CLAIMS

1. An apparatus for outputting virtual sound, the apparatus comprising: one or more microphones configured to detect an audio signal in an environment; and a processor coupled to the one or more microphones, the processor configured to: determine a sound characteristic of the audio signal; determine, based on the sound characteristic, whether a sound source of the audio signal is located in a first zone of the environment or a second zone of the environment, the first zone closer to the one or more microphones than the second zone; estimate a direction-of-arrival of the audio signal; estimate a location of the sound source based on the direction-of-arrival and at least one of the zones associated with the sound source; and insert a virtual sound into the environment, the virtual sound having one or more audio properties of a sound generated from the location of the sound source.

2. The apparatus of claim 1, wherein the processor is further configured to generate the virtual sound prior to insertion of the virtual sound into the environment.

3. The apparatus of claim 1, wherein the processor is further configured to: compare one or more acoustical characteristics of the environment to one or more entries in a memory, each entry associated with a different virtual sound; and retrieve the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

4. The apparatus of claim 1, wherein the processor is further configured to: classify the sound source as a particular object based on the audio signal; generate a virtual image of the particular object; and insert the virtual image at the location of the sound source in the environment.

5. The apparatus of claim 1, further comprising a camera coupled to the processor, the camera configured to capture the environment.

6. The apparatus of claim 5, wherein the camera includes a security camera.

7. The apparatus of claim 5, wherein the camera, the one or more microphones, and the processor are integrated into a headset.

8. The apparatus of claim 7, wherein the headset is configured to display a visual depiction of the environment.

9. The apparatus of claim 1, wherein the sound characteristic comprises at least one reverberation characteristic of the audio signal.

10. The apparatus of claim 9, wherein the at least one reverberation characteristic comprises a direct-to-reverberation ratio of the audio signal.

11. The apparatus of claim 1, wherein the processor is configured to: spatially filter a digital representation of a particular sound based on the one or more audio properties to generate a spatially-filtered audio file, the particular sound corresponding to the virtual sound; and send a spatially-filtered audio signal to a speaker, the spatially-filtered audio signal based on the spatially-filtered audio file, wherein the speaker projects the spatially-filtered audio signal as a spatially-filtered sound to insert the virtual sound into the environment.

12. The apparatus of claim 1, wherein the one or more audio properties correspond to a volume level of the sound generated from the location of the sound source, a direction-of-arrival of the sound generated from the location of the sound source, or both.

13. A method for outputting virtual sound, the method comprising: detecting an audio signal in an environment at one or more microphones; determining, at a processor, a sound characteristic of the audio signal; determining, based on the sound characteristic, whether a sound source of the audio signal is located in a first zone of the environment or a second zone of the environment, the first zone closer to the one or more microphones than the second zone; estimating a direction-of-arrival of the audio signal; estimating a location of the sound source based on the direction-of-arrival and at least one of the zones associated with the sound source; and inserting a virtual sound into the environment, the virtual sound having one or more audio properties of a sound generated from the location of the sound source.

14. The method of claim 13, further comprising generating the virtual sound prior to insertion of the virtual sound into the environment.

15. The method of claim 13, further comprising: comparing one or more acoustical characteristics of the environment to one or more entries in a memory, each entry associated with a different virtual sound; and retrieving the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

16. The method of claim 13, further comprising: classifying the sound source as a particular object based on the audio signal; generating a virtual image of the particular object; and inserting the virtual image at the location of the sound source in the environment.

17. The method of claim 13, further comprising capturing the environment via a camera.

18. The method of claim 17, wherein the camera includes a security camera.

19. The method of claim 17, wherein the camera, the one or more microphones, and the processor are integrated into a headset.

20. The method of claim 19, further comprising displaying a visual depiction of the environment at the headset.

21. The method of claim 13, wherein the sound characteristic comprises at least one

reverberation characteristic of the audio signal.

22. The method of claim 21, wherein the at least one reverberation characteristic comprises a direct-to-reverberation ratio of the audio signal.

23. The method of claim 13, wherein inserting the virtual sound into the environment comprises: spatially filtering a digital representation of a particular sound based on the one or more audio properties to generate a spatially-filtered audio file, the particular sound corresponding to the virtual sound; and sending a spatially-filtered audio signal to a speaker, the spatially-filtered audio signal based on the spatially-filtered audio file, wherein the speaker projects the spatially-filtered audio signal as a spatially-filtered sound to insert the virtual sound into the environment.

24. A non-transitory computer-readable medium comprising instructions for outputting virtual sound, the instructions, when executed by a processor, cause the processor to perform operations comprising: determining a sound characteristic of an audio signal in an environment, the audio signal detected at one or more microphones; determining, based on the sound characteristic, whether a sound source of the audio signal is located in a first zone of the environment or a second zone of the environment, the first zone closer to the one or more microphones than the second zone; estimating a direction-of-arrival of the audio signal; estimating a location of the sound source based on the direction-of-arrival and at least one of the zones associated with the sound source; and inserting a virtual sound into the environment, the virtual sound having one or more audio properties of a sound generated from the location of the sound source.

25. The non-transitory computer-readable medium of claim 24, wherein the operations further comprise generating the virtual sound prior to insertion of the virtual sound into the environment.

26. The non-transitory computer-readable medium of claim 24, wherein the operations further comprise: comparing one or more acoustical characteristics of the environment to one or more entries in a memory, each entry associated with a different virtual sound; and retrieving the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

27. The non-transitory computer-readable medium of claim 24, wherein the operations further comprise: classifying the sound source as a particular object based on the audio signal; generating a virtual image of the particular object; and inserting the virtual image at the location of the sound source in the environment.

28. The non-transitory computer-readable medium of claim 24, wherein the operations further comprise causing a camera to capture the environment.

29. The non-transitory computer-readable medium of claim 28, wherein the camera includes

a security camera.

30. The non-transitory computer-readable medium of claim 24, wherein inserting the virtual sound into the environment comprises: spatially filtering a digital representation of a particular sound based on the one or more audio properties to generate a spatially-filtered audio file, the particular sound corresponding to the virtual sound; and sending a spatially-filtered audio signal to a speaker, the spatially-filtered audio signal based on the spatially-filtered audio file, wherein the speaker projects the spatially-filtered audio signal as a spatially-filtered sound to insert the virtual sound into the environment.

31. An apparatus for outputting virtual sound, the apparatus comprising: means for detecting an audio signal in an environment; means for determining a sound characteristic of the audio signal; means for determining, based on the sound characteristic, whether a sound source of the audio signal is located in a first zone of the environment or a second zone of the environment, the first zone closer to the means for detecting than the second zone; means for estimating a direction-of-arrival of the audio signal; means for estimating a location of the sound source based on the direction-of-arrival and at least one of the zones associated with the sound source; and means for inserting a virtual sound into the environment, the virtual sound having one or more audio properties of a sound generated from the location of the sound source.

32. The apparatus of claim 31, further comprising: means for classifying the sound source as a particular object based on the audio signal; means for generating a virtual image of the particular object; and means for inserting the virtual image at the location of the sound source in the environment.

33. The apparatus of claim 31, wherein inserting the virtual sound into the environment comprises: spatially filtering a digital representation of a particular sound based on the one or more audio properties to generate a spatially-filtered audio file, the particular sound corresponding to the virtual sound; and sending a spatially-filtered audio signal to a speaker, the spatially-filtered audio signal based on the spatially-filtered audio file, wherein the speaker projects the spatially-filtered audio signal as a spatially-filtered sound to insert the virtual sound into the environment.

34. An apparatus for outputting sound, the apparatus comprising: one or more speakers configured to generate audio signals at different locations of an environment; one or more microphones configured to detect each audio signal; a memory; and a processor coupled to the memory, to the one or more speakers, and to the one or more microphones, the processor configured to: determine one or more acoustical characteristics at one or more locations of the environment based on detected audio signals generated by the one or more speakers; receive a user selection indicating a particular object; receive a user indication of a particular location; determine one or more audio properties of a sound generated from the particular location based on the one or more acoustical characteristics; and insert a virtual sound associated with the particular object into the environment, the virtual sound having the one or

more audio properties.

35. The apparatus of claim 34, wherein the processor is further configured to generate the virtual sound prior to insertion of the virtual sound into the environment.

36. The apparatus of claim 34, wherein the processor is further configured to: compare the one or more acoustical characteristics to one or more entries in a memory, each entry associated with a different virtual sound; and retrieve the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

37. The apparatus of claim 34, further comprising a wearable sensor coupled to the processor, the wearable sensor configured to: detect the particular location; and generate the user indication of the particular location; and send the user indication of the particular location to the processor.

38. The apparatus of claim 37, wherein the wearable sensor is further configured to be worn at a user's hand.

39. The apparatus of claim 34, wherein the one or more speakers are further configured to be worn at a user's hand.

40. A method for outputting sound, the method comprising: generating audio signals at different locations of an environment using one or more speakers; detecting each audio signal using one or more microphones; determining one or more acoustical characteristics at one or more locations of the environment based on detected audio signals generated by the one or more speakers; receiving a user selection indicating a particular object; receiving a user indication of a particular location; determining one or more audio properties of a sound generated from the particular location based on the one or more acoustical characteristics; and inserting a virtual sound associated with the particular object into the environment, the virtual sound having the one or more audio properties.

41. The method of claim 40, further comprising generating the virtual sound prior to insertion of the virtual sound into the environment.

42. The method of claim 40, further comprising: comparing the one or more acoustical characteristics to one or more entries in a memory, each entry associated with a different virtual sound; and retrieving the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

43. The method of claim 40, further comprising: detecting the particular location using a wearable sensor; generating, at the wearable sensor, the user indication of the particular location; and sending the user indication of the particular location from the wearable sensor to a processor.

44. The method of claim 43, wherein the wearable sensor is configured to be worn at a user's hand.

45. The method of claim 40, wherein the one or more speakers are configured to be worn at a user's hand.

46. A non-transitory computer-readable medium comprising instructions for outputting sound, the instructions, when executed by a processor, cause the processor to perform operations comprising: initiating generation of audio signals at different locations of an environment using one or more speakers; detecting each audio signal using one or more microphones; determining one or more acoustical characteristics at one or more locations of the environment based on detected audio signals generated by the one or more speakers; receiving a user selection indicating a particular object; receiving a user indication of a particular location; determining one or more audio properties of a sound generated from the particular location based on the one or more acoustical characteristics; and inserting a virtual sound associated with the particular object into the environment, the virtual sound having the one or more audio properties.

47. The non-transitory computer-readable medium of claim 46, wherein the operations further comprise generating the virtual sound prior to insertion of the virtual sound into the environment.

48. The non-transitory computer-readable medium of claim 46, wherein the operations further comprise: comparing the one or more acoustical characteristics to one or more entries in a memory, each entry associated with a different virtual sound; and retrieving the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

49. The non-transitory computer-readable medium of claim 46, wherein the user indication of the particular location is received from a wearable sensor.

50. The non-transitory computer-readable medium of claim 49, wherein the wearable sensor is configured to be worn at a user's hand.

51. The non-transitory computer-readable medium of claim 46, wherein the one or more speakers are further configured to be worn at a user's hand.

52. An apparatus for outputting sound, the apparatus comprising: means for generating audio signals at different locations of an environment; means for detecting each audio signal; means for determining one or more acoustical characteristics at one or more locations of the environment based on detected audio generated by means for generating; means for receiving a user selection indicating a particular object; means for receiving a user indication of a particular location; means for determining one or more audio properties of a sound generated from the particular location based on the one or more acoustical characteristics; and means

for inserting a virtual sound associated with the particular object into the environment, the virtual sound having the one or more audio properties.

53. The apparatus of claim 52, further comprising means for generating the virtual sound prior to insertion of the virtual sound into the environment.

54. The apparatus of claim 52, further comprising: means for comparing the one or more acoustical characteristics to one or more entries in a memory, each entry associated with a different virtual sound; and means for retrieving the virtual sound from a particular entry based on the comparison prior to insertion of the virtual sound into the environment.

55. The apparatus of claim 52, further comprising means for detecting the particular location, generating the user indication of the particular location, and sending the user indication of the particular location to the means for receiving the user indication.

56. The apparatus of claim 55, wherein the means for detecting the particular location is configured to be worn at a user's hand.

57. The apparatus of claim 52, wherein the means for generating the audio signals is configured to be worn at a user's hand.