A Survey Study of Today’s Monitoring Conditions

INTRODUCTION

When speakers are put in a room all sorts of acoustical interactions occur. The criteria that define the performance and quality of a speaker are completely different from those defining the acoustical behavior of a room and both must be taken into consideration when setting up speakers. After all, you cannot stop listening to the room effect with its reverberation time and diffuse reflections, or ignore any large dip in the speaker/room response at a specific frequency. The interaction of the speakers and the room are responsible for the overall monitoring condition and ultimately determine the quality of the recording process.

That is why a vital part of Genelec’s customer service is to provide on-site measurements and worldwide calibration of our large active monitor systems. Over the years Genelec has built a large database from this service. To get a detailed picture of current control room trends and quality, we analyzed measurements taken between 1997 and 2000 in 30 different countries. Focusing on 3-way systems we took a sample of 164 control rooms representing large recording studios, post-production houses, broadcast studios, private music composing studios and mobiles. All data analyzed were the speaker/room responses after the speakers had been calibrated and best possible results obtained.

As monitoring quality is becoming more and more critical, when multichannel materials become widely available and the number of channels increase, this study reveals that room acoustics still need improvement particularly in small rooms.

STANDARDS & RECOMMENDATIONS

Leading organizations has drawn plenty of excellent standards and recommendations on professional audio reproduction, but in the reality compromises often have to be made. The ITU, EBU, ISO, AES Technical Committee, German Surround Sound Forum and Japanese HDTV Surround Forum have all written extensively on the quality of monitoring conditions. Some of the important issues relating to our study are listed below.

- The magnitude response and its accepted deviation at the listening position.
- The dynamic range of the monitoring system, as well as the maximum room background noise level, the sound isolation between rooms, etc.
- The idealized speaker layout and positioning for stereo and multichannel configuration.
- The overall tolerance for the accepted reverberation time across the spectrum.
- The specifications on SPL and frequency response.

MEASUREMENT & ANALYSIS METHODS

All rooms were measured with a MLSSA measurement system that uses an MLS sequence (16’383 samples, 217ms long, 75kHz sampling rate). An impulse response was stored for each speaker, and the results presented here were extracted from these impulse responses.

The measurement procedure or apparatus did not change, and the monitor loudspeakers installed were all factory calibrated to have very precisely the same response. The only variable in our measurements is the monitoring room.

Signal-to-Noise Ratio and Listening Distance

Each impulse response contains the natural propagation delay before the impulse begins. This recorded section is an indication of the quality of the measurement (signal-to-noise ratio) and the measurement distance. The listening distance analysis only took into account monitors having their acoustical axis directed toward the listening position, and impulse response measurements recorded in a single position, without moving the microphone (250 measurements fulfilled this criteria). The distances estimated in this manner correspond to the precise actual listening distances.

Reverberation Time

The analysis of the reverberation time RT_{60} in small highly damped rooms is difficult because of sparse room modes and seldom a systematic decay as in large rooms. Having investigated various methods, a nonlinear fitting technique by Karjalainen et al. was used in both full-band and octave-band. High RT_{60} is not desirable in control rooms. Control rooms with very short RT_{60} were introduced in the 1960's, but most people find them uncomfortable and tiring when working long hours. With multichannel control rooms, the question of adequate RT_{60} opens again with direct energy and localization cues coming from all directions.

Room Operational Response Curve

A very interesting aspect of the monitoring conditions is the frequency response at the listening position. The 'room operational response curve' is defined as the third-octave smoothed magnitude response. The German Surround Sound Forum has proposed an acceptance window centered at the mean value calculated over the frequency range 50Hz-16kHz. If the amplitude spread of the distribution is small and stays within this window, it indicates that a control room is well designed and that monitor integration is excellent whatever the room type, volume and application. As the spread increases, the quality of the monitoring conditions worsens.

The level aligned magnitude values from all responses are collected. The median, 50% and 90% percentiles of this distribution are then extracted to demonstrate how frequently rooms comply with the acceptance window

Frequency Response Notches Analysis below 1kHz

Both comb filtering effects produced by boundary reflections and standing wave pressure minima are displayed as notches in the magnitude response. These notches can seriously deteriorate the monitoring conditions. A lot of essential music information is placed below 200Hz, and the loss of a specific frequency region there will destroy the musical tonal balance.

The notches were analyzed for frequencies between 40Hz and 1kHz (the search frequency band). To identify a notch, a detection level was increased until 10 deepest minimums were found. The center frequency and maximum depth of each notch was recorded.

Ideally we should not have any notches. This is the case in many well-designed rooms. However, others display serious notches at low frequencies.

Speaker Pair Comparisons

Many rooms have non-symmetrical layout, which leads to a different response from left and right speakers at listening position. Hence, the overall imaging and placement of sound sources is compromised and difficult to optimize. The stereo imaging can almost disappear or become unfocused. To assess the similarity of the magnitude responses of monitors in a room, we compared them, pair-wise, in three groups:

- The left-right stereo pairs in a room.
- The Left-Center-Right triplets compared to the Center channel (Front left-center-right triplets are found in multichannel rooms using typically two-way systems as surround speakers. As this study concentrates on three way systems, we did not include these surround speakers).
- In five-channel system, we compared the front monitors (left-center-right triplets) and the surround-left/surround-right pairs separately for clarity of results.

RESULTS OF ROOM-RELATED PARAMETERS

Signal-to-Noise Ratio & Listening Distance

The dynamic range in a measurement was typically 60dB (varying from less than 40dB to over 70dB). This measured dynamic range contains the noise contributions of both the measurement apparatus (insignificant here) and the room.

The listening distances for the studied three-way monitors range from 1.2 meters to 4.2 meters. The average listening distance is 2.49 meters (Fig. 1).

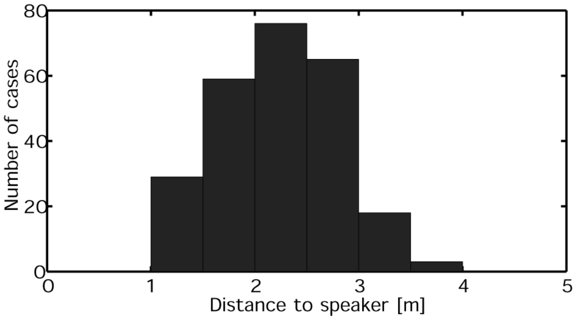

Fig. 1. The distribution of listening distances, all speakers including surround ($N = 250$).

Channel	*Stereo* R = 79	*L/C/R* R = 15	*5-channels* R = 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

Table 1. Mean listening distances of individual channels for left-right stereo rooms, left-center-right triplet rooms and five-channel surround rooms. Distances are in meters, and R is the number of rooms.

A typical listening distance of a three-way main monitor in a stereo configuration is 1.5–2.5 meters. This is shorter than listening distance for L/C/R-triplets and five-channel surround systems that are about 0.3 meters longer. The various standards for stereo and multichannel listening recommend distances between 2m to 3.4m for some, and 3m to 6m for others. Our measurements show that the practical listening distances are shorter than recommended.

Reverberation Time

Using the Nonlinear Fitting technique, the mean reverberation time RT_{60} is <u>380ms</u> from 200Hz to 4kHz (Fig. 2).

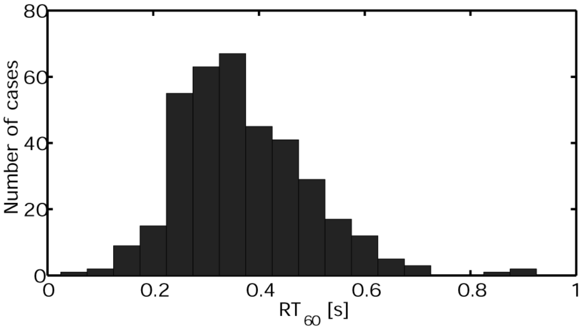

Fig. 2. Distribution of full bandwidth reverberation time RT_{60} ($N = 372$).

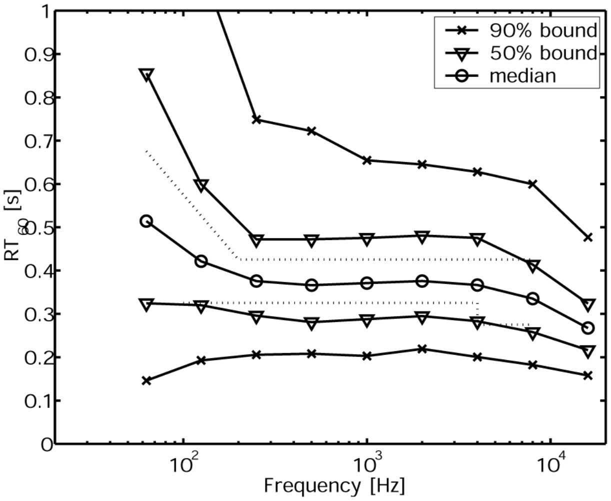

Fig 3. 90% and 50% bound limits and the median for reverberation time RT_{60} in octave bands ($N = 372$). German Surround Sound Forum limit shown (dotted line) centered at the median of mean RT_{60} levels.

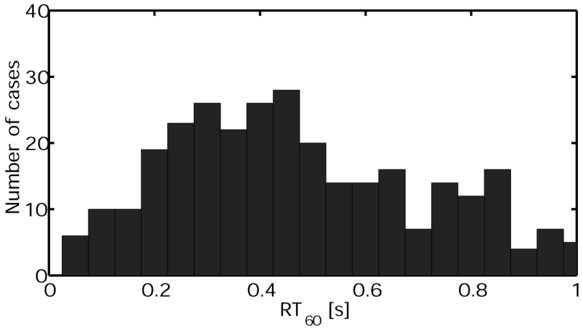

Fig. 4. Distribution of reverberation time RT_{60} of 63.5Hz octave band ($N = 372$).

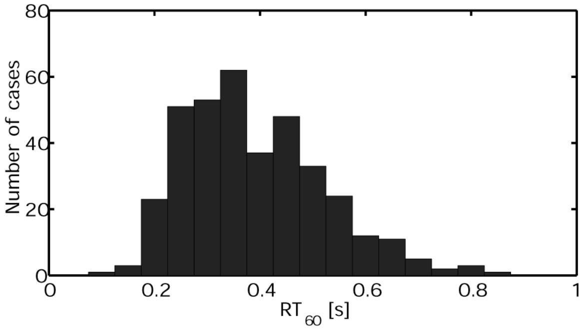

Fig. 5. Distribution of reverberation time RT_{60} of 1kHz octave band ($N = 372$).

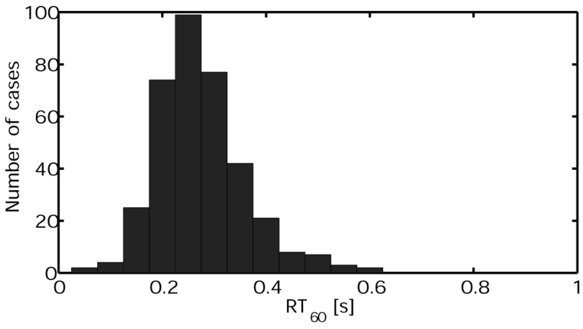

Fig. 6. Distribution of reverberation time RT_{60} of 16kHz octave band ($N = 372$).

Figures 4-6 display the RT_{60} for three extreme frequencies and show the difference in RT_{60} spread. As the frequency increases, the RT_{60} becomes more controlled in all rooms measured (164 rooms).

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 absorption or diffusion to control the reverberant decay field. Certain rooms show extremely long reverberation time at low frequencies (larger than 1sec), far beyond acceptable values for professional control room environment.

Magnitude Response

The overall listening position frequency balance for loudspeakers aimed at the listening position (250 speakers) is represented by third octave smoothed frequency response deviations (Fig. 7). To obtain these distributions, each frequency response is normalized to the mean level between 50Hz and 16kHz as proposed by the German Surround Sound Forum (SSF). The SSF proposed room operational curve limits, set relative to the 50Hz – 16kHz mean of the median of distributions, are also shown here. These suggested limits seem excellent guidelines to consider when judging the quality of a monitoring environment.

The 50% variation limit is within the proposed (SSF) limits for frequencies above 130Hz, and the 90% variations limit for frequencies above 400Hz. This was expected, since at low frequencies a speaker frequency response is dominated by the modal response of the room. Also, as quite a few rooms lack good acoustic design at low frequency, the spread of the 90% variation limit increases significantly below 300Hz. The presence of low frequency notches is displayed as a larger spread of the low frequency variation.

At higher frequencies, the variations are very well behaved, and within the proposed limits. However, only 5% of rooms show straight response up to 20kHz. The responses generally suffer loss of level above 16kHz, because most control rooms have excessive damping at high frequencies. This also demonstrates that it is important to very precisely aim speakers toward the listening position.

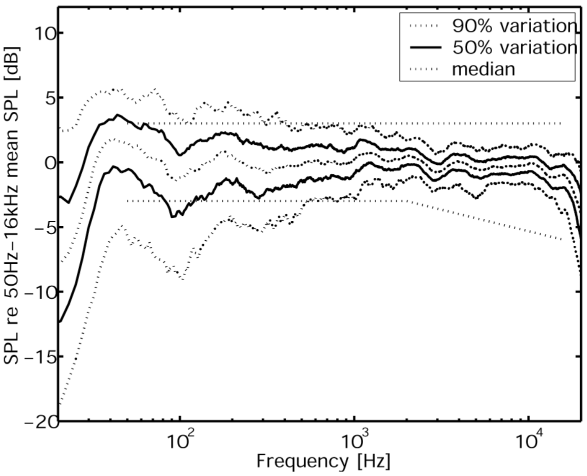

Fig. 7. Third octave smoothed sound pressure level measured at the listening position ($N = 250$) for speakers aimed at the listening position. 50Hz-16kHz mean level normalized to 0dB. Also shown are German Surround Sound Forum proposed limits (dotted line).

Frequency Response Notches below 1kHz

To demonstrate low frequency response notches, we studied measurements taken at the listening positions for those speakers having their acoustical axis directed toward this listening position.

Some reflections can be very strong, about the same level as the audio signal itself, hence the median notch depth is 14.2dB, but 30dB notches are not uncommon (Fig. 8). This is significant and the energy that is so totally missing will seriously alter the reproduction quality of the monitoring.

In our material the most typical notch frequency is 100Hz (Fig. 9). As this frequency is quite important in the music spectrum, it would be very important that room designers implement in-depth knowledge of the interaction between speakers and room in that frequency region.

Please note that our data contain information about the 10 deepest notches in each impulse response measurement, and not all notches. Some but not all notches up to -6dB level are included. Some rooms show deeper notches, and the shallower notches come from better rooms.

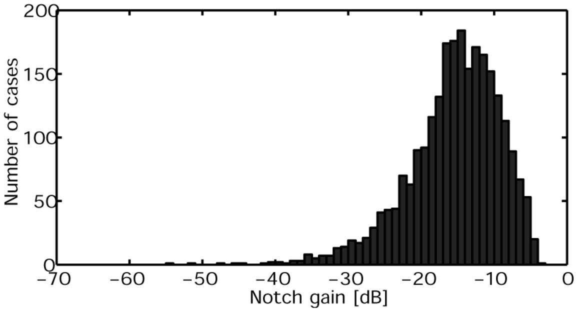

Fig. 8. Notch depth for the 10 deepest notches within a frequency band from 40Hz to 1kHz. Bin size is 1dB. Median notch depth is 14.2dB ($N = 250$).

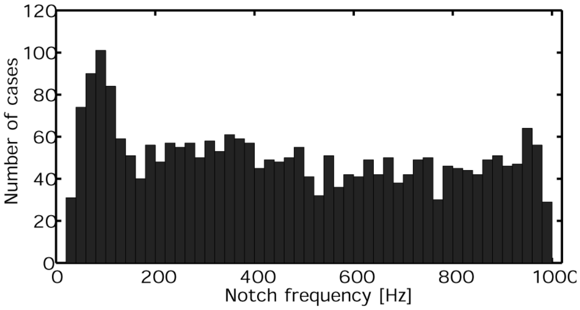

Fig. 9. Notch frequency distribution, bin size is 20Hz ($N = 250$).

RESULTS OF MAGNITUDE RESPONSE MATCH

To examine the magnitude response matching after calibration, we studied 105 rooms, out of the 164 rooms, that had speakers with acoustical axis aimed at the engineer's position. The pair-wise comparison included 243 speakers with 8 five-channel surround setups (40 speakers), 15 L/C/R triplets (45 speakers) and 79 L/R stereo pairs (158 speakers).

Left-Right Pairs

The level difference of left-right pairs in two-channel rooms (Fig. 10) shows a very good agreement, which confirms that most stereo rooms have a symmetrical layout, at least above 1kHz. Note that if there were a very strong and identical reflection interfering with the direct sound from both left and right speakers, the pair matching would display a close to 0dB level difference. In other words, the figure below does not represent the actual frequency response at listening position (for that see magnitude response).

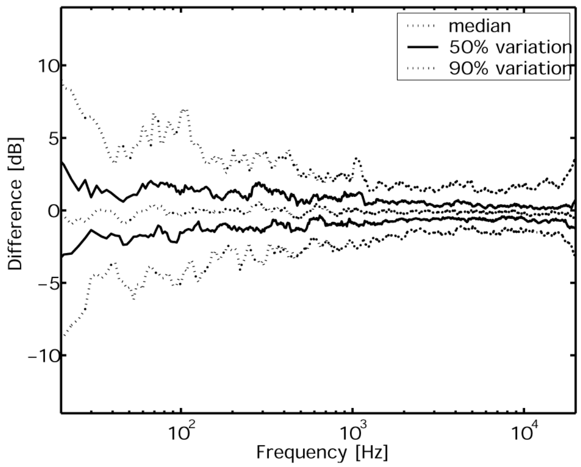

Fig. 10. Third octave smoothed sound pressure level difference of Left-Right pair in L/R stereo systems (*N* = 79).

The speakers used in this study have carefully controlled radiation characteristics at mid and high frequencies, minimizing room effects, and resulting in improved in-situ pair match. At low frequencies, where the room modal response dominates, the spread of the distribution increases.

In addition, we found 50% of systems show a mismatch of more than 2dB below 1kHz, which is likely to affect auditory imaging, localization ability and probably reduce the sharpness of stereo imaging. Once again, lower frequency modal responses are often poorly controlled in studios, and that will lead to mismatch in the above left-right pair analysis. Large object such as consoles, tables, large equipment racks, etc, will start to interfere with frequencies below 1kHz, where the mismatch becomes larger.

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 surround speakers.

The 90% variation for the L/C/R triplets increases below 400Hz and above 10kHz. The 50% variation is within a ±3dB window for frequencies above 1kHz (except for frequencies above 15kHz), and in a 6dB window below that frequency (shown here is the left-center pair match only).

Compared to stereo left-right pair match, L/C/R triplets produce a larger variation because the comparison is made with the center speaker. This is valid because in multichannel systems the center speaker is receiving an increasingly important role in forming the sound stage. However, in typical stereo installations the layout is fairly symmetrical for the left and right speakers, whereas the center speaker usually has differing radiating condition from the other two front speakers.

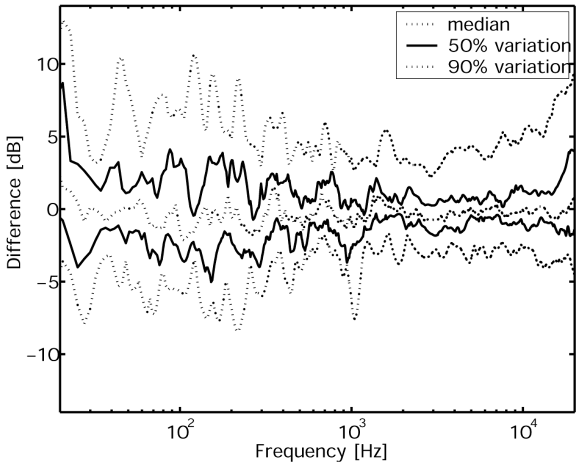

Fig. 11. Third octave smoothed sound pressure level difference of left-center pair in L/C/R systems (*N* = 15).

Five-Channel Surround Set-Ups

These use three-way speakers for all five audio channels, so control rooms are fairly large. A left-center-right (L/C/R) triplet in the five-channel reproduction system was calibrated to have a proper subjective balance during acoustical calibration. This was repeated on the paired surround-left and surround-right (SL/SR) speakers. The front channel match was studied separately and not compared with the surround channels.

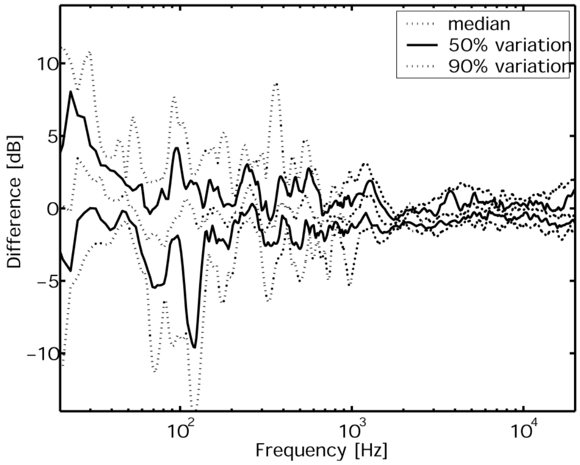

Fig. 12. Third octave smoothed sound pressure level difference of left-centre pair in five-channel surround systems.

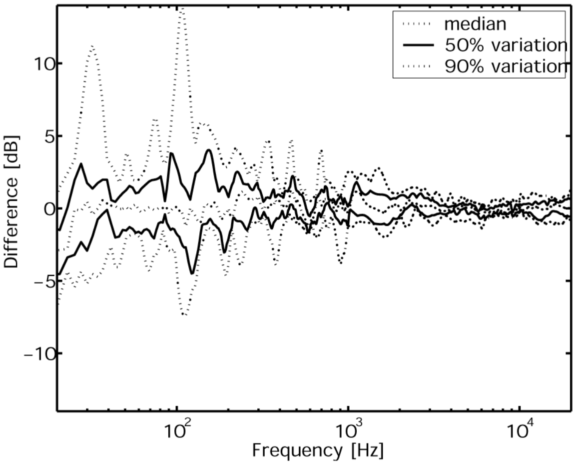

Fig. 13. Third octave smoothed sound pressure level difference of surround left-right pair in five-channel surround systems.

DISCUSSION

Loudspeaker Geometry

In many stereo control rooms, monitors are not focused on the listening position, so the engineer is not in the best place to monitor his/her work. Most standards recommend that speaker axis should be oriented toward the listening position. Some designers advise directing the acoustical axis behind the listening position, forcing monitoring with off-axis response. Even with modern constant directivity design approaches, monitor loudspeakers are typically optimized for on-axis response.

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

First, the widely accepted configuration for the 3/2-surround format places the Left and Right speakers at ±30^o and the surround speakers at ±100^o-120^o relative to the Center speaker. Though the three front speakers are almost always placed according to the suggested angles, the surround speakers are often placed beyond the 120^o from the center.

Second, all speakers are supposed to be on the same horizontal plane at the engineer's ear height (nominally 1.2 meters), or at maximum 10 degrees inclined above this level at an equal distance of 2…4m. Some recommendations allow the surround speakers to be inclined up to 15 degrees. The acoustical axis of the three-way speakers is rarely at the recommended 1.2m height, but typically elevated up to 15 degrees and pointing downwards. Most large rooms have monitors placed above the listening height, which brings the acoustical axis 0.5…1m above the listening level. This reduces low order floor reflections and is particularly relevant to flush-mounted speakers with large front baffle.

As for the surround speaker height, some standards recommend the same height and vertical tilt for surround and front speakers, while others recommend only surround speakers to be placed higher. In our survey surround speakers are frequently placed higher than front speakers. Very few rooms have five identical speakers in a surround setup. Surround speakers are typically not of the same type as front speakers.

When standards mention room size, the exact location of the speakers is not defined, although a minimum distance from neighboring walls is recommended and the importance of symmetric placement emphasized. The recommendation for speakers to be placed at least 1m away from walls appears too optimistic. In that case, if the wall behind the speaker has insufficient absorption, a quarter wavelength notch at 86Hz would be generated by the reflection off it. This would cause a severe irregularity in the free-standing system bass response. Increasing the distance would bring down the notch frequency. Together with the recommendation for the operational curve to extend down to 50Hz, a minimum distance would be 2m. This proves the value of flush mounting even for small two-way monitor systems as providing sufficient absorption to remove a wall reflection at low frequencies is difficult and expensive.

Speaker size or front baffle size is also not covered in the recommendations about the parameters affecting speaker placement, although it is recognized that large speakers may have to be placed high. The current recommendation of 1.2m height for the acoustical axis can lead to floor reflection notches in the 80Hz-120Hz frequency region, causing deterioration in the bass frequency response.

Reverberation Characteristics

The result of mean RT_{60} of 380ms is interesting. However, that single value does not reveal the increase in reverberation time toward low frequencies caused by insufficient bass trapping and improper room geometry. Also, some “professional” rooms with almost no acoustic treatment exhibited near constant reverberation time over the whole frequency band. Even if recommendations allow for larger RT_{60} below 200Hz, too many control rooms obtain too high RT_{60} in that region. This indicates that there is still a great need to improve the quality of monitoring room design.

Notches below 1kHz

Most of the notches found in our study are produced by low order reflections. Room modal response begins to dominate below 200Hz where the wavelength becomes large compared to size of objects in the room.

The reasons for notches in the frequency response are typically –

- First order floor reflections, often accentuated because of the placement of woofers relative to the floor.
- Interaction of the console with the first order floor reflection.
- Incorrect front wall design. For example discontinuities in the front wall, such as non-flushed windows, large TV screens or cavities.
- Insufficient front wall bass trapping in soft front wall construction, leading to reflections from the hard wall behind the speaker. This creates an LF notch at the frequency corresponding to the quarter-wavelength distance from the speaker’s front baffle to the hard (front) wall behind it. This distance can be larger than the depth of the speaker cabinet.
- With free-standing monitors, reflection as above from the (front) wall behind the monitor.
- Sustained standing waves. The notching can be severe and lead to irregularities in the response at the mix position.

Most of these issues could be solved with both better understanding of these phenomena and relevant solutions applied to the control room design.

Magnitude Response

The magnitude response of an individual monitor loudspeaker should be flat to within ±2dB in anechoic conditions using third octave smoothing. All monitors 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 [AES, EBU]. The 50% distribution limits in stereo rooms remain consistently within ±3dB bounds also at low frequencies. This is clearly the demonstration of a fairly good speaker/room interaction.

In multichannel rooms, the 50% distribution limits indicate a good frequency response control above 1kHz but an increasing distribution below this frequency. There are still frequent failures in low frequency design of monitoring spaces and the management of low order (early) reflections. In multichannel rooms the number of omni-directional low frequency radiators seriously complicate the room design at low frequency. The constraints on speaker placement and angles severely restrict the design of room geometry.

Most large modern control rooms can achieve adequate low frequency damping with properly designed acoustic treatment providing very high quality monitoring spaces. Small control rooms with free-standing monitor systems, compromised acoustical treatment and strong modal coloration do exhibit large variations at 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 [AES]. Monitor speakers included in the present study fulfill this requirement.

The pair match of the stereo pairs and surround left-right pairs is typically better than that of L/C/R triplets. One reason is 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 designs. It is difficult to design similar radiating conditions for the Center speaker in the middle axis of the room. Many installations have large computer screens, racks, etc, placed centrally near the engineer's position. These objects create strong reflections causing comb filtering at mid and high frequencies.

In the case of L/C/R triplets (with 2-way speakers as surround) we typically have small 5.1 rooms with short listening distances and compromised layout and equipment positioning. The compromise of the space and volume is clearly visible in the pair-match results.

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.

New approaches to multichannel monitoring room design are needed to produce environments 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.

The mismatch in tonal balance and spectrum between the various speakers in multichannel rooms seems an increasingly important issue. Each engineer and room designer should pay utmost attention to this so that the final material translates well outside the production rooms. As consumer’s reproduction systems are rapidly reaching a very good quality, professionals must continue to strive for the best in all aspects.