Studies on Main and Room Microphone Optimization

Helmut Wittek, Institut für Rundfunktechnik GmbH, Munich, Germany; (A)
Oliver Neumann and Markus Schäffler, Fachhochschule für Druck und Medien, Stuttgart, Germany; (B)
Céline Millet, Conservatoire de Paris, Cité de la Musique, France (C)

Theoretical and practical investigations in the optimization of microphone setup for 3/2-Stereo are presented.
The “Image Assistant”, an Internet-application for calculating and evaluating an arbitrary L-(C)-R microphone setup is introduced. L-C-R main microphone configurations were compared in listening tests revealing differences in directional imaging, sound color and focus.
Additionally, different room microphone configurations were compared regarding spaciousness and envelopment. Excerpts from the test objects are presented to demonstrate the existing differences.

A. DIRECTIONAL IMAGING USING “L-C-R STEREO MICROPHONES”

The recording of a sound ensemble in the 3/2-Stereo-Standard requires tools to optimize different basic aims of the sound engineer like e.g. :

- directional imaging between the loudspeakers L, C and R
- creation of room impression, spaciousness and envelopment
- creation of a suitable sound color

A common tool to optimize directional imaging is the use of a main microphone. Similar to 2/0-Stereo it is desired to enable balanced localization between the frontal loudspeakers. Adding the Center loudspeaker in a substantial use can produce stability and improve the representation of sound color and the focus of the phantom sources. Therefore a so-called “L-C-R stereo microphone” can be used to enable stable localization between the frontal loudspeaker bases. However, the design of a L-C-R stereo microphone has to pay attention to these essential points:

- enable balanced localization, i.e. produce a linear localization curve
- avoid colouring through interchannel crosstalk (same signal at all front loudspeakers)

The localization curve is the control indicator of the distribution of phantom sources produced by a stereo microphone. It describes the relationship between the angle of the input source in the recording room and the angle of the phantom source in the reproduction room. The relation should in general be linear, that means the situation in the recording room is reproduced proportional between the front loudspeakers.

A helpful tool to design microphones taking care of these characteristics is an application called “Image Assistant”.

A1. “Image Assistant”

This application which can be used online on the website www.hauptmikrofon.de [1] calculates localization curves. Its input parameters are:

- loudspeaker configuration L/R or L/C/R
- directional pattern of the microphones
- angular offset α of the microphones
- distances b and h of the microphones
- possibly additional input like level or time correction

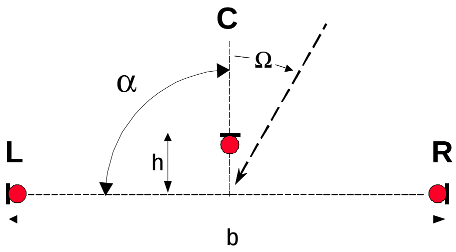

Fig. 1: Input parameters of a L-C-R stereo setup

The output of the Applet is the localization curve, resulting from the calculated signal differences at the microphones (= interchannel differences at the loudspeakers).

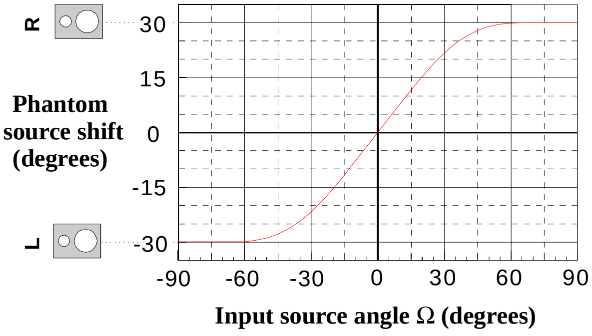

Fig. 2: Localization curve as calculated by the “Image Assistant” (example: ab, 2 omnis, b=40 cm)

The calculation is based on well known laws that describe the shift of a phantom source depending on the existing interchannel differences.

(e.g. ΔL = 6.5 dB or 0.4 ms lead to 50% phantom source shift, 18 dB or 1.0 ms lead to 100% shift, a 100% shift corresponds to half the loudspeaker distance)

In preliminary tests it was shown that these laws are valid both for the Standard-Stereo base L-R and the small loudspeaker bases L-C and C-R.

With the use of this data an approximation function can be built which allows the computation of the microphone setup:

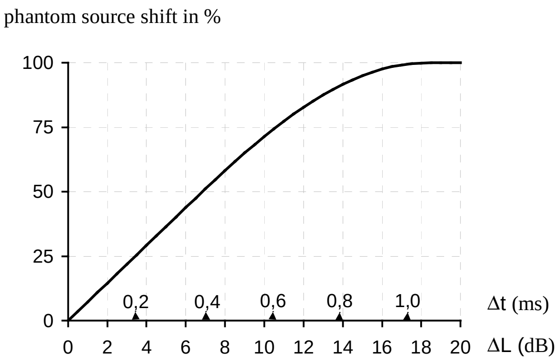

Fig. 3: Approximation function to calculate the phantom source shift depending on interchannel differences

The results of the computation can be used to optimize the microphone setup of two and three-channel configurations and to discuss their characteristics of directional imaging.

Two other examples for calculated localization curves:

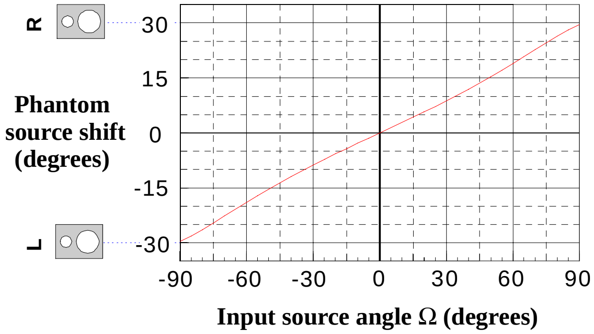

Fig. 4: XY with Cardioids, +/- 45° angular offset

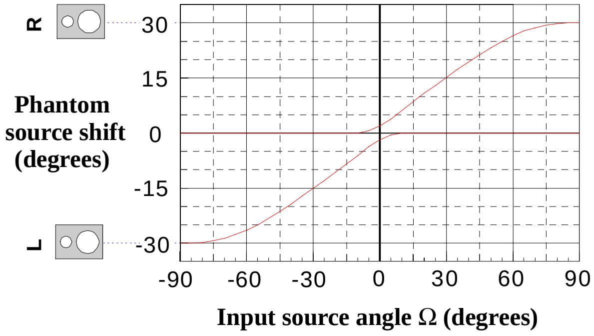

Fig. 5: OCT system, b=50 cm, OCT setup see Table1

A2. Listening test 1: directional imaging, optimal listening position

Listening tests were performed to investigate the directional imaging of stereo microphones. (see [2] and [3])

The test signals, derived from recordings through the participating stereo microphone configurations, were reproduced through the front loudspeakers of a 3/2-Stereo-configuration. The test participants were first placed in the optimal listening position.

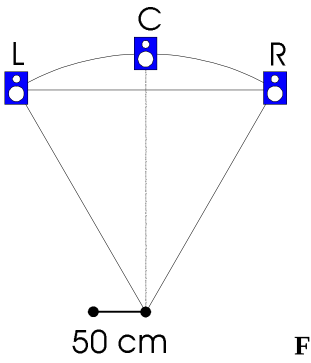

Fig. 6: Test setup

Three different microphone configurations were considered in the listening tests:

	OCT	INA 3	Quasi-ORTF
Mic L	Super Card.-90°	Cardioid, -60°	Cardioid, -30°
Mic C	Cardioid, 0°	Cardioid, 0°	-
Mic R	Super Card.-90°	Cardioid, 60°	Cardioid, 30°
Distance b	70 cm	92 cm	20 cm
Distance h	8 cm	26 cm	-
Rec. angle	120°	120°	120°
References	[4]	[5]	-

Table 1: participating Stereo setups, picture: see Fig.1

Figure 7 shows both computed and experimental data of the Quasi-ORTF stereo microphone.

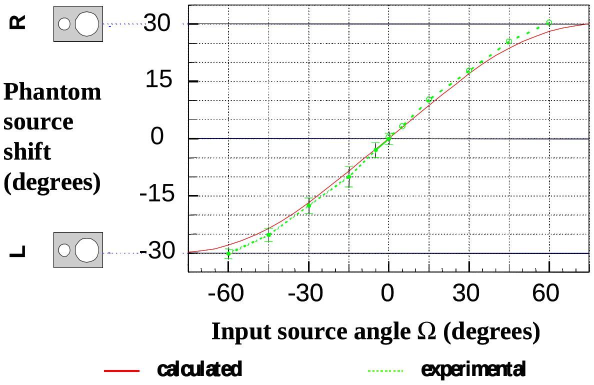

Fig. 7: Localization curves for Quasi-ORTF setup

The calculated prediction fits very well to the experimental curve. It can be seen that this stereo microphone leads to a very linear distribution of the phantom sources, which is desired for the three-channel microphones as well.

Figure 8 shows the data for the L-C-R setup INA 3:

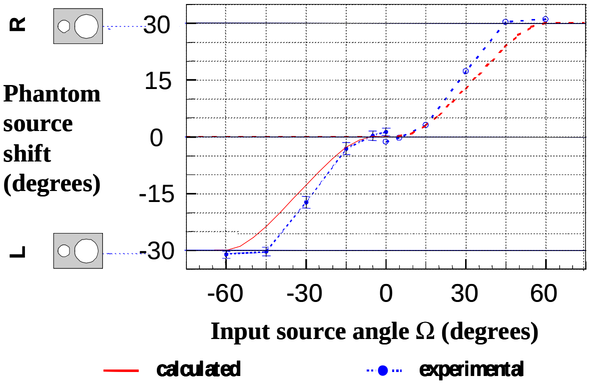

Fig. 8: Localization curves for INA 3 setup

Both the calculated and the experimental data show that the optimum of linear imaging is not obtained with this kind of three-channel microphone. This configuration leads to an agglomeration of phantom sources in the area of the Center channel.

Figure 9 shows the resulting curves for the OCT microphone setup:

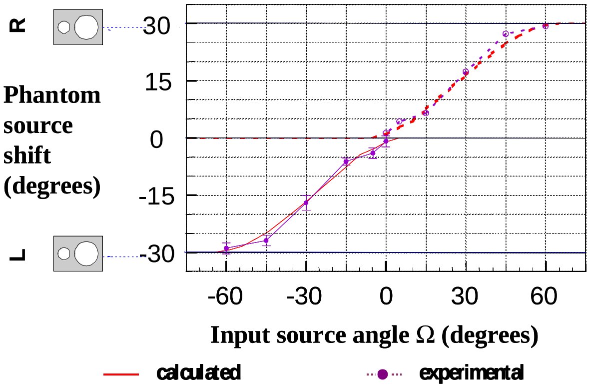

Fig. 9: Localization curves for OCT 70 setup

The imaging of this L-C-R microphone is quite linear. The directional image shows great similarity with the desired Quasi-ORTF image. The computed prediction matches well with the experimental data.

A3. Listening test 2: directional imaging, non-optimal listening position

Now the test participants were placed 50 cm left of the optimal listening position (see Fig. 6). The results are shown in Fig. 10:

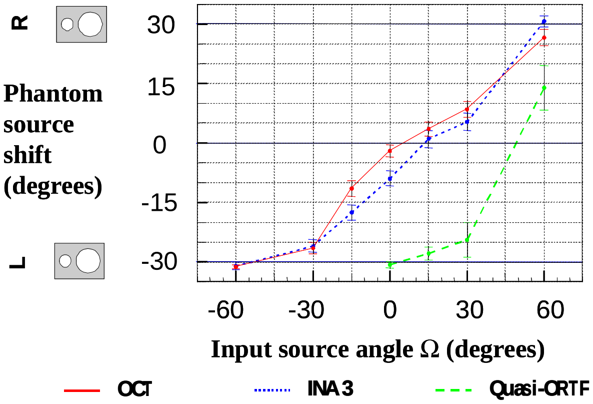

Fig. 10: Experimental localization curves for the non-optimal listening position

It is evident that the two-channel stereophonic representation cannot provide any more a satisfying distribution of phantom sources between the loudspeakers. The directional image of multichannel techniques that do not make substantial use of the Center channel will look similar.

The image of the L-C-R configuration INA 3 looks better. However, as a consequence of the relatively high level of the left channel also at desired representations within the Center area (interchannel crosstalk), the stereophonic image is moved towards the left loudspeaker.

The OCT representation is most close to a linear representation. Due to the high crosstalk damping, caused by the intelligent setup with Super cardioid microphones, the Center area imaging stays constant.

A4. Listening test 3: Focus and sound color of the phantom source

Now, the resulting phantom sources are compared with a reference, being a single loudspeaker positioned in the same direction as the phantom source. It was fed with a signal of one of the microphones and adjusted to the same level as the test signals. The test participants were asked to estimate the difference in the respective characteristic. Therefore a five-grade scale was offered:

- 1: not perceptible
- 2: slightly perceptible
- 3: perceptible
- 4: clearly perceptibly
- 5: very clearly perceptible

The results, shown in Figures 11-14, make clear that the OCT setup can improve sound quality through avoiding interchannel crosstalk.

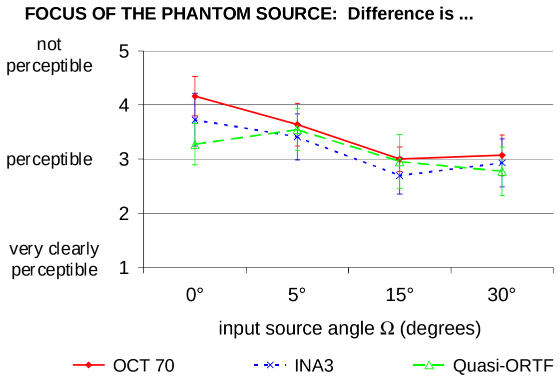

Fig. 11: OPTIMAL listening position: Focus

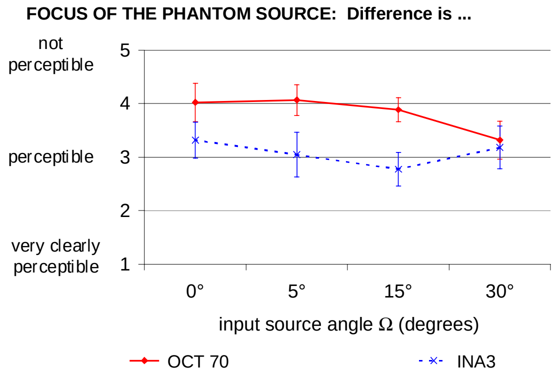

Fig. 12: NON-OPTIMAL listening position: Focus

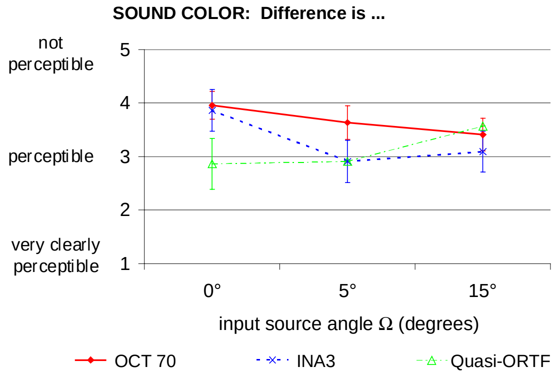

Fig. 13: OPTIMAL listening position: Sound Color

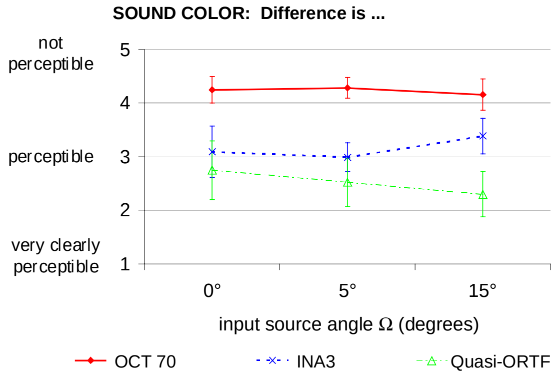

Fig. 14:
NON-OPTIMAL listening position: Sound Color

B. OPTIMIZING THE ROOM MICROPHONE CONFIGURATION

Listening tests were performed to optimize the room microphone configuration. ([6])
Theile [4] suggests among other possible solutions e.g. square arrangements shown in Fig. 16 to enable the appropriate imaging of side reflections. They have proven to be important for the perception of apparent source width (ASW), the distance of the source and spatial depth. In the test these square configurations were added to the L-C-R microphone OCT ([4]), responsible for the frontal directional imaging.

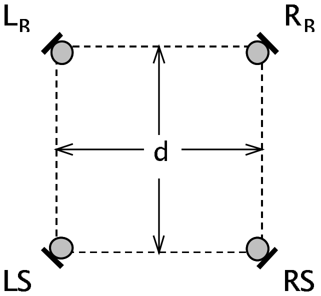

Fig 15: 4-channel square room microphone

The microphones L_R and R_R are discretely routed to the channels L and R, the microphones LS and RS to the channels LS and RS. So each pair of microphones L_R-R_R, R_R-RS, RS-LS and LS-L_R provides a stereophonic representation of early reflections and reverberation.
The degree of coherence of these pairs of microphones is connected directly to the perception of the reproduced sound field. It influences the perception of spatial depth, ASW, spatial impression and envelopment.
The recordings were made to investigate the effect of the interchannel coherence on the spatial impression.

B1. Test Setup

The OCT system was placed close to the musicians. At a distance of 8,50 m to the OCT a total of six room microphone square arrangements with different side lengths were set up. Three of them were built up with cardioid microphones and the other three with omnidirectional microphones. The side lengths of the squares were:

- cardioid microphones:
  d_1 = 15 cm; d_2 = 25 cm; d_3 = 300 cm
- omnidirectional microphones:
  d_1 = 40 cm; d_2 = 215 cm; d_3 = 500 cm

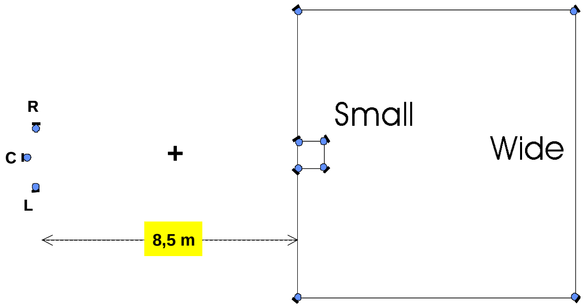

Fig. 16: Room setup for test signal recording

The gain of all room microphones was set equal to the gain of the microphones of the OCT. At a distance of 1 m behind the OCT a double layered molleton curtain was set up to prevent indirect sound from hitting the OCT.
The five (L-C-R plus two low-passed-omnis, see [4]) signals of the OCT and the four signals of each room microphone square arrangement had to be mixed down to the channels L, C, R, LS, RS.
It was found that in superior concert halls the first reflections followed the direct sound by about 20 ms. It is important that this arrival time gap doesn't get too long for the avoidance of echo effects.
In this recording the onset of the indirect sound was about 24 ms after the direct sound because the room microphones had a distance of 8,50 m to the OCT. With very percussive music there might have appeared echo effects. That is why the OCT was delayed by 10 ms. So the arrival time gap was 14 ms.

The levels of the room microphones were determined by listening so that the rear loudspeakers could not be localized. Because of their directivity cardioid microphones give less level in the diffuse sound field than omnidirectional microphones. That is why they had to be mixed with 4 dB higher level.

B2. Listening test results

2 different music examples were used for the tests:

- a slow excerpt of a string quintet
- a percussion trio

These results were found:
Though the omnidirectional microphones had a stronger low frequency response than the cardioid microphones the spatial impression created by the different microphones was quite similar.
However, these special observations were made listening to the wide square arrangements (215 cm, 300 cm, 500 cm):

- string quintet: Auditory "clouds" appeared in the regions of the loudspeakers. The spatial impression was not homogeneous. No ASW was observed.
- percussion trio: The ensemble sounded very close. The spatial impression was weak.

These observations were made listening to the small square arrangements (15 cm, 25 cm, 40 cm):

- string quintet: The spatial impression was very homogeneous and "natural" and resistant against movements of the listener. A strong ASW appeared.
- percussion trio: The ensemble sounded further away from the listener than with the wide arrangements. A real impression of space, distance and depth could be perceived which was also very stable.

An explanation for these observations could be that the small square arrangements produce a natural reflection pattern. ASW and perception of depth are closely related to the presence of early lateral reflections. Since these perceptions only occurred with the small square arrangements one may conclude that the wide arrangements did not produce a natural reflection pattern. Only by an adequate coherence of the signals the lateral stereophonic areas are exploited for reproduction of reflection patterns. If the signals are incoherent reflections will only be perceived from the directions of the loudspeakers.

C. SUPPLEMENTARY STUDY

The same multitrack recording was used in a further study ([7]). Two purposes were intended:

1. Assessment of the Binaural Room Scanning system (BRS [8]) with respect to the reproduction of spatial attributes of the stereophonic image.
2. Verification of results described above, using other reproduction device and different mixes of front and surround signals

This report summarises the second part of the study.

C1. Listening test

An A-B comparison test was used as in Study B. The main differences are the chosen examples, the mixing choices and the questionnaire.

Microphone setups :

Four microphone setups are used (see chapter B1): Two cardioid and two omnidirectional configurations. The 25 and 300 cm cardioid squares, respectively named Cs and Cb, (s for small and b for big) and the 40 and 215 cm omnidirectional squares. The cross-correlation was about 0.6 for the small configurations and about 0.1 for the big ones.

Examples :

Three test signals with strong character differences are extracted :

1) whip clicks played at musicians positions of the string quintet
2) end of the percussion piece (marimba, shakers, bells and djembe), loud character
3) viola solo phrase of the Pink Floyd’s piece during which the others instruments of the string quintet are playing legato the harmonie, soft character

Astonishingly, the A-B comparison was too difficult to conduct on the first example. In theory, it is the signal which enables the easiest listening of the reverberation but for the little differences in the manner to use the whip disconcerted the listeners.

Mixing :

The distance between main and room microphones (about 8 m) gives the optimum natural delay between direct sound and first reflections but not the optimum reverberation level. For each musical example, the levels for the configurations are adjusted to get the same loudness. Afterwards, the OCT L-C-R microphone level is increased from $-\infty$ to be just perceptible. Relative to this level 6 dB gain was set, in order to stabilize the sound source positions.

As a result, equal level balance of the direct / indirect sound was achieved for each configuration.

This mixing is very different from the one presented in the section B. The main differences can be resumed as following :

- No delay was set on the OCT signals.
- No equalization was done for the room microphones.
- The level of the OCT in relation with the level of the room microphones is about 7 dB lower.
- The level of the omnidirectional setups is about 2 dB lower than the level of the cardioid setups.
- The levels of the small setups on the percussion example are 1.5 dB lower than the level of the large setups.

Questionnaire :

The subjects are asked to assess B in comparison to A with a discrete scale from -2 to +2 for the following parameters :

- Room volume (smaller – bigger)
- Envelopment (more frontal – more surrounding)
- Envelopment homogeneity (more “auditory clouds” around the loudspeakers – more homogeneous)
- Distance of the musicians (nearer – farer)
- Apparent source width (larger – thinner)
- Localization of the musicians (more instable, difficult – more stable, easier)

A subjective question is added to know about the preference of the subject and why (because of timbre, sound source image, envelopment or other reason). Indeed, not only the spatial impression but also the timbre was changing. Therefore, a configuration could be more surrounding and yet never used because of a bass frequencies surplus.

Compared pairs :

Only the pairs within only a parameter is different are assessed. It means the ones with correlation change : Cs–Cb and Os–Ob and the ones with microphone characteristic change : Cs-Os and Cb-Ob.

Listening panel :

18 subjects took part. The results of 8 persons were eliminated because of their incapacity to recognize the A-A pairs.

C2. Results

Results were obtained by calculating the mean value, the median value, the first and the third quartile.

Comparison results:

- The small cardioid setup (Cs) provides more volume and envelopment than the wide setup (Cb). An improvement of the homogeneity was noticed only for the string quintett. Contradictory results concerning the sound source image (distance and ASW) were obtained depending on the musical example.
- The small omnidirectional setup (Os) provides less room volume and envelopment but is more homogeneous than the wide setup (Ob). The musicians seem to be nearer and the ASW thinner.
- Contradictory results were obtained for the localization ability depending on the musical example.
- Os provides more room volume and envelopment than Cs, the musicians seemed to be nearer and took more volume (larger ASW) but to the detriment of the localization ability.
- Similar results were obtained for Ob and Cb.

Listener’s preference:

- Cs is slightly preferred than Cb for percussion and strongly for string quintet. Cb is never preferred for the envelopment but especially for its timbre. Cs is preferred for its envelopment but for its timbre too.
- All subjects prefer the setup they assess with more precise (thinner), farer and more stable instruments.
- The timbre preference appears clearly for Cs and Os.

These results prove the importance of the timbre in the final judgement of listeners. In situations when the envelopment differs a lot between the two setups, the more enveloping setup is preferred. The coherence of distance of musicians, width of sound source and room volume is important.

C3. Discussion

By listening only to the room microphone setups without OCT, the differences become clearly perceptible.

- The large setups, because of their decorrelation, create auditory events near to each loudspeaker with sometimes notches between them.
- The small setups, because of their correlation and their distance to the main microphones create a more homogeneous spatial impression around the listener.

But room microphones are designed to be used in association with a main L-C-R system which gives a good sound source image. In this study, the OCT was used. Its principal quality is a really precise localization accuracy.

This study has shown that the listener doesn’t notice easily that the spatial sound is strongly decorrelated in the total mix.

This study allows to notice that the timbre is strongly affected by the room microphone setup designed for spatial impression reproduction. In our case, Cb has a lack and Os a surplus of bass frequencies. This can easily corrected by equalization.

References

[1] Wittek, H. 2000,1: "Image Assistant", JAVA-Applet and documentation on website www.hauptmikrofon.de

[2] Wittek, H., Theile G. 2000,2: "Investigations on directional imaging using L-C-R stereo microphones", 21.Tonmeistertagung 2000 (German), Proceedings ISBN 3-598-20362-4, p. 432–454

[3] Wittek, H. 2000,3: "Studies on directional imaging by means of L-C-R main microphones", Diplomarbeit IRT/ FH Düsseldorf, 9/2000

[4] Theile, G. 2000: "Multichannel Natural Music Recording Based On Pyschoacoustic Principles", AES Convention, 2000, Preprint 5156, supplementing extended version: www.irt.de/IRT/indexpubli.htm

[5] Herrmann, U., Henkels, V., Braun, D.: „Comparison of 5 surround microphone methods" (German), 20.Tonmeistertagung, 1998, Proceedings ISBN 3-598-20361-6, p. 508-517

[6] Neumann, O., Schäffler, M. 2001: "Optimizing the room microphone configuration for 3/2-stereo multichannel music recordings", Diplomarbeit, FH Stuttgart – Hochschule für Druck und Medien

[7] Millet, Céline 2001: „Studies on the performance of the BRS system concerning soundfield assessment", Diploma work, Conservatoire de Paris, Cité de la Musique, France, 2001

[8] Karamustafaoglu, A.; Horbach, U.; Pellegrini, R.; Mackensen, P.; Theile, G., 1999: "Design and application of a data-based auralisation system for surround sound", 106th AES Convention, 1999, Preprint 4976