video sample corresponding to the spherical representation of the scene; and wherein the processor is further configured to extract the first signaling information and the second signaling information from one or more syntax elements of the sample group.

33. The apparatus of claim 19, wherein: the media file is based on a media presentation description (MPD) format and includes one or more adaptation sets; each of the one or more adaptation sets includes one or more representations; and wherein the processor is further configured to: determine, based on one or more elements associated with the ROI included in the one or more representations, the first signaling information, the second signaling information and a link to the picture; and obtain the picture based on the link included in the media file.

34. The apparatus of claim 33, wherein the one or more representations are tile-based representations, and wherein the processor is configured to determine, based on the second signaling information, identifiers associated with tiles including the ROI included in the tile-based representations.

35. The apparatus of claim 19, wherein the spherical representation of the scene is projected onto the plane using a rectilinear projection.

36. The apparatus of claim 19, wherein the processor is further configured to extract pixels of multiple ROIs from the picture based on the first signaling information and the second signaling information.

37. The apparatus of claim 19, wherein the apparatus comprises a mobile device with one or more cameras to capture the 360-degree video data.

38. The apparatus of claim 19, wherein the apparatus comprises a display to render the viewport region.

39. A non-transitory computer-readable medium having stored thereon instructions that, when executed by one or more processors, cause the one or more processors to: obtain a media file associated with 360-degree video data, the 360-degree video data including a spherical representation of a scene, the media file including first signaling information and second signaling information of a viewport region corresponding to a region of interest (ROI) in the spherical representation, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; extract pixels corresponding to the viewport region from data of the picture based on the first signaling information and second signaling information; and provide the pixels to render the viewport region for display, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles,

that includes the viewport region, and wherein the minimum set is the least set, of the tiles of the plurality of tiles, that includes all of the pixels corresponding to the viewport region.

40. A method of processing video data, the method comprising: obtaining 360-degree video data, the 360-degree video data including a spherical representation of a scene; determining a region of interest (ROI) in the spherical representation of the scene; generating a media file including first signaling information and second signaling information of a viewport region corresponding to the ROI, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; and providing the media file for rendering the 360-degree video data or for transmission of a portion of the 360-degree video data including at least the ROI, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles, that includes the viewport region, and wherein the minimum set is the least set, of the tiles of the plurality of tiles, that includes all of the pixels corresponding to the viewport region.

41. An apparatus for processing video data, comprising: a memory configured to store 360-degree video data; and a processor configured to: obtain 360-degree video data, the 360-degree video data including a spherical representation of a scene; determine a region of interest (ROI) in the spherical representation of the scene; generate a media file including first signaling information and second signaling information of a viewport region corresponding to the ROI, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; and provide the media file for rendering the 360-degree video data or for transmission of a portion of the 360-degree video data including at least the ROI, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles, that includes the viewport region, and wherein the minimum set is the least set, of the tiles of the

plurality of tiles, that includes all of the pixels corresponding to the viewport region.

42. A non-transitory computer-readable medium having stored thereon instructions that, when executed by one or more processors, cause the one or more processor to: obtain 360-degree video data, the 360-degree video data including a spherical representation of a scene; determine a region of interest (ROI) in the spherical representation of the scene; generate a media file including first signaling information and second signaling information of a viewport region corresponding to the ROI, the first signaling information including a center position and a dimension of the viewport region measured in a spherical space associated with the spherical representation, and the second signaling information indicating a region, which includes the viewport region, of a picture formed by projecting the spherical representation including the ROI onto a plane, the picture comprising a plurality of tiles; and provide the media file for rendering the 360-degree video data or for transmission of a portion of the 360-degree video data including at least the ROI, wherein the second signaling information identifies a minimum set, of the tiles of the plurality of tiles, that includes the viewport region, and wherein the minimum set is the least set, of the tiles of the plurality of tiles, that includes all of the pixels corresponding to the viewport region.