VISUALIZATION OF IMPULSE RESPONSE OF VIRTUAL ACOUSTIC ENVIRONMENTS

Adam J. Sporka, Pavel Slavík

Department of Computer Science
Czech Technical University
Karlovo náměstí, 12000 Praha *2*
Czech Republic
sporkaa@cs.felk.cvut.cz, slavik@cslab.felk.cvut.cz
http://cs.felk.cvut.cz

ABSTRACT

There are many methods that allow generate realistic 3D acoustic scenes. By means of new technologies it is possible to achieve a high level of realism. In many applications, the pictures are enhanced by sounds generated in 3D environment. This creates more realistic impression for the user. There are many methods that allow us to generate 3D sound and in such a way create high degree of reality in synthetic scenes. Nevertheless when generating sound in 3D environment, it is necessary to have relatively detailed information about the sound characteristics of the scene. Due to the complexity of this information it is necessary to visualize the sound distribution in the space where the sound will be generated. This paper describes a system of visualization of the impulse response of given acoustic environment whose development is one of our current research projects. We notice here its purpose, principles of implementation and basic applications.

Keywords: geometric algorithms, computer graphics, visualization, virtual acoustics, impulse response

1. INTRODUCTION

With increasing importance of virtual 3D acoustics, it becomes necessary to provide instruments for interactive acoustic modeling.

In this paper, we would like to describe a visualization engine of impulse response, the product of one of our research projects. The objective of this project is to provide a system that visualizes the simulated sound field in acoustic environment, an instrument for interactive exploration of acoustic environments. With such a system, the user obtains a possibility of comfortable estimation of acoustical properties of the virtual scene.

We have chosen the impulse response as the fundamental source of information about the 3D scene. According to [Kolme80a], the *impulse*

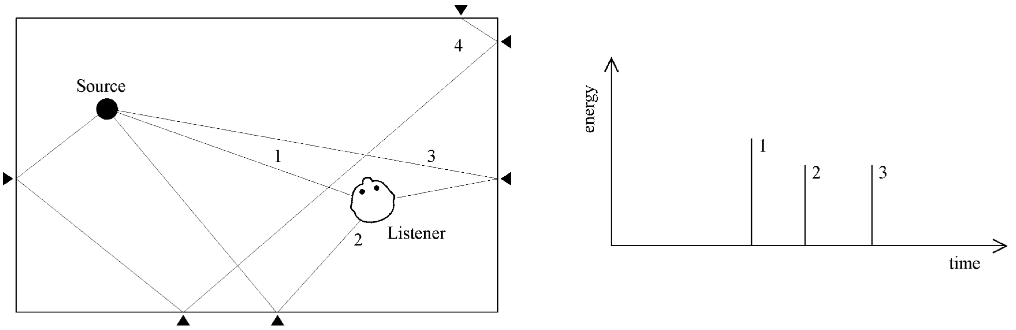


Figure 1: Introduction of the impulse response

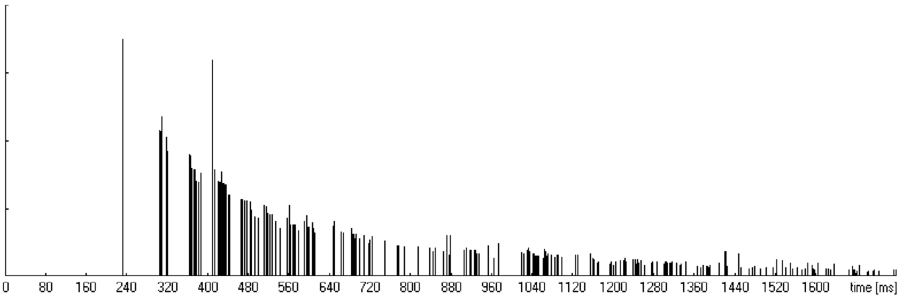


Figure 2: An impulse response of the large hall

response (IR) originates from multiple reflections on walls of acoustic impulse emanated from the sound source. It is a set of echoes of an impulse that are audible in a selected location for that is the IR generated. The IR is dependent on obstacle configuration and position of the sound source and the listener.

Fig. 1 displays the correspondence between sound echoes and the IR. As we can see, echoes 2 and 3 are attenuated compared to the direct sound (1) due to interaction with wall (small black triangles) and arrive later—this is caused by speed of the sound. As we can see, no record in the IR exists for echo 4, since this sound misses the target. Moreover, this echo is attenuated after its fourth reflection bellow a specified intensity level.

As another example, on fig. 2, we can see an IR of a large hall with concrete walls.

The IR is a great source of information about acoustic environment. For example, we can measure the *reverberation time* as time period between the first and last audible echo; using the IR as a convolutional filter for a sound signal of the source, we can generate very precious approximation to the real sound audible in given location. This process is called *spatialization.*

Having only single impulse response, we know acoustic properties of the environment in small area around the listener only.

To describe the whole environment, IRs form multiple locations are to be gathered. We can obtain them either from some acoustic measurement or using an output from a visualization engine. Since we consider the simulated acoustic environment, we shall discuss the latter approach in this text.

Having solved problem of IR calculation, another problem arises—the evaluation of the properties of the whole scene model. For example we are to design an auditorium. First, we design its model using a 3D scene editor, and then we generate a set of impulse responses for various positions of sound source and listener. We can create a virtual acoustic environment which user walks through while listening to spatialized sound using appropriate IR for his current position and the sound source.

Today, using relatively inexpensive technologies (A3D etc. [Aurea98a]), we can implement such a system with good approximation to the physical reality. Nevertheless, to evaluate the properties of the whole scene means to walk through the every corner and corridor of the virtual scene.

The better solution is to display such a set of impulse responses using some method of visualization.

2. BASIC PRINCIPLES OF VISUALIZATION ENGINE

In this section we discuss particular implementation problems. Our primary task is to calculate multiple impulse responses for a set of locations in the scene. We have many possibilities, which set of locations to choose.

A suitable selection is to use centers of the cells of the Cartesian or regular orthogonal grid. The result will be then a three-dimensional matrix of impulse responses that is four-dimensional matrix of floating-point numbers due to the IR storage method described later in this text.

To visualize the IR, we need only to visualize this matrix; and this task is relatively simple problem of computer graphics.

2.1. Representation and calculation of IR

Currently, there exist many methods of IR calculation. There are methods based on geometrical acoustics, methods of statistical acoustics, numerical methods, usage of neural networks etc.

Since our approach to the solution is geometrical, we use principles of geometrical acoustics, which describes the sound propagation using an abstraction—the *sound ray*.

For the first approach, we have decided to employ a geometrical method of ray tracing for ease of implementation and also to omit the influence of the sound diffraction and diffusion.

The sound ray has very similar properties as the ray of light as defined by 3D computer graphics. The ray is emanated from its source and travels along a straight path through the environment. In case of collision with an obstacle, it is partially absorbed and reflected in accordance to the energy conservation law. The most important difference from light rays is that we must take into account the sound velocity.

We can define the *reverberation path* as a sequence of sound rays where each ray is the reflection of the previous ray in this sequence except the first one, which is directly emanated from the source. The reverberation path is in fact the trajectory of appropriate sound echo (see fig. 1).

To calculate an impulse response for certain location in the environment, we have to find all the reverberation paths that end in the near vicinity of the location.

For each such path we calculate its total length and level of remnant energy carried by appropriate echo; as written, each reflection in the path causes the echo to loose an amount of its energy due to absorption of obstacles.

Then we can represent the impulse response using a list of vectors of the form (t, l), where t is time for which the sound ray has traveled along the reverberation path from the source to the spot and l is the amount of the remnant energy.

IR is more often represented as sample buffer where each sample gathers reverberation paths with range of time t determined by the buffer's temporal resolution. Such representation is called discrete impulse response. In this case we implement the storage of the vector (t, l) as superposition of value l over the level of the corresponding buffer sample.

For high-quality reverberators, the temporal resolution must be at least the same as temporal resolution of auralized sound signals, that is in order of 10μs. As we will see later, for purposes of the visualization, the satisfying resolution is from 1 to 10 milliseconds.

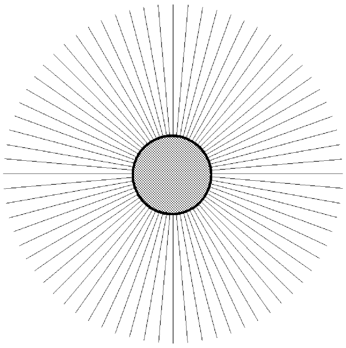


Fig. 3: Radial field approximation

To calculate a single impulse response for given position of the sound source and given point of IR, we can use the ray tracing method modified to purposes of acoustics. The major modification is in the ray generator. We create a finite set of rays that originate in position of the sound source. Their directions must be chosen so that they approximate a radial field (see fig. 3).

Then we trace each ray to the place of the first collision with an obstacle. There we perform attenuation of carried energy and generate a new ray with direction under proper angle of reflection. In fact, this way we construct the reverberation path for the echo carried by the ray. We perform this tracing until the amount of energy is above specified level and the ray has not hit the target—the vicinity of the point of IR. The vicinity can be defined as a simple spatial object, a sphere for example.

If the ray has hit the target, we write appropriate record into the calculated IR.

To calculate multiple impulse responses, we use a method very similar to the recent one. Besides the generation of multiple rays, we generate multiple targets as well. In the previous text we have mentioned our decision to use centers of cells of the Cartesian grid as the targets, since we calculate 3D matrix of the impulse responses. It is suitable to define the vicinity as the volume of the whole cell of this grid.

For each ray in the set we perform the described raytracing, but now, we will not inhibit any ray that has hit the target but we still keep the principle of small energy level ray inhibition. Having created the reverberation path, we must merge it into the IR matrix—that is, to check for all cells in the grid, whether the ray has stepped through them. For those, where it did, we determine, in what time it happened. Then we calculate the length of the intersection of the ray and the grid cell. If the time of arrival and leaving does not fall into the same IR buffer sample, we must divide this length into two aliquot parts for two successive samples of the same IR.

According to this length and the amount of energy carried by the ray, we calculate the increment to the impulse response. Finally, we add the increment to the appropriate sample in the IR buffer for the affected cell. Fig. 4 displays a view to the IR matrix in three successive ‘time frames'. The process of merging the reverberation path into the IR matrix is clearly visible. Affected cells in the given time are marked with thick borders and the gray-level represents the value of the increments.

Using terms of computer graphics, we draw antialiased line into the 4D voxel matrix.

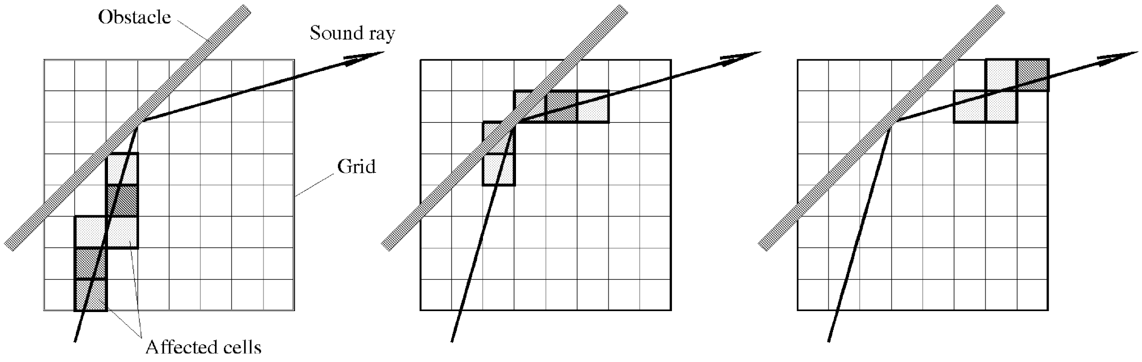


Fig. 4: Influence of the single ray to the IR matrix. Three succesive frames are displayed.

2.2. Visualization Methods

Having generated the 4D IR matrix, we must make a decision how to display it. Since our matrix has the time dimension, it is natural to consider it as animated voxel picture. Therefore we can consider this matrix as a sequence of 3D matrices.

From the physical point of view, the 3D matrix in given time shows the position of the wave front emanated from the sound source in zero time. Therefore, the animation shows us the distribution of the sound from its source.

According to previous paragraphs, our task is reduced to display a sequence of voxel matrices.

The simplest method is to display 2D cross-view along a chosen direction parallel with an axis of the orthogonal co-ordination system of the IR matrix. It is also necessary to generate a projection of the scene and paste it over the matrix image.

Figure 6 contains several frames resulting from the described visualization system. The higher value is in the matrix, the darker is the appropriate pixel.

Another idea is to display animated voxels of IR matrix directly within the model of the scene. Each voxel can be displayed as a certain geometrical object and its value can be represented as the size of the formation. User can freely move within the model of the scene, as he knows it and view very vivid animation of sound distribution. In our system we display co-ordinate system aligned cubes, each placed in the center of the appropriate grid cell.

Screen shots of our implementation of such system are in the figure 7.

However, our system has very large memory requirements. For example, to visualize the IR of a small auditorium of size, say 10×10×3 meters with 20 cm spatial resolution, we need to store 37500 voxels per one frame. Now, if each IR should have temporal resolution 0.001 s and maximal length 1.5 s, we need to store 1500 frames, that is about 55 millions of numbers.

To decrease the memory requirements, we can introduce several restrictions. For example, if we want know only what is the overall sound energy distribution, we can merge all reverberation paths into one frame only. The result matrix will describe the sum of all echoes audible in each cell of the grid. This is useful for the general acquaintance with sound properties of the scene (Fig. 5; figures 5, 6 and 7 contain visualization of the IR of the same acoustical environment.)

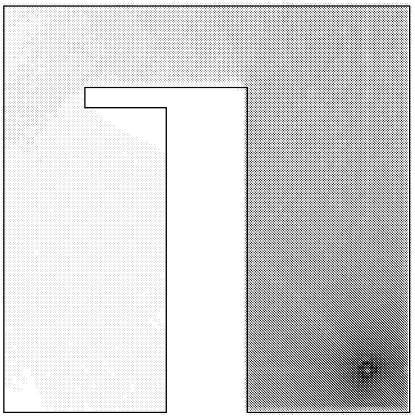
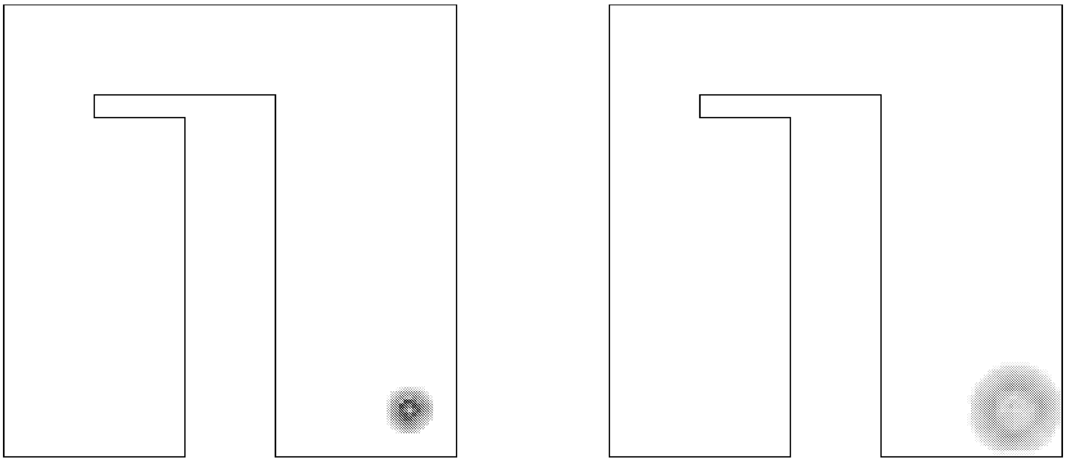


Fig 5: Sum of all echoes for each voxel of the IR matrix

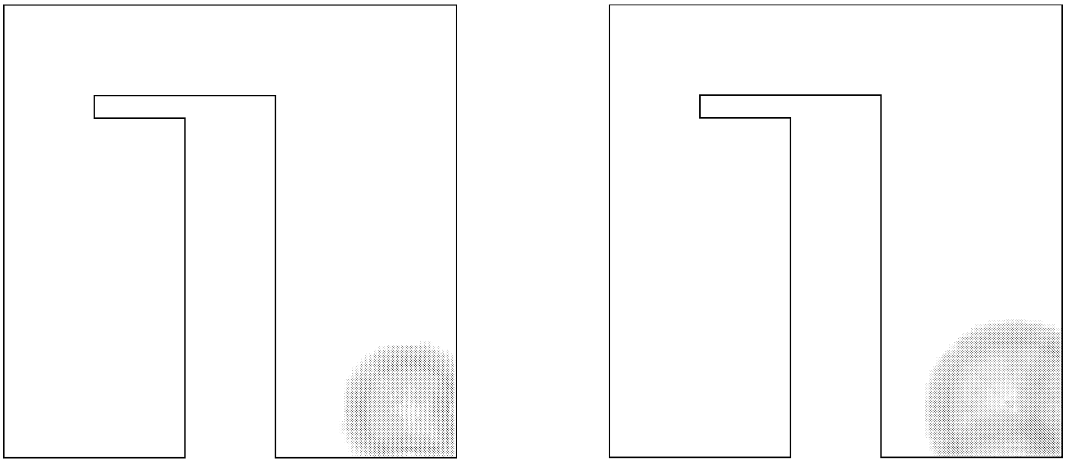
On the other hand, there could not be any need to cover the whole scene with the voxel grid; there may be only regions that we are interested into. For example, if we design an auditorium, we care only about the area, where the listeners sit, or even more—where their heads are located. Then we decrease the volume of the grid, but the spatial and temporal resolution can remain the same.

Since many acoustic phenomena are frequency dependent, we can add another dimension to our IR matrix. This would provide us a possibility to study qualitative properties of the sound field as well.



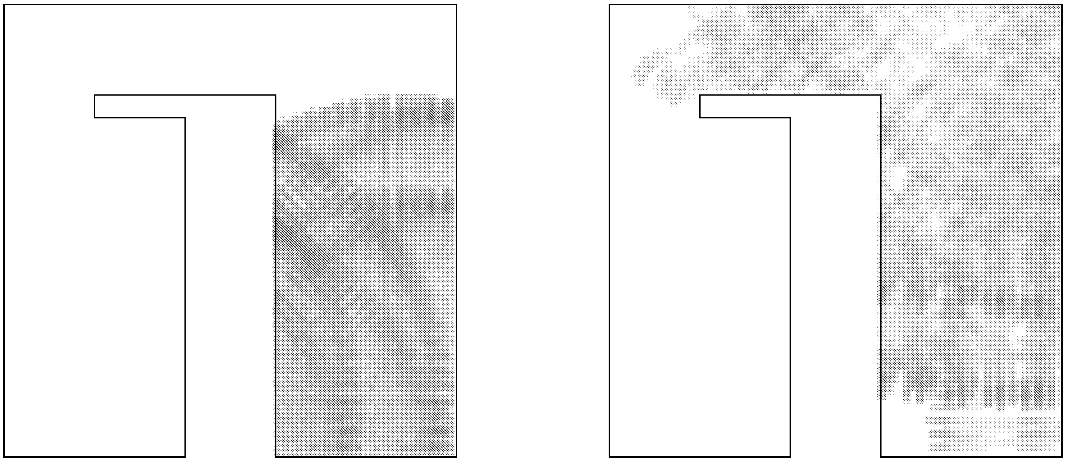
Frame 1, time 0.0018s

Frame 2, time 0.0036s



Frame 3, time 0.0054s

Frame 4, time 0.0072s



Frame 13, time 0.0234s

Frame 35, time 0.063s

Fig. 5: Sound distribution viewed as 2D cross-view.
Several frames of the animation. The longest walls of the room are 12 m long;
speed of sound is 340 m/s, the impulse occured at time t=0.

3. APPLICATIONS

There exist many possible applications of this system. The most apparent is the visual exploration of acoustical properties of virtual reality scenes.

As written earlier, we can extract much information from the IR, and even more when we have multiple impulse responses with possibility to easily compare each with other. Then we can for example compare reverberation time of different rooms in the scene, etc.

Such an exploration is necessary in cases when acoustic properties play an important role in multimedia applications where we enhance pictorial information with the 3D sound. Adding sound, we increase substantially the degree of realism. Another type of applications can be found in the field of software for blind and visually impaired persons where the high quality of sound information can give the user at least some idea about spatial arrangement of the scene.

By means of the visualization engine we can investigate the spatial characteristics of the scene and determine some critical parts. The system could be also used to investigate properties of public places where acoustics play an important role like concert halls, lecture rooms etc.

We can easily simulate and visualize the sound distribution in the scene for multiple positions and configurations of the sound source. Moreover, we can find quickly some critical spots that could be removed by some partial reconfiguration of the scene. A very good example could be application of the visualization engine in a theater. Here we can investigate the acoustic properties of the stage where various objects are located.

Here we can investigate spots where the audience can not hear actors or actors themselves can not hear properly the speech from their counterparts. With the increasing importance of the sound in computer applications, there will be more and more potential applications for the system developed.

4. CONCLUSION

We have implemented a simple impulse response visualization engine that employs classical algorithms of the computer graphics. This system allows user to visually evaluate acoustical properties of any virtual scene. Our system is useful also in applications where basic sound propagation laws are to be explained (demonstrations on seminars etc.) but also while developing new methods of IR calculation.

Nevertheless, much work has yet to be done in the acoustic simulation part and in the part of the visualization itself either.

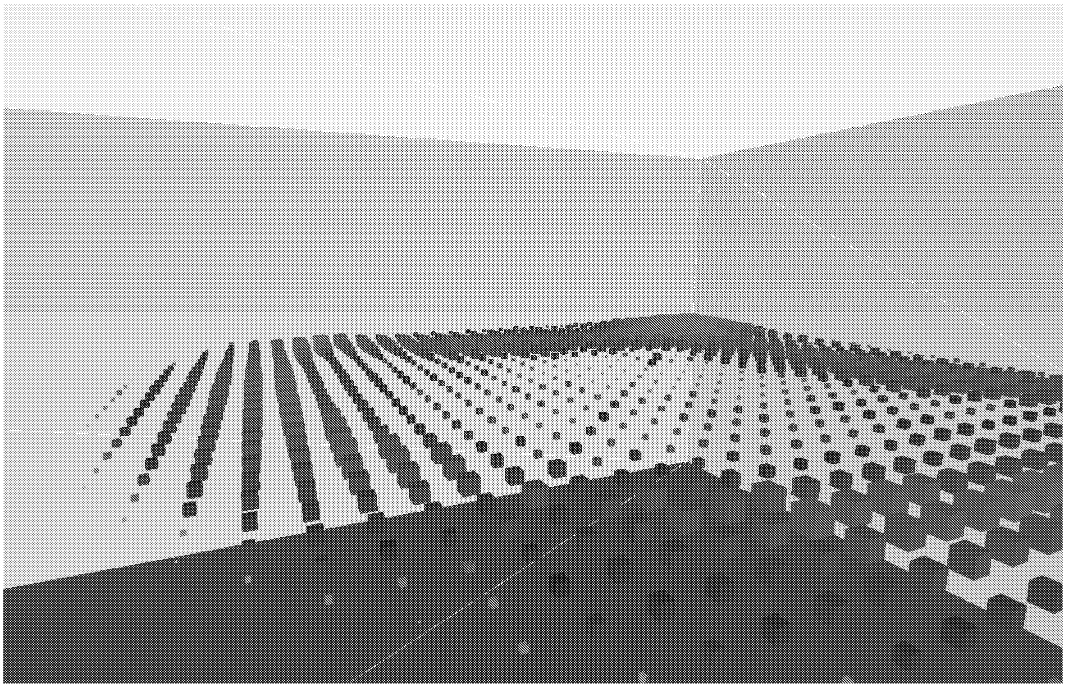


Fig. 6a: The 3D view to sound distribution. Frame 4, time 0.0072 s

Currently, we work on an implementation of the beam tracing algorithm within the described system. We expect increase of performance and robustness. Our system shall be enhanced to capability to deal with multiple and moving sound sources. Moreover, frequency dependent phenomena should be considered.

Nevertheless, even in current state of development our system provides useful results.

ACKNOWLEDGEMENT

This project has been sponsored by The Ministry of Education of The Czech Republic (Grant number OK 386).

REFERENCES

[Aurea99a] Aureal Semiconductor: A3D Software Development Kit

[Kolme80a] Kolmer, F., Kyncl, J.: Prostorová akustika (Spatial acoustics), *SNTL Prague*, 1980

[Spork98a] Sporka, A.: Rendering of Spatial Sound, *Bilateral Students Workshop, HTW-Dresden*, 1998

[Whitt80a] Whitted, T.: An Improved Illumination Model for Shaded Display. *ACM Communication on Graphics*, Vol. 26, No. 6, pp. 342-349, 1980

[Zara98a] Žára, J., Beneš B., Felkel, P.: Moderní počítačová grafika (Modern Computer Graphics), *Computer Press Prague*, 1998

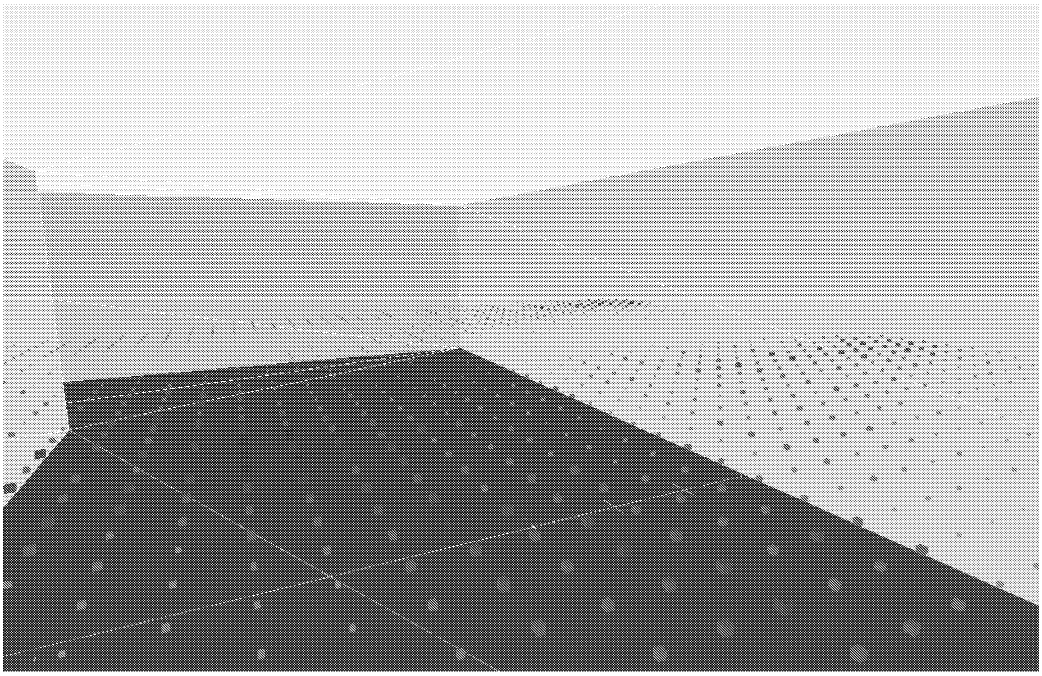


Fig. 6b: The 3D view to sound distribution. Frame 15, time 0.027s