

Apparatus and Method for Rendering a Sound Scene Comprising Discretized Curved Surfaces

5

Specification

The present invention relates to audio processing and, particularly, to audio signal processing for rendering sound scenes comprising reflections modeled by image sources in the field of Geometrical Acoustics.

10

Geometrical Acoustics are applied in auralization, i.e., real-time and offline audio rendering of auditory scenes and environments [1, 2]. This includes Virtual Reality (VR) and Augmented Reality (AR) systems like the MPEG-I 6-DoF audio renderer. For rendering complex audio scenes with six degrees of freedom (DoF), the field of Geometrical Acoustics is applied, where the propagation of sound data is modeled with models known from optics such as ray-tracing. Particularly, the reflections at walls are modeled based on models derived from optics, in which the angle of incidence of a ray that is reflected at the wall results in a reflection angle being equal to the angle of incidence.

15

20 Real-time auralization systems, like the audio renderer in a Virtual Reality (VR) or Augmented Reality (AR) system, usually render early specular reflections based on geometry data of the reflective environment [1,2]. A Geometrical Acoustics method like ray-tracing [3] or the image source method [4] is then used to find valid propagation paths of the reflected sound. These methods are valid, if the reflecting planar surfaces are large compared to the wave length of incident sound [1]. Furthermore, the distance of the reflection point on the surface to the boundaries of the reflecting surface must also be large compared to the wave length of incident sound.

25

If the geometry data approximates curved surfaces by triangles or rectangles, the classic Geometrical Acoustics methods are no longer valid and artifacts become audible. The resulting "disco ball effect" is illustrated in Figure 6. For a moving listener or a moving sound source the visibility of the image source will alternate between visible and invisible, resulting in a permanently switching localization, timbre, and loudness.

30

35 If a classic image source model is used, there is usually no mitigation technique applied for the given problem [5]. If diffuse reflections are modeled in addition to specular reflections, this will further reduce the effect, but cannot solve it. Summarizing, no solution for this problem is described in the state-of-the-art.

It is an object of the present invention to provide a concept for mitigating the disco ball effect in Geometrical Acoustics or to provide a concept of rendering a sound scene that provides an improved audio quality.

5

This object is achieved by an apparatus for rendering a sound scene of claim 1, a method of rendering a sound scene of claim 18, or a computer program of claim 19.

10 The present invention is based on the finding that the problems associated with the so-called disco ball effect in Geometric Acoustics can be addressed by performing an analysis of reflecting geometric objects in a sound scene in order to determine whether a reflecting geometric object results in visible zones and invisible zones. For an invisible zone, an image source position generator generates an additional image source position so that the additional image source position is placed between two image source positions being
15 associated with the neighboring visible zones. Furthermore, a sound renderer is configured to render the sound source at the sound source position in order to obtain an audio impression of the direct path and to additionally rendering the sound source at an image source position or an additional image source position depending on whether the listener position is located within a visible zone or an invisible zone. By this procedure, the disco
20 ball effect in Geometrical Acoustics is mitigated. This procedure can be applied in auralization such as real-time and offline audio rendering auditory scenes and environments.

25 In preferred embodiments, the present invention provides several components, where one component comprises a geometry data provider or a geometry pre-processor which detects curved surfaces such as "round edges" or "round corners". Furthermore, the preferred embodiments refer to the image source position generator that applies an extended image source model for the identified curved surfaces, i.e., the "round edges" or "round corners".

30 Particularly, an edge is a boundary line of a surface, and a corner is the point where two or more converging lines meet. A round edge is a boundary line between two flat surfaces that approximate a rounded continuous surfaces by means of triangles or polygons. A round corner or rounded corner is a point that is a common vertex of several flat surfaces that approximate a rounded continuous surfaces by means of triangles or polygons. Particularly,
35 when a Virtual Reality scene, for example, comprises an advertising pillar or advertising column, this advertising pillar or advertising column can be approximated by polygon-shaped planes such as triangle or other polygon-shaped planes, and due to the fact that

the polygon planes are not infinitesimally small, invisible zones between visible zones can occur.

Typically, there will exist intentional edges or corners, i.e., objects in the audio scene that are to be acoustically represented as they are, and any effects that occur due to the acoustical processing are intended. However, rounded or round corners or edges are geometric objects in the audio scene that result in the disco ball artefact or, stated in other words, that result in invisible zones that degrade the audio quality when a listener moves with respect to a fixed source from a visible zone into an invisible zone or when a fixed listener listens to a moving source that results in bringing the user into an invisible zone and then a visible zone and then an invisible zone. Or, alternatively, when both, the listener and the source move, it can be that a listener is at one point in time within a visible zone and at another point in time in an invisible zone that is only due because of the applied Geometrical Acoustics model, but has nothing to do with the real-world acoustical scene that is to be approximated as far as possible by the apparatus for rendering the sound scene or the corresponding method.

The present invention is advantageous since it generates high quality audio reflections on spheres and cylinders or other curved surfaces. The extended image source model is particularly useful for primitives such as polygons approximating cylinders, spheres or other curved surfaces. Above all, the present invention results in a quickly converging iterative algorithm for computing first order reflections particularly relying on the image source tools for modeling reflections. Preferably, a particular frequency-selective equalizer is applied in addition to a material equalizer that accounts for the frequency-selective reflection characteristic that typically is a high-pass filter that depends on a reflector diameter, for example. Furthermore, the distance attenuation, the propagation time and the frequency-selective wall absorption or wall reflection is taken into account in preferred embodiments. Preferably, the inventive application of an additional image source position generation "enlightens" the dark or invisible zones. An additional reflection model for rounded edges and corners relies on this generation of additional image sources in addition to the classical image sources associated with the polygonal planes. Preferably, a continuous extrapolation of image sources into the "dark" or invisible zones is performed preferably using the technology of frustum tracing for the purpose of calculating first order reflections. In other embodiments, the technology can also be extended to second or higher order reflection processing. However, performing the present invention for applying the calculation of first order reflections already results in high audio quality and it has been found out that performing higher order reflection calculation, although being possible, will not always justify the additional processing requirements in view of the additionally gained audio quality. The

present invention provides a robust, relatively easy to implement but nevertheless powerful tool for modeling reflections in complex sound scenes having problematic or specific reflection objects that would suffer from invisible zones without the application of the present invention.

5

Preferred embodiments of the present invention are subsequently discussed with respect to the accompanying drawings, in which:

Fig. 1 illustrates a block diagram of an embodiment of the apparatus for rendering a sound scene;

10

Fig. 2 illustrates the flowchart for the implementation of the image source position generator in an embodiment;

Fig. 3 illustrates a further implementation of the image source position generator;

15

Fig. 4 illustrates another preferred implementation of the image source position generator;

Fig. 5 illustrates the construction of an image source in Geometrical Acoustics;

20

Fig. 6 illustrates a specific object resulting in visible zones and invisible zones;

Fig. 7 illustrates a specific reflection object where an additional image source is placed at an additional image source position in order to "enlighten" the invisible zones;

25

Fig. 8 illustrates a procedure applied by the geometry data provider;

Fig. 9 illustrates an implementation of the sound renderer for rendering the sound source at the sound source position and for additionally rendering the sound source at an image source position or an additional image source position depending on the position of the listener;

30

Fig. 10 illustrates the construction of the reflection point R on an edge;

Fig. 11 illustrates the quiet zone related to a rounded corner; and

35

Fig. 12 illustrates the quiet zone or quiet frustum of related to a rounded edge of e.g. Fig. 10.

10.

Fig. 1 illustrates an apparatus for rendering a sound scene having reflection objects and a sound source at a sound source position. In particular, the sound source is represented by a sound source signal that can, for example, be a mono or a stereo signal and, in the sound scene, the sound source signal is emitted at the sound source position. Furthermore, the sound scene typically has an information on a listener position, where the listener position comprises, on the one hand, a listener location within a, for example, three-dimensional space or where the listener position incurs, on the other hand, a certain orientation of the head of the listener within a three-dimensional space. A listener can be positioned, with respect to her or his ears, at a certain location in the three-dimensional space resulting in three dimensions, and the listener can also turn his head around three different axes resulting in additional three dimensions so that a six degree of freedom's Virtual Reality or Augmented Reality situation can be processed. The apparatus for rendering a sound scene comprises a geometry data provider 10, an image source position generator 20 and a sound renderer 30 in a preferred embodiment. The geometry data provider can be implemented as a preprocessor for performing certain operations before the actual runtime or the geometry data provider can be implemented as a geometry processor doing its operation also at runtime. However, performing the calculations of the geometry data provider in advance, i.e., before the actual Virtual Reality or Augmented Reality rendering will free a processing platform from the corresponding geometry preprocessor tasks.

The image source position generator relies on the source position and the listener position and, particularly due to the fact that the listener position will change in runtime, the image source position generator will operate in runtime. The same is true for the sound renderer 30 that additionally operates in runtime using the sound source data, the listener position and additionally using the image source positions and the additional image source positions if required, i.e., if the user is placed in an invisible zone that has to be "enlightened" by an additional image source determined by the image source position generator in accordance with the present invention.

Preferably, the geometry data provider 10 is configured for providing an analysis of the reflection object of the sound scene to determine a specific reflection object that is represented by a first polygon and a second adjacent polygon. The first polygon has associated a first image source position and the second polygon has associated a second image source position, where these image source positions are constructed, for example, as illustrated in Fig. 5. These image sources are the "classical image sources" that are mirrored at a certain wall. However, the first and second image source positions result in a sequence comprising a first visible zone related to the first image source position, a second

visible zone related to the second image source position and an invisible zone placed between the first and the second visible zone as illustrated in Figs. 6 or 7, for example. The image source position generator is configured for generating the additional image source position such that the additional image source located at the additional image source position is placed between the first image source position and the second image source position. Preferably, the image source position generator additionally generates the first image source and the second image source in a classical way, i.e., by mirroring, for example, at a certain mirroring wall or, as is the case in Fig. 6 or Fig. 7, when the reflecting wall is small and does not comprise a wall point where the rectangular projection of the source crosses the wall, the corresponding wall is extended only for the purpose of image source construction.

The sound renderer 30 is configured for rendering the sound source at the sound source position in order to obtain the direct sound at the listener position. Additionally, in order to also render a reflection, the sound source is rendered at the first image source position, when the listener position is located within the first visible zone. In this situation, the image source position generator does not need to generate an additional image source position, since the listener position is such that any artefacts due to the disco ball effect do not occur at all. The same is true when the listener position is located within the second visible zone associated with the second image source. However, when the listener is located within the invisible zone, then the sound renderer uses the additional image source position and does not use the first image source position and the second image source position. Instead of the "classical" image sources modeling the reflections at the first and the second adjacent polygons, the sound renderer only renders, for the purpose of reflection rendering, the additional image source position generated in accordance with the present invention in order to fill up or enlighten the invisible zone with sound. Any artefacts that would otherwise result in a permanently switching localization, timbre and loudness are avoided by means of the inventive processing using the image source position generator generating the additional image source between the first and the second image source position.

Fig. 6 illustrates the so-called disco ball effect. Particularly, the reflecting surfaces are sketched in black and are denoted by 1, 2, 3, 4, 5, 6, 7, 8. Each reflecting surface or polygon 1, 2, 3, 4, 5, 6, 7, 8 is also represented by a normal vector indicated in Fig. 6 in a normal direction to the corresponding surface. Furthermore, each reflecting surface has associated a visible zone. The visible zone associated with a source S at a source position 100 and reflecting surface or polygon 1 is indicated at 71. Furthermore, the corresponding visible zones for the other polygons or surfaces 2, 3, 4, 5, 6, 7, 8 are illustrated in Fig. 6 by reference numbers 72, 73, 74, 75, 76, 77, 78, for example. The visible zones are generated in such a

way that only within the visible zone associated with a certain polygon, the condition of the incidence angle being equal to the reflection angle of a sound emitted by the sound source S is fulfilled. For example, polygon 1 has a quite small visible zone 71, since the extension of polygon 1 is quite small, and since the angle of incidence being equal to the angle of reflection can only be fulfilled for reflection angles within the small visible zone 71.

Furthermore, Fig. 6 also has a listener L located at a listener position 130. Due to the fact that the listener L is placed within the visible zone 74 associated with polygon number 4, the sound for the listener L is rendered using the image source 64 illustrated at S/4. This image source S/4 indicated at 64 in Fig. 6 is responsible for modeling the reflection at reflecting surface or polygon number 4, and since the listener L is located within the visible zone 74 associated with the image source for the certain wall, no artefacts would occur. However, should there be a movement of the listener in the quite zone between visible zones 73 and 74 or into the invisible zone between visible zones 74 and 75, i.e., when the listener moves upwards or downwards, then a classical renderer would stop rendering using image source S/4, and since the listener is not located in visible zone 73 or visible zone 75 associated with image source S/3 63 or S/565, then the renderer would not render any reflections without the present invention.

In Fig. 6, the disco ball effect is illustrated and the reflecting surfaces are sketched in black, gray areas mark the regions where the n-th image source "Sn" is visible, and S marks the source at the source position, and L marks the listener at the listener position 130. The reflecting object in Fig. 6 being a specific reflection object could, for example, be an advertising pillar or advertising column watched from the above, the sound source, could, for example, be a car located at a certain position fixed relative to the advertising color, and the listener would, for example, be a human walking around the advertising pillar in order to look what is on the advertising pillar. The listening human will typically hear the direct sound from the car, i.e., from position 100 to the human's position 130 and, additionally, will hear the reflection at the advertising pillar.

Fig. 5 illustrates the construction of an image source. Particularly, and with respect to Fig. 6, the situation of Fig. 5 would illustrate the construction of image source S/4. However, the wall or polygon 4 in Fig. 6 does not even reach until the direct connection between the source position 100 and the image source position 64. The wall 140 illustrated in Fig. 5 as being a mirroring plane for the generation of the image source 120, based on the source 100, is not existent in Fig. 6 at the direct connection between the source 100 and the image source 120. However, for the purpose of constructing image sources, a certain wall, such as polygon 4 in Fig. 6, is extended in order to have a mirroring plane for mirroring the source

at the wall. Furthermore, in classical image source processing, assumptions are made, in addition to an infinite wall, that the source emits a plane wave. However, this assumption is not material for the present invention, and the same is true for the infinity of the wall, since, for the purpose of mirroring the wall, an infinite wall actually is only required for explaining the underlying mathematical model.

Furthermore, Fig. 5 illustrates the condition of having same angles of incidence on the wall and of the reflection from the wall. Furthermore, the path length for the propagation path from the source to the receiver is maintained. The path length from the source to the receiver is exactly the same as the path length from the image source to the receiver, i.e., $r_1 + r_2$, and the propagation time is equal to the quotient between the total path length and the sound velocity c . Furthermore, a distance attenuation of the sound pressure p being proportional to $1/r$ or a distance attenuation of the sound energy being proportional to $1/r^2$ is typically modeled by the renderer rendering the image source.

Furthermore, a wall absorption/reflection behavior is modeled by means of the wall absorption or reflection coefficient α . Preferably, the coefficient α is dependent on the frequency, i.e., represents a frequency-selective absorption or reflection curve $H_w(k)$ and typically has a high-pass characteristic, i.e., high frequencies are better reflected than low frequencies. This behavior is accounted for in preferred embodiments. The strength of the image source application is that subsequent to the construction of the image source and the description of the image source with respect to the propagation time, the distance attenuation and the wall absorption, the wall 140 will be completely removed from the sound scene and is only modeled by the image source 120.

Fig. 7 illustrates a problematic situation, where the first polygon 2 having associated the first image source position $S/2$ 62 and the second polygon 3 having associated therewith the second image source position 63 or $S/3$ are placed with a short angle in between, and the listener 130 is placed in the invisible zone between the first visible zone 72 associated with the first image source 62 and the second visible zone 73 associated with the second image source $S/3$ 63. In order to "enlighten" the invisible zone 80 illustrated in Fig. 7, an additional image source position 90 being placed between the first image source position 62 and the second image source position 63 is generated. Instead of modeling the reflection by means of the image source 63 or the image source 62 that is constructed as illustrated in Fig. 5 for the classical procedure, the reflection is now modeled using the additional image source position 90 that preferably has the same distance to the reflection point at least in a certain tolerance.

For the additional image source position 90, the same path length, propagation time, distance attenuation and wall absorption is used for the purpose of rendering the first order reflection in the invisible zone 80. In a preferred embodiment, a reflection point 92 is determined. The reflection point 92 is at the junction between the first polygon and the second polygon when watched from above, and typically is in a vertical position, for example in the example of the advertising pillar that is determined by the height of the listener 130 and the height of the source 100. Preferably, the additional image source position 90 is placed on a line connecting the listener 130 and the reflection point 92, where this line is indicated at 93. Furthermore, the exact position of the additional sound source 90 in the preferred embodiment is at the intersection point of the line 93 and the connecting line 91, connecting the image source positions 62 and 63 that have visible zones adjacent to the invisible zone 80.

However, the Fig. 7 embodiment only illustrates a most preferred embodiment, where the path of the additional image source position is exactly calculated. Furthermore, the specific position of the additional sound source position on the connecting line 92, depending on the listener position 130, is also calculated exactly. When the listener L is closer to the visible zone 73, then the sound source 90 is closer to the classical image source position 63 and vice versa. However, locating the additional sound source position in any place between the image sound sources 62 and 63 will already improve the entire audible impression very much compared to simply suffering from the invisible zones. Although Fig. 7 illustrates the preferred embodiment with an exact position of the additional sound source position, another procedure would be to locate the additional sound source at any place between the adjacent sound source positions 62 and 63 so that a reflection is rendered in the invisible zone 80.

Furthermore, although it is preferred to exactly calculate the propagation time depending on the exact path length, other embodiments rely on an estimation of the path length as depending on a modified path length of image source position 63, or a modified path length of the other adjacent image source position 62. Furthermore, with respect to the wall absorption or wall reflection modeling, for the purpose of rendering the additional sound source position 90, either the wall absorption of one of the adjacent polygons can be used, or an average value of both absorption coefficients if they are different from each other can be used, and even a weighted average can be applied depending on whether the listener is closer to which visible zone, so that a certain wall absorption data of the wall having the visible zone to which the user is located closer receives a higher weighting value in a weighted addition compared to the absorption/reflection data of the other adjacent wall having the visible zone being further away from the listener position.

Fig. 2 illustrates a preferred implementation of the procedure of the image source position generator 20 of Fig. 1. In a step 21, it is determined, whether the listener is in an visible zone such as 72 and 73 of Fig. 7 or in an invisible zone 80. In case it is determined that the user is in the visible zone, the image source position such as S/2 62 when the user is in zone 72 or the image source position 63 or S/3 if the user is in the visible zone 73 is determined. Then, the information on the image source position is sent to the renderer 30 of Fig. 1 as is illustrated in step 23.

Alternatively, when step 21 determines that the user is placed within the invisible zone 80, the additional image source position 90 of Fig. 7 is determined and as soon as same is determined as illustrated in step 24, this information on the additional image source position and if applicable, other attributes such as a path length, a propagation time, a distance attenuation or a wall absorption/reflection information as also sent to the renderer as illustrated in step 25.

Fig. 3 illustrates a preferred implementation of step 21, i.e., how in a specific embodiment, it is determined whether the listener is in an visible zone or in an invisible zone. To this end, two basic procedures are envisioned. In one basic procedure, the two neighboring visible zones 72 and 73 are calculated as frustums based on the source position 100 and the corresponding polygon and, then it is determined, whether the listener is in one of those visible frustums. When it is determined that the listener is not located within one of the frustums, as it is indicated in step 26, then a conclusion is made that the user is in the invisible zone. Alternatively, instead of calculating two frustums describing the visible zones 72 and 73 of Fig. 7, another procedure is to actually determine the invisible frustum describing the invisible zone 80, and if the invisible frustum is determined, then it is decided that the listener is within the invisible zone 80, when the listener is placed within the quit frustum. When it is determined that the listener is in the invisible zone as is the result of step 27 and step 26 of Fig. 3, then the additional image source position is calculated as illustrated in step 24 of Fig. 2 or step 24 of Fig. 3.

Fig. 4 illustrates a preferred implementation of the image source position generator for calculating the additional image source position 90 in a preferred embodiment. In a step 41, the image source positions for the first and the second polygons, i.e., image source position 62 and 63 of Fig. 7 are calculated in a classical or standard procedure. Furthermore, as illustrated in step 42, a reflection point on the edge or corner as has been determined by the geometric data provider 10 as being a "rounded" edge or corner is determined. The determination of the reflection point 92 in Fig. 7, for example, is on the crossing line between

the two polygons 2 and 3 and, in case of an exact rendering also in the vertical dimension, the vertical dimension of the reflection point is determined in step 42 depending on the height of the listener and the height of the source and other attributes such as the distance of the listener and the distance of the source from the reflection point or line 92.

5 Furthermore, as illustrated in block 43, a sound line is determined by connecting the listener position 130 and the reflection point 92 and by extrapolating this line further into the region where the image source positions are located and have been determined in block 41. This sound line is illustrated by reference number 93 in Fig. 7. In step 44, a connection line between the standard image sources as determined by block 41 is calculated, and then, as

10 illustrated in block 45, the intersection of the sound line 93 and the connection line 91 is determined to be the additional sound source position. It is to be noted that the order of steps as indicated in Fig. 4 is not compulsory. Since the result of a step 41 is only required before the step 44, the steps 42 and 43 can already be calculated before calculating step 41 and so on. The only requirement is that, for example, the step 42 has to be performed

15 before step 43 so that the sound line, for example, can be established.

Subsequently, further procedures are given in order to illustrate a further procedure of calculating the additional image source position. The extended image source model needs to extrapolate the image source position in the “dark zone” of the reflectors, i.e. the areas

20 between the “bright zones” in which the image source is visible (see Figure 1). In a first embodiment of this method, a frustum is created for each round edge and it is checked, if the listener is located within this frustum. The frustum is created as follows: For the two adjacent planes of the edge, namely the left and the right plane, one computes the image sources S_L and S_R by mirroring the source on the left and the right plane. From these points

25 together with the beginning and the end point of the edge one can define four planes $k \in [1,4]$ in Hesse-Normal form where the normal vectors $\vec{N}_k$ are pointing inside of the frustum,

$$\vec{N}_k \vec{X} - d_k = 0. \quad (1)$$

30 If the distance

$$l_k = \vec{N}_k \vec{L} - d_k \quad (2)$$

is greater than or equal zero for all 4 planes, then the listener is located within the frustum that defines the coverage area of the model for the given round edge. The invisible zone frustum is illustrated in Fig. 12 additionally showing the source position 100 and the image sources 61 and 62 belonging to the respective polygons 1 and 2. The frustum starts on the

35

edge between polygons 1 and 2 and opens towards the source position out from the drawing plane and into the drawing plane.

- 5 In this case, one can determine the reflection point on the round edge as follows:

Let $\vec{P}_S$ be the orthogonal projection of the source position $\vec{S}$ onto the edge and $\vec{P}_L$ be the orthogonal projection of the listener position $\vec{L}$ onto the edge. This yields the reflection point $\vec{R}$ as follows:

10

$$d_S = |\vec{P}_S - \vec{S}| \quad (3)$$

$$d_L = |\vec{P}_L - \vec{L}| \quad (4)$$

15

$$\vec{R} = \frac{d_S}{d_S + d_L} \vec{P}_L + \frac{d_L}{d_S + d_L} \vec{P}_S \quad (5)$$

The construction of the reflection point is illustrate in Fig. 10 showing the listener position L, the source position S, the projections Ps and Pl and the resulting reflection point,

- 20 The computation of the coverage area of the round corners is very similar. Here, the k adjacent planes yield k image sources which together with the corner position result in a frustum that is bounded by k planes. Again, if the distances of the listener to these planes are all greater than or equal zero, the listener is located within the coverage area of the round corner. The reflection point $\vec{R}$ is given by the corner point itself.

25

This situation, i.e., the invisible frustum or a round corner is illustrated in Fig. 11 illustrating four image sources 61, 62, 63, 64 belonging to the four polygons or planes 1, 2, 3, 4. In Fig. 11, the source is located in a visible zone and not in the invisible zone starting with its tip at the corner and opening away from the four polygons.

30

For higher-order reflections, one can extend this method according to the frustum-tracing method where one splits up each frustum into sub-frustums whenever one hits a surface, round edge, or round corner.

- 35 Fig. 8 illustrates a further preferred implementation of the geometric data provider. Preferably, the geometric data provider operates as a true data provider that generates, during runtime, pre-stored data on objects in order to indicate that an object is a specific

reflection object having a sequence of visible zones and an invisible zone in between. The geometric data provider can be implemented as using a geometry pre-processor that is executed once during initialization, as it does not depend on the listener or source positions. Contrary thereto, the extended image source model as applied by the image source position
5 generator is executed at run-time and determines edge and corner reflections depending on the listener and source positions.

The geometric data provider may apply a curved surface detection. The geometry data provider also termed to be the geometry-processor calculates the specific reflection object
10 determination in advance, in an initialization procedure or a runtime. If, for example, a CAD software is used to export the geometry data, as much information about curvatures as possible is preferably used by the geometry data provider. For example, if surfaces are constructed from round geometry primitives like spheres or cylinders or from spline
15 interpolations, the geometry pre-processor / geometry data provider is preferably implemented within the export routine of the CAD software and detects and uses the information from the CAD software.

If no a priori knowledge about the surface curvature is available, the geometry preprocessor or data provider needs to implement a round edge and round corner detector by using only
20 the triangle or polygon mesh. For example, this can be done by computing the angle Φ between two adjacent triangles 1, 2 or 1a, 2a as illustrated in Fig. 8. Particularly, the angle is determined to be a "face angle" in Fig. 8, where the left portion of Fig. 8 illustrates a positive face angle and the right portion in Fig. 8 illustrates a negative face angle. Furthermore, the small arrows illustrate the face normal in Fig. 8. If the face angle is below
25 a certain threshold, the adjacent edge in both adjacent polygons forming the edge are considered to represent a curved surface section and is marked as such. If all edges that are in connection to a corner are marked as being round, the corner is also marked as being round, and as soon as this corner becomes pertinent for the sound rendering, the functionality of the image source position generator for generating the additional image
30 source position is activated. When, however, it is determined that a certain reflection object is not a specific reflection object but a straight forward object, where any artifacts are not expected or are even intended by a sound scene creator, the image source position generator is only used for determining the classical image source positions, but any determination of an additional image source position in accordance with the present
35 invention is deactivated for such a reflection object.

Fig. 9 illustrates a preferred embodiment of the sound renderer 30 of Fig. 1. The sound renderer 30 preferably comprises a direct sound filter stage 31, the first order reflection filter

stage 32 and, optionally, a second order reflection filter stage and probably one or more higher order reflection filter stages as well.

Furthermore, depending on the output format required by the sound renderer 30, i.e.,
5 depending on whether the sound renderer outputs via headphones, via loudspeakers or just for storage or transmission in a certain format, a certain number of output adders such as a left adder 34, a right adder 35 and a center adder 36 and probably other adders for left surround output channels, or for right surround output channels, etc. are provided. While the left and the right adders 34 and 35 are preferably used for the purpose of headphone
10 reproduction for virtual reality applications, for example, any other adders for the purpose of loudspeaker output in a certain output format can also be used. When, for example, an output via headphones is required, then the direct sound filter stage 31 applies head related transfer functions depending on the sound source position 100 and the listener position 130. For the purpose of the first order reflection filter stage, corresponding head related transfer
15 functions are applied, but now for the listener position 130 on the one hand and the additional sound source position 90 on the other hand. Furthermore, any specific propagation delays, path attenuations or reflection effects are also included within the head related transfer functions in the first order reflection filter stage 32. For the purpose of higher order reflection filter stages, other additional sound sources are applied as well.

20
If the output is intended for a loudspeaker set up, then the direct sound filter stage will apply other filters different from head related transfer functions such as filters that perform vector based amplitude panning, for example. In any case, each of the direct sound filter stage 31, the first order reflection filter stage 32 and the second order reflection filter stage 33
25 calculates a component for each of the adder stages 34, 35, 36 as illustrated, and the left adder 34 then calculates the output signal for the left headphone speaker and the right adder 35 calculates the headphone signal for the right headphone speaker, and so on. In case of an output format that is different from a headphone, the left adder 34 may deliver the output signal for the left speaker and the right adder 35 may deliver the output for the
30 right speaker. If only two speakers in a two-speaker environment are there, then the center adder 32 is not required.

The inventive method avoids the disco-ball effect, that occurs when a curved surface, approximated by a discrete triangle mesh, is auralized using the classical image sound
35 source technique [3, 4]. The novel technique avoids invisible zones, making the reflection always to be audible. For this procedure it is necessary to identify approximations of curved surfaces by threshold face angle. The novel technique is an extension to the original model, with special treatment faces identified as a representation of a curvature.

Classical image sound source techniques [3, 4] do not consider that the given geometry can (partially) approximate a curved surface. This causes dark zones (silence) to be casted away from edge points of adjacent faces (see Fig. 1). A listener moving along such a surface
5 observes reflections to be switched on/off depending where he/she is located (enlighted/invisible zone). This causes unpleasant audible artifacts, also diminishing the degree of realism and thus the immersion. In essence, classical image source techniques fail to realistically render such scenes.

10 References

[1] Vorländer, M. "Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality." *Springer Science & Business Media*, 2007.

[2] Savioja, L., and Svensson, U. P. "Overview of geometrical room acoustic modeling
15 techniques." *The Journal of the Acoustical Society of America* 138.2 (2015): 708-730.

[3] Krokstad, A., Strom, S., and Sørsdal, S. "Calculating the acoustical room response by the use of a ray tracing technique." *Journal of Sound and Vibration* 8.1 (1968): 118-125.

20

[4] Allen, J. B., and Berkley, D. A. "Image method for efficiently simulating small room acoustics." *The Journal of the Acoustical Society of America* 65.4 (1979): 943-950.

[5] Borish, J. "Extension of the image model to arbitrary polyhedra." *The Journal of the
25 Acoustical Society of America* 75.6 (1984): 1827-1836.

Claims

1. Apparatus for rendering a sound scene having reflection objects and a sound source at a sound source position, comprising.
- 5
- a geometry data provider (10) for providing an analysis of the reflection objects of the sound scene to determine a reflection object represented by a first polygon (2) and a second adjacent polygon (3) having associated a first image source position (62) for the first polygon and a second image source position (63) for the second
- 10
- polygon, wherein the first and second image source positions result in a sequence comprising a first visible zone (72) related to the first image source position (62), an invisible zone (80) and a second visible zone (73) related to the second image source position (63);
- 15
- an image source position generator (20) for generating an additional image source position (90) such that the additional image source position (90) is placed between the first image source position and the second image source position; and
- 20
- a sound renderer (30) for rendering the sound source at the sound source position and, additionally
- for rendering the sound source at the first image source position, when a listener position (130) is located within the first visible zone,
- 25
- for rendering the sound source at the additional image source position (90), when the listener position is located within the invisible zone (80), or
- for rendering the sound source at the second image source position, when the listener position is located within the second visible zone.
- 30
2. Apparatus of claim 1, wherein the geometry data provider (10) is configured to retrieve pre-stored information on the reflection objects stored during an initialization stage, and wherein the image source position generator (20) is configured to generate the additional image source position (90) in response to the pre-stored
- 35
- information indicating the reflection object.
3. Apparatus of claim 1 or 2, wherein the geometry data provider (10) is configured to detect, during runtime or during an initialization stage and using geometry data on

the sound scene delivered by a computer added design (CAD) application, the reflection object.

4. Apparatus of one of the preceding claims, wherein the geometry data provider (10)
5 is configured to detect, during runtime or during an initialization stage, as the reflection object, an object having a round geometry, a curved geometry, or a geometry derived from a spline interpolation.

5. Apparatus of one of claims 1 or 2, wherein the geometry data provider (10) is
10 configured

to compute an angle between two adjacent polygons of a reflection object and to mark the two adjacent polygons as a specific pair of polygons, when the angle is below a threshold,

15 to compute a further angle between two further adjacent polygons of the reflection object and to mark the two further adjacent polygons as a further specific pair of polygons, when the further angles below the threshold, and

20 to detect the reflection object, when the further adjacent polygons and the adjacent polygons have an edge in common, or belong to the same corner.

6. Apparatus of one of the preceding claims,

25 wherein the image source position generator (20) is configured to analyze, whether the listener position (130) is in the invisible zone (80), and to generate the additional image source position (90) only when the listener position (130) is located in the invisible zone (80).

30 7. Apparatus of claim 6, wherein the image source position generator (20) is configured to determine a first geometrical range associated with the first polygon or a second geometrical range associated with the second polygon, or a third geometrical range between the first geometrical range and the second geometrical range,

35 wherein the first geometrical range determines the first visible zone or wherein the second geometrical range determines the second visible zone, or wherein the third geometrical range determines the invisible zone (80), and

wherein the first or the second geometrical range is determined such that a condition that an incidence angle from the source position to the first or the second polygon is equal to a reflection angle from the first or the second polygon is fulfilled for a position in the first or the second geometrical zone, or

5

wherein the third geometrical range is determined such that the condition of a reflection angle being equal to the incidence angle is not fulfilled for a position in the invisible zone (80).

10 8. Apparatus of claim 6 or 7,

wherein the image source position generator (20) is configured to calculate (26) a first frustum for the first polygon and to determine (27), whether the listener position is located within the first frustum, or

15

wherein the image source position generator (20) is configured to calculate (26) a second frustum for the second polygon and to determine (27), whether the listener position (130) is located within the second frustum, or

20 wherein the image source position generator (20) is configured to calculate (26) an invisible zone frustum and to determine (27), whether the listener is located within the invisible zone frustum.

25 9. Apparatus of claim 8, wherein the image source position generator (20) is configured to define four planes having normal vectors pointing inside the first frustum, the second frustum or the invisible zone frustum, and

30 wherein the image source position generator (20) is configured to determine (27), whether a distance of the listener position (130) to each plane is greater than or equal to 0, and to detect that the listener is located within a frustum of the first frustum, the second frustum or the invisible zone frustum, when the distance of the listener to each plane is greater than or equal to 0.

35 10. Apparatus of one of the preceding claims,

wherein the image source position generator (20) is configured to calculate the additional image source position (90) as a position between the first image source position (62) and the second image source position (63).

11. Apparatus of claim 10, wherein the image source position generator (20) is configured to calculate the additional image source position (90) on a connection line (91) between the first image source position (62) and the second image source position (63).
5
12. Apparatus of claim 10, wherein the image source position generator (20) is configured to calculate the additional image source position (90) as a position on a circular arc with radius r_1 around reflection point (92), where r_1 denotes the distance between the source position (100) and the reflection point (92).
10
13. Apparatus of claim 10 or 11 or 12, wherein the image source position generator (20) is configured to calculate the additional image source position (90), so that a distance between the additional image source position (90) and the second image source position (63) is proportional to a distance of the listener position (130) to the second visible zone (73), or so that a distance between the additional image source position (90) and the first image source position (62) is proportional to a distance of the listener position (130) to the first visible zone (72).
15
14. Apparatus of claim 11 or 12 or 13,
20
- wherein the image source position generator (20) is configured to determine a reflection point (92) using an orthogonal projection of the vector for the sound source position (100) and an orthogonal projection of the vector for the listener position (130) with respect to the first polygon (2) or the second polygon (3) or the adjacent edge between the first polygon (2) and the second polygon (3) or to determine a point where the first polygon (2) and the second polygon (3) are connected to each other as the reflection point (92) and,
25
- wherein the image source position generator (20) is configured to determine a section point of a line (93) connecting the listener position (130) and the reflection point (92) and the connection line (91) between the first image source position (62) and the second image source position (63) as the additional image source position (90).
30
15. Apparatus of one of the preceding claims,
35

wherein the image source position generator (20) is configured to calculate the first image source position (62) by mirroring the sound source position (100) at a plane (2) defined by the first polygon, or

5 wherein the image source position generator (20) is configured to calculate the second image source position (63) by mirroring the sound source position (100) at a plane (3) defined by the second polygon.

10 16. Apparatus of one of the preceding claims,

wherein the sound renderer (30) is configured to render the sound source so that a sound source signal is filtered using a rendering filter (31, 32, 33) defined by at least one of a distance between a corresponding image sound source position to the listener position and a delay time incurred by the distance, and an absorption coefficient or a reflection coefficient associated with the first polygon or the second polygon, or a frequency-selective absorption or reflection characteristic associated with the first polygon or the second polygon.

20 17. Apparatus of one of the preceding claims,

wherein the sound renderer (30) is configured to render the sound source using the sound source signal and the sound source position (100) and the listener position using a direct sound filter stage (31), and to render the sound source using the sound source signal and a corresponding additional sound source position and the listener position (130) as a first order reflection in a first order reflection filter stage, wherein the corresponding image sound source position comprises the first image sound source position, or the second image sound source position or the additional image sound source position (90).

30 18. Method of rendering a sound scene having reflection objects and a sound source at a sound source position, comprising.

35 providing an analysis of the reflection objects of the sound scene to determine a reflection object represented by a first polygon (2) and a second adjacent polygon (3) having associated a first image source position (62) for the first polygon and a second image source position (63) for the second polygon, wherein the first and second image source positions result in a sequence comprising a first visible zone

(72) related to the first image source position (62), an invisible zone (80) and a second visible zone (73) related to the second image source position (63);

5 generating an additional image source position (90) such that the additional image source position (90) is placed between the first image source position and the second image source position; and

rendering the sound source at the sound source position and, additionally

10 rendering the sound source at the first image source position, when a listener position (130) is located within the first visible zone,

rendering the sound source at the additional image source position (90), when the listener position is located within the invisible zone (80), or

15 rendering the sound source at the second image source position, when the listener position is located within the second visible zone.

19. Computer program for performing, when running on a computer or a processor, the
20 method of claim 18.