A Survey Study Of In-Situ Stereo And Multi-Channel Monitoring Conditions

Aki V. Mäkivirta and Christophe Anet
Genelec OY
Olvitie 5, Iisalmi, Finland
aki.makivirta@genelec.com chris.anet@genelec.com

ABSTRACT

The in-situ responses of a total of 372 loudspeakers in 164 professional monitoring rooms around the world have been measured after acoustical calibration. All measured rooms have been equipped with factory calibrated three way monitors and acoustically calibrated with standardized apparatus. The results provide a thorough understanding of typical monitoring conditions for stereo and multi-channel rooms, distribution in room parameters and quality of reproduced audio. Results are compared to current standards and recommendations.

INTRODUCTION

The audio quality of monitoring determines the resulting quality of the recording process. The monitor loudspeakers strongly effect the aesthetic decisions of artists and engineers producing audio material. The installation of the speakers and the monitoring room itself has typically an even stronger influence, and an increasing number of standards and recommendations have been issued to ensure consistency and high quality of the monitoring process both with and without an accompanying picture [1-5]. These give recommendations of aspects of in-situ reproduction quality such as

- magnitude response at the listening location
- dynamic range and noise level
- time-of-flight difference
- idealized layout configuration
- nominal listening level
- reverberation and reflection in the room
- quality of the reproduction system

METHODS

All rooms included in the present study are equipped with a factory calibrated 3-way main monitors, produced by one manufacturer[1].

Measurement Apparatus and Method

The measurement method used an MLS sequence[2] of period 16383 samples (217ms) at a sampling rate of 75.47kHz. An impulse response of length 16383 samples was stored for each loudspeaker.

The measurement apparatus was laboratory calibrated to a reference measurement system before and after the measurements. The microphone[3] was calibrated by producing a separate calibration file that was used during analysis to equalize the magnitude response calculated from an impulse response. To ensure high consistency of the measurement process one person used the measurement appa-

---

[1] Genelec OY, Iisalmi, Finland

[2] MLSSA measurement system, DRA Laboratories.

[3] Neutrik Type 3382.

ratus throughout the study to perform all on-site calibrations and measurements.

Detection of impulse response start

The onset of an impulse response is determined as 10 samples before the impulse response power estimate exceeds 3% (–30dB) of its peak value. This corresponds closely to an impulse response start definition given in standards of room reverberation measurement [6].

Signal-to-Noise Ratio and Listening Distance

The stored impulse response contains the natural time-of-flight before the impulse begins. This section in the recorded signal represents the total system plus room noise level, and it is used as an indication of the quality of measurement (signal-to-noise ratio) and measurement distance.

The signal-to-noise ratio D of a measurement is calculated as a ratio of the recorded impulse response data $a(t)$ peak value to the noise level rms value within the time-of-flight period $[t_0,\ t_1]$ before the onset of the actual impulse response

$$D = \frac{\max\left|a(t)\right|}{\sqrt{\dfrac{1}{T}\int_{t_0}^{t_1} a^2(t)\,dt}} \tag{1}$$

The listening distance is calculated from the time-of-flight assuming the speed of sound in air as 344m/s.

Reverberation Time

Nonlinear fitting technique due to Karjalainen et al. [8] is used to estimate the reverberation time in both the full-band and octave-band cases. This method is accurate in the presence of noise and works well also in sparse reverberation fields typical of high quality control rooms.

Karjalainen's method fits three parameters into time envelope data. These are the initial amplitude A, decay time constant τ and final constant level A_n modeling the noise level of a measurement within the time span $[t_1,\ t_2]$ or interest. The modeling is formulated as a least-squares minimization task[4]

$$\min_{A,\tau,A_n} \int_{t_1}^{t_2} \left[a(t) - y(t)\right]^2 dt \tag{2}$$

The 60-dB decay time T_{60} is related to the decay time constant τ by

$$T_{60} = -\frac{1}{\tau}\ln(10^{-3}) \approx \frac{6.908}{\tau} \tag{3}$$

The initial and final levels A and A_n are not recorded in our study. The octave bands included in the study are from 63Hz to 16kHz. Before performing the fitting, the start of fitting time span t_1 is updated to point to the maximum value in an impulse response if the maximum value of an octave-band impulse response does not occur at the detected impulse response start.

Early-Late Energy Ratio

The early-late ratio C_x describes the energy level ratio before and after a certain point in time, and is linked to the reverberation time in a space. The early-late ratio C_x is defined as

$$C_x = 10\log_{10}\frac{\int_0^x p^2(t)dt}{\int_x^\infty p^2(t)dt} \tag{4}$$

To assess the level of late energy, several early-late energy ratios were calculated (C_{35}, C_{50}, C_{80}, C_{95}) for each impulse response.

Room Operational Response Curve

The target for the magnitude response at the listening location, or the room operational response curve, is defined as the third-octave smoothed magnitude response. The target for this is given as an acceptance window function centered at the mean value in this response calculated over the frequency range 50Hz-16kHz.

Sliding third octave smoothed magnitude responses are calculated for each impulse response. Each magnitude response is level aligned by setting the mean value calculated over the frequency range 50Hz-16kHz to one (or zero dB).

At each frequency, level aligned magnitude values from all responses are collected to form the magnitude response distribution.

The median, 50% and 90% percentiles of this distribution are extracted.

Notches at $f < 1$kHz

Boundary reflections and room modes can produce comb filtering effects and standing wave cancellation in the monitoring space displayed as notches in the magnitude response. These effects are quantified for frequencies 40Hz $< f <$ 1kHz (the search frequency band). Ten deepest notches are sought in the raw magnitude response within the specified frequency band.

Notches are identified with the following process. A detection level is moved up starting from the global minimum in the data within the search frequency band. A monotonously increasing neighborhood of an identified minimum is recorded. The detection level is increased until 10 largest minimums are found.

All minimums found in this manner are notches, but the monotonously increasing neighborhood for some of them is quite narrow, and therefore Q-value calculation for the notches is not attempted. The center frequency and maximum depth of the notches is recorded.

Speaker Pair Comparisons

To assess the similarity of magnitude responses of monitor speakers in a room, the magnitude responses of monitor speakers are compared pair-wise in three groups

- left-right stereo pairs in a room
- left-center-right front speaker triplets in a room
- five-channel systems in a room

The difference between Left-Right stereo pair is calculated.

The Left-Center-Right triplets are compared to the Center channel.

The five-channel system analysis compares separately third octave smoothed responses of front monitors (left-center-right triplet) and rear left-right triplet because the front-back balancing is not performed as a part of system alignment.

LOUDSPEAKER MATERIAL

A total of 372 loudspeaker in 164 professional monitoring rooms around the world have been measured after acoustical calibration. All rooms included in the present study are equipped with a factory calibrated 3-way main monitor system, produced by one manufacturer[5].

Of the 372 recorded impulses, 277 (75%) have been measured at the engineer's position (Tables 1-3). Of the measurements at the engineer's position, 250 (90%) have also been measured on the acoustical axis of the loudspeaker, implying that a speaker's acoustical axis is aimed toward the engineer's position.

---

[4] Implementation in Matlab uses the `lsqcurvefit` function.

[5] Genelec Oy, Iisalmi, Finland

Additional data has been recorded for the impulse response files, tabulated for several factors known to affect measured responses. The factors included are –

- installation type (free-standing, flush-mounted in a soffit)
- measurement location (at or outside the engineer's position)
- measurement taken on a speakers acoustical axis (yes/no)
- a speaker vertically tilted toward the listening point (yes/no)
- monitor height (below/at/higher than the listening point)
- room size (small/medium/large)
- ceiling (soft/hard)
- side walls (soft/hard)
- back wall (soft/hard)
- room geometry (square, rectangular but not square, other)

Loudspeaker type is given in Table 2, indicating also the type of speaker installation. The 1037A, 1037B and 1038A are mostly used free-standing, but even loudspeakers intended for in-wall installation (1034, 1035, 1036, 1039) were sometimes used free-standing. The detailed specifications of all monitor speakers included in the study can be found on the web site of the manufacturer.[6]

ROOM-RELATED PARAMETERS

Signal-to-Noise Ratio

The dynamic range in a measurement is typically 60dB (Fig. 1) and varies from less than 40dB to over 70dB.

The dynamic range measured in this way contains the noise contributions of both the measurement apparatus and the room in which the measurement was taken. Typically the contribution of the measurement apparatus is not significant.

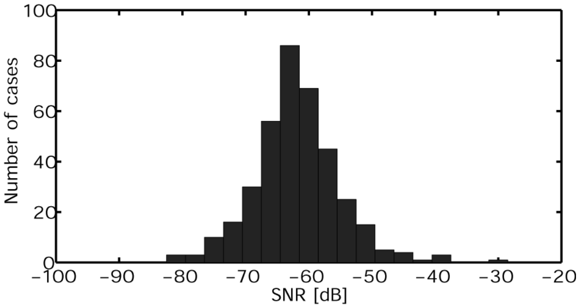

Fig. 1. Signal-to-noise ratio of impulse responses ($N = 372$).

Listening distance

The distribution of listening distance has been analyzed for those monitor speakers having their acoustical axis directed toward and the impulse responses recorded at the listening position. Of all recorded impulse responses, 250 were measured at the listening position, and were analyzed for distance.

The distances for studied three-way monitor systems range from 1.2 meters to 4.2 meters. The average listening distance is 2.49 meters (Fig. 2). The distances estimated in this manner correspond to actual listening distances and can be taken to indicate the listening distance distribution for main monitors including both front and rear monitors in multichannel audio configurations. The mean distance of individual channel location is given in Table 5.

The measurement signal-to-noise ratio correlates only weakly to the listening (measurement) distance (Fig. 2).

Table 1. Type of installation and measurement axis.

Measured at engineer's position	*Type of installation*	*Measured on acoustical axis*		*Total*
		Yes	No	
Yes	In-wall	129	23	152
	Free standing	121	4	125
No	In-wall	66	0	66
	Free standing	29	0	29
Total		345	27	372

Table 2. Speaker height and vertical tilt in monitoring rooms of various sizes.

Room Size	*Speaker Height*				*Total*
	Ear Level		*Higher*		
	Vertical Tilt		*Vertical Tilt*		
	yes	no	yes	no	
n.a.			2	4	6
Small		10	8	12	30
Medium	2	11	56	33	102
Large		34	130	70	234
Total	2	55	196	119	372

Table 3. Acoustical axis orientation at the engineer's position.

Room Size	*At Engineer's position*		*Total*
	On-Axis	*Off-Axis*	
n.a.	6		6
Small	24		24
Medium	66	10	76
Large	154	17	171
Total	250	27	277

Table 4. Loudspeaker type and type of installation.

Monitor Type	*Type of Installation*		*Total*
	In-Wall	*Free Standing*	
1033A	10		10
1034A, 1034B	21	10	31
1035A, 1035B	34	2	36
1036A	10	4	14
1037A, 1037B	29	78	107
1038A, 1038AC	81	52	133
1039A	33	8	41
Total	218	154	372

---

[6] www.genelec.com

Table 5. Mean listening distances of individual channels for left-right stereo rooms, left-center-right triplet rooms and five-channel full surround rooms. Distances are in meters.

Channel	*Stereo* N = 79	*L/C/R* N = 15	*5-chan.* N = 8
L	2.36	2.66	2.68
C		2.66	2.71
R	2.39	2.66	2.72
SL			2.70
SR			2.69

The distribution in listening distances (Figs. 3-13) is the widest, as can be expected, in stereo pairs.

The minimum listening distance for a three-way system is about 1 meter. However, it is clear that listening distances this short may significantly modify the frequency response as the listening is done in the near field, and 3-way loudspeaker products included in the study are not intended for near-field listening.

A typical listening distance of a three-way main monitor in a stereo configuration is 1.5–2.5 meters. This is shorter than that of L/C/R-triplets (Figs. 6-8). The mean listening distance of L/C/R-triplets and five-channel surround systems (Figs. 9-13) is about 0.3 meters longer.

An electronic delay device may be use to equalize the time-of-flight delay differences. Since the measurement system used for alignment is connected straight to the speaker input, any electronic delay devices are excluded from this analysis.

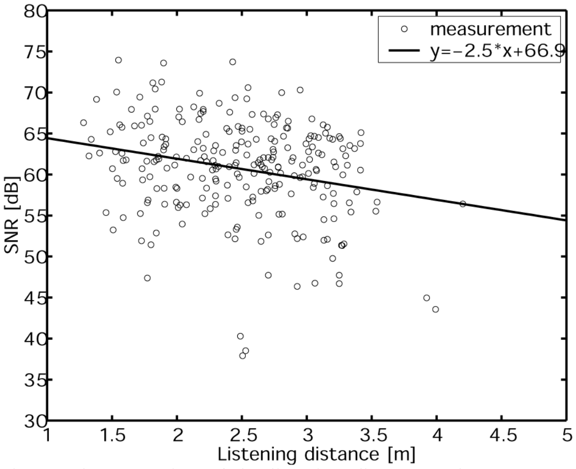

Fig. 2. The regression of the listening distance and measurement signal-to-noise ratio estimated in impulse responses ($N = 250$).

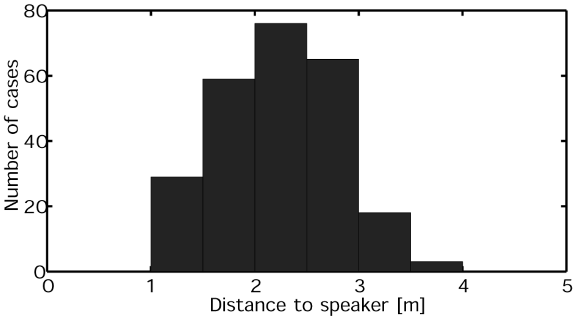

Fig. 3. The distribution of listening distance, all loudspeakers including surround. Mean distance to a loudspeaker is 2.49 meters ($N = 250$).

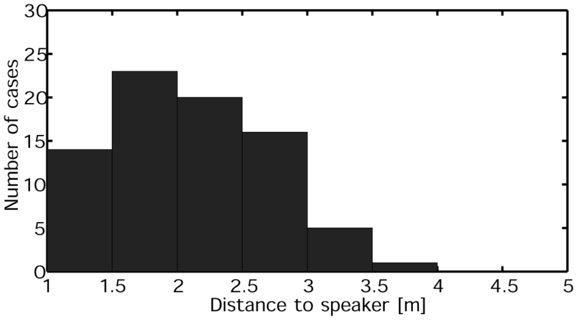

Fig. 4. Listening distance to left speaker in left-right stereo pairs, (mean = 2.36m, $N = 79$).

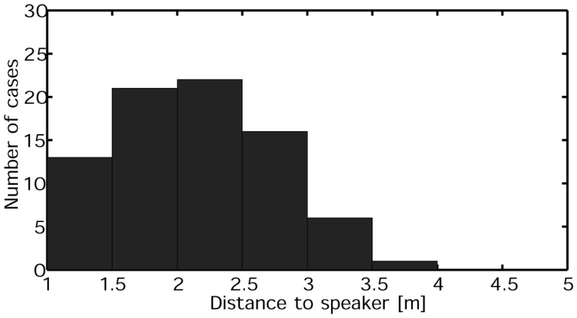

Fig. 5. Listening distance to right speaker in left-right stereo pairs, (mean = 2.39m, $N = 79$).

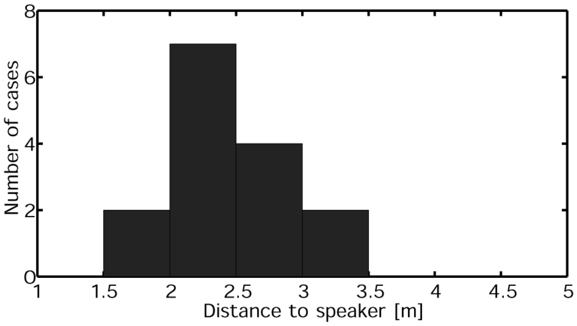

Fig. 6. Listening distance to center speaker in left-center-right triplets, (mean = 2.66m, $N = 15$).

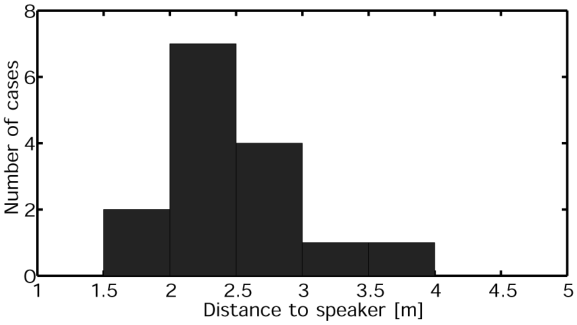

Fig. 7. Listening distance to left speaker in left-center-right triplets, (mean = 2.66m, $N = 15$).

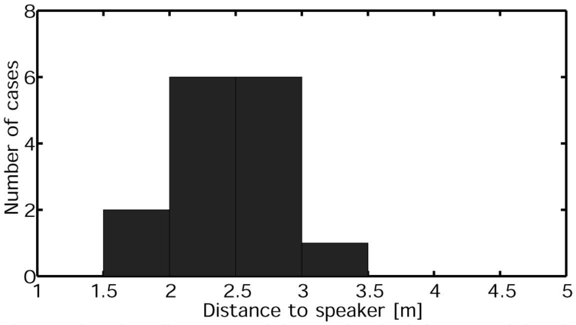

Fig. 8. Listening distance to right speaker in left-center-right triplets, (mean = 2.66m, $N = 15$).

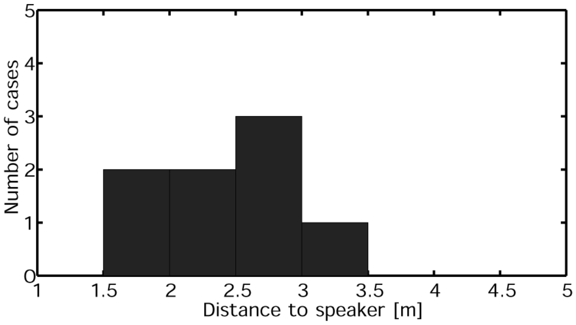

Fig. 9. Listening distance to front left speaker in five-channel surround systems, (mean = 2.68m, $N = 8$).

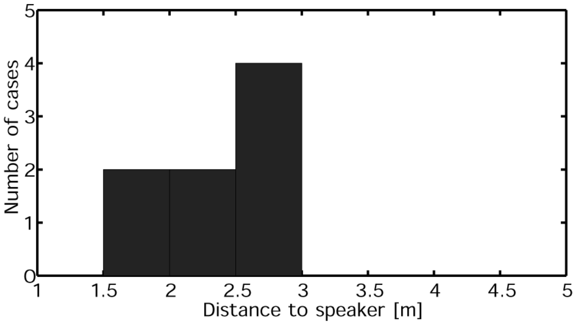

Fig. 10. Listening distance to front center speaker in five-channel surround systems, (mean = 2.71m, $N = 8$).

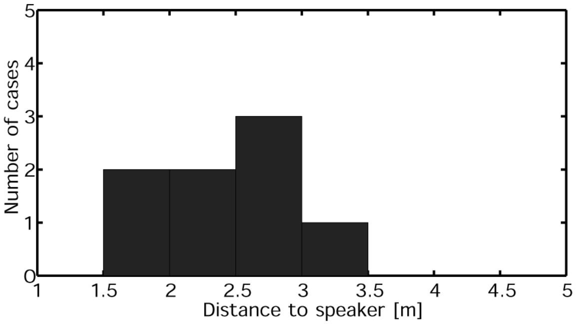

Fig. 11. Listening distance to front right speaker in five-channel surround systems, (mean = 2.72m, $N = 8$).

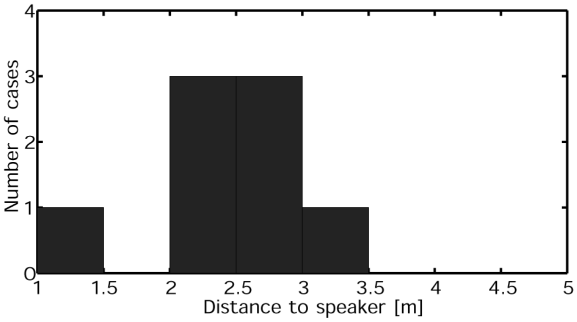

Fig. 12. Listening distance to rear left speaker in five-channel surround systems, (mean = 2.72m, $N = 8$).

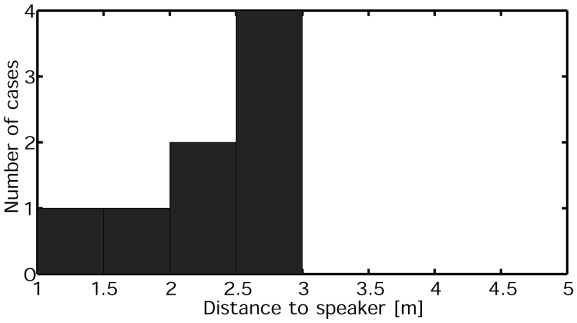

Fig. 13. Listening distance to rear right speaker in five-channel surround systems, (mean = 2.72m, $N = 8$).

Reverberation Time

Distributions of reverberation time are given (Figs. 15-24) for the full bandwidth and octave-band measurements.

At each octave band, the median, 50% and 90% bounds are given in Fig. 14. The mean reverberation time RT_{60} is 380ms from 200Hz to 4kHz.

Most rooms show reverberation times that conforms to present standards and recommendations for high quality monitoring rooms, but there are large differences, related to the use of either absorption or diffusion to control the reverberant decay field.

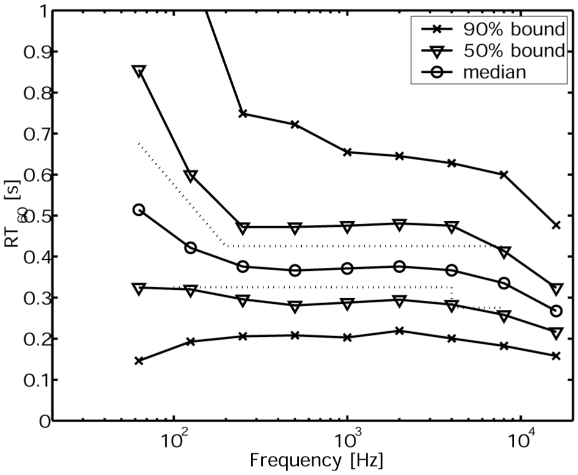

Fig 14. The minimum, maximum, 90% and 50% limits and the median for reverberation time RT_{60} in octave bands in measured rooms (N = 372). German Surround Sound Forum limit shown centered at the median of mean RT_{60} levels.

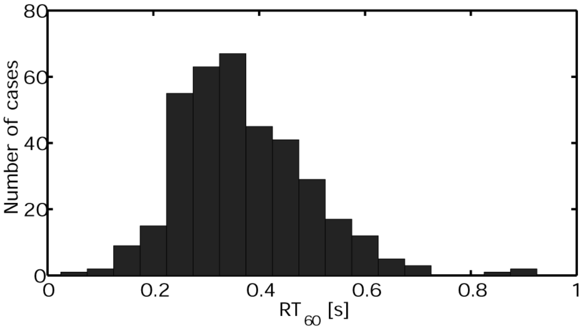

Fig. 15. Distribution of full bandwidth reverberation time RT_{60} ($N = 372$).

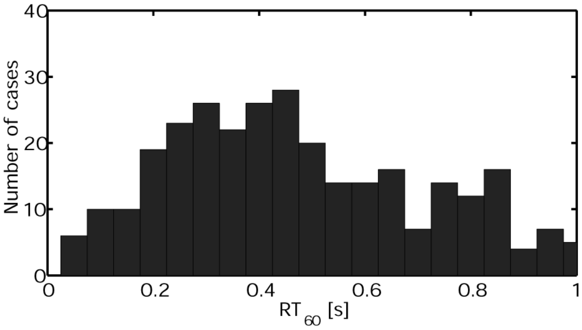

Fig. 16. Distribution of reverberation time RT_{60} of 63.5Hz octave band ($N = 372$).

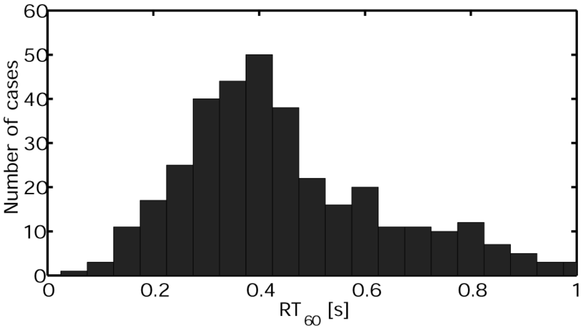

Fig. 17. Distribution of reverberation time RT_{60} of 125Hz octave band ($N = 372$).

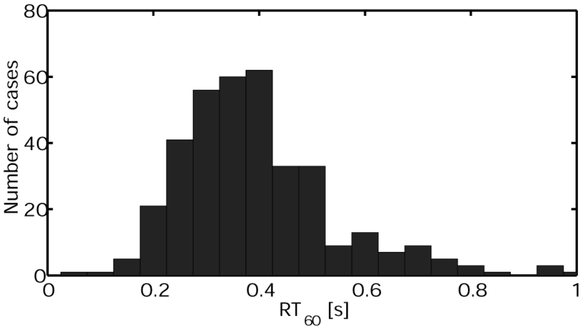

Fig. 18. Distribution of reverberation time RT_{60} of 250Hz octave band ($N = 372$).

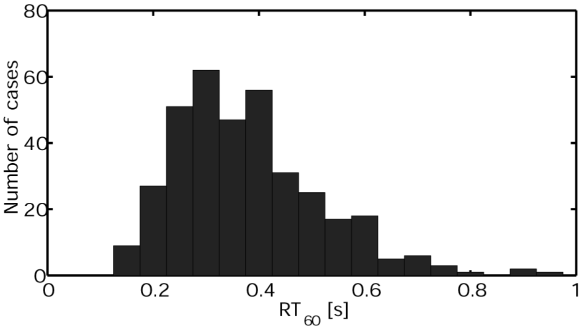

Fig. 19. Distribution of reverberation time RT_{60} of 500Hz octave band ($N = 372$).

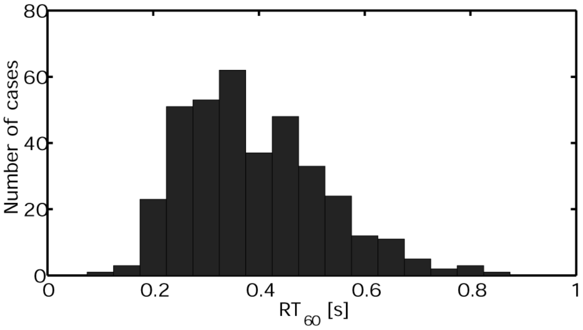

Fig. 20. Distribution of reverberation time RT_{60} of 1kHz octave band ($N = 372$).

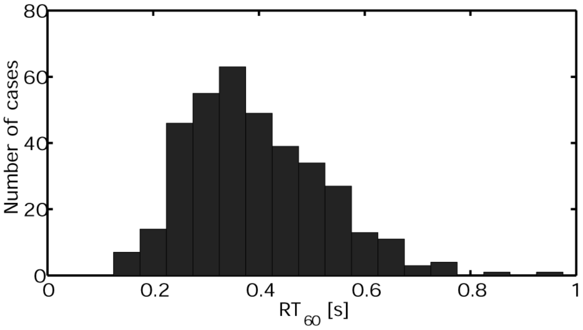

Fig. 21. Distribution of reverberation time RT_{60} of 2kHz octave band ($N = 372$).

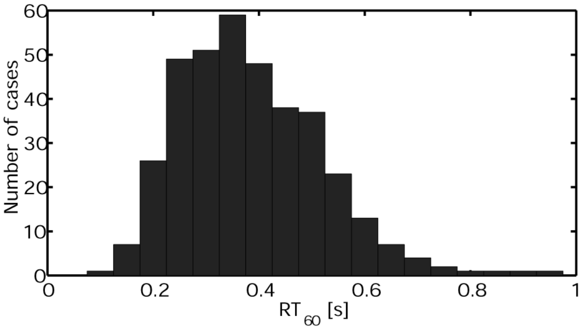

Fig. 22. Distribution of reverberation time RT_{60} of 4000Hz octave band ($N = 372$).

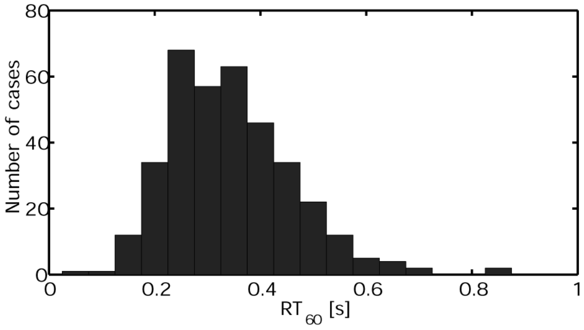

Fig. 23. Distribution of reverberation time RT_{60} of 8kHz octave band (N = 372).

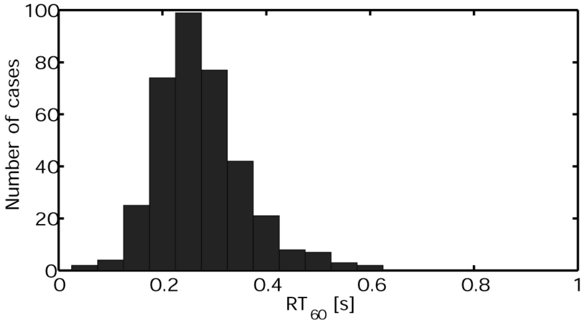

Fig. 24. Distribution of reverberation time RT_{60} of 16kHz octave band (N = 372).

Early-Late Energy Ratio

C_{35}, C_{50}, C_{80}, C_{95} early-late energy ratios have been estimated but only the extremes are presented for clarity in Fig. 25 and 26. The C_{35} represents the level of early energy relative to the direct sound in the monitoring room. The C_{95} represents the level of late reverberation relative to the combined direct sound and early energy.

The reverberation characteristics of rooms can also be viewed by C_{35} and C_{95}. These show the ratio of power integrated in the interval 0…35ms and after 35ms, or 0…95ms and after 95ms, respectively. Some rooms have strong resonant behavior leading to low C_{35} and C_{95} figures.

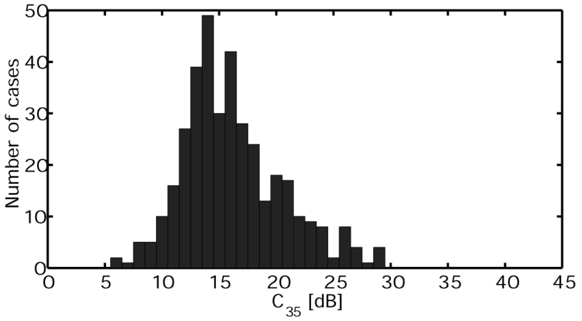

Fig. 25. The distribution of early-late energy ratio of energy earlier and later than 35ms after onset of the impulse (C_{35}) in the measured rooms (N = 372).

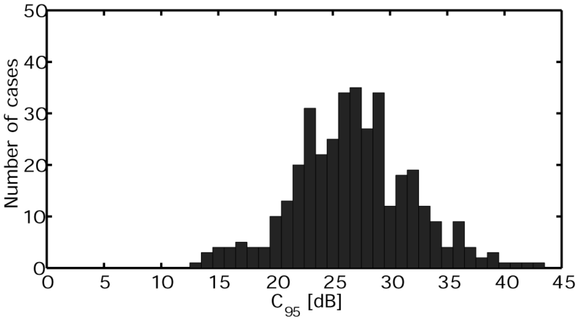

Fig. 26. The distribution of early-late energy ratio of energy earlier and later than 95ms after onset of the impulse (C_{95}) in the measured rooms (N = 372).

Magnitude Response

The overall frequency balance at the listening position for loudspeakers aimed at the listening position (N = 250) is represented by third octave smoothed frequency response deviations (Fig. 27) at the listening position for loudspeakers aimed at the listening position (N = 250). To obtain the distributions, each frequency response is normalizing to the mean level between 50Hz and 16kHz as proposed by the German Surround Forum [5]. Also shown in are the room operational curve limits as proposed in [5] set relative to the 50Hz – 16kHz mean of the median of distributions.

The 50% variation limit is within the proposed limits [5] for frequencies f > 130Hz, and 90% of rooms for frequencies f > 400Hz. We can see the responses generally suffering loss of level above 16kHz. Only 5% of rooms show straight response up to 20kHz.

As expected, the frequency response at the listening position for loudspeakers aimed at the listening position shows a decreasing variation toward high frequencies. Notches in the magnitude responses are displayed as a larger spread toward negative values.

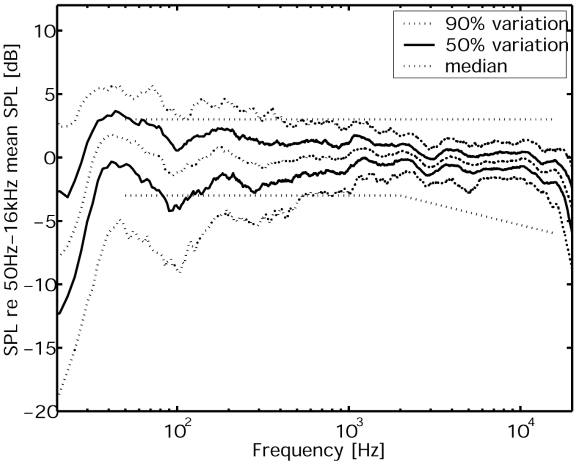

Fig. 27. Third octave smoothed sound pressure level measured at the listening position (N = 250) for speakers aimed at the listening position. 50Hz-16kHz mean level normalized to 0dB. Also shown are German Surround Forum proposed limits.

Response Notches for $f < 1$kHz

To demonstrate notches in the frequency response recorded at the primary listening position for speakers that have been aimed optimally toward the listener, we studied measurements taken at the listening positions for those speakers having their acoustical axis directed toward the listening position

The median notch depth is 14.2dB, but 30dB notches are not uncommon (Fig. 28).

In our material the most typical notch frequency (Fig. 29) is 100Hz, but deeper notches appear at higher frequencies (Fig. 30).

It is important to note that our data contains information of 10 deepest notches in each impulse response measurement, not all notches. As can be seen, also notches up to 6dB level are included, but this should not be taken to mean that all notches up to this level are invariably included. Some rooms show deeper notches, and the less deep notches come simply from better rooms.

MAGNITUDE RESPONSE MATCH

In order to look at the magnitude response matching after calibration, we have studied those 105 rooms of the total of 164 rooms that had speakers with acoustical axis aimed at the engineer's position. The total number of speakers measured at and aimed toward the engineer's position was 250.

The pair-wise comparison included 243 speakers with eight five-channel surround setups (40 speakers), 15 L/C/R triplets (45 speakers) and 79 L/R stereo pairs (158 speakers).

Three of the remaining seven speakers were single speakers, implying that other speaker(s) in the room have not been measured at or aimed toward the engineer's position. These speakers have been excluded because there was no pair to match.

Four of the remaining speakers have been measured in a five-channel surround room but the surround speakers have not been measured at the engineer's position, and therefore this set of speakers was excluded from the pair-wise comparison.

Left-Right Pairs

The left-right stereo pairs in two-channel rooms (Fig. 31, number of rooms N=79) show a very good agreement.

The 50% bounds are about ±2dB above 1kHz and ±4dB below that frequency. The 90% distribution limits are ±4dB for $f > 150$Hz, ±8dB for $50\text{Hz} < f < 150$Hz.

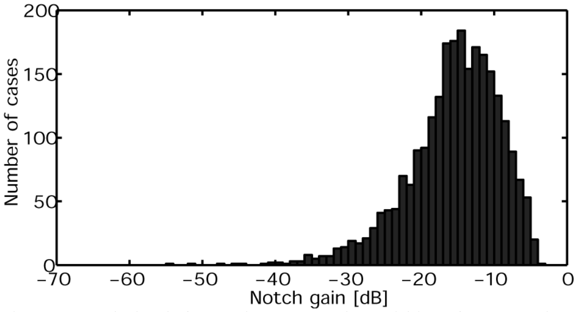

Fig. 28. Notch depth for 10 deepest notches within a frequency band 50Hz to 1kHz. Size of a bin is 1dB, median notch depth is 14.2dB ($N = 250$).

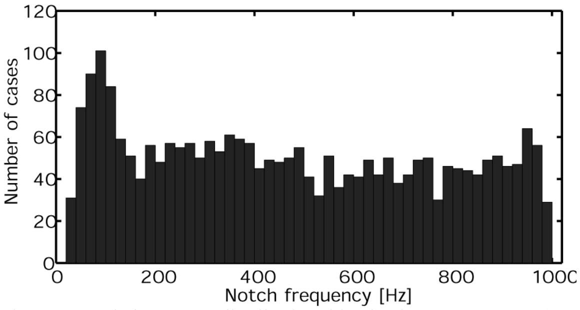

Fig. 29. Notch frequency distribution, bin size is 20Hz ($N = 250$).

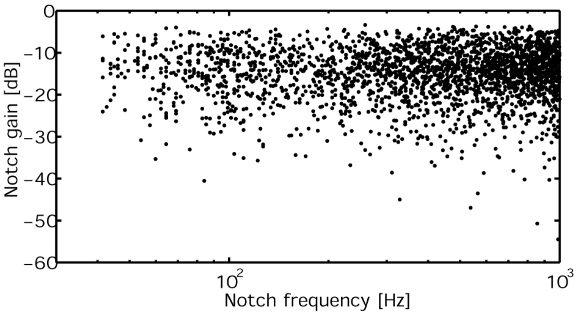

Fig. 30. The scatter of notch frequency versus gain ($N = 250$).

The agreement is better than that of the L/C/R triplets when the responses are compared to the center channel, probably because typical installations are symmetrical for the Left-Right pair whereas the center speaker usually has differing radiating condition from those of left and right front speakers

The typical increase in distribution toward low frequencies is seen. The speakers used in this study have carefully controlled radiation characteristics at mid and high frequencies, minimizing room effects, and resulting in improved pair-match except for low frequencies. Despite that 50% of systems show a mismatch of more than 2dB below 1kHz, and a mismatch of this magnitude is likely to affect auditory imaging, and probably result in reduced sharpness of stereo imaging.

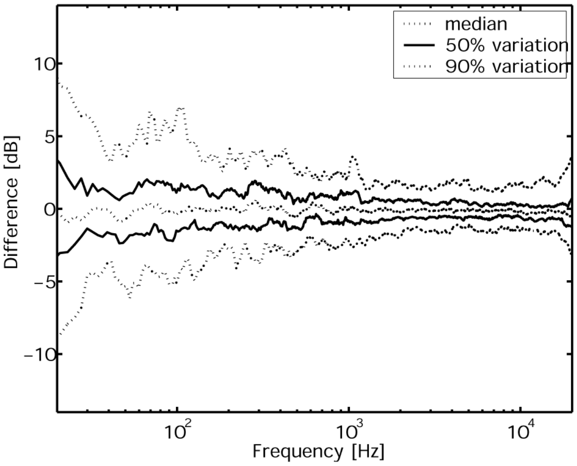

Fig. 31. Third octave smoothed sound pressure level difference of Left-Right pair in L/R stereo systems ($N = 79$).

Left-Center-Right Triplets

Left/center/right-triplets (L/C/R) are found as front speakers in multichannel reproduction rooms using typically two-way systems as rear speakers.

The 90% distribution for the L/C/R triplets (Fig. 32 and 33, number or rooms $N = 15$) increases below 400Hz and above 10kHz.

The 50% distribution is within a ±3dB window for frequency $f >$ 1kHz (except for $f >$ 15kHz) and in a 6dB window below that frequency.

Comparing to stereo left-right pair match, the larger distribution is produced because the comparison is made to the center speaker. This is valid in the sense that in multichannel systems the center speaker is receiving an increasingly important role in forming the sound stage.

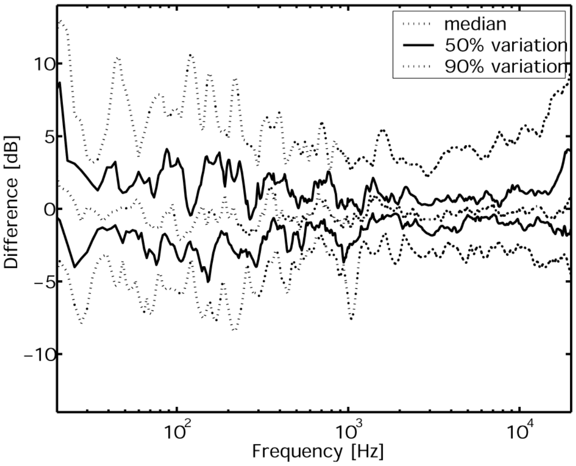

Fig. 32. Third octave smoothed sound pressure level difference of left-center pair in L/C/R systems ($N = 15$).

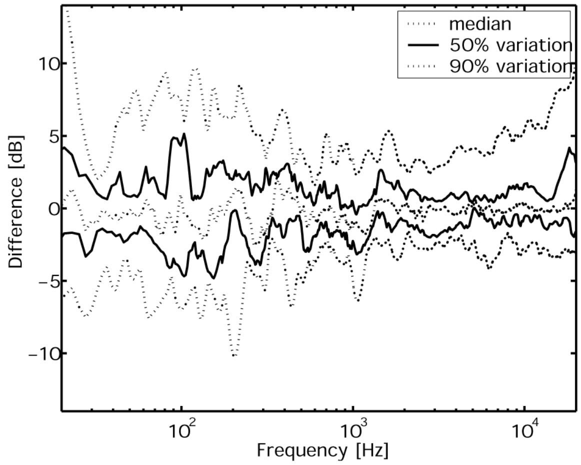

Fig. 33. Third octave smoothed sound pressure level difference of right-center pair in L/C/R systems ($N = 15$).

Five-Channel Surround Set-Ups

Five-channel surround systems use three-way speakers for all five audio channels. Rooms with such setups are fairly large.

A left-center-right (L/C/R) triplet in the five-channel reproduction system is calibrated to have a proper subjective balance during acoustical calibration. Also the rear left-right (SL/SR) speaker pair is calibrated to correct subjective balance. However, the front-rear balance is not adjusted as a part of the acoustical calibration due to the differing operational standards in monitoring rooms included in the study. Therefore the SL/SR level is not compared with the L/C/R level. The front channel match is studied separately from the rear channels.

In five-channel setups third octave-smoothed responses (Figs. 34, 35) of left and right speakers with respect to the center speaker show smaller response deviations than for the left/center/right triplets (Figs. 32, 33). Also the rear left-right difference is small (Fig. 36). The number of rooms in this category is rather small.

For the studied five-channel setups the mean distance from the measurement position to the monitor speakers is the same for all speakers within few centimeters.

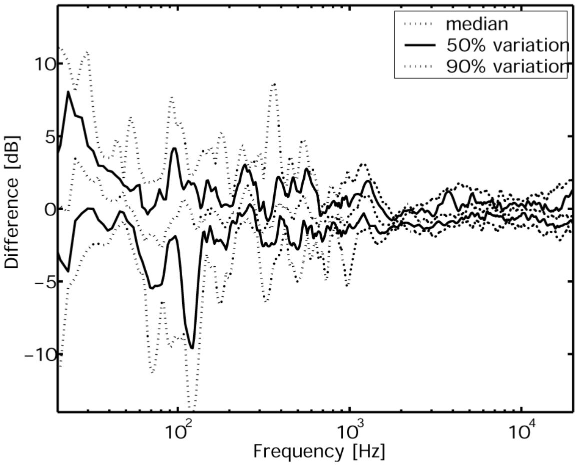

Fig. 34. Third octave smoothed sound pressure level difference of left-centre pair in five-channel surround systems.

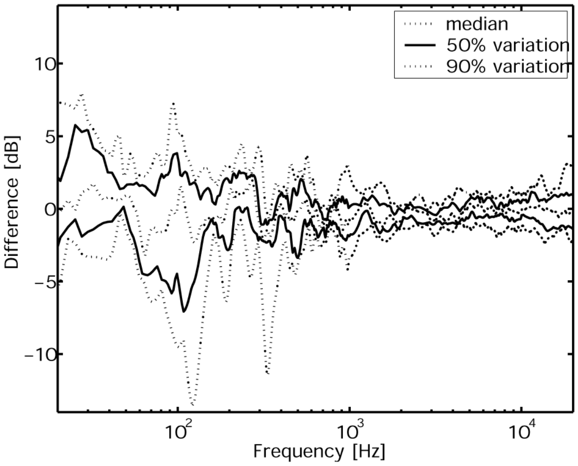

Fig. 35. Third octave smoothed sound pressure level difference of right-centre pair in five-channel surround systems.

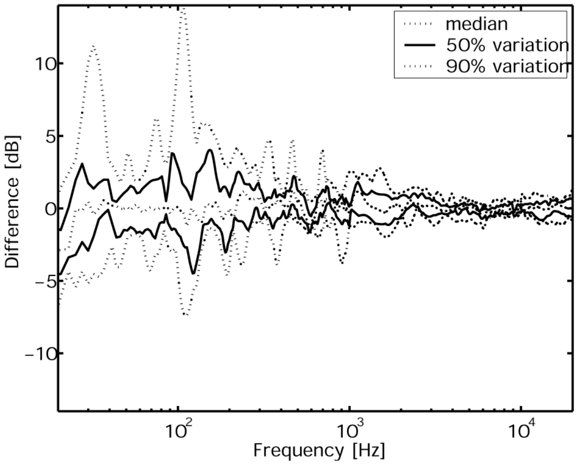

Fig. 36. Third octave smoothed sound pressure level difference of rear left-right pair in five-channel surround systems.

DISCUSSION

Loudspeaker Geometry

There is a significant discrepancy between the recommendations for speaker placement for surround monitoring and what actually happens.

The widely accepted reference configuration for a 3/2 format surround sound reproduction places the Left and Right speakers at ±30° and the surround speakers at ±100…120° relative to the Center speaker. All speakers are on the same horizontal plane at the engineer's ear height (nominally 1.2 meters) [1,2,5] or at maximum 10 degrees inclined above this level [3,4] at an equal distance of 2…4m [1,2,5,8]. Some recommendations allow the rear speakers to be inclined more, up to 15 degrees [1,5,8].

The room size is discussed [3,8] but exact location of the speaker set in the room is not discussed, although a minimum distance to neighboring walls is given and the aim of a symmetric placement is stressed.

The recommendation for speakers to be placed at least 1m away [4,8,9] from walls appears too optimistic. In that case, if the wall behind a speaker has insufficient absorption, a quarter wavelength notch at 86Hz can be produced as a result of a reflection off the wall behind the speaker, causing severe irregularity of a free-standing system bass response.

Increasing the distance brings down the notch frequency. A minimum distance in concert with the recommendation for the operational curve to extend down to 50Hz would be 2m, demonstrating the value of flush mounting even for small two-way monitor systems. Providing sufficient absorption to remove a wall reflection at low frequencies can be difficult and expensive.

Listening distances for three-way systems range from 1.2 meters to 4.2 meters. An average listening distance is 2.5 meters, agreeing well with recommendations. However, in measured rooms the acoustical axis of three-way speakers is rarely placed at the recommended 1.2m height.

The acoustical axis is typically elevated (pointing downwards to the engineer's position) up to 15 degrees. Most large rooms have monitors placed above the listening height with the bottom of the speaker at the height of 1.2m, bringing the acoustical axis 0.5…1m above the listening level. This reduces problems with low order floor reflections and is particularly relevant to flush-mount speakers with a large front baffle.

Horizontally, the speaker acoustical axis is typically oriented toward the listening position. This is also mostly recommended in the literature [1,2,5,8] although there is even advice to direct the acoustical axis past the listening position [9], forcing monitoring with off-axis response. Even with modern constant directivity design approaches monitor loudspeakers are typically optimized for on-axis response, and off-axis response can be significantly degraded.

Speaker size or front baffle size is not discussed as a parameter affecting speaker placement in recommendations although it is recognized that large speakers may have to be placed high. Current recommendation of 1.2m height for the acoustical axis can lead to notches in the 80...120Hz frequency region, causing a deterioration in the bass frequency response.

There is also some discrepancy concerning the surround speaker height. EBU recommends [4] the same height and vertical tilt for surround and front speakers while other recommendations [1,5,8] allow only surround speakers to be placed higher. In our material surround speakers are frequently placed higher than front speakers. Very few rooms have five identical speakers in a surround setup. Rear speakers are typically not of the same type as front speakers.

Reverberation Characteristics

Most rooms show reverberation time that conforms to present recommendations for high quality monitoring rooms.

The decay properties of monitoring rooms are typically not similar to those of larger rooms, but show a reflection-free time zone and after that an evoking dense reflection field, requiring a careful consideration of the measurement of reverberation time for these spaces.

An increase in the reverberation time can be seen when strong diffusion is used. Insufficiency of bass trapping in some rooms is indicated by an increase in the reverberation time distribution limits toward low frequencies. On the other hand, some rooms exhibited very exactly constant reverberation time over the whole frequency band, including low bass frequencies. This indicates that there is still work to be done to increase the quality of monitoring room design.

Notches below *f* = 1kHz

Most of the notches found in our study are produced by low order reflections. As we look at the notch distribution and spread, we see notches around 100Hz more frequently than on other frequencies.

The reflections can be very strong, essentially the same level as the audio signal, producing notch depths up to –40dB.

Room modal resonances begin to dominate listening location notch characteristics below 200Hz where the wavelength typically becomes large compared to size of objects in a room. At higher frequencies notches are created because of low-order reflections and diffractions in the room.

The reasons for notches in the frequency response are typically –

- First order floor reflections.
- Interaction of the console size and shape with the first order floor reflection.
- Incorrect front wall design, e.g. discontinuities in the front wall, such as non-aligned windows, large TV screens or cavities and recesses.
- In case of a soft front wall, if front wall bass trapping is not sufficient, LF notches can be created due to reflections from the hard wall behind a speaker. This creates a quarter-wavelength notch at the frequency corresponding to the distance from the speaker front baffle to the hard (front) wall behind the speaker. This distance can be larger than the depth of the speaker cabinet.
- With free-standing monitors, notches are typically created because of a reflection from the (front) wall behind a monitor.

- Sustained standing waves or modes in a room can cause severe notching and irregularity in the response at the mix position.

Magnitude Response

The magnitude response of an individual monitor loudspeaker should be flat to within ±2dB in anechoic conditions using third octave smoothing. All loudspeakers included in this study fulfill this requirement.

In-situ frequency response measured at the engineer's position using third-octave smoothing should be flat within ±3dB from 50Hz to 16kHz with some level reduction allowed at high frequencies [3,8].

In multichannel rooms, the sound pressure level 50% distribution limits indicate a good frequency response control above 1kHz but an increasing magnitude response distribution below this frequency. Opposite to this, the 50% distribution limits in stereo rooms remain consistently within ±3dB bounds also at low frequencies.

The 90% distribution limits of magnitude response indicate that some rooms have problems in their frequency response below 1kHz. This is primarily a product of room design and loudspeaker installation. There are still frequent failures in low frequency design of monitoring spaces and the management of low order (early) reflections.

Most large modern control rooms can achieve adequate low frequency damping and have a properly designed acoustic treatment enabling a very high quality monitoring. Small control rooms with free-standing monitor systems and compromised acoustical treatment can exhibit large variations in low frequencies.

Magnitude Response Match

According to recommendations the magnitude response difference between front loudspeakers in anechoic conditions should be less than 0.5dB within 250Hz…2kHz [8]. Monitor speakers included in the present study fulfill this requirement. Random loudspeaker pair match after factory calibration is small enough to render production tolerance insignificant to the pair match measured in this study.

The measured median difference is larger than 2dB for all cases studied within this frequency band. The L/C/R triplets show a median difference of 4dB. This demonstrates that room effects play an important role in determining the in-situ frequency response.

The pair match of the stereo pairs and surround left-right pairs is typically better than that of L/C/R triplets. One reason for this is the fact that the Center channel is exposed to different radiating conditions than the Left and Right speakers, and the match was calculated by comparing the Left and Right speakers to the Center speaker. The Left and Right speakers have very similar and symmetrical radiation conditions in modern monitoring room designs while it may be very difficult, if not impossible, to design similar radiating conditions for the Center speaker being typically in the middle of the room. Many installations have large objects such as computer screens, furniture, racks etc. placed centrally near the engineer's position. These objects create strong reflections resulting in comb filtering in mid and high frequencies.

In the case of L/C/R triplets we typically have two-way speakers as surround speakers. This implies smaller rooms and smaller listening distances. Many small 5.1 rooms have non-ideal layout and equipment positioning. This is displayed in the L/C and R/C pair match for the L/C/R triplets where mid and high frequency distribution is higher than in the five-channel systems.

The five-channel systems included in the study use three-way speakers also as surround speakers, implying large monitoring spaces. The directivity control of the waveguide structures incorporated in the three-way speakers is apparent in measurements of the five-channel setups. The pair match of the L/C and R/C pairs is very good above 500Hz demonstrating minimal low order reflections at the engineer's position.

The present study can not answer questions regarding the psychoacoustic significance of mismatch in sound level between the speakers, but it is clear that this mismatch becomes an increasingly acute problem in the future with fast spreading of multichannel audio. New approaches are necessary in monitoring room design to produce rooms capable of accurate reproduction. Better control of directivity in the loudspeaker may also decrease problems in poorly designed environments, but will never be a substitute for a carefully designed room. Flush mounting proves once again a valuable method of decreasing low frequency problems due to reflections off the nearest walls.

References

1. ITU Recommendation ITU-R BS.775-1, "Multichannel Stereophonic Sound System with and without an accompanying Picture", Geneva 1992-1994.
2. ITU Recommendation ITU-R BS.1116-1, "Methods for the Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems", Geneva 1994.
3. EBU Document Tech. 3276-1998 (2nd ed.), "Listening Condition for the Assessment of Sound Programme Material: Monophonic and Two-Channel Stereophonic", 1982.
4. EBU Document Tech. 3276-1999, "Supplement 1: Multichannel Sound", 1999.
5. Surround Sound Forum Recommended Practice SSF01/98, "Listening Conditions and Reproduction Arrangements for Multichannel Stereophony", `http://www.tonmeister.de/foren/surround/surround.htm`, 1998.
6. Surround Sound Forum Recommended Practice SSF-02/1-E-2, "Multichannel Recording in 3/2 Format", `http://www.tonmeister.de/foren/surround/surround.htm`, 1999.
7. ISO 3382, "Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters", International Standards Organization, Geneva, 1997.
8. AES Technical Committee on Multichannel and Binaural Audio Technology (TC-MBAT), "Multichannel Surround Sound Systems and Operations", Technical Document (version 1.5), 2001.
9. AES Recommended Practice for Professional Audio, "Subjective Evaluation of Loudspeakers", AES20-1996, published in *J. Audio Engineering Soc.*, Vol. 44, No. 5, May 1996.
10. M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, “Estimation of Modal Decay Parameters from Noisy Response Measurements”, *Proc. AES 110th Convention*, Amsterdam, The Netherlands, 2001 May 12–15, preprint 5290.