

Claims

1. An apparatus for microphone positioning, comprising:

5 | a spatial power density determiner ~~(10; 21)~~ for determining a spatial power density indicating power values for a plurality of locations of an environment based on sound source information indicating one or more power values and one or more position values of one or more sound sources located in the environment, and

10 | a spatial information estimator ~~(20; 22)~~ for estimating a position of a microphone based on the spatial power density,

15 | wherein the spatial information estimator ~~(20; 22)~~ comprises a sound scene center estimator ~~(41)~~ for estimating a position of a center of a sound scene in the environment,

20 | wherein the spatial information estimator ~~(20; 22)~~ furthermore comprises a microphone position calculator ~~(42; 44)~~ for determining the position of the microphone based on the position of the center of the sound scene,

25 | wherein the spatial information estimator ~~(20; 22)~~ comprises an orientation determiner ~~(45)~~ for determining an orientation of the microphone, wherein the orientation determiner ~~(45)~~ is adapted to determine the orientation of the microphone based on the spatial power density,

30 | wherein the spatial power density determiner ~~(10; 21)~~ is adapted to determine the spatial power density by applying the formula

$$\Gamma(x, y, k, n) = \sum_{i=1}^N \text{power}_i(k, n) \cdot g(\gamma_i, x - x_{\text{ESS}i}, y - y_{\text{ESS}i}, k, n),$$

35 | to calculate the power values $\Gamma(x, y, k, n)$ for the plurality of locations in the environment for a time-frequency bin ~~(k, n)~~, wherein k denotes the frequency index and n denotes the time index, wherein N denotes a number of the sound sources, wherein x, y denote coordinates of one of the plurality of locations, wherein $\text{power}_i(k, n)$ denotes the power value at an i -th sound source for time-frequency bin ~~(k, n)~~, wherein $x_{\text{ESS}i}, y_{\text{ESS}i}$ denote coordinates of the i -th sound source, wherein γ_i is

a scalar value and wherein g is a function depending on x , y , x_{ESSi} , y_{ESSi} , k , n and γ_i , when the environment is a two-dimensional environment, or

wherein the spatial power density determiner ~~(10; 21)~~ is adapted to determine the spatial power density by applying the formula

$$\Gamma(x,y,z,k,n) = \sum_{i=1}^N \text{power}_i(k,n) \cdot g(\gamma_i, x - x_{ESSi}, y - y_{ESSi}, z - z_{ESSi}, k, n),$$

to calculate the power values $\Gamma(x,y,z,k,n)$ for the plurality of locations of the environment for a time-frequency bin ~~(k, n)~~, wherein k denotes the frequency index and n denotes the time index, wherein N denotes a number of the sound sources, wherein x , y , z denote coordinates of one of the plurality of locations, wherein $\text{power}_i(k,n)$ denotes the power value at an i -th sound source for time-frequency bin ~~(k, n)~~, wherein x_{ESSi} , y_{ESSi} , z_{ESSi} denote coordinates of the i -th sound source, wherein γ_i is a scalar value and wherein g is a function depending on x , y , z , x_{ESSi} , y_{ESSi} , z_{ESSi} , k , n and γ_i , when the environment is a three-dimensional environment.

2. An apparatus according to claim 1, wherein the orientation determiner ~~(45)~~ is adapted to determine the orientation of the microphone such that the microphone is oriented towards the center of the sound scene.

3. An apparatus according to claim 1 or 2, wherein the microphone position calculator ~~(42; 44)~~ is adapted to calculate the position of the microphone, wherein the microphone is a virtual spatial microphone.

4. An apparatus according to one of the preceding claims, wherein the sound scene center estimator ~~(41)~~ is adapted to calculate a center of gravity of the spatial power density for estimating the center of the sound scene.

5. An apparatus according to one of claims 1 to 3,

wherein the sound scene center estimator ~~(41)~~ is configured to determine a power delay profile based on the spatial power density and to determine a root mean squared delay based on the power delay profile for each one of a plurality of locations in the environment, and

wherein the sound scene center estimator ~~(41)~~ is configured to determine the location of the plurality of locations as the center of the sound scene, which has the minimum root mean squared delay of the root mean squared delays of the plurality of locations.

- 5
6. An apparatus according to one of the preceding claims,

wherein the microphone position calculator ~~(42; 44)~~ is adapted to determine a broadest-width line of a plurality of lines through the center of the sound scene in the environment,

wherein each of the plurality of lines through the center of the sound scene is associated with an energy width, and wherein the broadest-width line is defined as the line of the plurality of lines through the center of the sound scene having the largest energy width,

wherein the microphone position calculator ~~(42; 44)~~ is adapted determine the position of the microphone such that a second line, which passes through the center of the sound scene and the position of the microphone is orthogonal to the broadest-width line.

7. An apparatus according to claim 6, wherein energy width of a considered line of the plurality of lines indicates a largest width of a segment on the considered line, such that the first point of the segment limiting the segment, and such that a different second point of the segment limiting the segment, have both a power value indicated by the spatial power density, that is greater than or equal to a predefined power value.

8. An apparatus according to one of claims 1 to 5, wherein the microphone position calculator ~~(42; 44)~~ is configured to apply a singular value decomposition to a matrix having a plurality of columns,

wherein the columns of the matrix indicate positions of locations in the environment relative to the center of the sound scene, and

wherein the columns of the matrix only indicate the positions of locations having power values indicated by the spatial power density that are greater than a predefined threshold value, or the columns of the matrix only indicate the positions

of locations having power values indicated by the spatial power density that are greater than or equal to a predefined threshold value.

9. An apparatus ~~(991)~~ for generating a virtual output signal, comprising:

an apparatus for microphone positioning ~~(992)~~ according to one of claims 1 to 8, wherein the microphone position calculator ~~(993)~~ of the apparatus for microphone positioning is configured to calculate the position of a microphone as a calculated microphone position, and

an apparatus ~~(994)~~ for generating an audio output signal as the virtual output signal to simulate a recording of a virtual microphone at the calculated microphone position, wherein the apparatus for generating an audio output signal comprises:

a sound events position estimator ~~(110)~~ for estimating a sound source position indicating a position of a sound source in the environment, the sound source emitting a sound wave, wherein the sound events position estimator ~~(110)~~ is adapted to estimate the sound source position based on a first direction information provided by a first real spatial microphone being located at a first real microphone position in the environment, and based on a second direction information provided by a second real spatial microphone being located at a second real microphone position in the environment; and

an information computation module ~~(120)~~ for generating the audio output signal based on a first recorded audio input signal being recorded by the first real spatial microphone, based on the first real microphone position and based on the calculated microphone position.

10. A method for microphone positioning, comprising:

determining a spatial power density indicating power values for a plurality of locations of an environment based on sound source information indicating one or more power values and one or more position values of one or more sound sources located in the environment, and

estimating a position of a microphone based on the spatial power density, and

determining an orientation of the microphone,

wherein estimating the position of the microphone based on the spatial power density is conducted by estimating a position of a center of a sound scene in the environment, and by determining the position of the microphone based on the position of the center of the sound scene,

wherein the spatial information estimator ~~(20; 22)~~ comprises an orientation determiner ~~(45)~~ for determining an orientation of the microphone, wherein the orientation determiner ~~(45)~~ is adapted to determine the orientation of the microphone based on the spatial power density,

wherein the spatial power density determiner ~~(10; 21)~~ is adapted to determine the spatial power density by applying the formula

$$\Gamma(x,y,k,n) = \sum_{i=1}^N \text{power}_i(k,n) \cdot g(\gamma_i, x - x_{\text{ESS}i}, y - y_{\text{ESS}i}, k, n),$$

to calculate the power values $\Gamma(x,y,k,n)$ for the plurality of locations in the environment for a time-frequency bin ~~(k, n)~~, wherein k denotes the frequency index and n denotes the time index, wherein N denotes a number of the sound sources, wherein x, y denote coordinates of one of the plurality of locations, wherein $\text{power}_i(k,n)$ denotes the power value at an i -th sound source for time-frequency bin ~~(k, n)~~, wherein $x_{\text{ESS}i}, y_{\text{ESS}i}$ denote coordinates of the i -th sound source, wherein γ_i is a scalar value and wherein g is a function depending on $x, y, x_{\text{ESS}i}, y_{\text{ESS}i}, k, n$ and γ_i , when the environment is a two-dimensional environment, or

wherein the spatial power density determiner ~~(10; 21)~~ is adapted to determine the spatial power density by applying the formula

$$\Gamma(x,y,z,k,n) = \sum_{i=1}^N \text{power}_i(k,n) \cdot g(\gamma_i, x - x_{\text{ESS}i}, y - y_{\text{ESS}i}, z - z_{\text{ESS}i}, k, n),$$

to calculate the power values $\Gamma(x,y,z,k,n)$ for the plurality of locations of the environment for a time-frequency bin ~~(k, n)~~, wherein k denotes the frequency index and n denotes the time index, wherein N denotes a number of the sound sources, wherein x, y, z denote coordinates of one of the plurality of locations, wherein $\text{power}_i(k,n)$ denotes the power value at an i -th sound source for time-frequency bin ~~(k, n)~~, wherein $x_{\text{ESS}i}, y_{\text{ESS}i}, z_{\text{ESS}i}$ denote coordinates of the i -th sound source,

wherein γ_i is a scalar value and wherein g is a function depending on $x, y, z, x_{ESSi}, y_{ESSi}, z_{ESSi}, k, n$ and γ_i , when the environment is a three-dimensional environment.

11. Computer program for implementing the method according to claim 10 when the
5 computer program is executed on a computer or processor.