

CLAIMS

1. A method for synthesizing sound, comprising:
receiving one or more first inputs, each first input comprising a respective monaural audio track;
receiving one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs;
assigning to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and elevation coordinates of the respective 3D source locations; and
synthesizing left and right stereo output signals by applying the respective left and right filter responses to the first inputs, wherein the one or more first inputs comprise a first plurality of audio input tracks, and wherein synthesizing the left and right stereo output signals comprises: spatially upsampling the first plurality of the input audio tracks in order to generate a second plurality of synthesized inputs, having synthesized 3D source locations with respective coordinates different from the respective 3D source locations associated with the first inputs; filtering the synthesized inputs using the filter response functions computed at the azimuth and elevation coordinates of the synthesized 3D source locations; and after filtering the first inputs using the respective left and right filter responses, summing the filtered synthesized inputs with the filtered first inputs to produce the stereo output signals.
2. The method according to claim 1, wherein the one or more first inputs comprise a plurality of first inputs, and wherein synthesizing the left and right stereo output signals comprises applying the respective left and right filter responses to each of the first inputs to generate respective left and right stereo components, and summing the left and right stereo components over all of the first inputs.
3. The method according to claim 2, wherein summing the left and right stereo components comprises applying a limiter to the summed components in order to prevent clipping upon playback of the output signals.
4. The method according to claim 1, wherein at least one of the second inputs specifies a 3D trajectory in space, and
wherein assigning the left and right filter responses comprises specifying, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and
wherein synthesizing the left and right stereo output signals comprises sequentially applying to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory.
5. The method according to claim 1, wherein the filter response functions comprise a notch at a given frequency, which varies as a function of the elevation coordinates.
6. The method according to claim 1, wherein spatially upsampling the first plurality of the input audio tracks comprises applying a wavelet transform to the input audio tracks to generate respective spectrograms of

the input audio tracks, and interpolating between the spectrograms according to the 3D source locations to generate the synthesized inputs.

7. The method according to claim 6, wherein interpolating between the spectrograms comprises computing an optical flow function between points in the spectrograms.

8. The method according to claim 1, wherein synthesizing the left and right stereo output signals comprises extracting low-frequency components from the first inputs, and wherein applying the respective left and right filter responses comprises filtering the first inputs after extraction of the low-frequency components, and then adding the extracted low-frequency components to the filtered first inputs.

9. The method according to claim 1, where the 3D source locations have range coordinates that are to be associated with the first inputs, and wherein synthesizing the left and right stereo outputs comprises further modifying the first inputs responsively to the associated range coordinates.

10. A method for synthesizing sound, comprising: receiving one or more first inputs, each first input comprising a respective monaural audio track; receiving one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs; assigning to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and elevation coordinates of the respective 3D source locations; and synthesizing left and right stereo output signals by applying the respective left and right filter responses to the first inputs, wherein at least one of the second inputs specifies a 3D trajectory in space, and wherein assigning the left and right filter responses comprises specifying, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and wherein synthesizing the left and right stereo output signals comprises sequentially applying to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory, and wherein receiving the one or more second inputs comprises: receiving a start point and a start time of the trajectory; receiving an end point and an end time of the trajectory; and automatically computing the 3D trajectory between the start point and the end point such that the trajectory is traversed from the start time to the end time.

11. The method according to claim 5, wherein automatically computing the 3D trajectory comprises calculating a path over a surface of a sphere that is centered at an origin of the azimuth and elevation coordinates.

12. Apparatus for synthesizing sound, comprising: an input interface configured to receive one or more first inputs, each first input comprising a respective monaural audio track, and to receive one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs; and a processor, which is configured to assign to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and

elevation coordinates of the respective 3D source locations, and to synthesize left and right stereo output signals by applying the respective left and right filter responses to the first inputs, wherein the one or more first inputs comprise a first plurality of audio input tracks, and wherein the processor is configured to spatially upsample the first plurality of the input audio tracks in order to generate a second plurality of synthesized inputs, having synthesized 3D source locations with respective coordinates different from the respective 3D source locations associated with the first inputs, to filter the synthesized inputs using the filter response functions computed at the azimuth and elevation coordinates of the synthesized 3D source locations, and to sum the filtered synthesized inputs with the filtered first inputs to produce the stereo output signals.

13. The apparatus according to claim 12, and comprising an audio output interface, comprising left and right speakers, which are configured to play back the left and right stereo output signals, respectively.

14. The apparatus according to claim 12, wherein the one or more first inputs comprise a plurality of first inputs, and wherein the processor is configured to apply the respective left and right filter responses to each of the first inputs to generate respective left and right stereo components, and to sum the left and right stereo components over all of the first inputs.

15. The apparatus according to claim 14, wherein the processor is configured to apply a limiter to the summed components in order to prevent clipping upon playback of the output signals.

16. The apparatus according to claim 12, wherein at least one of the second inputs specifies a 3D trajectory in space, and wherein the processor is configured to specify, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and to sequentially apply to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory.

17. The apparatus according to claim 12, wherein the filter response functions comprise a notch at a given frequency, which varies as a function of the elevation coordinates.

18. The apparatus according to claim claim 12, wherein the processor is configured to spatially upsample the first plurality of the input audio tracks by applying a wavelet transform to the input audio tracks to generate respective spectrograms of the input audio tracks, and interpolating between the spectrograms according to the 3D source locations to generate the synthesized inputs.

19. The apparatus according to claim 18, wherein the processor is configured to interpolate between the spectrograms using an optical flow function computed between points in the spectrograms.

20. The apparatus according to claim 12, wherein the processor is configured to extract low-frequency components from the first inputs, to apply the respective left and right filter responses to the first inputs after extraction of the low-frequency components, and then to add the extracted low-frequency components to the

filtered first inputs.

21. The apparatus according to claim 12, where the 3D source locations have range coordinates that are to be associated with the first inputs, and wherein the processor is configured to further modify the first inputs responsively to the associated range coordinates.

22. Apparatus for synthesizing sound, comprising: an input interface configured to receive one or more first inputs, each first input comprising a respective monaural audio track, and to receive one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs; and a processor, which is configured to assign to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and elevation coordinates of the respective 3D source locations, and to synthesize left and right stereo output signals by applying the respective left and right filter responses to the first inputs, wherein at least one of the second inputs specifies a 3D trajectory in space, and wherein the processor is configured to specify, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and to sequentially apply to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory, and wherein the processor is configured to receive a start point and a start time of the trajectory and an end point and an end time of the trajectory, and to automatically compute the 3D trajectory between the start point and the end point such that the trajectory is traversed from the start time to the end time.

23. The apparatus according to claim 22, wherein the 3D trajectory comprises a path over a surface of a sphere that is centered at an origin of the azimuth and elevation coordinates.

24. A computer software product, comprising a non-transitory computer-readable medium in which program instructions are stored, which instructions, when read by a computer, cause the computer to receive one or more first inputs, each first input comprising a respective monaural audio track, and to receive one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs, wherein the instructions cause the computer to assign to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and elevation coordinates of the respective 3D source locations, and to synthesize left and right stereo output signals by applying the respective left and right filter responses to the first inputs, and wherein the one or more first inputs comprise a first plurality of audio input tracks, and wherein the instructions cause the computer to spatially upsample the first plurality of the input audio tracks in order to generate a second plurality of synthesized inputs, having synthesized 3D source locations with respective coordinates different from the respective 3D source locations associated with the first inputs, to filter the synthesized inputs using the filter response functions computed at the azimuth and elevation coordinates of the synthesized 3D source locations, and to sum the filtered synthesized inputs with the filtered first inputs to produce the stereo output signals.

25. The product according to claim 24, wherein the one or more first inputs comprise a plurality of first

inputs, and wherein the instructions cause the computer to apply the respective left and right filter responses to each of the first inputs to generate respective left and right stereo components, and to sum the left and right stereo components over all of the first inputs.

26. The product according to claim 25, wherein the instructions cause the computer to apply a limiter to the summed components in order to prevent clipping upon playback of the output signals.

27. The product according to claim 24, wherein at least one of the second inputs specifies a 3D trajectory in space, and wherein the instructions cause the computer to specify, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and to sequentially apply to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory.

28. The product according to claim 24, wherein the filter response functions comprise a notch at a given frequency, which varies as a function of the elevation coordinates.

29. The product according to claim 24, wherein the instructions cause the computer to spatially upsample the first plurality of the input audio tracks by applying a wavelet transform to the input audio tracks to generate respective spectrograms of the input audio tracks, and interpolating between the spectrograms according to the 3D source locations to generate the synthesized inputs.

30. The product according to claim 29, wherein the instructions cause the computer to interpolate between the spectrograms using an optical flow function computed between points in the spectrograms.

31. The product according to claim 24, wherein the instructions cause the computer to extract low-frequency components from the first inputs, to apply the respective left and right filter responses to the first inputs after extraction of the low-frequency components, and then to add the extracted low-frequency components to the filtered first inputs.

32. The product according to claim 24, where the 3D source locations have range coordinates that are to be associated with the first inputs, and wherein the instructions cause the computer to further modify the first inputs responsively to the associated range coordinates.

33. A computer software product, comprising a non-transitory computer-readable medium in which program instructions are stored, which instructions, when read by a computer, cause the computer to receive one or more first inputs, each first input comprising a respective monaural audio track, and to receive one or more second inputs indicating respective three-dimensional (3D) source locations having azimuth and elevation coordinates to be associated with the first inputs, wherein the instructions cause the computer to assign to each of the first inputs respective left and right filter responses based on filter response functions that depend upon the azimuth and elevation coordinates of the respective 3D source locations, and to synthesize left and right

stereo output signals by applying the respective left and right filter responses to the first inputs, and wherein at least one of the second inputs specifies a 3D trajectory in space, and wherein the instructions cause the computer to specify, at each of a plurality of points along the 3D trajectory, filter responses that vary over the trajectory responsively to the azimuth and elevation coordinates of the points, and to sequentially apply to the first input that is associated with the at least one of the second inputs the filter responses that are specified for the points along the 3D trajectory, and wherein the instructions cause the computer to receive a start point and a start time of the trajectory and an end point and an end time of the trajectory, and to automatically compute the 3D trajectory between the start point and the end point such that the trajectory is traversed from the start time to the end time.

34. The product according to claim 33, wherein the 3D trajectory comprises a path over a surface of a sphere that is centered at an origin of the azimuth and elevation coordinates.