- reduced computational complexity compared to a “real” ER calculation
- adaptation of the ER pattern dependent on the physical room properties
- do not require any specific sound engineering skill and experience to set necessary parameters
- distinct ER patterns for indoor and outdoor
- no additional side information needed (for an encoder/bitstream/renderer scenario including transmission of a bitstream), in the case that the predefined patterns are calculated within the renderer
- very little additional side information needed (for an encoder/bitstream/renderer scenario including transmission of a bitstream), in the case that the predefined patterns are calculated in the encoder from the scene geometry

This is achieved by using parameterizable but fixed spatial ER patterns that do not depend on the exact geometry of the room. In a preferred embodiment of the invention, the pattern also does not depend on the listener position in the room. Instead, only one (or a few) global characteristic parameters are used to configure the ER pattern. In this way, the pattern can be rendered extremely efficiently.

In the following newly invented ER patterns, specifically room acoustic parameters are used like RT60, predelay time, room dimensions or room volume, frequency dependency of RT60 for pattern configuration. The ER pattern is defined in a way to produce a (temporally) smooth transition between the direct sound and the late reverb. It should be of neutral timbre. It is dependent on room volume and surface. It is not dependent on the position of the source and receiver in the room.

It is the objective of the invention to produce a plausible and convincing perception by the listener, fitting to the overall room acoustical parameters. This is sufficient for most use cases, especially since the listener has no possibility for a direct comparison with a rendering of the “real” physically correct ER.

Summary of the Invention

In accordance with a first embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that the early reflections depend on a relationship between a source position and a listener position. The inventors found, that it is possible to consider a source position independent ER pattern without, e.g., floor reflection; so that ER

rendering gets easier while the rendering result is still pretty good. The early reflection portion of the room impulse response used for the rendering, is exclusively determined by an early reflection pattern. A spatial relationship between a sound source and the listener is not considered for the early reflection portion of the room impulse response. Further the early reflection positions in the early reflection pattern are invariant with respect to changes in a listener head orientation. This is based on the finding that the same ER pattern can be used for determining the early reflection portion of the room impulse response independent whether the listener looks to the sound source or in any other direction.

Accordingly, in accordance with a first embodiment of the present disclosure, an apparatus for sound rendering is configured to receive information on a listener position and a sound source position. The apparatus is configured to render an audio signal of the sound source using a room impulse response whose early reflection portion is exclusively determined by an early reflection pattern. The early reflection pattern is indicative of a constellation, e.g. constellation shall denote a set of positions along with defining their mutual placement in terms of the angles between the lines connecting the positions; a synonymous term shall be “pattern”, of early reflection positions. The early reflection pattern is positioned at the listener position in a manner so that the early reflection positions are located around the listener position and at angular directions from the listener position which are invariant with respect to changes in a listener head orientation, i.e. the constellation is translatorily placed at the listener position.

In accordance with a second embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that the early reflection patterns for outdoor environments are highly individual and dependent on the physical setup of the scene. The inventors found, that ER pattern generated using moderate analysis of an environment can result into an acoustically convincing, but computationally moderate ER rendering result.

Accordingly, in accordance with a second embodiment of the present disclosure, an apparatus for determining an early reflection pattern for sound rendition is configured to perform a geometric analysis of an acoustic environment by, at each of one or more analysis positions, determining a function indicative, for each of different distances from the respective analysis position, a value representative of an early reflection contribution; and by inspecting the function or a further function derived therefrom with respect to one or more maxima to derive one or more control parameters. Additionally, the apparatus is configured to determine an early reflection pattern, which is indicative of a constellation of early reflection positions, by placing the early reflection positions using the one or more control parameters.

In accordance with a third embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that a transmission of early reflection patterns of the audio scenes for the rendering may result in high signaling costs. The inventors found, that ER pattern can be generated by use of bitstream hints resulting into an acoustically convincing, but computationally moderate ER rendering result. By using only hints in the bitstream, the signaling costs can be reduced, since it is not necessary to transmit the complete ER pattern.

Accordingly, in accordance with a third embodiment of the present disclosure, an apparatus for sound rendering is configured to receive first information on a listener position and a sound source position. The apparatus is configured to receive a bitstream comprising, e.g. and read therefrom, a representation of an audio signal of a sound source positioned at the sound source position and one or more early reflection pattern parameters. For example, the bitstream is audio bitstream with the early reflection parameter inside a header or metadata field of the bitstream, or a file format stream with the early reflection parameter inside a packet of the file format stream and a track of the file format stream comprising an audio bitstream representing the audio signal. Additionally, the apparatus is configured to determine an early reflection pattern, which is indicative of a constellation of early reflection positions, depending on the one or more early reflection pattern parameters. Further, the apparatus is configured to render the audio signal of the sound source using a room impulse response whose early reflection portion is determined by an early reflection pattern. The early reflection pattern is indicative of a constellation, e.g. constellation shall denote a set of positions along with defining their mutual placement in terms of the angles between the lines connecting the positions; an synonymous term shall be “pattern”, of early reflection positions. The early reflection pattern is positioned at the listener position in a manner so that the early reflection positions are located around the listener position and at angular directions from the listener position which are invariant with respect to changes in listener head orientation, i.e. the constellation is translatorily placed at the listener position.

In accordance with a fourth embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that a tremendous amount of computation has to be spent to determine each reflection from the source to the listener, taking into consideration the geometry of walls, occluding objects and other effects to compute a physically accurate reflection pattern. The inventors found, that simple room acoustical parameters, like room dimension, room volume or predelay, can be used to determine the number of early reflection

positions within an early reflection pattern. It is not needed to analyze the real early reflection of the scene, since the early reflections can be approximated dependent on a room acoustical parameter. The inventors found that ER pattern generation by ER number dependency on room acoustical parameter results into an acoustically convincing, but computationally moderate ER rendering result.

Accordingly, in accordance with a fourth embodiment of the present disclosure, an apparatus for determining an early reflection pattern for sound rendition is configured to receive at least one room acoustical parameter which is representative of an acoustical characteristic of an acoustic environment. The apparatus is configured to determine an early reflection pattern, which is indicative of a constellation of early reflection positions, in a manner so that a number of the early reflection positions depend on the at least one room acoustical parameter.

In accordance with a fifth embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that each source is associated with a different early reflection pattern. The inventors found, that it is not necessary to use different ER pattern for signals of different sources. This is based on the idea that the signals can be weighted and summed dependent on a source listener relationship, so that only the weighted sum of the audio signals is rendered based on the ER patter. The inventors found that ER rendition by use of a ER pattern for more than one sound source results into acoustically convincing, but computationally moderate ER rendering result.

Accordingly, in accordance with a fifth embodiment of the present disclosure, an apparatus for sound rendering is configured to receive information on a listener position, a first sound source position and a second sound source position. The apparatus is configured to render audio signal of the two sound sources using a room impulse response whose early reflection portion is determined by an early reflection pattern. The early reflection pattern is indicative of a constellation, e.g. constellation shall denote a set of positions along with defining their mutual placement in terms of the angles between the lines connecting the positions; an synonymous term shall be “pattern”, of early reflection positions. The early reflection pattern is positioned at the listener position in a manner so that the early reflection positions are located around the listener position and at angular directions from the listener position which are invariant with respect to changes in listener head orientation, i.e. the constellation is translatorily placed at the listener position. The apparatus is configured to render the audio signals of the two sound sources by forming a weighted sum of a first audio signal of a first sound source positioned at the first sound source position and a second audio signal of a second sound source positioned

at the second sound source position. The weighted sum weights the first audio signal more than the second audio signal, if a first distance between the first sound source position and the listener position is smaller than a second distance between the second sound source position and the listener position, and weights the second audio signal more than the first audio signal, if the first distance is larger than the second distance. Additionally, the apparatus is configured to render the audio signals of the two sound sources by generating early reflection contribution loudspeaker signals relating to the early reflection portion of the room impulse response by rendering the weighted sum from the early reflection positions.

In accordance with a sixth embodiment of the present disclosure, the inventors of the present application realized that one problem encountered when trying to use early reflection (ER) rendering of audio signal stems from the fact that a tremendous amount of computation has to be spent to determine each reflection from the source to the listener, taking into consideration the geometry of walls, occluding objects and other effects to compute a physically accurate reflection pattern. The inventors found, that simple room acoustical parameters, like room dimension, room volume or predelay, can be used to parametrize function defining a position of the early reflections. It is not needed to analyze the real early reflection of the scene, since the early reflections can be approximated dependent on the room acoustical parameter. Further it was found that spiral functions provide a good distribution of the early reflection positions. The inventors found that ER pattern generation using one or more spiral functions results into an perceptually convincing, but computationally moderate ER rendering result.

In accordance with a first aspect of the present application, an apparatus for determining an early reflection pattern for sound rendition is configured to receive at least one room acoustical parameter which is representative of an acoustical characteristic of an acoustic environment and determine an early reflection pattern, which is indicative of a constellation of early reflection positions, by parameterizing one or more spiral functions centered at the listener position, and place the early reflection positions using the one or more spiral functions.

In accordance with a second aspect of the present application, an apparatus for sound rendering, is configured to receive first information on a listener position and a sound source position; render an audio signal of the sound source using a room impulse response whose early reflection portion is determined by an early reflection pattern, which is indicative of a constellation of early reflection positions, and which is positioned at the listener position in a manner so that the early reflection positions are located around the listener position and at angular directions from the listener position which are invariant with respect to changes in

listener head orientation, the apparatus comprising an apparatus for determining the early reflection pattern according to the first aspect.

In accordance with a third aspect of the present application, a bitstream for being subject to sound rendition according to the second aspect is provided.

In accordance with a fourth aspect of the present application, a digital storage medium storing a bitstream for being subject to sound rendition according to the third aspect is provided.

In accordance with a fifth aspect of the present application, a method for determining an early reflection pattern for sound rendition, comprises receiving at least one room acoustical parameter which is representative of an acoustical characteristic of an acoustic environment; determining an early reflection pattern, which is indicative of a constellation of early reflection positions, by parameterizing one or more spiral functions centered at the listener position, and placing the early reflection positions using the one or more spiral functions.

In accordance with a sixth aspect of the present application, a method for sound rendering, comprises receiving first information on a listener position and a sound source position; rendering an audio signal of the sound source using a room impulse response whose early reflection portion is determined by an early reflection pattern, which is indicative of a constellation of early reflection positions, and which is positioned at the listener position in a manner so that the early reflection positions are located around the listener position and at angular directions from the listener position which are invariant with respect to changes in listener head orientation, the method comprising the method for determining the early reflection pattern according to fifth aspect.

In accordance with a seventh aspect of the present application, a computer program for causing a computer, when executing the computer program, to perform the method of the fifth or sixth aspect.

Brief Description of the Drawings

The drawings are not necessarily to scale, emphasis instead generally being placed upon illustrating the principles of the invention. In the following description, various embodiments of the invention are described with reference to the following drawings, in which:

Fig. 1 shows an embodiment of an early reflection pattern;