CLAIMS

What is claimed is:

1. An audio processing device for processing an event by using a spatial audio position transfer function dataset, the device comprising:

an audio rendering module configured to position 1st and 2nd audio signals respectively comprising at least a voice communication stream and a media stream in selected ones of at least a first spatial audio position and a second spatial audio position, each of the first and second spatial audio positions rendered by using respective first and second transfer functions from the spatial audio position transfer function dataset;

a monitoring module to monitor initiation of a voice communication event, the event comprising receipt of a phone call, and upon initiation of the phone call, processing the first and second audio signals by positioning the voice communication to the first spatial audio position and the media stream to the second spatial position; and

an output module configured for rendering the resultant audio over two output channels to a coupled pair of headphones.

2. The audio processing device as recited in claim 1 wherein the spatial audio position transfer function dataset is one of an individualized Head Related Impulse Response "HRIR" dataset or an individualizd Binaural Room Impulse Response "BRIR" dataset which are datasets customized for an individual.

3. The audio processing device as recited in claim 2 further comprising a second processor configured to extract image based properties for the individual from an input image and for transmitting the image based properties to a selection processor configured to determine the individualized HRIR or BRIR datasets from a memory having a candidate pool of a plurality of HRIR or BRIR datasets that have been provided for a population of individuals, the HRIR or BRIR datasets each associated with their corresponding image based properties.

4. The audio processing device as recited in claim 3 wherein the selection processor determines the individualized BRIR dataset by accessing the candidate pool by comparing the extracted image based properties for the individual against the extracted properties for the candidate pool, to identify one or more BRIR datasets based on a closeness metric and the

processing strategies used are one of multiple match; multiple-recognizer types; and cluster based.

5. The audio processing device as recited in claim 2 wherein the first and second spatial audio positions from the determined individualized BRIR dataset are derived by interpolation or other computation methods from a captured dataset in the memory and wherein the first and second spatial audio positions comprise respectively foreground and background positions.

6. The audio processing device as recited in claim 5 wherein upon determination by the individual listener that the voice call is of a lower priority and generation of a corresponding control signal, the voice call is directed to the background position and the music to the foreground position.

7. The audio processing device as recited in claim 2 wherein upon determination by the individual listener that the voice call is of a lower priority and generation of a corresponding control signal, the apparent distance of the voice call is increased and the apparent distance of the music is decreased using individualized BRIRs corresponding to different distances for the same directions.

8. The audio processing device as recited in claim 2 wherein the positioning of the voice communication to the first spatial audio position and the media stream to the second spatial audio position from their respective initial positions is performed in an abrupt manner.

9. The audio processing device as recited in claim 2 further including a portable image capturing device configured for acquiring the input image and wherein the audio processing device is one of a mobile phone, communication device, or tablet that acquires the image and extracts the image-based properties.

10. The audio processing device as recited in claim 1 wherein the audio processing device is configured to reposition the media stream to the first spatial audio position upon termination of the voice communication stream.

11. The audio processing device as recited in claim 1 wherein the media stream comprises music.

12. The audio processing device as recited in claim 1 wherein the apparent distance of the voice call is increased and the apparent distance of the music is decreased using respective

first and second spatial audio position sound transfer functions from the individualized BRIRs corresponding to different distances for the same directions.

13. The audio processing device as recited in claim 1 wherein the output module is coupled to the headphone via one of a wireless connection and a wired connection.

14. The audio processing device as recited in claim 1 wherein the output module includes a digital to analog converter and the coupling to the headphones is through an analog port.

15. The audio processing device as recited in claim 1 wherein the output module is configured to pass a digital signal to the headphones and the headphones include a digital to analog converter.

16. The audio processing device as recited in claim 1 further comprising a user interface configured for choosing a location for at least one of the first spatial audio position and the second spatial audio position.

17. A method for processing audio streams to a set of headphones comprising:

positioning 1st and 2nd audio signals respectively comprising at least a voice communication stream and a media stream in selected ones of at least a first spatial audio position and a second spatial audio position, each of the first and second spatial audio positions rendered by using respective first and second transfer functions from a spatial audio position transfer function dataset;

monitoring initiation of a voice communication event, the event comprising receipt of a phone call, and upon initiation of the phone call processing the first and second audio signals by positioning the voice communication to the first spatial audio position and the media stream to the second spatial audio position wherein at least an associated room impulse response is present for the second spatial audio position; and

rendering the resultant audio over two output channels to a coupled pair of headphones.

18. The method as recited in claim 17 wherein the spatial audio position transfer function dataset is one of an HRIR dataset or a BRIR dataset customized for an individual.

19. The method as recited in claim 18 wherein the customization includes extracting image-based properties for the individual from an input image and transmitting the image based properties to a selection processor configured to determine an individualized HRIR or BRIR dataset from a memory having a candidate pool of a plurality of HRIR or BRIR datasets that

have been provided for a population of individuals, the HRIR or BRIR datasets each associated with their corresponding image based properties.

20. The method as recited in claim 19 wherein determining the individualized BRIR dataset comprises interpolation between existing BRIR datasets in the candidate pool.