Preprint 2887 (A3)

# On the Performance of Two-Channel and Multi-Channel Stereophony

Günther Theile,

Institut für Rundfunktechnik GmbH (IRT),
Munich, West Germany

**Presented at the 88th Convention 1990 March 13–16 Montreux**

AES



**AN AUDIO ENGINEERING SOCIETY PREPRINT**

# ON THE PERFORMANCE OF TWO-CHANNEL AND MULTI-CHANNEL STEREOPHONY

Günther Theile
Institut für Rundfunktechnik GmbH (IRT), Munich, FRG

ABSTRACT

**Psychoacoustic principles are considered in order to enhance the performance of loudspeaker stereophony. It is found that the simulation of depth and space in a conventional two-loudspeaker stereo arrangement can be optimized by means of the socalled room-related recording technique, which applies natural interaural signal differences instead of pure intensity or time differences for stereophonic imaging. However, for creating the natural perception of spatial impression and enlarging the listening area at the same time, additional loudspeakers must be applied. The benefit from additional front channels and surround channels of a future HDTV sound system is discussed not only with respect to picture accompanying sound but also with respect to high-quality sound without picture.**

## 1. INTRODUCTION

A particularly large number of studies have been published during the last few years with the goal of improving the capabilities of current stereophony. This applies to microphone techniques as well as to mixing and reproduction techniques, and major overall progress can be expected. Below, those possible developments in stereophonic recording and reproduction technique are to be described which may improve the "naturalness" of the stereophonic sound image and enlarge the listening zone in the playback room.

Firstly: How can the desired naturalness of the stereophonic sound image be defined? The simplest theorem would be: The reproduced sound image must correspond to the original sound image. This definition appears to be problematic because identity can definitely not be required, in principle, as a goal for optimizing the stereophonic recording and reproduction technique. Identity may be appropriate for head-referred stereophony (dummy-head stereophony), or perhaps for reproduction of a speaker's voice through loudspeakers, but it is probably appropriate to a limited extent only for reproduction of the sound of a large orchestra through loudspeakers. Aesthetic irregularities in the orchestra, poor conditions of room acoustics as well as the necessity of creating a sound image "suitable for a living room" - i.o.w., the essential problems of loudspeaker stereophony - actually force a deviation from identity. The desired natural stereophonic sound image should therefore meet two requirements: It should satisfy aesthetically and it should match the tonal and spatial properties of the original sound at the the same time.

Both requirements will undoubtedly be contradictory in many situations. However, the compromise, i.e., optimization by the sound engineer, will be the better, the more flexible the stereophonic recording technique is and the more accurately the psychoacoustic principles are understood and taken into account from the technical and artistic points of view.

Under these aspects, possibilities for improving the capabilities of loudspeaker stereophony will be outlined in the following chapters. Optimizations of the stereophonic main microphone and mixing technique were already proposed during the preceding sound engineers' conferences in 1984 and 1986 in Germany /1/, /2/. First practical study results are now available /3/, /4/. However, certain handicaps of conventional two channel loudspeaker stereophony cannot be overcome without increasing the number of reproduction channels. The question therefore arises: why should the very high data reduction of modern source coding methods (a CD could provide at least 16 sound channels for the same duration and sound quality) and the enormous possibilities offered by modern chip technology not be used to considerably improve the capabilities of loudspeaker stereophony by the utilization of additional reproduction channels?

## 2. LOUDSPEAKER STEREOPHONY WITH TWO REPRODUCTION CHANNELS

### 2.1 What can it do?

Which stereophonic loudspeaker signals does the ear require so that a natural sound image is achieved? Which quality of stereophonic presentation of the direction, distance and spatial impression[1] is at all possible, in principle, in the case of conventional two-

---

1

The term "spatial impression" comprises two attributes of the sound image /5/, /6/: The first is "reverberance" (a temporal slurring of auditory events /6/ which is caused by late reflections and reverberation. The second is "auditory spaciousness" (a spatial spreading of auditory event /6/) which is caused by early reflections in the range of 10 to 80 ms delay.

channel loudspeaker reproduction? Fundamental statements have already been derived in earlier papers by means of the association model /7/ for spatial hearing:

1. The distance of the phantom sound source /6/ is equal to the (mean) distance from the two stereo loudspeakers. The spatial perspective can only be represented, in the simulation plane between the loudspeakers in a manner similar to the perspective presentation in the visual area /8/ (see fig. 1). The real distance from the loudspeakers corresponds to the real distance from the picture.

2. The spatial perspective in the simulation plane is achieved the better, the more accurately the interaural signal differences during natural listening are imitated by the loudspeaker signal differences /8/. Due to an inverse filtering process postulated in /7/, /8/, /9/, the auditory system recognizes the relations between the left and right loudspeaker signals independent of the binaural crosstalk and evaluates them according to the listening experience.

Thus in principle, optimum presentation of direction, distance and in the simulation plane are made possible by the stereophonic signal differences generated by dummy head /8/. A dummy head signal which produces the head-referred, three-dimensional perception of space during headphone reproduction generates, during loudspeaker reproduction[2], an equivalent loudspeaker-referred presentation of the spatial perspective in the simulation plane, which is comparable to the spatial perspective of a picture.

In order to verify this important statement, suitable experiments can be carried out. When the dummy head signals are compared in a listening test with stereophonic signals which do not provide the head-specific interaural signal differences with sufficient accuracy, a relatively high degree of sensitivity of the ear to such interaural "irregularities" is noted during playback through headphones: The quality of the perceived spatial image suffers in some way and to some degree. The result of the same listening comparison during playback through loudspeakers is both surprising and impressive: the quality of the percived spatial image (in the simulation plane) suffers in the similar way and to almost the same degree! Two examples are shown in fig. 2:

---

2

"Loudspeaker reproduction" does not include the technique of "biphonal reproduction", i.e., loudspeaker reproduction techniques which aim to simulate headphone reproduction by compensating the interaural crosstalk portions (a survey will be found in /10/). The biphonal reproduction methods can not be considered as a possiblity to improve the capability of loudspeaker stereophony because the listening area is always minimal.

a) A dummy head recording of a sound source (e.g. a speaker located on the right side of the dummy head, see fig. 2a) will produce a correspondent image on the right sight of the headphone listener, and, in the case of loudspeaker reproduction, a sharp image located close to the right loudspeaker, due to maximum magnitudes of the interaural signal differences of the dummy head (case A). When the maximum natural interaural time difference of 0.74 ms is enlarged to the unnatural value of about 1 ms by means of a delay device (case B), the stereophonic quality drops distinctly. This is true even in the case of playback through loudspeakers. The sound event in a more blurred and vague manner in the case of the unnatural interaural time difference of 1 ms.

b) When comparing a dummy head recording (A) and a coincidence microphone recording (B) of an orchestra by means of headphones, the superiority of the dummy head is obvious. The dummy head signal produces a head-referred natural spatial impression, but the coincidence microphone signal produces a poor spatial impression. It is important that a corresponding stereophonic quality difference can be observed in the case of loudspeaker reproduction: The dummy head signal generates a loudspeaker-referred presentation of the spatial perspective in the simulation plane (according to fig. 1), but the coincidence microphone signal produces a "flat distribution" of sound sources between the two loudspeakers in front of the listener without simulating spatial perspective. The coincidence microphone signal which does not provide any head-specific interaural signal difference fails not only in generating a head-referred presentation of the authentic spatial impression and depth but also in generating a loudspeaker-referred simulation of the spatial impression and depth.

Summarizing, loudspeaker stereophony according to the association model is based on introducing corresponding physical attributes of the ear signals (which correlate with phenomena of natural spatial hearing) into the stereophonic signals /8/. This is contradictory to the consequence of summing localization theories, which attempt to introduce them into the resulting ear signals of the listener. On basis of summing localization theories it is even today tried to assess stereophonic techniques (see e.g. /11, 12/). As a recent example, Lipshitz has concluded that coincidence microphone techniques are most advantageous for getting a natural spatial impression /11/:

> "I believe that spaced-microphone recording techniques are fundamentally flawed, although highly regarded in some quarters, and that coincident-microphone recordings are the correct way to go".

His arguments are based on an analysis of the resulting interaural characteristics of listener's ear signal, according to the principle of summing localization.

> "The level and time (or phase) differences at the listener's ears are <u>not</u> the same as those at the loudspeakers ... It is important that, as far as possible, the two

loudspeaker signals combine at the listener's ears to produce cues which are compatible with natural hearing."

However, natural interaural attributes of the listener's ear signals are only possible by using the dummy-head technique ("head-referred imaging"). In contrast to this, the conventional two-loudspeaker stereophony is a "loudspeaker-referred imaging" technique, and it is important that, as far as possible, the two loudspeaker signals contain natural interaural attributes rather than the resultant listener's ear signals in the playback room. This requirement is not met by pure intensity or time stereophony, and the stereophonic quality is disadvantageous with respect to depth/space imaging (intensity stereophony) or to localization (time stereophony) in comparison to dummy head, OSS, or ORTF technique as found in practical comparison tests on the performance of main microphones in different concert halls /3/.

On this basis, there are possibilities for optimizing the stereophonic presentation of direction, distance and spatial impression through two loudspeakers. The socalled "sphere" microphone /2/ and the "room-related" balancing technique /1/, /4/, /13/ represent appropriate optimization proposals for the recording end:

### 2.1.1 Sphere microphone

In /2/ a microphone system had been proposed where two boundary layer microphones are placed on the sides of a sphere, as shown in fig. 3. This socalled "sphere microphone" produces stereophonic signals which are composed of natural interaural differences, quite similar to dummy head signals, as demanded above. However, in contrast to dummy head, it has a linear frontal frequency response (see fig. 4, upper curve). The sphere microphone signal does not contain those dummy head specific spectral cues which are used for front-back orientation during headphone listening /6/, /9/ (fig. 4, lower curve), but which are not used during loudspeaker-referred presentation, and would therefore cause coloration problems /2/. Presuming an exact diffuse-field equalization /7/, the sphere microphone could thus provide sound colour neutrality both for direct sound as well as for indirect sound.

At present, prototypes of the sphere microphone are being proved in differing situations and compared with other main microphones. First results confirm that the sphere microphone in fact combines favourable imaging characteristics with respect to spatial perspective, accuracy of localization and sound colour.

### 2.1.2 Room-related balancing

Theoretical considerations /1/ and practical tests /13/ show that spot microphone signals can be added to the sphere microphone signal without disturbing the perception of spatial perspective. A disturbing effect occurs in the case of conventional panpot-balancing as demonstrated in fig. 5: The signal picked up by a spot microphone is reproduced earlier than the corresponding main microphone signal. Thus, the ear interprets the spot microphone signal as the direct sound /1/, /14/, and the favourable imaging characteristics of the sphere microphone (or any appropriate main microphone) are lost. Such recordings sound flat, without spatial depth.
It is common practice to moderate this space-disturbing effect by artificial reverberation, and/or compensating the delay of the main microphone signal (e.g. /14/). However, those techniques are not satisfying, because the stereophonic quality of the direct sound is not improved by this method. In practice a delay compensation leads to "notching" effects which are particularly disturbing when the musicians move about near the spot microphone. Experiments have shown /13/ that a pure panpot-balancing can even be preferred in comparison to delay compensated panpot-balancing, depending on the recording situation and the desired balancing gain.[3]

In order to preserve the perception of spatial perspective due to main microphone signal, and at the same time achieve a high balancing gain, the spot-microphone signal should be delayed much more than necessary for the compensation, so as to fall within the region of the early reflections. It has been proposed to achieve the desired increase in volume by adding the sound energy from artificially generated reflections /1/. This socalled "room-related" balancing technique has been tested and optimized recently /13/, with the aid of an appropriate audio processing unit /15/. It was found that a loss in depth can be avoided satisfactorily by generating just two artificial reflections from the spot microphone signal (according to fig. 6), and that the greater the required balancing gain, the more room-related balancing technique is favourable.

Principally, the room-related balancing algorithm would be implemented into digital mixing desks, so that it could be used alternatively to conventional panpot-balancing. However, first of all further optimization is useful to minimize the signal processing effort and to introduce improvements, such as distance equalization (taking account of change in spectrum by absorbtion at the room boundaries), additional artificial reverberation (generated from the spot microphone signal in accordance with the artificial reflections), etc.

---

3
Balancing gain is the level of the balancing signal with reference to the balancing signal's threshold level, which is measured at the threshold of perception of the balancing signal in the main-microphone signal.

### 2.1.3 Summary

It can be concluded that current two-channel stereophony recording techniques can be improved with regard to the naturalness of the stereophonic sound image defined above. A consistent consideration of new knowledge and understanding on psychoacoustic principles, particularly on spatial hearing, leads first of all to the principal result that the two-channel stereo presentation of direction, distance and space is only possible as a presentation of spatial perspective in the simulation plane between the loudspeakers. The distance from the loudspeakers is the real distance from the simulation plane.

Considering that, only an optimization of the techniques for simulating spatial perspective could be hoped to get. This is successful by using natural interaural signal differences instead of pure intensity or time differences, or a mixture of them. The room-related recording technique is consequently based on this knowledge. This technique can be employed to the well-tried main/spot microphone methods, which means that the optimum main microphone is the sphere microphone (or a corresponding microphone generating natural interaural signal differences) and that any spot microphone signal added to the main microphone signal represents additional, artificial reflections from the recording room. As the generation of natural interaural signal differences and simulation of artificial reflections and reverberation from a spot microphone signal is possible with the aid of modern computer technology, room-related recording technique principally can also be employed to poly-microphony. This would result in a stereophonic presentation of any artificially created space in the simulation plane.

## 2.2 What is not achievable with two loudspeakers?

The theory, however, does not only point out possibilities for improving the natural features of a stereophonic sound image; the limits of loudspeaker stereophony become clear at the same time: The direct sound generates the stereophonic sound image in the narrow imaging range ± 30° and in the distance of the loudspeakers. The listener's acoustic environment is determined solely by the listening room and suited to a very limited extent only for the presentation of the original room /16/. Any limitation to two reproduction channels therefore narrows down the performance of loudspeaker stereophony in several respects (see also /10/). The principal weakness of two-channel reproduction are as follows:

1. As shown in section 2.1, it is not possible to generate the natural perception of spatial impression and spatial depth: rather the spatial perspective in the simulation plane is presented /1, 8/.

2. Reflections and reverberation of the listening room rarely support the stereophonic presentation of the original room, but they are the only parameters to generate the listener's acoustic environment /16, 10/. A contradictory spatial impression which is generated by the listening room could affect imaging of the original room adversely.

3. Localizable auditory events with any directions outside the imaging range $\Theta = \pm 30°$ (e.g., genuine surround imaging) cannot be presented.

4. The listening area is extremely narrow, especially in the case of loudspeakers with a high degree of directivity /17/ and in the case of playback setups which ensure a high localization focus /18/.

By extending the playback setup to 2+n channels, the performance of loudspeaker stereophony can basically be improved with respect to all four of the above points. An attempt has been made in the chapter below to estimate the benefit of various additional reproduction channels according to practical and theoretical aspects.

## 3. LOUDSPEAKER STEREOPHONY WITH 2+N REPRODUCTION CHANNELS

Various sound transmission systems are known - and some of them had been tested or used in practice - which attempt to avoid one or several of the four deficiencies by employing more than two reproduction channels. Some of these systems are:

| | | | | | |
|---|---|---|---|---|---|
| Ambiophony | (1960) | /19/ | ABC trapezoid system | (1982) | /23/ |
| Quadrophony | (1969) | /20/ | Triphonic | (1982) | /24/ |
| Tetraphony | (1971) | /21/ | Ambisonic | (1983) | /25/ |
| Eidophony | (1977) | /22/ | Dolby Surround | (1985) | /26/ |

The list is incomplete: especially in the movie industry, film sound systems operating with 4 - 7 sound channels (CINERAMA, CINEMASCOPE, TODD-AO, DOLBY-STEREO) /27/ have been in existence for some 35 years. Still, the primary goal in movie sound systems is definitely not the goal defined in the beginning, i.e., to achieve optimum naturalness of the sound image. These systems are primarily designed to be able to produce the most convincing direction and space effects possible for a large audience in the most economic manner possible. In the process, sound carrier characteristics, technical sound diffusion and operating aspects narrow down the leeway considerably. Film sound systems must meet different requirements than home audio systems.

But even the multichannel sound systems provided for home entertainment (essentially the systems /19 to 26/ listed above) are only capable to a very limited extent to overcome some of the weakness of two-channel loudspeaker stereophony. In many instances, limitations of the stereophonic quality which are specific to the system are accepted because of the advantages gained; this applies, for instance, to localization focus, perceived direction stability, or perceived direction resolution. One cause for this can be found in the endeavour to ensure compatibility with the existing two transmission channels (radio, television, storage). Most of the time, three, four or more loudspeaker signals are combined to a conventional stereo signal by means of matrixing and recovered after the transmission (3-2-3, 4-2-4). The resultant inadequate channel separation either leads to quality losses or limits the application range (e.g., to presentation of effects). Even psychoacoustically favorable matrixing and complex decoders do not assure satisfactory channel separation and quality; cf. for instance /29/.

The marginal condition "two-channel compatible transmission", in particular, has certainly contributed to the fact that multichannel sound systems have not been able so far to take the lead. One exception is the Dolby-Surround system because it is used to decode at home the Dolby-Stereo encoded sound of videos or movie films broadcast over TV stations and thus to reproduce the direction and space effects designed for the cinema as well as the "stable middle".[4] Due to 4-2-4-matrixing, the desired naturalness of the sound image can, however, not be achieved by this system. It is not designed for pure sound presentation without picture, and therefore unsuited as universal home audio system.

## 3.1 A new audio home standard

The influence of movie sound technology on the design of the audio systems of future TV standards is unmistakable. The Dolby Surround system has already been proposed for future narrowband and wideband HDTV systems. The same reproduction arrangement is used in a Japanese system proposal, involving four discrete channels /28, 30/. Another CCIR-proposal, the "substituting source concept" /16, 30/ which provides five channels just for the presentation of picture related sound (imaging range $\Theta = \pm 16°$). As in movie sound technology, special emphasis with respect to future TV standards is placed on the ability

4

The Dolby Surround reproduction arrangement consists of the conventional two stereo loudspeakers, a center loudspeaker (front center), and one or several real or lateral loudspeakers /26/.

to represent the picture related sound in a stable direction within the picture area. This is not possible in the narrow listening area of two-channel stereophony. Moreover efforts are made at preserving the possibilities offered by movie sound with respect to the presentation of sound outside the picture area (off sources, acoustic atmosphere). The development work going on worldwide at this time /30/ permits the expectation that HDTV home reproduction with a large picture and at least four-channel sound will be reality soon.

Thus, today's television set will be replaced in a few years by a "home movie" which will include a relatively expensive sound reproduction system in addition to the picture display. The loudspeaker system will be able to overcome or reduce some weaknesses of conventional stereo systems. Since the conventional stereo loudspeakers and the loudspeakers for the TV sound are not likely to stand side by side in the living room, we are faced with the problem of defining a new loudspeaker reproduction standard which, on the one hand, is able to satisfy the requirements for HDTV /30/ and, on the other hand, can provide optimum stereophonic quality.
Careful studies will have to be undertaken to clarify if and in which way additional loudspeakers are able to enlarge the listening zone and improve at the same time the naturalness of the stereophonic sound image, as defined in the beginning of the paper. In other words, it is necessary to ascertain which HDTV loudspeaker setup is useful not only with respect to picture accompanying sound but also with respect to high-quality sound without picture. For instance, it should be possible to reproduce classical music with optimum naturalness, with and without accompanying picture. (When listening to music at home, it could even be a strong desire to close the eyes to concentrate just on the auditory event instead of being irritated by a more or less appropriate picture).

The basic requirements upon a suitable HDTV sound system listed in CCIR Report 1072 /30/ are:

- downward compatibility to more simple reproduction for less demanding situations;
- system compatibility for programme material originally prepared for single or multi-channel sound presentation;
- good localization of sound sources and correspondence with the visual image as far as possible for viewers and listeners in optimal and extreme seating positions;
- imaging capability of auditory ambience of the scene (e.g. reverberation, environment, and off-picture sound sources);
- reproduced sound ideally corresponding to the subjective quality given by a 16-bit-linear digital encoding with a sampling frequency of 48 kHz;
- maximum economy in all respects, including transmission efficiency.

Several aspects on satisfying these requirements will be discussed below in a greater detail, however, with particular respect to optimum stereophonic presentation to listeners within a sufficiently large listening area. The goal is a universal multichannel stereo home standard which provides

- compatibility to sound production technique commonly used in practice;
- flexibility to allow practical, commercial and aesthetical considerations for producer and consumer;
- headroom to enhance stereophonic quality and size of listening area with the aid of modern audio signal processing.

Having in mind these guidelines, it is the basic intention to make use of the psychoacoustic principles in order to get maximum benefit from additional front channels and surround[5] channels.

## 3.2 Size of listening area

Lateral displacement of the listener from the center results in distortions in the presentation of direction in the imaging plane ("imaging distortions"[6]). If the listener shifts laterally just 10 cm from the center of the conventional stereo loudspeaker setup, a time difference of 300 us results. This leads to a direction displacement of the center phantom sound source by about 12°, corresponding to an imaging distortion of V = 20 %.

Fig. 7 is an illustration of calculated curves (basis: phantom sound source displacement 4°/100 $\mu$s) for the limit of listening area at various permissible imaging distortions V as a function of the listening distance. It can be seen, for instance in the stereo standard setup, that the listening area width is only ± 5 cm for a permissible imaging distortion of V = 10 %, regardless of whether the imaging width B is 1 m, 2 m or 3 m. However, the listening area width varies with the relativ distance of the listener; it will be smaller

---

5 The term "surround channels" originates from the movie industry and is used there mainly to represent the acoustic environment outside the picture area. Surround loudspeakers are to be defined in this context as well as loudspeakers outside the front stereophonic imaging area. This does not imply that it is the aim to provide a full surround imaging area (surround moving events).

6 The imaging distortion V is defined as the direction deviation of an auditory event from the setpoint, referred to the overall imaging width B in %.

if the listeners position is closer to the loudspeakers than stereo standard and will be enlarged up to 10 cm or more if the listeners position is farer away from that point.

Theoretically, a smaller loudspeaker base does not result in a correspondingly narrow listening area (cf. Fig. 8). This seems to contradict practical experience (e.g., with nearfield loudspeakers). However, the cause is to be found in the effect of the reflected sound in the reproduction room on the propagation of the phantom sound source. The localization focus is particularly high in the near field of the loudspeakers /31/, and so is the ear's sensitivity to imaging distortions /17, 18/.

It would certainly be wrong to enlarge the listening area by means of particularly non-directional loudspeakers at the expense of the localization focus. In most cases, the reflected sound would thereby affect adversely the imaging of the original room.

### 3.2.1 Stereophonic subareas

An obvious approach to enlarge the listening area is the split-up of the entire imaging area into "stereophonic subareas" (if the stereophonic imaging area is formed by 2+n loudspeakers, (2+n)-1 stereophonic subareas are formed). When the correlation between the listening area width and imaging distortions (Fig. 7) is applied to the subareas, useful listening area widths are obtained for as few as two or three subareas. As an example, the resultant listening areas for a permissible imaging distortion V = 10 % are shown in Fig. 9 and are comparable to the listening areas in Fig. 8. In /31/ localization curves (image direction as a function of panpot level difference) of all subareas at several listening points have been measured. They are in congruence with the estimation made in Fig. 9.

Fig. 10 indicates that the listening area width increases out of proportion to the number of loudspeakers. Since the listening area width theoretically does not depend on the imaging width B (cf. Fig. 8), the number of loudspeakers should increase as B increases if the ratio of imaging width to listening area width is to be about constant.

In practice, the listening areas are somewhat wider because the room portion in the stereophonic signal increases the largeness of the phantom sound source (presentation of auditory spaciousness /5, 6/, cf. chapter 2.1. The associated decrease in the localization focus corresponds to the natural sound image and is intentional. On the other hand, the "imaging lack of focus" /32/ of the reproduction system, which is caused by the indirect sound in the playback room /18/, is not desired.

Moreover, the permissible imaging distortion which is assumed to be V = 10 % may be quite far below the tolerance threshold, if stereophonic imaging takes place without picture. Experience shows that the listener does not normally keep his head in the center with an accuracy of ± 5 cm (cf. Figs. 7 and 8): - Another aspect is that the imaging distortion V = 10 % is a limit value which refers to the most critical case of the phantom sound source in the center of a stereophonic subarea. All other auditory event directions, especially those near the loudspeakers, are more stable (cf. /31/).

### 3.2.2 Considering HDTV requirement

In the case of HDTV the viewing angle at the reference point will be 33° (distance 3 x height of picture, aspect ratio 9 : 16). The minimum distance of the viewer to the screen is assumed to be 2 x height (2H), the maximum width should be equal to the screen width. Within this area, proper image localization should be possible, providing sufficient sharpness and directional fidelity. In particular, a small sized auditory image monotonously moving across the stage without jumping has been assumed to be the worst-case presentation.

Within certain tolerance limits, congruity of auditory and visual event should be achievable independent of viewer's/listener's displacement (from the axis) at the reference distance. If an auditory event is accompanied by a corresponding visual event, the "perceptual fusion" phenomenon is effective. This phenomenon will be most effective for a strong relationship between both events (i.e. a speaking person).

In this context, experiments on the perceptibility of direction differences between the viewing and listening event had been carried out /33/. The perceived directions of short light impulses were compared with the perceived directions of short gaussian shaped noise impulses of the same duration. The results lead to the guideline that the influence of listener's displacement from the axis on the perceived direction of auditory sensation (without accompanied picture) should be smaller than ± 4°, in order to ensure complete congruity of visual and auditory sensation. This means, the permissable imaging distortion V as defined above without accompanied picture should be less than 7 %, in order to ensure complete congruity.

On the other hand, it seems more reasonable to consider the listener's/viewer's tolerance threshold instead of the perceptibility threshold to imaging distortions. In a recent study on this matter /34/ the acceptable extend of angular displacement between visual and auditory images had been investigated by means of normal program material (i.e. by including the distruction by the program content, the associative link between visual and auditory images, the source largeness). The results show that the acceptable displacement is about 11° (V = 18 %).

Considering also worst case situations (e.g. moving images), it can be concluded from all results that an imaging distortion of V = 10 % is the appropriate basis for estimating the width of the listening area even in the case of accompanying picture according to Fig. 11. At the reference distance 3H the width of the listening area is sufficient. It is therefore proposed to use the 4-loudspeaker configuration shown in Fig. 11 in the front imaging area as one part of a future audio home standard (maximum performance level).

### 3.2.3 Recording technique

The loudspeaker set-ups shown in Fig. 9 resp. Fig. 11 principally allow to make use of two differing recording techniques. First, according to the name "subareas", it seems obvious to apply the psychoacoustic principles of loudspeaker stereophony in analogy. This means that, idealy, for each subarea an individual stereophonic subsignal has to be created, and that only neighbouring loudspeaker signals should be correlated.

A pure panpot-mix in this way is well-known and common practice in the film industry since about four decades /35/, however, problems arise when for instance the main/spot-microphone technique should be used in order to achieve an optimum presentation of the original room in the simulation plane as described in chapter 2. Although room-related recording/mixing principals could be transferred consistently to subarea-stereophony with the aid of appropriate audio signal processors (e.g. in the case of poly-microphony), this technique has not been developed yet. Thus, there is a headroom to accommodate future developments in the studio, bearing in mind an optimum realism of spatial presentation within a sufficiently large listening area.

Another well-tried mixing technique for film sound applications (e.g. /27/, for TV see also /36/) is employed mainly in the 3-loudspeaker configuration (Fig. 9). Important (picture related) sources such as dialogue are assigned to the center channel. Additional (not picture related) images, e.g. ambience, effect, music are panned between the two outer basic stereo channels without making use of the center channel. Contradictory to the first approach the neighbouring loudspeaker signals are not correlated. With the exeption of rapidly moving effects correlation of the two outer loudspeaker signals is also avoided for appropriate presentation of ambience and music. Since phantom sound sources are rarely applied, this technique is unproblematic with respect to directional stability.

The 4-loudspeaker configuration (Fig. 11) enables further development of this technique. It seems interesting to replace the "mono-center" signal by a "stereo-center" signal, in order to provide directional imaging of the picture related sources. Thus, two independent stereophonic imaging areas are available: The inner area, which provides sufficient congruity of auditory and visual event as described in /31/ and chapter 3.2.2, and the basic stereo area between the outer loudspeakers, which is preferably suitable for presentation of uncorrelated sound. The resulting characteristics of the sound image (discrete sources only in the center region, non-discrete effects in the outer region) are quite similar to those achieved in common praxis in TV two-channel stereo sound, however, 4-channel presentation has the advantage of directional stability.

This second method of utilizing the 4-loudspeaker configuration could be an alternative to the stereophonic subarea concept. The choice will depend on the application, on the program, and on the equipment available in the studio. Futhermore, practical experience and individual conceptions might result in different methods of combining advantages of both techniques.

## 3.3 Space imaging

The listener's acoustic environment can principally be designed by employing additional surround loudspeakers. Thus, the impression of space does not have to be presented exclusively as spatial perspective in the simulation plane (as chapter 2). A largely natural, real spatial impression can be produced by reproducing reflections and reverberation through loudspeakers outside the stereophonic imaging area, particularily through loudspeakers at the side and behind the listening area. Numerous realizations of this well-known method (e.g. /10, 19, 27, 38/) has been developed in close connection to fundamental investigations (e.g. /6, 37, 39, 40, 41, 42/) on room acoustics in concert halls, with the common objective of creating lateral sound in the listening room.

Experiences were not only made on film sound systems but also on several home sound systems (e.g., ambiophony /19/, quadrophony /43/, tetraphony /21/, trapezoidal arrangement /10, 44/) and have resulted in new experiments again and again. Application of quadrophony to the transmission of serious music was proposed, for instance, in /43/ and /45/ Four-channel test transmissions by DRS (Switzerland) - the "room sound" was transmitted by a second transmitter - are reported in /46/. Recently a digital audio signal processor for home use has been developed /38/ which generates lateral reflections and reverberation from a usual 2-channel stereo signal for creating the spatial impression /5, 6/ (4 surround speakers are applied), under consideration of results on room acoustic research.

### 3.3.1 Psychoacoustical basis

In principle, the psychoacoustic laws for perceiving spaciousness in the concert hall are applied to the generation of the "surround sound" . In addition to reverberation, the socalled "early lateral sound" which reaches the listener's ears from lateral directions about 10 to 80 ms after the direct sound[7] is particularly important. It induces an interaural decorrelation of the two ear input signals which is specific to the particular room, and thus auditive spaciousness: Auditory events fill larger amounts of spaces than in a free sound field under comparable conditions.

The dependence of spaciousness on delay time, level, angle of incidence, and spectrum of the early lateral reflections has been investigated. Also, the overall level of the direct sound and of the reflections has been found to be of central importance /39, 40/. Various investigators attempted to define an index of spaciousness as a function of objective parameters of the sound field (cf. /6/). The index depends on either the interaural degree of correlation or the ratio of lateral sound energy to frontal or total energy ariving at the listener. In /38/ it is suggested that it is the pattern of hall reflections themself which cause listener preference, and not the resulting low interaural correlation. (This coincides with the knowledge derived from the association model /7/).

Moreover, reproduction of the lateral reflections could result in perception of spatial depth. It was found already in the room related balancing technique /4, 13/ that the presentation of distances in the simulation plan /8/ succeeds if even one left and onre right lateral reflection can be imitated. The stereophonic quality of simulated spatial depth (cf. chapter 2.1) changes to a quality of real perception of space depth if the lateral reflections are emitted by surround loudspeakers and actually arrive from lateral directions.

As regards the room-related balancing technique /1, 4, 13/, it should thus be possible to add the artificial reflections and artificial reverberation not only to the front stereo signal but to the surround signals as well, or not add them at all. It could be expected that such lateral reproduction of balancing signals represents a convenient form for the room-related balancing technique in the case of multichannel stereophony. However, several questions relating to the recording/mixing technique need to be answered if surround speakers are applied for generating the natural perception of spatial impression.

---

7 The optimum delay of the first lateral reflexions is in the range 20 ... 30 ms.

### 3.3.2 Arrangement of surround loudspeakers

Experiments have been carried out in different laboratories to determine the optimum location and number of additional surround loudspeakers. One study on the design of an HDTV sound system was made in Japan /28/; it covered, a.o., the effect of a pair of surround loudspeaker with differing locations on the evaluation of the stereophonic sound image during reproduction of music (evaluation aspect: "sensation of reality"). Fig. 12 shows the results, relativ to the loudspeaker configuration A (without surround speakers). It demonstrates that lateral (C) or lateral/back (D) location of one pair of surround speakers is preferable.

Experiments on concert-hall simulation are described in /38/. It was found, a.o., that rear loudspeakers are next in importance after side loudspeakers, and that the optimum placement of the surround speakers depends widely on the situation (i.e. on the characteristics of the listening room, on the program etc.). As one result, it has been recommended to apply two pairs of surround loudspeakers, as shown in Fig. 13, in order to maximize the effect of spatial impression.

Experiments of our own were made to verify these results by means of special multi-channel music recordings, with the goal of including the effect of various surround loudspeaker arrangements on the sound image, namely in direct comparison with the effect of various main microphones on the sound image /3/ (evaluation aspect: "sensation of concert hall atmosphere"). Microphone arrangement, loudspeaker arrangement and results are shown in Fig. 14.

The five sound images A .... E which were compared with one another were generated as follows:

- A: Stereo loudspeakers only ± 30°, signal: main microphone XY (120° angle of aperture)
- B: Stereo loudspeakers only ± 30°, signal: main microphone AB (1.1 m)
- C: As B, plus surround loudspeakers ± 120° with omnidirectional microphone 2 (10 m behind main microphone)
- D: As B, plus surround loudspeakers ± 90° with omnidirectional microphone 1, plus surround loudspeakers ± 120° with omnidirectional microphone 2.
- E: As B, plus surround loudspeaker ± 90° with omnidirectional microphone 1, plus surround loudspeakers ± 120° with omnidirectional microphone 2.

The result indicates preference for the sound images C, D, E, i.e., the surround loudspeakers achieve a distinctly better imaging of the acoustic atmosphere in a concert hall, when compared with conventional technique (sound images A, B). Preference is particularly pronounced in the case of the surround constellations D, C. The concert hall atmosphere was present in the test signal "music" (horn solo and orchestra) by the spatial impression, and in the test signal "effect" (auditorium during intermission) by the acoustic environment. When A is compared with B, the superiority of the main microphone AB over XY is demonstrated again (cf. /3/), however, in relation to additional surround speakers, the difference is judged as small.

An important prerequisite for achieving natural spatial impression is that care is taken not to localize the surround loudspeakers. This can be primarily accomplished at the recording by delaying the surround signals appropriately with respect to the main microphone signals. In the test signals were used, this was done by selecting the omnidirectional microphone positions as shown in Fig. 14. Due to the delays, the surround signals act like lateral reflections from the hall, and the "law of the first wave front" /6, 37/ is effective. At the same time, all loudspeakers involved which are distributed in the listening room generate a naturally diffuse sound field in the listening area, because the microphone signals appear non-correlated due to the great distances between the microphones ("room-microphones"). As a result, a physical imitation of first reflections and reverberation and thus natural spatial impression is obtained. The listener not only perceives a spatial perspective in the front imaging plane, but feels included in the acoustic event.

Looking at these principles it seems obvious that the achievable quality of the spatial impression depends on the number of applied surround loudspeakers. Fortunately, the increase of quality is out of proportion to the number of speakers (see Fig. 14). However, in congruence to the results in /38/, it is proposed to provide two pairs of surround loudspeakers according to Fig. 14E for a multichannel sound system with maximum performance level. Four surround speakers instead of two or even only one are appropriate for several reasons:

1. Localization of the surround speakers can be avoided more effectively, particularily for listeners at the border of the listening area. The level of a single reflection is limited due to the law of the first wave front, on the other hand enough early lateral energy is necessary for creating the natural degree of spaciousness.

2. The generation of an artificial diffuse sound field within the complete listening area for creating reverberance is more successful.

3. The location of the surround speakers is less critical with respect to both distance and direction. The quality of the spatial impression will be more independent of listener's location within the listening area.

4. The spatial impression which is generated by the room-acoustical characteristics of the listening room, and which could affect imaging of the original room adversely, can be masked the better the more loudspeakers are distributed in the listening room, radiating reflections and reverberation from the original room.

5. The presentation of directional effects (HDTV: "off-sources"), when surround loudspeaker generate the first wave front, can be done with higher direction resolution. As lateral phantom sound sources are instable /6, 47/, mainly the surround loudspeaker locations can be used for directional imaging outside the front imaging area.

It should be pointed out that two pairs of surround loudspeakers are proposed for a sound system providing maximum performance level. One pair of surround speakers would certainly be suitable for the system of lower level, however, one single surround loudspeaker should not be applied. According to several proposals (e.g. /19, 26/), a mono-surround signal should radiated by means of at least two loudspeakers, eventually after having decorrelated the signal to a certain degree with the aid of appropriate signal processors (see also /38/).

The proposed arrangement of surround loudspeakers thus does not enlarge the stereophonic imaging range $\theta = \pm 30°$; rather an acoustic environment is added to the frontal stereophonic presentation of directions and distances (simulation plane). This environment can be designed by the audio engineer, for instance along the lines of optimum reproduction of the three-dimensional impression or in order to generate a new artificial room, or to create surround sound effects. As the acoustic environment is supplemented to improve the spatial impression, there is great similarity to stereo ambiophony /19/ and also to the ABC trapezoidal system /23, 44/, but not to quadrophony /20/.

Although full stereophonic surround imaging would be possible with at least six loudspeakers in a circular arrangement /47/, this could only be accomplished at the expense of a very narrow listening zone. Full surround imaging is neither practical nor meaningful.

## 4. CONCLUSION

The performance of two-channel loadspeaker stereophony should be enhanced by applying additional loudspeakers. The speaker arrangement in Fig. 15 would enlarge the listening area and improve at the same time the naturalness of the stereophonic sound image. As a maximum level of performance, four front loudspeakers are used for directional imaging, and four surround loudspeakers for imaging of space and effects.

It is assumed that the maximum performance level can be reduced in a compatible way by lowering the number of front channels and/or surround channels as convenient for the consumer, within a defined hierarchy of performance levels (numbers of channels). Considering the practical limitations due to furniture and space in the appartment it is important that the location of surround loudspeakers is largely uncritical and can easily be adapted to the individual situation (in contrast to Quadrophony). Also the dimensions of the front and surround channels can be small, particularily if a common subwoofer /27, 38/ is applied.

As regards the maximum number of transmission channels, current work on low bit-rat coding of high-quality audio signals (e.g. /48/) has to be considered. As the bit-rate of a CD audio signal can be reduced to less than 100 kbit/s without any subjective quality loss, ten audio channels would consume less than 1 % of the bit-rate provided for the transmission of the wideband HDTV picture (125 Mbit/s). The costs for a decoder chip will be low, especially when compared with the picture decoder.

Furthermore, taking into account a multilingual capability desired in /30/, an HDTV sound system should provide an appropriate number of transmission channels for this purpose. A flexible use of transmission channels is assumed to be a goal, providing either a very high performance multichannel sound presentation or, alternatively, a reduced number of reproducing channels plus additional channels for alternative languages.

At the latest with the start of HDTV home reception, there will be a new sound reproduction standard which will be much more powerful than the one in use so far. It should be usable optimally for high-quality reproduction of music, because it probably will not take much longer before new storage media containing more than two channels will be available. In light of this prospect, careful consideration should be given once more in the planning of a future digital radio system to the question if a long-term commitment to two transmission channels is really appropriate or necessary. High-quality stereophony will mean multi-channel stereophony in the future.

**Bibliography**

/1/ Theile, G.: Hauptmikrofone und Stützmikrofone - neue Gesichtspunkte für ein bewährtes Aufnahmeverfahren (Main microphone and spot microphones - new aspects for a proven recording technique). 13th Audio engineers' conference 1984, conference minutes pp. 170-184

/2/ Theile, G.: Das Kugelflächenmikrofon (The sphere microphone). 14th Audio engineers' conference 1986, conference minutes pp. 277-293

/3/ Wöhr, M.; Nellesen, B.: Untersuchungen zur Wahl des Hauptmikrofonverfahrens (Studies on the selection of the main microphone method). 14th Audio engineers' conference 1986, conference minutes pp. 106-120

/4/ Wöhr, M.; Goeres, J.; Pösselt, C.; Theile, G.: Raumbezogene Stütztechnik - Möglichkeit zur Optimierung der Aufnahmequalität (Room-related balancing technique - Possibility for optimizing the recording quality). 15th Audio engineers' conference 1988, conference minutes pp. 302-315

/5/ Kuhl, W.: Räumlichkeit als Komponente des Höreindrucks (Spaciousness as component of the auditory impression). Acoustica (40), 1978, pp. 167-181

/6/ Blauert, J.: Spatial Hearing - The Psychophysics of Human Sound Localization. The MIT-Press, Cambridge, Mass.; 1983, ISBN 0 - 262 - 02190 - 0

/7/ Theile, G.: Zur Theorie der optimalen Wiedergabe von stereofonen Signalen über Lautsprecher und Kopfhörer (On the theory of optimum reproduction of stereophonic signals through loudspeakers and earphones). Rundfunktechn. Mitt. 25, 1981, pp. 155-170

/8/ Theile, G.: On the stereophonic imaging of natural spatial perspective via loudspeakers: theory. In: Perception of Reproduced Sound, 1987. ISBN 87 - 982562-1-1, pp. 135-146

/9/ Theile, G.: On the standardization of the frequency response of high quality studio headphones. J. Augio Eng. Soc. 34, 1986, p. 956-969.

/10/ Steinke, G.: Stand und Entwicklungstendenzen der Stereofonie (Current situation and development trends in stereophony), parts 1 and 2. Techn. Mitt. des RFZ 28, 1984, pp. 1-10 and 25-32

/11/ Lipshitz, S.P.: Stereo Microphone Techniques ... Are the Purists wrong? J. Audio Eng. Soc. 34, 1986, p. 716-744.

/12/ Bennett, J.C.; Barker, K.; Edeko, F.O.: A New Approach to the Assessment of Stereophonic Sound System Performance. J. Audio Eng. Soc. 33, 1985, p. 314-321.

/13/ Wöhr, M.; Theile, G.; Goeres, H.J.; Persterer, A.: Room-Related Balancing Technique - A Method for Optimizing Recording Quality. 887h AES Convention, March 1990.

/14/ Jecklin, J.: Musikaufnahmen: Grundlagen, Technik, Praxis. (Music Recordings: Foundations, Technique, Practice). München: Franzis-Verlag 1980, ISBN 3-7723-6701-1, pp. 55

/15/ Richter, F. and Persterer, A.: Design and Applications of a Creative Audio Processor. 86th AES Convention, Hamburg 1989, Preprint 2782

/16/ Plenge, G.: Überlegungen zur Stabilität und Leistungsfähigkeit verschiedener stereofoner Übertragungsverfahren (Reflections on the stability and performance of various stereophonic transmission methods). 13th Audio engineers' conference 1984, conference minutes pp. 158-169

/17/ Kuhl, W.: Über eine Lautsprecheranordnung zur Wiedergabe stereofoner Schallaufnahmen (A loudspeaker arrangement for reproducing stereophonic sound recordings). Rundfunktechn. Mitt., 1959, pp. 170 ff.

/18/ Gaäl, D.: Calculation of the Stereophonical Localization Area. 53rd AES Convention 1976, Preprint B - 4

/19/ Keibs, L.: Möglichkeiten der Stereo-Ambifonen Schallübertragung auf 2 Kanälen (Possiblities for transmitting sound by stereo ambiphony over two channels). Acustica, 1962, pp 118-124

/20/ AES: An anthology of articles on Quadrophonic Sound of the J. Audio Eng. Soc. Vol. 17 - Vol. 23 (1969 - 1975), October 1975

/21/ Gerzon, M.A.: Experimental Tetrahedral Recording (3 parts). Studio Sound: August, September, October 1971)

/22/ Scherer, P.: Ein neues Verfahren der raumbezogenen Stereophonie mit verbesserter Übertragung der Rauminformation (A new method of room-related stereophony with improved transmission of room information). Rundfunktechn. Mitteilungen, 1977, 196 ... 204

/23/ Kowalgin, J.A.: Stereofonie: Entwicklungsprobleme und -richtungen (Stereophony: Development problems and trends). Technika Kino i Televidenja 26. 1982, pp 3 ... 11

/24/ Torick, E.: Triphonic Sound System for Television Broadcasting. 124th Techn. Conf. of the SMPTE, New York, Nov. 1982

/25/ Gerzon, M.A.: Ambisonics in multichannel broadcasting and video. J. Audio Eng. Soc. (33), 1985, pp. 859-871

/26/ Julstrom, S.: A High-Performance Surround Sound Process for Home Video. J. Audio Eng. Soc. (35).

/27/ Blake, L.: Film Sound Today. Library of Congress Cataloging in Publication Data, Reveille Press Hollywood 1984, ISBN 0-917937-00-7

/28/ Ohgushi, K.; Komiyama, S.; Kurozumi, K.; Morita, A.; Ujihara, J.: Subjective evaluation of multichannel stereophony for HDTV. IEEE Transactions on Broadcasting (33), 1987, pp. 197-202

/29/ Theile, G.: Quadraphonie mit dem SQ-System? (Quadrophony with the SQ system) Funktechnik (29), 1974, pp. 315-320 and 359-362

/30/ CCIR Report 1072 (1): Suitable sound systems to accompany high-definition and enhanced television systems. Draft Revision 1989, Doc. 10/369

/31/ Rebscher, R.; Theile, G.: Enlarging the Listening Area by Increasing the Number of loudspeakers. 88th AES convention, March 1990

/32/ Ripka, v.A.; Theile, G.: Die Beurteilung verschiedener stereofoner Wiedergabeeinrichtungen bezüglich der Abbildungsschärfe (Evaluation of various stereophonic reproduction systems with respect to the imaging focus). Fortschritte der Akustik, DAGA '87, Bad Honnef: DPG-GmbH 1987, pp. 585-588

/33/ Thomas, M.; Plenge, G.: Probleme der bildbezogenen Tondarstellung bei HDTV-Großbildwiedergabe (Problems of picture-related sound presentation in HDTV large-picture reproduction), 14th Audio engineer's conference 1986, conference minutes pp. 338-350

/34/ Komiyuma, S.: Subjective Evaluation of Angular Displacement between Picture and Sound Directions for HDTV sound systems. J. Audio Eng. Soc. 37, 1989, pp. 210--214

/35/ Snow, W.: Basic Principles of Stereophonic Sound. Journ. of the SMPTE 61, 1953, pp. 567-589

/36/ Torick, E.: Triphonic Sound System for Television Broadcasting. EBU Doc. GT V3 1073, 1983.

/37/ Cremer, L.; Müller, H.A.: Principles and applications of room-acoustics, vol. 1, Applied Science Publishers, Essex, England, 1982.

/38/ Griesinger, D.: Theory and Design of a Digital Audio Signal Processor for Home Use. J. Audio Eng. Soc. 37, 1989, pp. 40-50

/39/ Barron, M.; Marshall, H.A.: Spatial Impression Due to Early Lateral Reflections in Concert Halls. J. Sound Vibr. 77, 1981, p. 211-232.

/40/ Plenge, G.; Lehmann, R.; Wettschurek, R.; Wilkens, H.: New methods in architectural investigations to evaluate the acoustic qualities of concert halls. J. Acoust. Soc. Amer. 57, 1975, p. 1291-1299.

/41/ Blauert, J.: Some fundamentals of auditory spaciousness. In: Perception of Reproduced Sound, 1987, ISBN 87 - 982562 - 1 - 1, pp. 33-40

/42/ Lehmann, P.; Wilkens, H.: Zusammenhang subjektiver Beeurteilungen von Konzertsälen mit raumakustischen Kriterien. (Relationship of subjective judgements of concert halls to spatial acoustic criteria) Acustica 45, 1980, pp. 256-268

/43/ Rothe, E.: Anwendung der Quadrafonie bei der Übertragung ernster Musik (Application of quadrophony during transmission of classical music). Fernseh- und Kinotechnik (30), 1976 pp. 89-91

/44/ Condamines, R.: Microphone and loudspeaker arrangements for tetraphony. EBU Review-Technical, No. 143, 1974, pp. 12-18

/45/ Jecklin, J.: Quadrofonie (Quadrophony). Electronic Sound No. 6, 1984, pp. 34-36

/46/ Breh, K.: Es steht im Raum (It's in space). Stereoply 11, 1985, pp. 42-45

/47/ Theile, G.; Plenge, G.: Localization of lateral phantom - sources. J. Audio Eng. Soc. (25), 1977, pp. 196-200

/48/ Theile, G.; Stoll, G.; Link, M.: Low bit-rate coding of high-quality audio signals. An introduction to the MASCAM system. EBU-Review-Technical No. 230, 1988, pp. 158-181

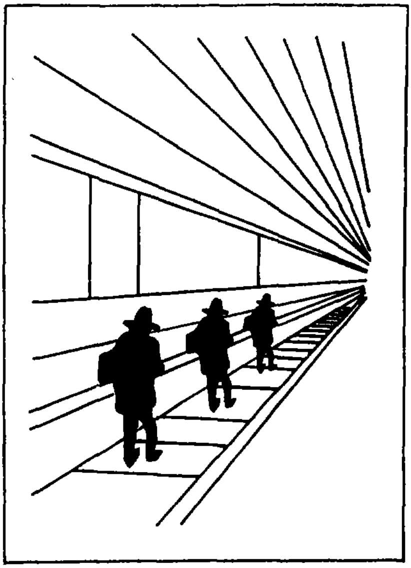

Fig. 1: The distance of this picture can be compared with the distance of stereo loudspeakers. The visual perspective, which is simulated by applying phenomena of spatial v·i sion, can be compared to the stereophonic perspective, which can be simulated by applying corresponding phenomena of spatial hearing.

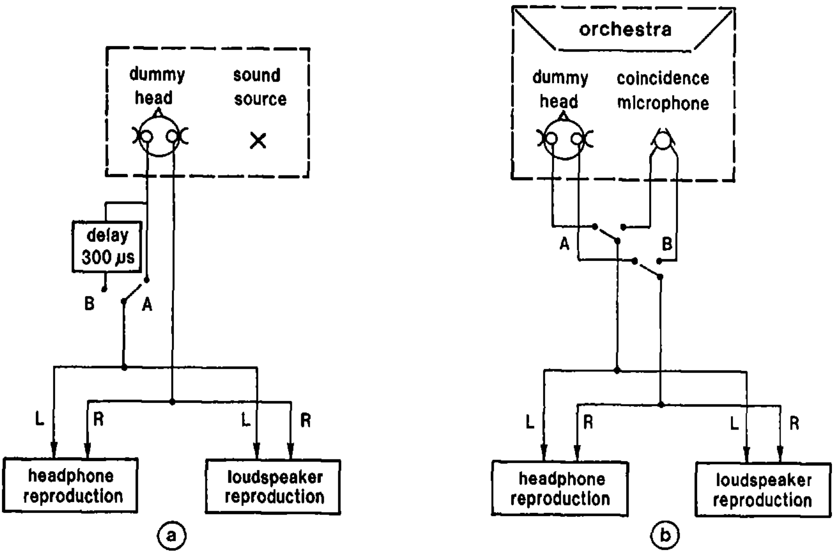


Fig. 2: Two examples for demonstrating stereophonic quality differences in the case of headphone reproduction and loudspeaker reproduction

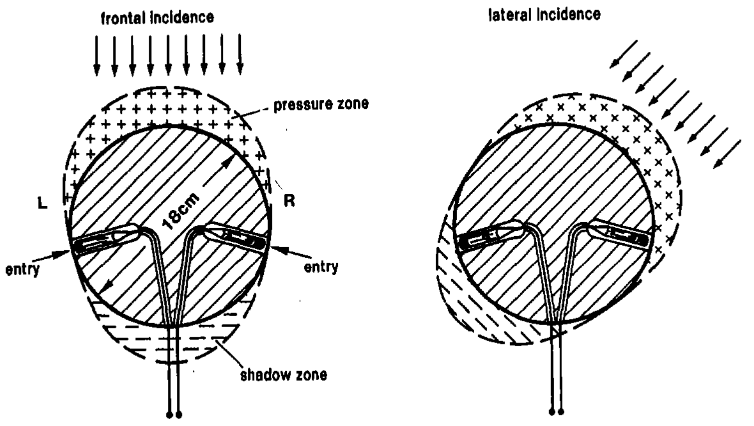


**Fig. 3.:** Principal function of the sphere microphone

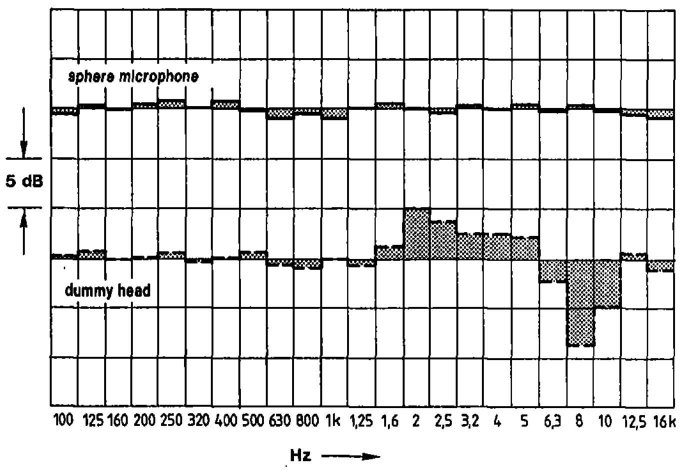


**Fig. 4:** Frequency response of the sphere microphone and dummy head (0°-free-field minus diffusefield)

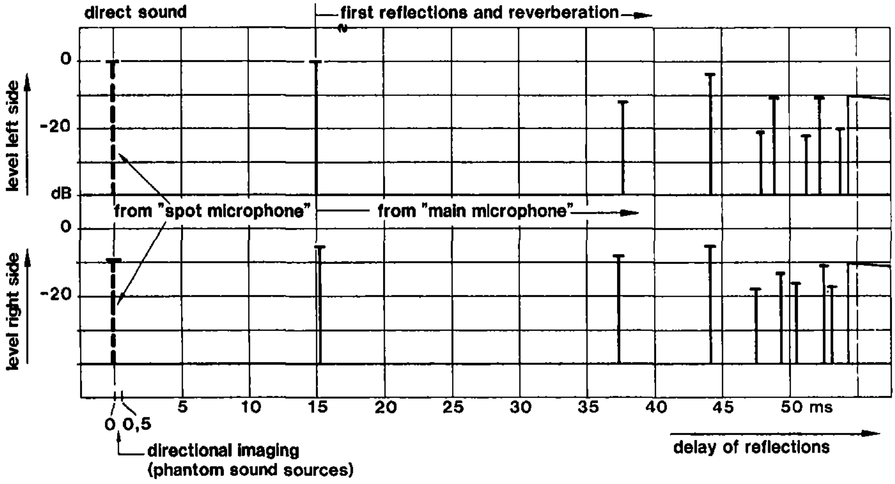


Fig. 5: Stereophonic signal for brief impulses. Main microphone signal plus spot microphone signal without time delay (panpot-balancing)

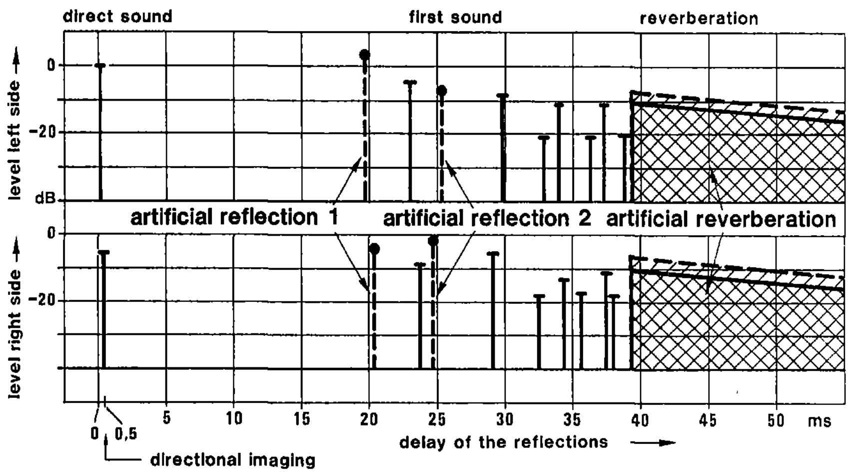


Fig. 6: Stereophonic signal for brief impulses. Main microphone signal plus artificial reflections and reverberation (room-related balancing)

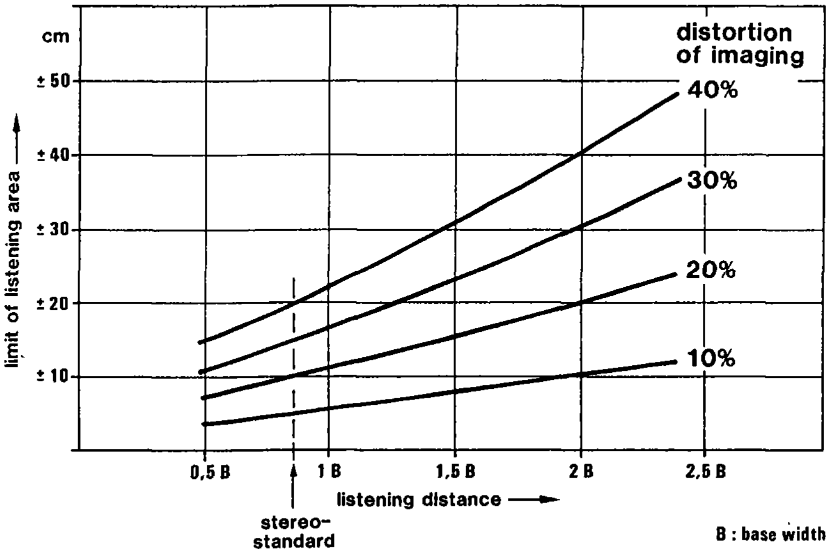


Fig. 7: Maximum width of the listening area in loudspeaker stereophony. Parameter: permissible deviation of perceived direction (imaging distortion)

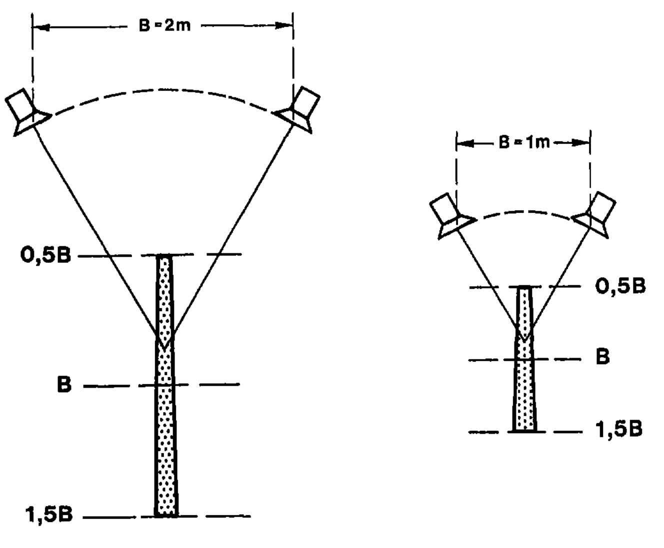


Fig. 8: Listening areas in two-channel stereophony (Permissible imaging distortion V = 10 %)

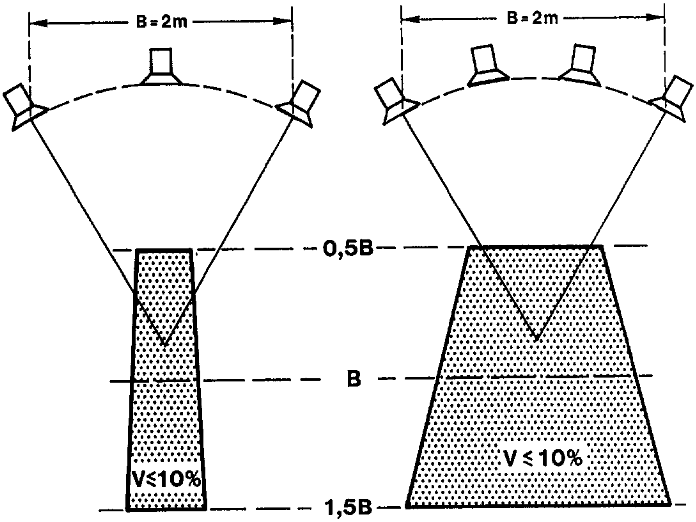


**Fig. 9:** Listening areas in multi-channel stereophony (Permissible imaging distortion V = 10 %)

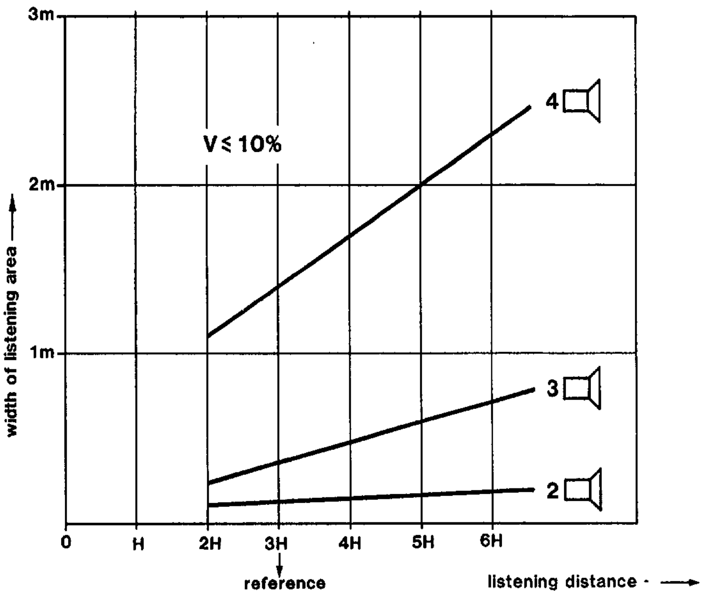


**Fig. 10:** Width of listening area for 2-, 3-, or 4-channel stereophony (Permissible imaging distortion V = 10 %)

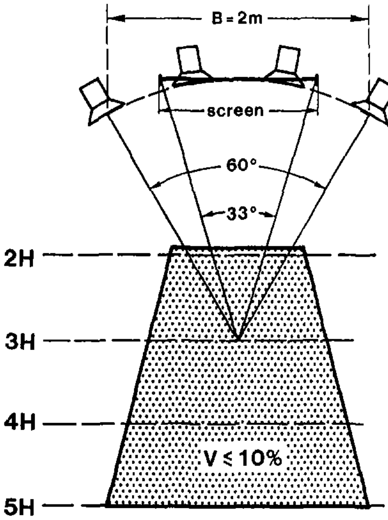


Fig. 11: Application of 4-channel stereophony to HDTV (aspect ratio W : H = 16 : 9)

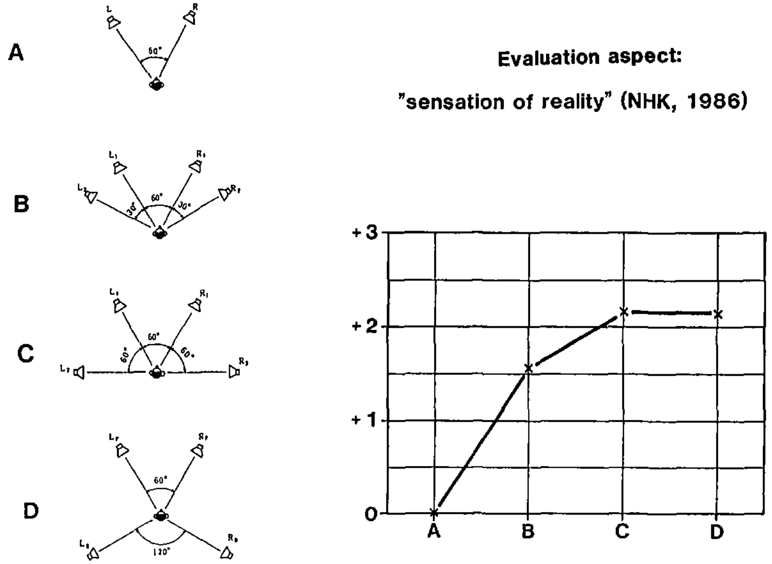


Fig. 12: Effect of various arrangements of two surround loudspeakers (Adapted from Ohgushi et al /28/).

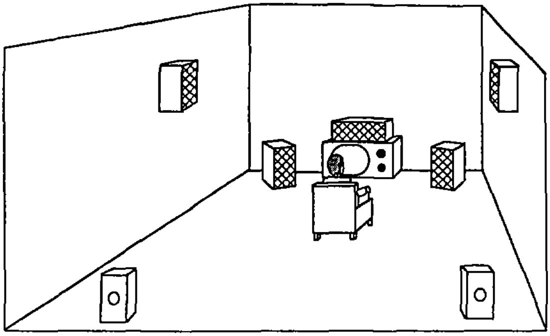

**Fig. 13:** Loudspeakers placement for film and music, proposed by Griesinger (copy from /38/)

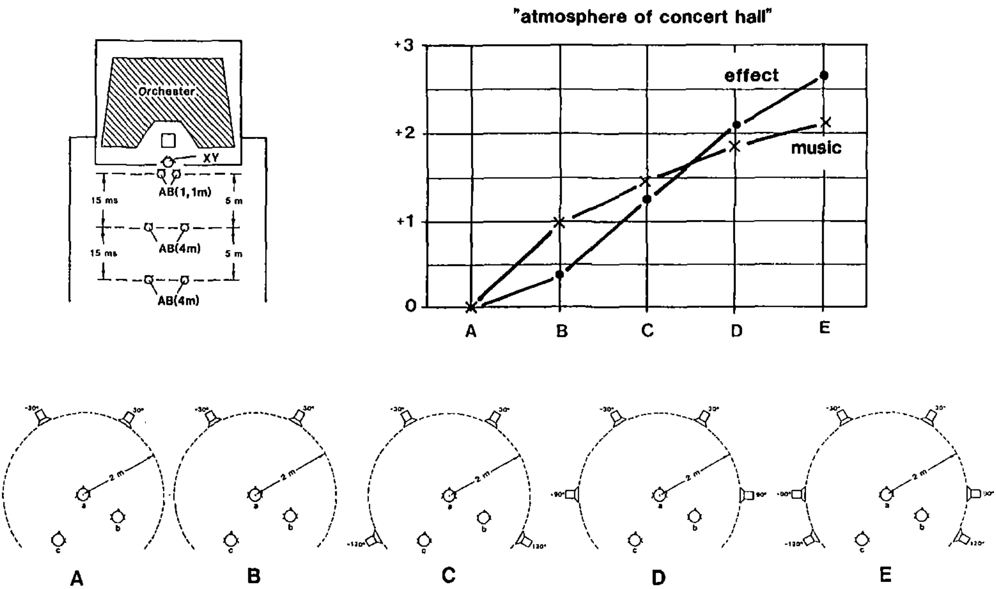


**Fig. 14:** Recording arrangement, reproduction arrangements, and results of comparison tests (evaluation aspect: "atmosphere of concert hall")

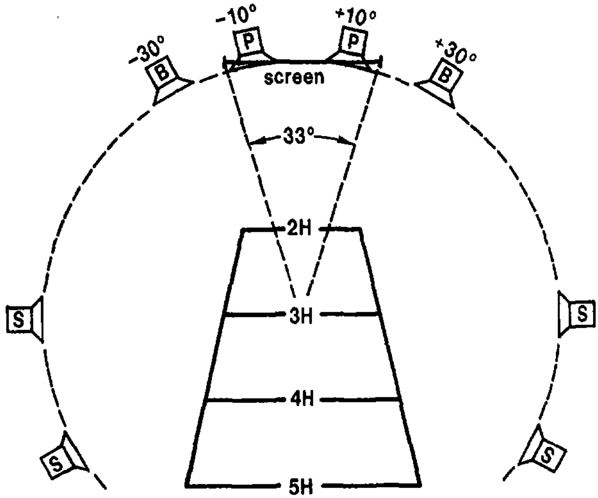


**Fig. 15:** Multi-channel sound system, suitable to accompany HDTV systems (Maximum performance level, basic loudspeaker arrangement). The location of the surround speakers is highly variable.