DOI: 10.22032/dbt.39973

DOI (proceedings): 10.22032/dbt.39936

Abstract Reviewed Paper at ICSA 2019

Presented by VDT.

3D audio for live events

L. Nipkow

Silent Work GmbH, Zurich, Switzerland, Email: lasse.nipkow@silentwork.com

Abstract

Many applications for 3D audio are aimed at live events. Several manufacturers of PA systems now offer solutions. These are not optimized for native 3D audio, but achieve an immersive impression due to placement of objects and added reverberation.

To operate native 3D audio for live events, there are two key questions to consider:

- How should the loudspeaker layout be chosen so that the audience gets the most impressive listening experience at a reasonable cost?
- How should 3D audio content be designed so that it reaches the audience as effectively as possible through such PA systems?

Two phenomena can be used for 3D audio: Envelopment and projection of sound sources in the front. The most impressive case is provided by a loudspeaker setup using imaginary connecting lines between the loudspeakers, which results in a volume. Vertical phantom sound sources presented in the front sound more natural than horizontal ones. The number of vertical front loudspeakers determins the resolution of the image.

The wedge-shaped loudspeaker setup, that means with height loudspeakers in the front, meets both requirements minimally. The greater the listening area, the more loudspeakers are needed. The greater the distances of the loudspeakers to each other, the more audible holes are perceived, depending on the acoustics of the playback room.

1. Introduction

3D audio has enormous potential: listeners can experience previously unheard sound at events. This is especially the case, as long as 3D audio setups are rarely found in homes, cinemas and cars and thus consumers do not have access to it.

To make the most out of 3D audio, event organizers must meet some requirements; these are very complex:

- Guarantee optimal room acoustics.
- Provide a sufficiently high density of loudspeakers in the reproduction room.
- Set up a meaningful loudspeaker layout so that psycho-acoustical effects such as vertical stereophony and the cocktail party effect can be used.
- Produce suitable content that uses the loudspeaker setup highly natively.

Depending on the situation, such as open-air situation or addressing a hall, requires different measures to achieve a convincing result.

2. Loudspeaker layout and psycho-acoustics

2.1. Initial situation for the playback of 3D audio

Sound engineers create reproducible 3D audio content primarily in studio and movie rooms. These rooms sound very good in contrast to large event halls. This raises the question of how to best scale up mobile PA systems for 3D audio, in order to lose the least noticeable added value of 3D audio. Therefore, it is crucial to look at the appearing phenomena of 3D audio and to investigate which perceptible sound changes occur when the scaling factor changes. For this purpose, the loudspeaker layout of the two components sound sources (direct sound) and envelopment (room sound) are first considered in isolation from each other and then combined.

2.2. Loudspeaker layout for envelopment

Envelopment is one of the most important features of 3D audio. Spatiality comes from sound, which is reflected in rooms at the boundary surfaces; the sound of the room meets the listener distributed equally from all directions. Convincingly appearing envelopment therefore requires addressing of the listener with room sound from different directions.

2.2.1. Envelopment for the horizontal plane

For the simplest case of an envelopment by PA, four loudspeakers are used; they stand at the same distance around the listener on the horizontal plane, see **Fig. 1**.

This works very convincingly in the sweet spot for listening situations in control rooms and living rooms. However, if the volume of the room, and thus the area of the loudspeaker setup is much larger, i.e. this means more than 5m between

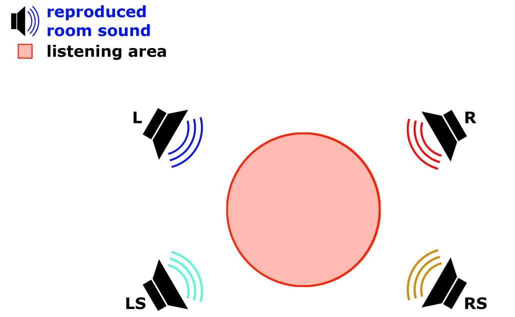

Fig. 1: View from above on the loudspeaker setup: The listener already feels impressively enveloped by four loudspeakers that reproduce the room sound. This requires relatively small distances between the loudspeakers (up to about 5m x 5m). Due to the small number of loudspeakers, the resulting listening area is small. When a listener approaches one of the loudspeakers, the level rises from that direction. If the level exceeds 10 dB, the remaining loudspeakers are no longer effective for an envelopment.

adjacent loudspeakers, noticeable holes between the loudspeakers occur outside the sweet spot. This can be avoided by placing the loudspeakers in a sufficiently high resolution around the listener; thus the mentioned perceptible holes do not appear, see **Fig. 2**.

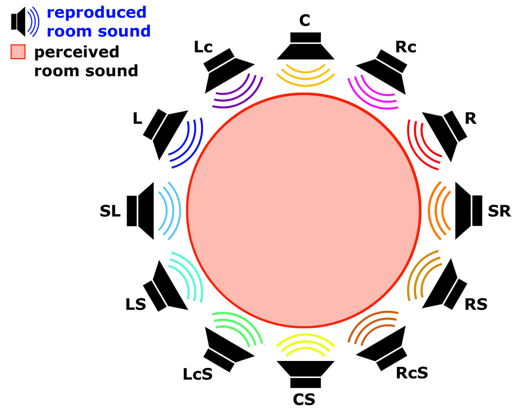

Fig. 2: View from above onto the loudspeaker setup: When loudspeakers reproduce room sound around the listener from all horizontal directions, the listener feels a very impressive envelopment. Since many loudspeakers reproduce room sound from all directions, the room sound is distributed very regularly in the room. A listener can get much closer to a loudspeaker before its sound exceeds the 10dB threshold.

On the one hand, the cost of a high-resolution loudspeaker setup increases with each additional loudspeaker. On the other hand, the resolution also has an influence on the perceptibility of the individual loudspeakers as sound sources. The room sound is very balanced in the center of the listening area. This means from that position, with a favorable signal selection each loudspeaker signal is perceptible. The more a listener approaches a loudspeaker, the louder its signal becomes in relation to the other loudspeakers. If, at the location of the listener, the level of the nearest loudspeaker exceeds 10 dB compared to other loudspeaker signals, a clear misbalance of the spatiality occurs [1]; the listener perceives only the loudest loudspeaker, compare cocktail party effect [2], see **Fig. 3**.

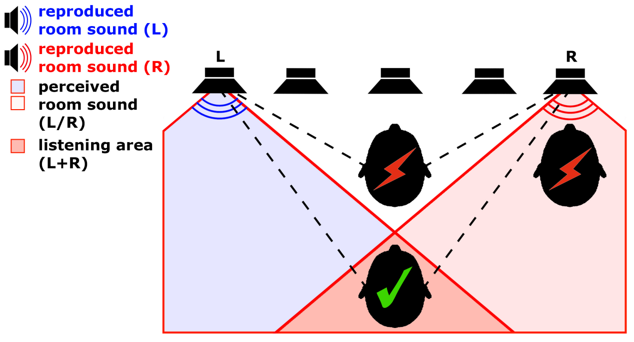

Fig. 3: View from above on a part of a 3D audio loudspeaker setup: If a listener is in the listening area, he / she hears both room sound signals from left and right about equally loud and at a relatively small angle. If the person is close to a loudspeaker, he / she practically only hears that signal. If the listener is very close to the front between two loudspeakers, the angle between the two loudspeakers becomes very large: In this case, the listener perceives a hole in the middle.

The optimal resolution or optimal distance between the adjacent loudspeakers can be determined on the basis of the radiation characteristic (dispersion): The narrower the dispersion angle, the closer the loudspeakers must be to one another. Furthermore, the minimum possible distance between a person and the loudspeakers can be influenced: with a doubling of the resolution, the minimum possible distance is halved.

The expense of up-scaling such a loudspeaker arrangement and at the same time avoiding perceptible holes becomes greater as the listening area increases. **Fig. 4** shows a case of a square loudspeaker setup with a width of 20m or a listening area of about 400m^2.

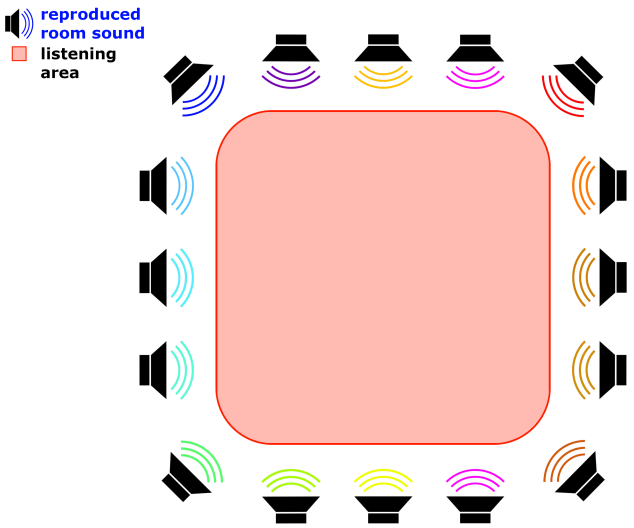

Fig. 4: View from above onto a large loudspeaker setup with a listening area of about 400m^2: The loudspeaker ring produces a homogeneous sound field with room sound at a sufficiently high resolution.

2.2.2. Envelopment for 3D audio

The effect of the envelopment becomes more impressive when the loudspeaker setup spans a volume rather than a horizontal plane. In this case, the listener has the impression of being in the recording room when using meaningfully selected loudspeaker signals, see **Fig. 5**.

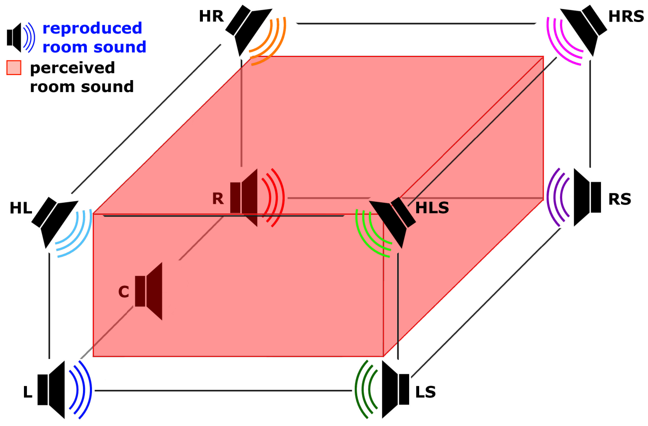

Fig. 5: The reproduced room sound between the front and rear loudspeakers should be balanced in level to ensure an impressive envelopment. In this case, a listener can be placed almost anywhere in the setup and has the impression of being in the recording room.

The expense of up-scaling such a loudspeaker arrangement and at the same time avoiding perceptible holes increases with a larger listening area as well here. The following chapters show that due to psychoacoustic laws, a loudspeaker arrangement requires much fewer loudspeakers than expected from a geometrical point of view.

2.3. Signal quality for optimal envelopment

Perceived envelopment occurs during loudspeaker reproduction, when all loudspeakers in the setup reproduce room signals of similar levels and a correlation around zero [1]. The correlation has a considerable influence on the perceived spatiality; it increases prominent at a degree of coherence of less than k = 0.2 (ear signals), see **Fig. 6** [2].

The degree of coherence of the sound source signals is usually different from that of the ear signals. Incoherent sound source signals result in incomplete incoherent, i.e. in partially coherent ear signals [2]. This means that the correlation of loudspeaker signals must be less than the degree of coherence in ear signals in order to produce the maximum possible spatiality; the optimum is around zero.

A correlation of 1, however, is very unfavorable for room signals; correlation 1 corresponds to a sound with identical loudspeaker signals. There may be strong audible comb filtering effects. Spatiality, on the other hand, occurs primarily in reflected sound from side walls [3].

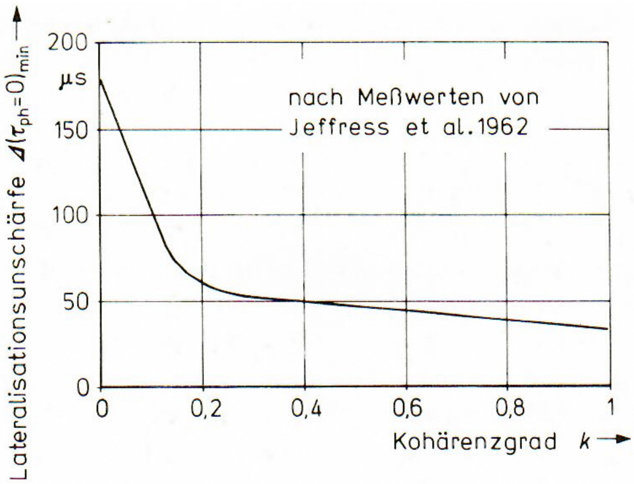

Fig. 6: Lateral blurring Δ (τph = 0)min as a function of the degree of coherence of the ear signals. Low pass noise fg = 2kHz, level about 90dB, 7 subjects. The measured value for k = 0 corresponds to complete uncertainty in the experimental setup used [2].

Sound engineers should avoid such signals for the side loudspeakers in particular; in many cinemas with a surround sound system, however, this is exactly what is done by distributing few signals by the audio playback systems to many surround loudspeakers.

2.4. Reduction of the number of loudspeakers

Especially in a mobile PA system, the outlay must be in a sensible relationship to the benefit. Therefore, loudspeakers which provide no significant added value in the set position in the loudspeaker setup should be omitted.

2.4.1. Reduction of height loudspeakers

At many live events there is a stage or a clear orientation of the audience to the front. This means that side loudspeakers lead to a two-sided lateral addressing of the audience with room sound.

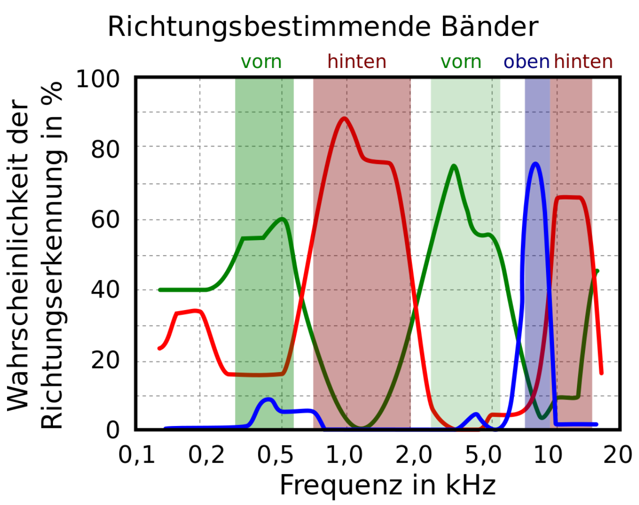

Fig. 7: The directional bands [2] make it possible to omit loudspeakers in the setup without compromising on sonority: Listeners perceive both sides of laterally reproduced room sound with an increase in high frequencies. As a result, the room sound not only appears from the sides but also from both above and from behind.

Lateral sound addressing on both sides of the listener leads in turn to emphasis of high frequencies [4] among others also in the range around 8kHz similarly to height loudspeakers of a 9.1 loudspeaker setup. This means that the directional bands elevate room sound with a correlation around zero, see **Fig. 7**.

Since these are signals with a correlation around zero and not identical signals, the elevation effect between bottom and top is blurred; this contributes to the filling of the above-mentioned holes. The balance between the left and right sides may be around 15dB when used with a side-positioned loudspeaker; at higher level differences, the elevation effect decreases rapidly, and the listener can only perceive the sound of the louder loudspeaker as a real sound source.

Due to comb filtering effects lateral addressing also leads to elevation effects with lower intensity even at low frequencies at 600 Hz [5], see **Fig. 7**.

2.4.2. Reduction of the rear loudspeakers

Listeners who are located far in front of the listening area can not perceive room sound from behind because of the masking due to direct sound and lateral addressing due to room sound; the level difference is much higher than 10dB, compare cocktail party effect [2].

As described above, both-sided lateral addressing of listeners leads to an emphasis of high frequencies. In addition to 8kHz higher frequencies around 10kHz are raised and lead to a perception from behind. Thus, listeners have the impression of perceiving reverberation also from the top and rear when using side loudspeakers, see **Fig. 7**.

An addressing from the rear is associated with risks: Listeners who are located far back in the listening area and thus near the rear loudspeakers, primarily hear the sound of those loudspeakers. As a result, sound from all other directions, and especially from the front, may be heavily masked.

Room sound from behind leads to a narrowing of the stereo panorama for room sound and an increase in the degree of coherence of the ear signals respectively. Therefore, such signals do not contribute to the spatiality but only to an avoidance of perceptible holes.

Furthermore, concertgoers in halls or churches and in the vicinity of the back wall do not perceive room sound from that direction. Due to their individual position in the room, they perceive room sound primarily from the direction where the largest proportion of the room volume lies.

All of these reasons argue that loudspeakers should not be used on the back of the loudspeaker setup regarding room sound. This results in the following structure for the setup, see **Fig. 8**.

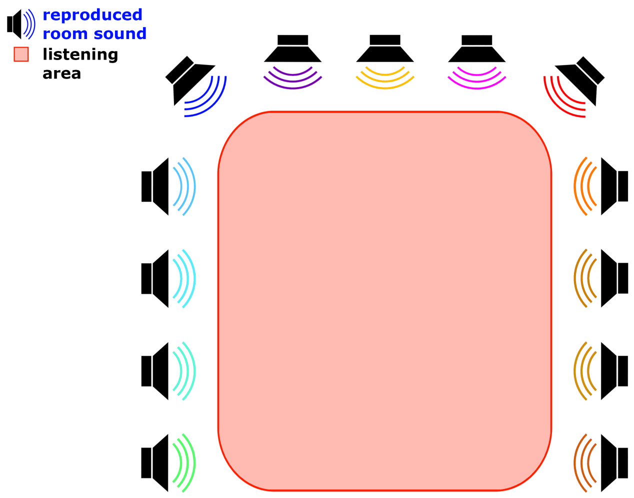

Fig. 8: View from above of a loudspeaker setup for a 3D audio PA system without height loudspeaker as well as without rear loudspeakers. Due to psychoacoustic phenomena, the room sound appears both from the back and from above for listeners within the loudspeaker setup at a favorable signal selection.

2.4.3. One-sided spatiality

One-sided spatiality occurs when room sound reaches the listener only from one side. If, instead of a real sound source, several loudspeakers are arranged on a line to the side of the listener, the reproduced sound covers an angle of sound incidence of up to approximately ±70°. Since in this case several loudspeakers reproduce sound with correlation around zero, the listener does not perceive a real sound source anymore, but sound that largely corresponds to side reflections. This type of sound addressing also leads to elevation effects. At least 4 loudspeakers are necessary for this, so that on the one hand the angle of sound incidence is sufficiently large and on the other hand no perceptible holes occur between the lateral loudspeakers, see **Fig. 9**.

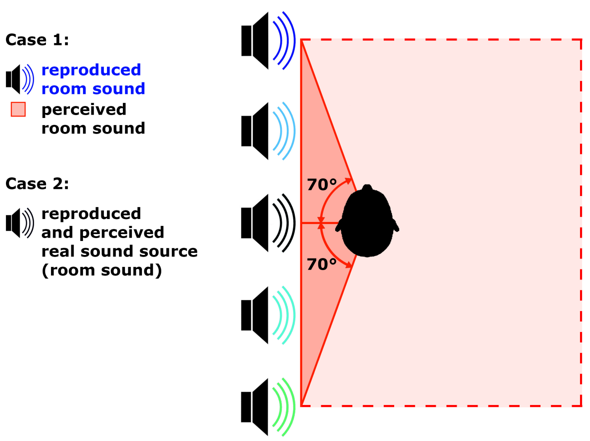

Fig. 9: View from above onto a loudspeaker setup. In case 1, 5 loudspeakers reproduce room sound with a correlation around zero. The listener perceives a one-sided spatiality on the left, extending from the front to the rearmost loudspeaker. This corresponds to slightly less than half (about 40%) of an envelopment (100%) on the horizontal plane. In case 2 only one loudspeaker provides the listener from the side with sound. The listener perceives a real sound source instead of a spatiality; this corresponds to a point on the horizontal plane.

2.5. Loudspeaker layout for sound sources

For the simplest case of imaging by sound addressing a stereo configuration is used. On a wide stage with 10m or even 20m, the resulting stereo base width is too large; it would create a middle hole. In addition, in the case of phantom sources outside of the middle axis, some serious distortions of the image arise: the farther a listener moves away from the middle axis, the further the image shifts in the direction of the closer loudspeaker [6]. Therefore, the solution is similar to room sound: a sufficiently high number of loudspeakers within the stereo base.

2.5.1. Combination of real sound and phantom sources

An extension of a 2-channel stereo setup with additional loudspeakers raises the question as to what criteria the sound engineer should use to assign signals to the loudspeakers.

In the simplest case, the loudspeakers within the stereo base represent real sound sources. The direct sound of mono-captured or strongly directed instruments such as vocals, trumpet, bass drum, e-bass, etc. has no stereophonic width. This sound content is therefore predestined to be used as real sound sources. In addition, the transparency of real sound sources is much greater than that of phantom sources [6]. This has a dramatic effect, especially at high frequencies.

Acoustic instruments with resonating bodies such as grand pianos, strings and large bodies of sound such as choirs that are captured stereophonically require a stereo width in the reproduction in order to sound as natural as possible.

In general, the wider the stereo base for stereophonic microphones is, the more impressive the horizontal stereo reproduction becomes. Conversely, the narrower the chosen stereo base is, the more powerful / compact the sound sources will be. If the number of loudspeakers in the front is sufficiently high, stereo bases of different widths can be created and also several narrow stereo bases next to each other, see **Fig. 10**. Signals from stereo microphones and synthesizers or similar sources can be used. When using different adjacent narrow stereo bases, the instruments spatially delimit clearer from each other without sounding mono.

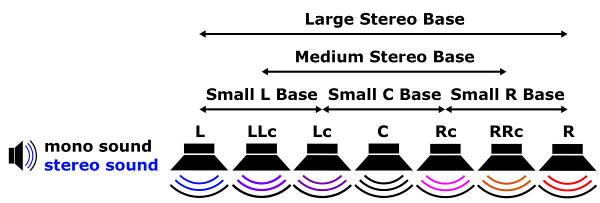

Fig. 10: View from above onto the loudspeaker setup: If there are enough loudspeakers in the front, different widths of stereo bases can be created and also several narrow stereo bases next to each other. Each loudspeaker can represent a real sound source, see center loudspeaker ‚C'.

In some situations, it makes sense to combine phantom and real sound sources for instruments. A typical case is the bass drum: The attack should appear in the center channel and the boost distributed over several loudspeakers in order to achieve a spatial extent of the instrument.

2.5.2. Projection of sound sources in the front

In the context of direct sound, early reflections lead to a distance impression of sound sources [7]. This occurs when direct sound and early reflections are reproduced from the same direction. If the lower loudspeakers in the front L-C-R reproduce direct sound predominantly and the upper loudspeakers in the front HL-HR reproduce early reflections, this results in an audible connection of the lower and upper front levels; this is what the author describes as a projection of the sound body in the front [8], see **Fig. 11**.

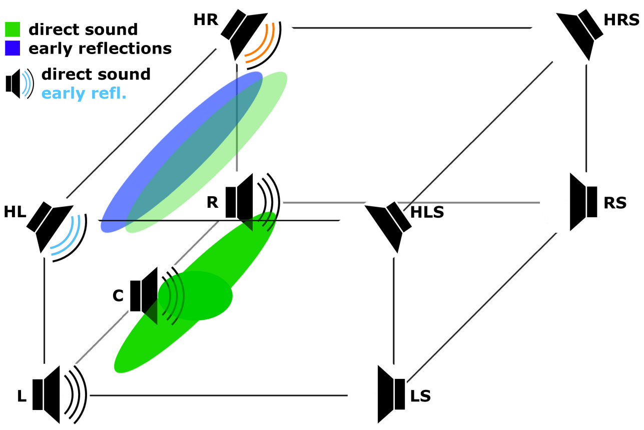

Fig. 11: If the lower loudspeakers in the front L-C-R reproduce direct sound predominantly and the upper loudspeakers in the front HL-HR reproduce early reflections, this results in an audible connection of the lower and upper front levels: projection in the front.

The instruments are mainly located at the lower loudspeakers due to direct sound. However, they sound more natural than without the early reflections; it corresponds approximately to the conditions in a concert hall near the stage.

2.5.3. Vertically reproduced direct sound

If one of the lower loudspeakers and the loudspeaker vertically above it reproduce stereophonic direct sound with sound content correlation [8], such as L-HL, this results in a similar natural effect as in early reflections, see **Fig. 12**.

In this case, the image is not clearly perceived from the lower loudspeaker but is located between the two loudspeakers involved. Noteworthy in this case is the localization sharpness in the horizontal plane, which corresponds to that of a real sound source [6]. The combination of a very sharp localization with a simultaneously strong naturalness of the instrument and impressiveness of a synthetic sound means that the sound source is perceived very direct and therefore attracts attention. This is especially true for sound sources with a high proportion of high frequencies [8].

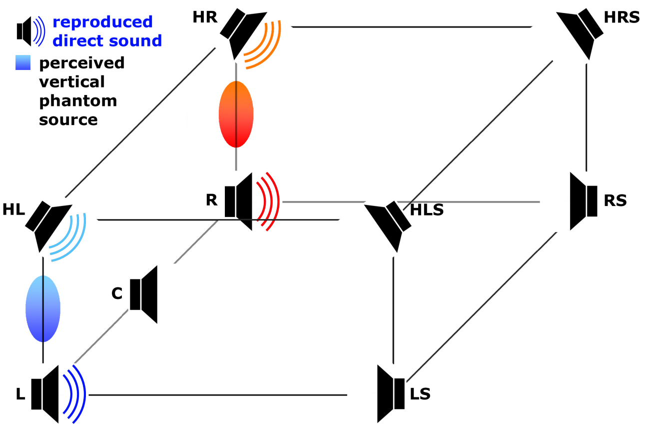

Fig. 12: In this 9.1 loudspeaker setup, two stereo pairs of two vertical instruments on the front edge reproduce two instruments / sounds. The transparency of a mix remains the same as if real sound sources were used for a surround sound reproduction.

The lower and upper loudspeakers must sound similar and have a wide dispersion characteristic both vertically and horizontally for vertical stereophony to work reliably. Only in this way it is possible for a listener to perceive both loudspeaker signals equally loud and, on the other hand, balanced levels between several vertical stereo pairs, largely independent of their position in the reproduction room.

Furthermore, vertical stereophony works optimally only if the lower and upper loudspeakers are perpendicular to each other and have no horizontal components. Otherwise, partially horizontal phantom sound sources arise.

As described above, vertical stereo pairs in the front edges of the loudspeaker setup are not enough at a 10m or 20m stereo base; there are noticeable holes in the image. The solution lies in a loudspeaker layout with two identical height levels of loudspeakers. Thus, both the above-described requirements for vertical arrangement and high resolution are simultaneously fulfilled, see **Fig. 13**.

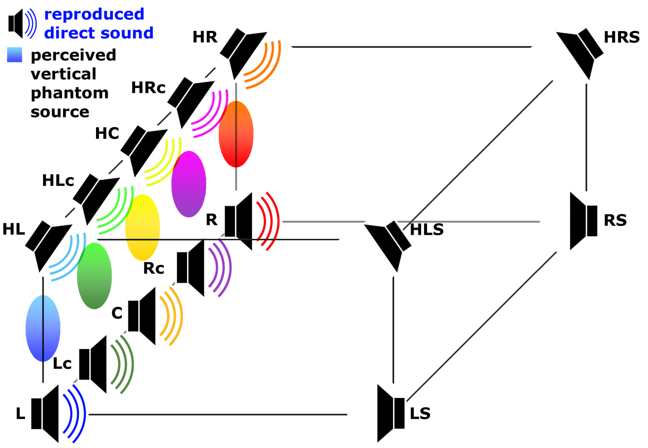

Fig. 13: In this extended 9.1 loudspeaker setup, five vertically and equidistantly arranged stereo pairs in the front represent instruments / sounds. The transparency is very high in this constellation compared to horizontal phantom sources in the front.

2.6. Loudspeaker layout for sound sources and envelopment

Sound sources (direct sound) and envelopment (room sound components) always occur at the same time and overlap each other in a PA system. Therefore, the result for a PA system with 3D audio is the combination of the two partial results, see **Fig. 14**.

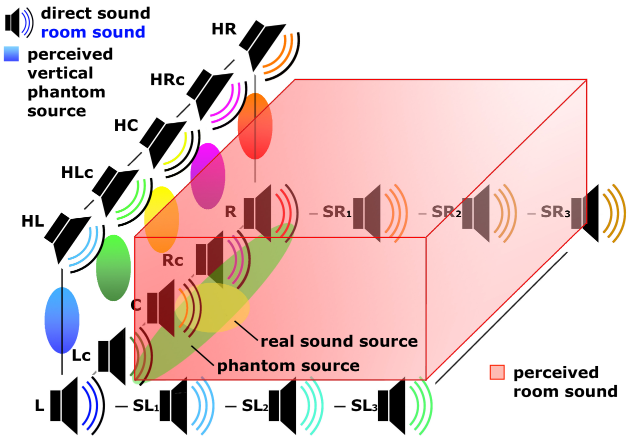

Fig. 14: In this extended 9.1 loudspeaker setup, five stereo and equidistantly arranged stereo pairs in the front represent instruments / sounds. All loudspeakers reproduce room sound and thus lead to an impressive envelopment.

If loudspeakers are placed at head height in the reproduction room, the listeners closest to the loudspeakers cover the sound reproduced by the loudspeakers, in particular of the high frequencies. As a result, only the front listeners can hear the front lower loudspeakers, and only the rear listeners can hear the rear lower loudspeakers unlimited. Thus, an imbalance occurs for most listeners. It is therefore more appropriate to arrange the lower loudspeakers slightly above the heads of the listeners to avoid such shading effects.

3. Applied psychoacoustics in the signal assignment for a 3D audio PA system

The previous chapters explain the optimal design for an economical 3D audio loudspeaker layout and basic considerations for the required loudspeaker signals. The next step is to consider how sound engineers have to assign microphones signals, samples, etc. to the many loudspeakers from a psychoacoustic point of view.

3.1. Signals for the envelopment

3.1.1. Native captured room sound

In many cases, captured room sound is the best choice for a plastic-sounding envelopment. The following conditions exist for this:

- The recording room must be acoustically excellent. Room sound from small rooms or rooms with acoustic problems are much more noticeable if reproduced by a 3D audio system than with mono and stereo. This means that such signals lead to bad-sounding recordings.
- The room signals must not have direct sound, especially at high frequencies. Otherwise, especially in the back of the listening area, listeners perceive sound sources from the direction of the direct sound components. In many cases, this leads to double and multiple images or increased muddiness.

3.1.2. Artificially generated room sound

It is not easy to provide room microphone signals with correlation around zero in the required number, especially for larger loudspeaker setups. Room sound processors and upmixing tools represent an alternative. As with the room microphone signals, important conditions must also be considered:

- Simple reverb units and plug-ins have a comparatively low quality for confounding real room sound. Only the best reverberation processors generate satisfactorily appearing room sound [8].
- Upmixing tools have the challenge of extracting spatial components from every possible existing sound material. Therefore, they are limited in the generation of plastic sounding room sound.

3.1.3. Combination of native room sound and upmixing

Especially in the case of already realized recordings, there are in many cases fewer captured room signals than are needed for sound addressing in 3D. One solution is the combination of existing room sound signals with an upmix of the original signal, compare [8].

3.1.4. Pad and noise-like sounds

In synthetically produced music there is usually no room sound. In most cases, room sound is also out of the question for that kind of music. Alternatively, pad and noise-like sounds can be used for the envelopment [6]. In many cases, it is sufficient to use multiple instances of the same sound to generate signals in sufficiently high numbers of channels with correlation around zero [8].

3.2. Assigning sound sources to loudspeakers

In this article, the author limits himself to the case of direct sound from the front. Direct sound from other directions as well as panning direct sound is not discussed here.

3.2.1. Mono sound sources

As described in chapter 2.5.1., mono sound sources are especially suitable for use as real sound sources. There are two cases:

- Cocktail party effect: When two or more real sound sources are present at different locations in the room, they are clearly distinguishable from each other and do not cover each other or only weakly [6].

- Sense of depth: If two or more mono sound sources appear at the same point in space, masking and merging effects occur [6]. Depending on the level ratios and spatial proportion of individual sounds, a pronounced sense of depth occurs [9].

3.2.2. Stereo sound sources

Most acoustic instruments and in particular string instruments radiate direct sound with different spectra depending on the direction. In addition, resonant bodies represent their own room, which behaves acoustically similar to a recording room [10]. Stereo close-miked instruments of this type therefore always have partly direct sound, some room sound, which cannot be separated easily. The stereo information of the signal is retained if the assignment to the loudspeakers is two-channelled. There are basically two possibilities for an assignment:

- Horizontal phantom source: At this assignment, the imaged sound source acts as an audible connection between instruments due to the expansion of horizontal phantom sources [8]. At the same time, masking and merging effects occur, allowing sense of depth. As described in chapter 2.5.1., different widths of stereo bases can be used during the mixing process.
- Vertical phantom source: At this assignment, the imaged sound source is spatially delimited from those with other directions. As a result, on the one hand, the transparency is very high. On the other hand, vertically stereo reproduced instruments sound more natural than horizontal phantom sources [8].

4. Conclusion

The cost of a live event PA system with 3D audio is very high compared to a 9.1 loudspeaker setup due to the even higher number of required loudspeakers and playback channels. By applying psychoacoustic phenomena, however, it can be kept to a reasonable level.

The spatial resolution for a natural sounding image is phenomenal, especially at larger loudspeaker setups with many front loudspeakers or vertical stereo pairs. So there is not just the chance of obtaining a large part of the psychoacoustically occurring phenomena in an upscaling, but due to the much higher resolution of sound sources in the front, the reproduction of music using 3D audio is also much closer to natural hearing with a maximum localization sharpness of about 1°!

5. References

[1] Nipkow, L.: Properties of Auro 3D room signals. 27. Tonmeistertagung, Cologne (2012), Proceedings, ISBN 978-3-9812830-3-7; 757-770

[2] Blauert, J.: Räumliches Hören. S. Hirzel Verlag, Stuttgart, 1974, ISBN 3-7776-0250-7

[3] Blauert, J.: Räumliches Hören. Nachschrift: neue Ergebnisse und Trends seit 1972. S. Hirzel Verlag, Stuttgart, 1985, ISBN 3-7776-0410-0

[4] Lee, H.: 'Sound Source and Loudspeaker Base Angle Dependency of the Phantom Image Elevation Effect' Journal of the Audio Engineering Society, 65 (9), pp. 733-748 (2017). ISSN 1549-4950

[5] Lee, H.: 'Phantom Image Elevation Explained'. In: Audio Engineering Society the 141st International Convention, Los Angeles, USA (2016)

[6] Nipkow, L.: Sound Design in 3D for Dance Music. 4. ICSA, Graz (2017), Proceedings

[7] Potratz, U.: Untersuchung der Gestaltungsmöglichkeiten früher Reflexionen mit Hilfe eines raumakustischen Modells. Diplomarbeit, Erich-Thienhaus-Institut der Hochschule für Musik Detmold (2005), http://www.eti.hfm-detmold.de/lehraktiv/diplomarbeiten/diplomarbeitenordner/da-potratz.pdf

[8] Nipkow, L.: The Importance of 3D in Immersive Audio. 30. Tonmeistertagung, Cologne (2018), Proceedings, ISBN 978-3-9812830-9-9

[9] Nipkow, L.: Sense of depth in 3D for Pop Music. 29. Tonmeistertagung, Cologne (2016), Proceedings

[10] Nipkow, L.: Room signals – properties and influence on Auro 3D recordings. 2. ICSA, Erlangen (2014), Proceedings, ISBN 978-3-98 12830-4-4, 96-101