CLAIMS

1. An apparatus, comprising:

an interface system; and

a logic system configured for:

receiving, via the interface system, audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations;

receiving, via the interface system, reproduction environment data comprising an indication of a number of reproduction speakers of an actual three-dimensional reproduction environment and an indication of the location of each reproduction speaker within the actual reproduction environment;

mapping the audio reproduction data created with reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and

rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

2. The apparatus of claim 1, wherein the actual reproduction environment comprises a sound system environment of a cinema.

3. The apparatus of claim 1, wherein the actual reproduction environment comprises a 7.1 configuration.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 19 of 27 was completed at 28.08.2013 15:37:36
Received at the EPO on Aug 28, 2013 15:39:36. Page 19 of 27

4. The apparatus of claim 1, wherein the actual reproduction environment data includes reproduction speaker layout data indicating reproduction speaker locations.

5. The apparatus of claim 1, wherein the actual reproduction environment data includes reproduction speaker zone layout data indicating reproduction speaker locations.

6. The apparatus of claim 5, wherein the metadata includes information for mapping an audio object position to a single reproduction speaker location.

7. The apparatus of claim 1, wherein the rendering involves creating a gain based on one or more of a desired audio object position, a distance from the desired audio object position to a reference position, a velocity of an audio object or an audio object content type.

8. The apparatus of claim 1, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

9. The apparatus of claim 1, wherein the rendering involves imposing speaker zone constraints including data for disabling selected reproduction speakers.

10. The apparatus of claim 1, wherein the actual reproduction environment comprises a screen for the projection of video images; wherein the audio reproduction data is synchronized with the video images; and wherein the rendering involves applying screen-to-room balance control according to screen-to-room balance control data received from the user input system.

11. The apparatus of claim 1, further comprising a display system, wherein the logic system is configured to control the display system to display a dynamic three-dimensional view of the actual reproduction environment.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 20 of 27 was completed at 28.08.2013 15:37:52
Received at the EPO on Aug 28, 2013 15:39:36. Page 20 of 27

12. The apparatus of claim 1, wherein the rendering involves controlling a spread in one or more of three dimensions of an audio object over multiple reproduction speakers.

13. The apparatus of claim 1, wherein the rendering involves dynamic object blobbing in response to speaker overload, by directing audio energy to an increased number of neighboring reproduction speakers while maintaining overall constant energy.

14. The apparatus of claim 1, wherein the rendering involves mapping audio object positions to planes of speaker arrays of the actual reproduction environment.

15. The apparatus of claim 1, further comprising a memory device, wherein the interface system comprises an interface between the logic system and the memory device.

16. The apparatus of claim 1, wherein the interface system comprises a network interface.

17. The apparatus of claim 1, wherein the logic system is configured to determine whether to apply panning rules for an audio object position to a plurality of speaker locations or to map an audio object position to a single speaker location.

18. The apparatus of claim 17, wherein the logic system is configured to smooth transitions in speaker gains when transitioning from mapping an audio object position from a first single speaker location to a second single speaker location.

19. The apparatus of claim 17, wherein the logic system is configured to smooth transitions in speaker gains when transitioning between mapping an audio object position to a single speaker location and applying panning rules for the audio object position to the plurality of speaker locations.

20. The apparatus of any of claims 1–19, wherein the logic system is further configured to compute speaker gains corresponding to the plurality of speaker zones.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 21 of 27 was completed at 28.08.2013 15:38:08
Received at the EPO on Aug 28, 2013 15:39:36. Page 21 of 27
AMENDED SHEET

21. The apparatus of any of claim 20, wherein the logic system is further configured to compute speaker gains for audio object positions along a one-dimensional curve between virtual speaker positions.

22. A method, comprising:

receiving audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations;

receiving reproduction environment data comprising an indication of a number of reproduction speakers in an actual reproduction environment and an indication of the location of each reproduction speaker of the three-dimensional actual reproduction environment;

mapping the audio reproduction data created with reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and

rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

23. The method of claim 22, wherein the actual reproduction environment comprises a sound system environment of a cinema.

24. The method of claim 22, wherein the rendering involves creating a gain based on one or more of a desired audio object position, a distance from the desired audio object

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 22 of 27 was completed at 28.08.2013 15:38:25
Received at the EPO on Aug 28, 2013 15:39:36. Page 22 of 27
AMENDED SHEET

position to a reference position, a velocity of an audio object or an audio object content type.

25. The method of claim 22, wherein the rendering involves imposing speaker zone constraints which include data for disabling selected reproduction speakers.

26. A non-transitory medium having software stored thereon, the software including instructions for performing the following operations:

receiving audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations;

receiving reproduction environment data comprising an indication of a number of reproduction speakers in an actual reproduction environment and an indication of the location of each reproduction speaker of the three-dimensional actual reproduction environment;

mapping the audio reproduction data created with reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and

rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

27. The non-transitory medium of claim 26, wherein the actual reproduction environment comprises a sound system environment of a cinema.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 23 of 27 was completed at 28.08.2013 15:38:41
Received at the EPO on Aug 28, 2013 15:39:36. Page 23 of 27

28. The non-transitory medium of claim 26, wherein the rendering involves creating a gain based on one or more of a desired audio object position, a distance from the desired audio object position to a reference position, a velocity of an audio object or an audio object content type.

29. The non-transitory medium of claim 26, wherein the rendering involves imposing speaker zone constraints which include data for disabling selected reproduction speakers.

30. The non-transitory medium of claim 26, wherein the rendering involves dynamic object blobbing in response to speaker overload, by directing audio energy to an increased number of neighboring reproduction speakers while maintaining overall constant energy.

31. An apparatus (2100) for authoring an audio object, the apparatus (2100) comprising:

an interface system (2105);

a user input system (2135);

a display system (2130); and

a logic system (2110) configured for:

receiving audio data via the interface system;

displaying a virtual reproduction environment in a graphical user interface on the display system (2130); wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations;

receiving user input regarding a position of an audio object via the user input system;

determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space according to the user input received via the user input system, wherein the determining involves constraining the time-variable position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio data;

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 24 of 27 was completed at 28.08.2013 15:38:57
Received at the EPO on Aug 28, 2013 15:39:36. Page 24 of 27
AMENDED SHEET

displaying an audio object trajectory according to the trajectory data in the graphical user interface; and

creating metadata associated with the audio data object; wherein the metadata includes the trajectory data.

32. The apparatus of claim 31, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

33. The apparatus of claim 31, wherein the trajectory data comprises a set of positions within the three-dimensional space at multiple time instances.

34. The apparatus of claim 31, wherein the trajectory data comprises an initial position, velocity data and acceleration data.

35. The apparatus of claim 31, wherein the trajectory data comprises an initial position and an equation that defines positions in three-dimensional space and corresponding times.

36. The apparatus of claim 31, further comprising a sound reproduction system, wherein the logic system is configured to control the sound reproduction system, at least in part, according to the metadata.

37. The apparatus of claim 31, wherein the plurality of speaker zones correspond to reproduction speakers of an actual three-dimensional reproduction environment comprising reproduction speakers, or wherein the plurality of speaker zones correspond to virtual speakers of a virtual surround sound environment.

38. The apparatus of claim 31, wherein an increased elevation of the audio object is indicated in the graphical user interface by an increase in a diameter of a circle that represents the audio object in the graphical user interface.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 25 of 27 was completed at 28.08.2013 15:39:12
Received at the EPO on Aug 28, 2013 15:39:36. Page 25 of 27
AMENDED SHEET

39. A method, comprising:

receiving audio data;

displaying a virtual reproduction environment in a graphical user interface on a display system; wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations;

receiving user input regarding a position of an audio object;

determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space, wherein the determining involves constraining the position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio data;

displaying an audio object trajectory according to the trajectory data in the graphical user interface; and

creating metadata associated with the audio object; wherein the metadata includes the trajectory data.

40. The method of claim 39, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

41. A non-transitory medium having software stored thereon, the software including instructions for performing the following operations:

receiving audio data;

displaying a virtual reproduction environment in a graphical user interface on a display system; wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations;

receiving user input regarding a position of an audio object;

determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space, wherein the determining involves constraining the position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio data;

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 26 of 27 was completed at 28.08.2013 15:39:28
Received at the EPO on Aug 28, 2013 15:39:36. Page 26 of 27
AMENDED SHEET

displaying an audio object trajectory according to the trajectory data in the graphical user interface; and

creating metadata associated with the audio object; wherein the metadara includes the trajectory data.

42. The non-transitory medium of claim 41, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

uration: 28.08.2013 15:32:33 - 28.08.2013 15:39:36. This page 27 of 27 was completed at 28.08.2013 15:39:36
Received at the EPO on Aug 28, 2013 15:39:36. Page 27 of 27
AMENDED SHEET

What is claimed is:

1. An apparatus, comprising: an interface system; and a logic system configured for: receiving, via the interface system, audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations; receiving, via the interface system, reproduction environment data comprising an indication of a number of reproduction speakers of an actual three-dimensional reproduction environment and an indication of the location of each reproduction speaker within the actual reproduction environment; mapping the audio reproduction data created with reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

2. The apparatus of claim 1, wherein the rendering involves creating a gain based on one or more of a desired audio object position, a distance from the desired audio object position to a reference position, a velocity of an audio object or an audio object content type.

3. The apparatus of claim 1, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

4. The apparatus of claim 1, wherein the rendering involves imposing speaker zone constraints including data for disabling selected reproduction speakers.

5. The apparatus of claim 1, wherein the actual reproduction environment comprises a screen for the projection of video images; wherein the audio reproduction data is synchronized with the video images; and wherein the rendering involves applying screen-to-room balance control according to screen-to-room balance control data received from the user input system.

6. The apparatus of claim 1, wherein the rendering involves controlling a spread in one or more of three dimensions of an audio object over multiple reproduction speakers.

7. The apparatus of claim 1, wherein the rendering involves dynamic object blobbing in response to speaker overload, by directing audio energy to an increased number of neighboring reproduction speakers while maintaining overall constant energy.

8. The apparatus of claim 1, wherein the rendering involves mapping audio object positions to planes of speaker arrays of the actual reproduction environment.

9. The apparatus of claim 1, wherein the logic system is configured to determine whether to apply panning rules for an audio object position to a plurality of speaker locations or to map an audio object position to a single speaker location.

10. The apparatus of claim 9, wherein the logic system is configured to smooth transitions in speaker gains when transitioning from mapping an audio object position from a first single speaker location to a second single speaker location.

11. The apparatus of claim 9, wherein the logic system is configured to smooth transitions in speaker gains when transitioning between mapping an audio object position to a single speaker location and applying panning rules for the audio object position to the plurality of speaker locations.

12. A method, comprising: receiving audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations; receiving reproduction environment data comprising an indication of a number of reproduction speakers in an actual reproduction environment and an indication of the location of each reproduction speaker of the three-dimensional actual reproduction environment; mapping the audio reproduction data created with

reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

13. A non-transitory medium having software stored thereon, the software including instructions for performing the following operations: receiving audio reproduction data comprising one or more audio objects and associated metadata; wherein the associated metadata includes trajectory data for at least one of the one or more audio objects indicating a time-variable audio object position of the at least one audio object within the three-dimensional space; wherein the audio object position is constrained to a two-dimensional surface; wherein the audio reproduction data has been created with respect to a virtual reproduction environment comprising a plurality of speaker zones at different elevations; receiving reproduction environment data comprising an indication of a number of reproduction speakers in an actual reproduction environment and an indication of the location of each reproduction speaker of the three-dimensional actual reproduction environment; mapping the audio reproduction data created with reference to the plurality of speaker zones of the virtual reproduction environment to the reproduction speakers of the actual reproduction environment; and rendering the one or more audio objects into one or more speaker feed signals based, at least in part, on the associated metadata, wherein each speaker feed signal corresponds to at least one of the reproduction speakers within the actual reproduction environment.

14. An apparatus (2100) for authoring an audio object, the apparatus (2100) comprising: an interface system (2105); a user input system (2135); a display system (2130); and a logic system (2110) configured for: receiving audio data via the interface system; displaying a virtual reproduction environment in a graphical user interface on the display system (2130); wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations; receiving user input regarding a position of an audio object via the user input system; determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space according to the user input received via the user input system, wherein the determining involves constraining the time-variable position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio

data; displaying an audio object trajectory according to the trajectory data in the graphical user interface; and creating metadata associated with the audio data object; wherein the metadata includes the trajectory data.

15. The apparatus of claim 14, wherein the two-dimensional surface comprises one of a spherical surface, an elliptical surface, a conical surface, a cylindrical surface or a wedge.

16. The apparatus of claim 14, wherein the trajectory data comprises a set of positions within the three-dimensional space at multiple time instances.

17. The apparatus of claim 14, wherein the plurality of speaker zones correspond to reproduction speakers of an actual three-dimensional reproduction environment comprising reproduction speakers, or wherein the plurality of speaker zones correspond to virtual speakers of a virtual surround sound environment.

18. The apparatus of claim 14, wherein an increased elevation of the audio object is indicated in the graphical user interface by an increase in a diameter of a circle that represents the audio object in the graphical user interface.

19. A method, comprising: receiving audio data; displaying a virtual reproduction environment in a graphical user interface on a display system; wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations; receiving user input regarding a position of an audio object; determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space, wherein the determining involves constraining the position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio data; displaying an audio object trajectory according to the trajectory data in the graphical user interface; and creating metadata associated with the audio object; wherein the metadata includes the trajectory data.

20. A non-transitory medium having software stored thereon, the software including instructions for performing the following operations: receiving audio data; displaying a virtual reproduction environment in a graphical user interface on a display system; wherein the virtual reproduction environment comprises a plurality of speaker zones at different elevations; receiving user input regarding a position of an audio

object; determining trajectory data indicating a time-variable position of the audio object in a three-dimensional space, wherein the determining involves constraining the position to a two-dimensional surface within the three-dimensional space; wherein the audio object comprises the audio data; displaying an audio object trajectory according to the trajectory data in the graphical user interface; and creating metadata associated with the audio object; wherein the metadata includes the trajectory data.