Preprint 2886 (A2)

Room-Related Balancing Technique
A Method for Optimizing Recording Quality

Martin Wöhr, Günther Theile, Hans-Jürgen Goeres, Alexander Persterer

Bayerische Rundfunk
Munich, West Germany

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

AN AUDIO ENGINEERING SOCIETY PREPRINT

ROOM-RELATED BALANCING TECHNIQUE

A METHOD FOR OPTIMIZING RECORDING QUALITY

by

Martin Wöhr, Günther Theile, Hans-Jürgen Goeres, and Alexander Persterer

ABSTRACT

In contrast to natural sounds, in the case of stereophonic reproduction the ears localize the loudspeakers as the source of the sound. The spatial depth of a recording cannot therefore be reproduced adequately with conventional microphone balancing techniques. This paper describes how to simulate spatial depth artificially. Listening tests demonstrated that stereophonic reflections clearly improve the quality of a sound image, making it more natural. It was shown that the favourable imaging characteristics found for an appropriate main microphone can be transferred consistently to the spot-microphone signals with the aid of modern computer technology.

1. INTRODUCTION

The stereophonic reproduction of a piece of music should aim to be as natural as possible. Besides the artistic content and the quality of performance, essential ingredients of a good recording are transparency, presence, and fidelity of tone colour of the various instruments, as well as localization accuracy and directional stability of the individual sound sources. Equally, the acoustic characteristics of the room where the recording is made should be taken into account to an appropriate extent in the loudspeaker reproduction.

With the improved technical quality of electro-acoustic reproduction equipment, it has become possible to hear further and further into the sound image. The spatial aspect of a recording gains more and more importance in the process.

In order to create a spatial impression, not only the direct sound but also suitable reflections must be reproduced (e.g. /1,2,3/). If such spatial information is missing from the acoustical signals recorded, it should be made possible to generate these components artificially, so that the ears can recognize both the position and the nature of a sound when it is played back /4/.

A series of tests was carried out to check the extent to which theoretical considerations on 2-channel stereo spatial imaging /5,6,7/ can be implemented in practice. They were based on intensive preoccupation with the psycho-acoustics of binaural hearing, especially a knowledge of the effect that direct sound, early reflections, and reverberation have on the acoustic impression. So as to be able to investigate the important criteria for the human ear, it was necessary to be able to reproduce the natural conditions artificially, to make deliberate use of them, and if necessary to alter them. In this paper we show that spot-micro-phone signals must not disturb the favourable imaging characteristics found for an appropriate main microphone /6,8/. We succeeded in enhancing the loudness of an instrument that is too quiet, without having to put up with the usual loss in depth. First of all, a few remarks on some basic considerations.

2. SPATIAL PERCEPTION

In binaural hearing, the localization of a sound source is determined by directional and distance phenomena.

The perception of direction is based mainly on the time difference between the sound arriving at the left and right ear, but also on the differences in intensity in the case of an off-centre sound source. These differences in sound level arise from the shadow effect of the head, and are strongly dependent on frequency as well as direction.

Whereas directional hearing works only with two ears, significant characteristics for perceiving distance exist even when one hears with only one ear. Experiments have shown that the impression of distance is strongly influenced by acquired listening experience /3,9,10/. For example, differing distances from a sound source lead to a change in intensity; the brain can deduce a difference in distance from the relation between intensity and sound character in the case of familiar sound sources.

Besides the direct sound, significant components for localizing a source of sound in a closed room are the time and frequency dependence of the early reflections. These are influenced by the size of the room, the materials and shape of the bounding surfaces, and the directional pattern of the various instruments. It is fairly easy to estimate the distance in closed rooms, but in the open there are considerable errors because we lack the necessary long-term experience - after all, we spend most of our lives indoors.

These physical and psycho-acoustical facts mean that a listener in a concert hall can very quickly recognize the character of the room and the arrangement of the sound sources without opening his eyes. There is a basic difference with loudspeaker reproduction: here the perceived distance is equal to the distance from the loudspeakers. The ears localize the speakers as the sound sources, and different distances can be represented purely in the"simulation plane" /6, 7/ between them. Spatial depth in the 2-channel stereo sound image can be simulated only with differing elements of form /6/:

- The first wave-front contains depth information, since the ears associate an intensity/timbre relationship with the signal level/spectrum relationship, based on past experience, and thus deduce the distance to the source of sound.
- The relationship between direct and indirect components of the sound field contains important distance information which the ears interpret on the basis of past experience; this provides the sound image with spatial perspective.
- The early reflections and the reverberation produce particular interaural signal differences.

These elements should be present to an appropriate extent in a stereo loudspeaker signal /6/.

3. MAIN MICRIPHONE AND SPOT MICROPHONES

An extended sound source, for example a symphony orchestra, is usually recorded using a stereo main microphone and several mono spot microphones.

The main microphone should be capable of representing the dynamics and timbre of the sound source, as well as its direction and distance, as faithfully as possible for the human ear. The quality and positioning of the main microphone provide a significant proportion of the overall sound image of a recording.

In 1986 the ARD (German broadcasting organizations) , the IRT (Institut für Rundfunktechnik) and the VDT (Verband Deutscher Tonmeister) carried out joint experiments to investigate whether the performance of a main microphone can be adequately described for varying types of application /8/. The listening tests showed, for example, that in the matter of spatial impression there was a clear preference for main-microphone techniques that produced differences in both sound level and propagation times (ORTF, OSS, dummy head). Main microphones which supplied only differences in level (intensity stereo , XY,.MS) turned out to be less good at spatial representation; techniques which took account purely of propagation-time differences (AB) have a reduced directional stability. There was thus a subjective preference for techniques which provide signals whose degree of correlation corresponds approximately to that of the natural ear signal in the concert hall. Figure 1 shows the impulse pattern of an equivalence microphone /5/ (e.g. ORTF, OSS) subjected to sound incident from half-left.

The time displacement and differing signal level between the left and the right channel in the direct sound lead jointly to a directional image. The time sequence of direct signal and early reflections provide the listener - in the case of both natural sounds and stereo loudspeaker reproduction - with characteristic clues that the brain uses to form a picture of the nature of a sound field. The later the first reflections occur after the direct signal, the further the reflecting walls must be from the sound source. The time elapsing between the arrival of the direct signal and the early reflections conveys an impression of the size of the room, amongst other things. Subsequent reflections follow at ever closer intervals; they merge into reverberation.

The spot microphone has the task of making up for deficiencies in the sound image supplied by the main microphone. It is best placed near the instruments to be balanced. The resulting signals thus contain a large amount of direct sound and little spatial sound; that is to say, no information on the spatial arrangement of the sound sources in the sense of direction and distance. The appropriate directions are assigned when the signals are summed using panorama potentiometers (panpots). The gain in loudness is achieved in a mixing process which is called intensity balancing technique.

4. PROBLEMS WITH THE INTENSITY BALANCING TECHNIQUE

A simple mixing of the signals from the main microphone and spot microphone destroys the reflection pattern of the original recording room, as imaged by the main microphone.

The signal picked up by a spot microphone is reproduced by the loudspeaker earlier than the corresponding main-microphone signal. Thus, the ear interprets the sound reaching the spot microphone as the direct sound /5, 11/. The signal from the main microphone is then erroneously recognized as the first reflection, and subsequent reflections follow too late. The result is that the natural reflection pattern of the concert hall is suddenly distorted. The correct imaging characteristics of a good main microphone are lost. In consequence, such recordings sound too close and "flat".

The negative effect of intensity balancing, the "issuing from the loudspeaker" effect, is frequently moderated by artificial reverberation. However, the association with the original recording room is then fuzzy. Another possibility is to reduce the level of the balancing signal to such an extent that the signal is perceived only slightly. Then the advantage of the balancing, the desirable increase in the loudness of particular sounds, is not achieved any more.

4.1 Intensity balancing with propagation time compensation

An improvement in retaining the good characteristics of a concert hall and at the same time achieving sufficient presence is possible using the following approach

/12/. The mixing of main microphone and spot microphone is time-matched; i.e. the propagation-time difference between the two microphones is compensated. However, the desirable situation could not be achieved satisfactorily: in practice this compensation technique leads to "notching" effects. These are particularly disturbing when the musicians move about near the spot microphone.

4.2 Room-related balancing technique

To avoid notching errors, and at the same time achieve a high balancing gain, the spot-microphone signal must be delayed much more than necessary for the compensation as described in paragraph 4.1. It should fall within the region of the early reflections, and is thus less damaging to the overall effect /5/. The sequence of sound impulses, as shown in Figure 2, is as follows:

1. direct sound from the main microphone with appropriate directional and distance information;
2. first reflections, picked up by the main microphone;
3. "balancing reflections" to increase the volume (balancing gain);
4. further reflections and reverberation, picked up by the main microphone.

The favourable imaging characteristics of a main microphone, as regards spatial impression, are scarcely altered by the room-related balancing technique. Recordings made with the aid of careful room-related balancing of the main signal and a sensible amount of delayed reverberation show a definite improve-ment in quality, in the sense of a natural stereo sound image.

5. OPTIMIZING THE ROOM-RELATED BALANCING TECHNIQUE

An improvement in present-day stereo recording techniques is conceivable if the balancing signal does not simply consist of individual coherent first reflections, according to Fig. 2, but exhibits the same stereo quality as the natural reflections picked up by the main microphone. As an example, Fig. 3 shows two artificial balancing reflections. Their stereophonic represen-tation is generated from a monophonic spot-microphone signal by using a digital audio processor.

6. LISTENING TESTS

6.1 Hardware

The technological progress in special-purpose microprocessors for digital signal processing opens up the possibility of putting this idea to practical use. Bavarian Radio and the IRT carried out a joint series of tests on stereophonic balancing reflections in 1988. They made use of the CAP 340 M audio processor (AKG) /13/ with appropriately modified software.

The core of the system is a freely programmable processor with signal paths that can be configured at will. The design permits various types of processing such as recursive and non-recursive filtering, delays, reverberation and summing. The performance is 340 Mflops maximum; the entire signal processing is carried out in 32-bit floating-point format. The system is controlled by a host computer via a host interface in two stages:

- setting up the desired system structure,
  i.e. configuring the signal paths;
- passing the variable parameters,
  such as level or transfer functions, to the current program

Figure 4 shows one possible configuration.

The input signal is delayed so as to implement the time delays of the balancing reflections, and then convoluted with the transfer functions of the main micro-phone, in accordance with the direction of the reflections to be simulated. Finally the two left-hand and the two right-hand signals are added, and appear at outputs 1 and 2. A user shell was implemented at the monitor screen of the host computer, in order to carry out tests on the room-related balancing technique. The symbols for the eight channels of the panpotless mixer can be seen in Figure 5.

The uppermost line specifies which of the available analogue inputs is connected to which channel. Azimuth and elevation can be set in the second and third lines, the delay time in the fourth, while the fifth line represents the level control. The next line is for programming the channel mode: the options are binaural processing, conventional panpot (Δ L), and thirdly a panpot that generates time differences (Δ t). A channel name can be entered in the lowest line. The individual elements are activated by means of a mouse.

When the extensive listening material was being prepared, a program package was developed which uses more rationally the possibilities offered by the system, and permits direct access to the objective itself. A plan view of the recording configuration is displayed on the screen (Figure 6).

The individual elements (main microphone, spot microphones, and reflecting surfaces) are positioned with the aid of the mouse; absorption factors can also be entered. The program takes the geometrical data and by using a mirror-image method automatically calculates the necessary parameters for the acoustic configuration (direction of incidence of the sound, propagation time, signal level); it controls the processor appro-priately, including generating a suitable reverberation.

The equipment was adapted to interface with a conventional studio mixer via insert jacks.

6.2 Method

The experimental phase was divided into several sub-tests and one main test. The various sound material was recorded in the Philharmonic Hall in Munich (Gasteig cultural centre), and subsequently mixed using a multi-channel digital recorder in accordance with the requirements of each test.

The following balancing techniques were compared and assessed in various tests:

(1)	Intensity balancing	simple panpot summation of main microphone and spot microphone
(2)	Delayed intensity balancing	as (1), with propagation time compensation
(3)	Binaural balancing	room-related delay plus convolution of the spot-microphone signal with the dummy-head spectral transfer function
(4)	ORTF balancing	room-related delay plus convolution of the spot-microphone signal with the transfer function of an ORTF microphone
(5)	ORTF balancing with distance equalization	as (4), taking account of the change in spectrum by absorption at the room boundaries (e.g. high-frequency absorption).

Balancing technique (1) corresponds to Fig. 1, and techniques (3), (4), (5) correspond to Fig. 3.

An initial test was necessary to verify the room-related balancing techniques; that is to say, two questions had to be settled:

1. What delay values have to be set for a given impression of distance?

2. How many balancing reflections are necessary?

The tests showed that spot-microphone signals must not be delayed by more than 27 ms after the direct signal (as picked up by the main microphone), and that a loss in depth can be avoided satisfactorily by using just two reflections. All subsequent tests were based on these preliminary results.

The values for a balancing gain were determined in a further series of preliminary tests. In order to be able to compare the increase in sound volume of the various balancing techniques in the main test, the treshold of perception of the balancing signals in the main-microphone signal first of all had to be determined. This value of balancing gain stands for 0 dB. Three values were then defined with reference to that threshold:

+ 3 dB = slight balancing gain
+ 6 dB = moderate balancing gain
+ 9 dB = large balancing gain.

Three values of balancing gain have been chosen in order to verify that a slight increase in loudness of an instrument to be balanced would if not necessarily have to be processed binaurally. If not the associated processing capacity could have been used more sensibly somewhere else.

The main test consisted of a comparison between the intensity balancing technique (1) with each of the other techniques (2) to (5) in turn. The main-microphone signal (dummy-head signal) without balancing signal was taken as a reference. The listeners were asked which of the two examples came closer to the reference with respect to the perceived distance of the instrument. The test material included not just the different techniques and balancing gains, but also differing sound samples from alternating directions. For the test results we have chosen two examples:

- testsignal 1: trio clarinet, flute. oboe
- testsignal 2: duet trumpet, trombone

In each case of both recordings one instrument (clarinet, trumpet) was located close to the main microphone while playing a phrase "forte"; the other instruments, at the rear of the stage, accompanied it "piano". The object was to reproduce the trombone resprective the flute at higher volume by means of mixing the according spot-microphone signal.

6.2.1 Listening test 1

The test sequence was A-B-C-B-C, where A represents the unbalanced main-microphone signal, B the according intensity balanced signal (1), and C represents the balanced signals by using methods (2) to (5). Each test sequences compared exhibited the same balancing gain (3 dB, 6 dB, 9 dB). The listeners were asked to make a relative assessment between B and C in respect to the distance, given by the reference signal A . The evaluation was done with the help of the seven-grade comparison scale according to CCIR - Rec. 562-2.

6.2.2 Listening test 2

A control experiment was carried out in each case, so as to confirm the results. The listeners were asked to make an absolute assessment of technique (1) to (5) with respect to the main-microphone reference signal. The object of this test was also to classify the subjective perception of depth. The five-grade quality scale according to CCIR-Rec. 562-2 was used.

6.3 Results

6.3.1 Listening test 1

The graphs in Figure 7a and 7b show that the room-related balancing techniques were preferred. In a comparison between the three room-related techniques, the ORTF balancing technique was assessed as similar to the binaural. This is no surprise, since the brain possesses only limited ability to analyse reflections, which means that the complete binaural signal relationships are not essential for reproducing the balancing reflections. We also see that distance compensation achieves no appreciable improvement for the examples used here. It is probable, however, that when other sound sources have to be balanced - instruments with a large noise component at close quarters, or critical signals from a solo singer - then frequency-dependent corrections would be effective. Further tests are needed on this subject.

6.3.2 Listening test 2

At a uniform balancing gain of 6 dB, the results of the control experiments (Figure 8a and 8b) was that distance can be perceived significantly better with a room-related balancing technique than with simple intensity balancing or delayed intensity balancing.

6.3.3 Summary

Two important results are illustrated in Figure 9. Figure 9a shows the quality assessment with regard of distance imaging of delayed intensity balancing (2) in comparison to intensity balancing (1) . It demonstrates that an exact compensation of the propagation time difference between main- and spot microphone leads to results which are not better or even worse than achieveable with pure panpot technique, dependent on the balancing gain.

In contrast to that the room-related balancing technique allowes to preserve the desired presentation of spatial depth obtained by a appropriate main microphon. Figure 9b, showing the quality assessment of ORTF-balancing technique (4) in comparison to panpot balancing technique (1), verifies this statement. Furthermore it can be deduced that the greater the necessary balancing gain, the more room-related balancing technique is preferred.

7. CONCLUSION

Music is made in rooms and concert halls. Recording techniques should not be restricted to simple directional imaging of stereo sounds, but should also provide a depth image.

In the room-related balancing technique, by means of simulating reflections, the necessary increase in volume is achieved by adding the sound energy from artificially generated reflections. Too quiet an instrument can be enhanced for our ears without making it appear to spill out of the loudspeaker. The perception of depth is not lost, and important interaural signal differences are retained in the stereo signal.

The high signal processing effort initially considered necessary for binaural balancing can be considerably reduced by using the ORTF balancing. That will reduce the cost of future acquisitions. The tests showed that the favourable imaging characteristics of a good equivalence main microphone can be consistently applied to the balancing technique. An ideal application would be for a digital mixing console, because its internal signal processing is insensitive with respect to S/N ratio.

Thinking ahead, this balancing technique could also be used for poly-microphony. The requirements made in the beginning, to reproduce the original spatial characteristics of the sound source, must then be dropped. They would be replaced by the simulation of any desired artificially created space.

BIBLIOGRAPHY

/1/ Keibs, L.; Möglichkeiten der Stereo-Ambifonen Schallübertragung auf 2 Kanalen (Possiblities for transmitting sound by stereo ambiphony over two channels). Acustica, 1962, pp. 118 - 124

/2/ Baron, M. and Marshall, H.A.:Spatial impression Due to Early Lateral Reflections in Concert Halls. J. Sound Vibr. 77, 1981, pp. 211 - 232

/3/ Blauert, J.: Spatial Hearing - The Psychophysics of Human Sound Localization. The MIT-Press, Cambridge, Mass.: 1983, ISBN 0 -262 - 02190 - 0

/4/ Griesinger, D.: Theory and Design of a Digital Audio Signal Processor for Home Use. J. Audio Eng. Soc. 37, 1989, pp. 40 - 50

/5/ Theile, G.: Hauptmikrofone und Stützmikrofone - neue Gesichtspunkte für ein bewährtes Aufnahmeverfahren (Main microphones and spot microphones - new aspects for a proven recording technique). 13th Audio engineers' conference 1984, conference minutes pp. 170 -184

/6/ Theile, G.: On the stereophonic imaging of natural spatial perspective via loudspeakers: theory. In: Perception of Reproduced Sound, 1987. ISBN 87 - 982562-1-1, pp. 135 - 146

/7/ Theile, G.: Das Kugelflächenmikrofon (The sphere microphone). 14th Audio engineers' conference 1986, conference minutes pp. 277 - 293

/8/ Wöhr, M.: Nellesen, B.: Untersuchungen zur Wahl des Hauptmikrofonverfahrens (Studies on the selection of the main microphone method). 14th Audio engineers' conference 1986, conference minutes pp. 106 - 120

/9/ Laws, P.: Zum Problem des Entfernungshörens und der Im-Kopf-Lokalisation von Hörereignissen (On the problem of distance hearing and the localization of auditory events inside the head). Dissertation, Techn. Hochsch. Aachen, 1972.

/10/ Plenge, G.: On the difference between localization and lateralization. J. Acoust. Soc. Amer. 56, 1974, 944 - 951.

/11/ Jecklin, J.: A Different Way to Record Classical Music. J. Audio Eng. Soc. 29, 1981, pp. 329 - 332

/12/ Jecklin, J.: Musikaufnahmen: Grundlagen, Technik, Praxis. (Music Recordings: Foundations, Technique, Practice). München: Franzis-Verlag 1980, ISBN 3 - 7723 - 6701 - 1, pp. 55

/13/ Richter, F. and Persterer, A.: Design and Applications of a Creative Audio Processor. 86th AES Convention, Hamburg 1989, Preprint 2782

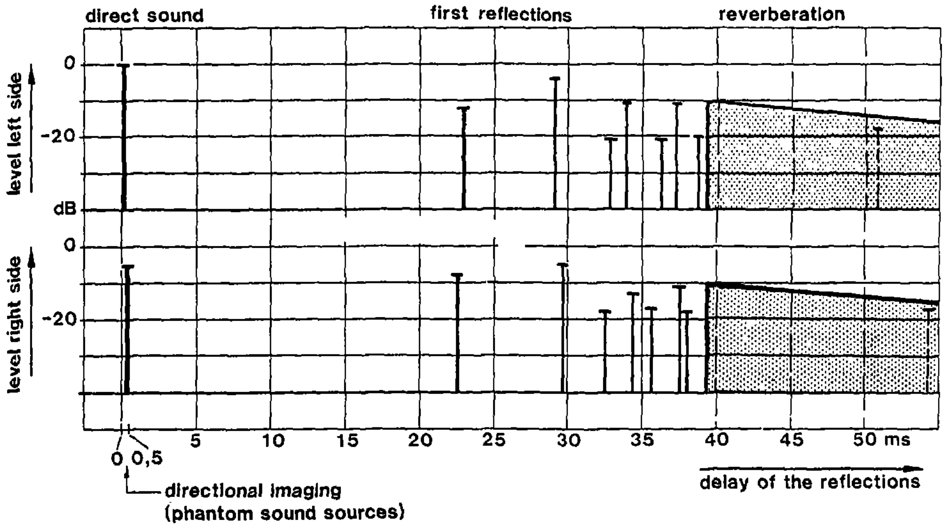


Fig. 1 Impulse pattern of an equivalence microphone

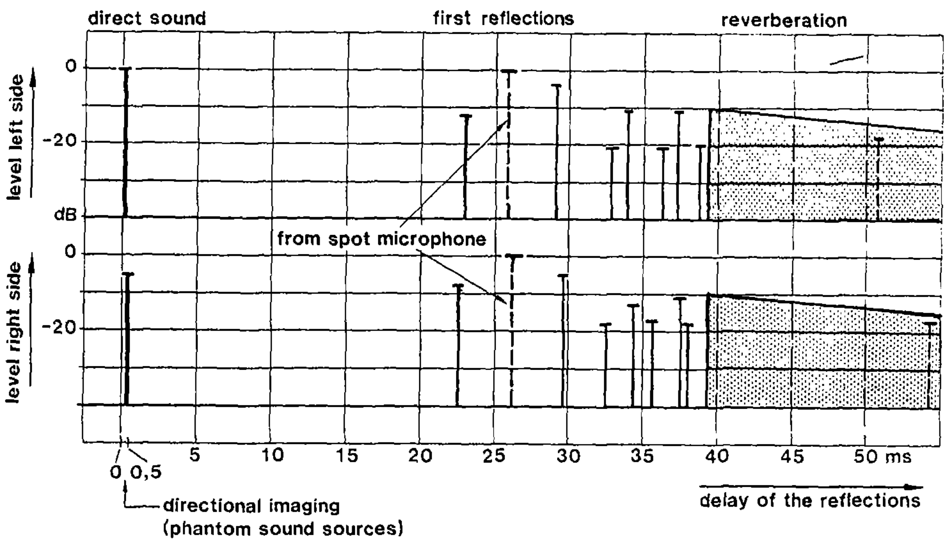


Fig. 2 Impulse pattern using room-related balancing

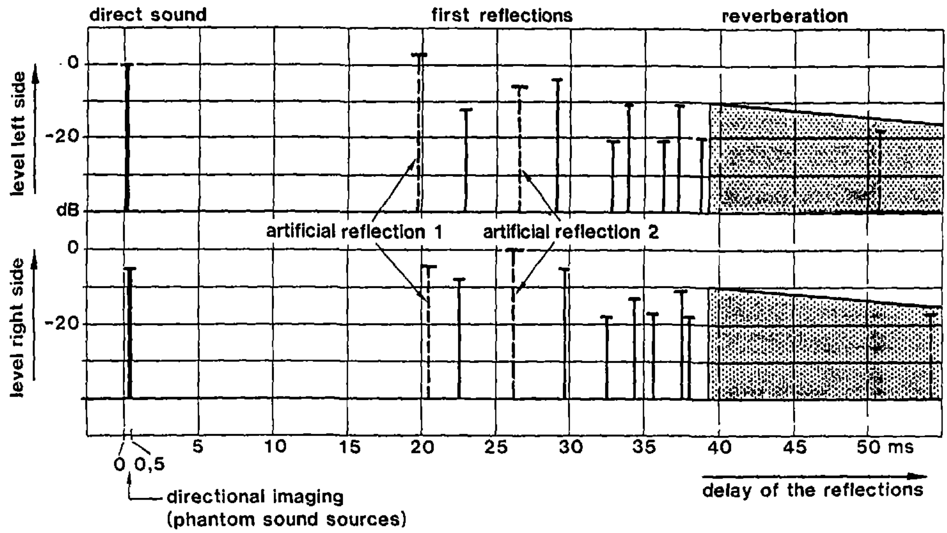


Fig. 3 Impulse pattern of optimized room-related balancing
(stereophonic representation of 2 artificial balancing reflections)

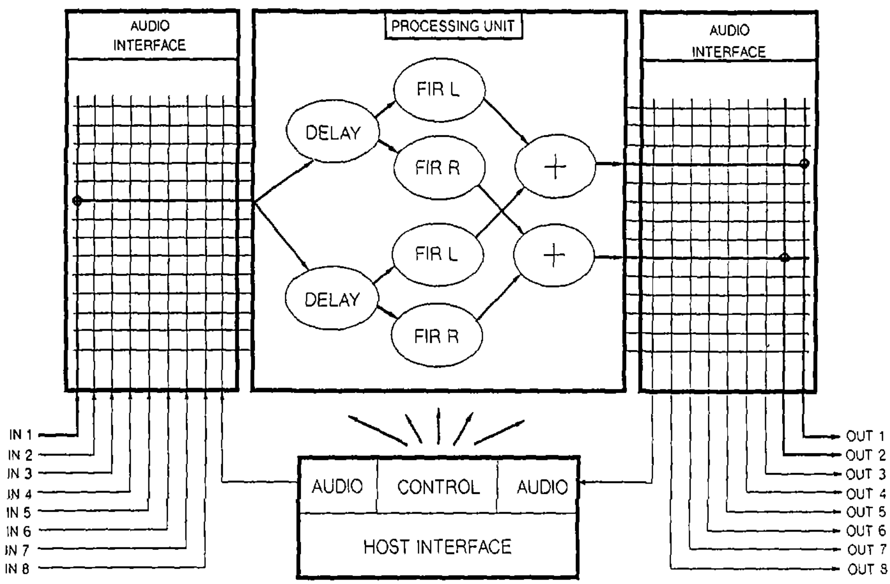


Fig. 4 System structure of the CAP 340 M
(the processing unit generates 2 artificial reflections in the present application)

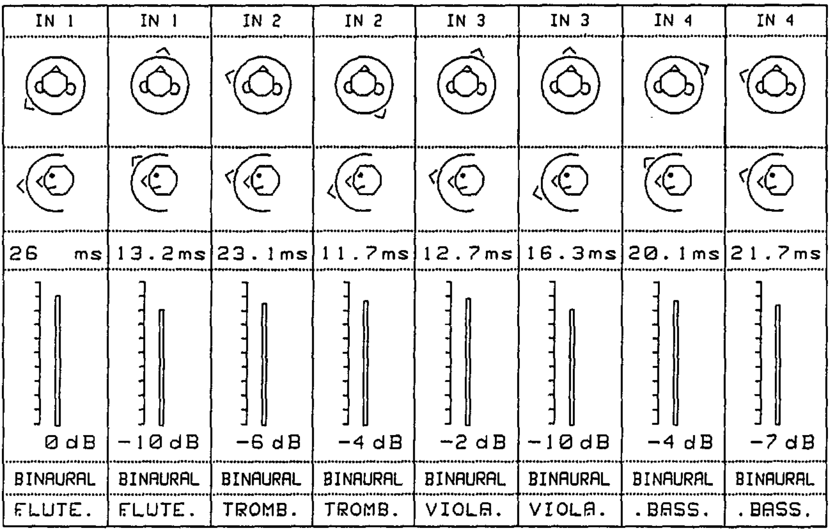


Fig. 5 Screen shot showing in this moment the user shell

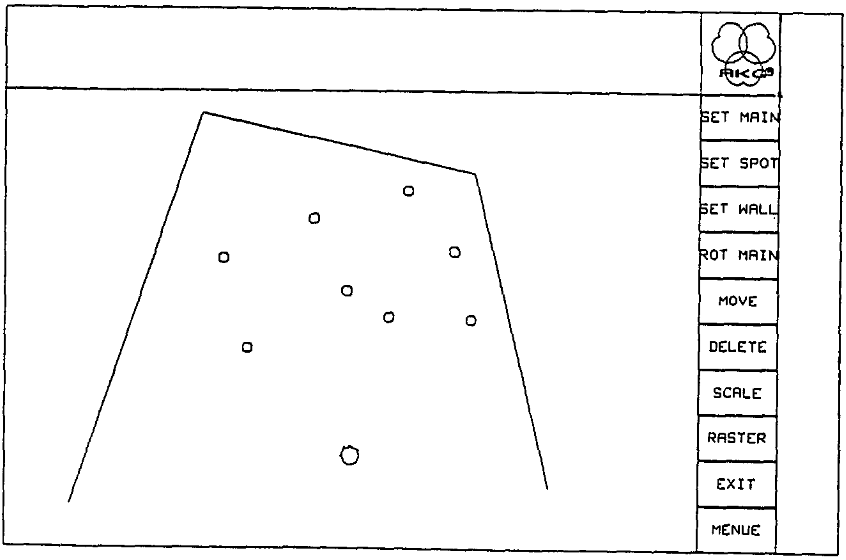


Fig. 6 Setting-up display for the positions of the main microphone, several spot microphones, and reflecting surfaces

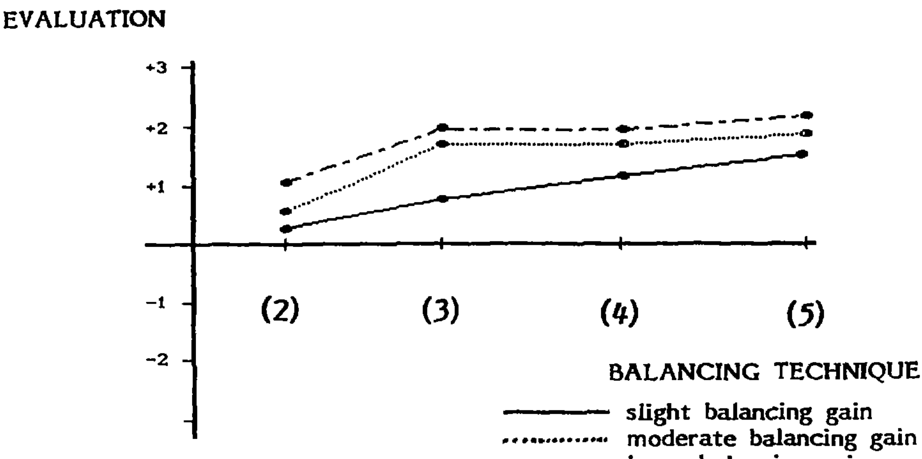
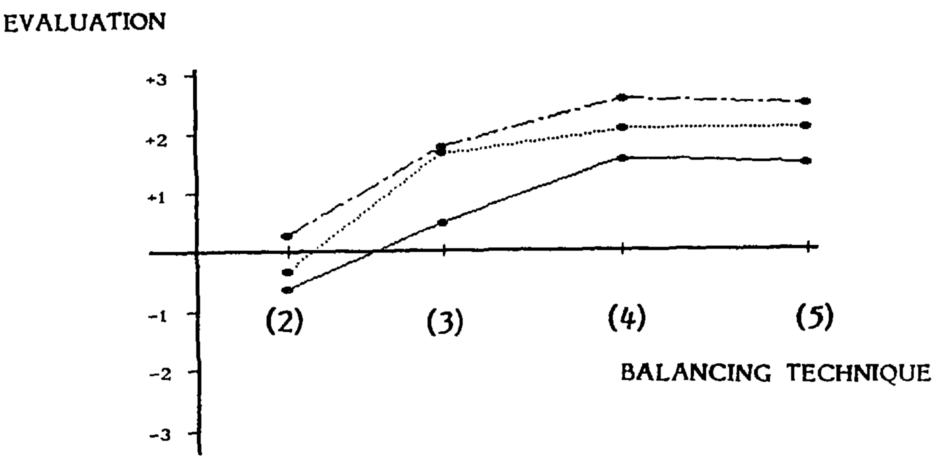


Fig. 7 Judgements of the comparison test

a) testsignal 1 ; trio clarinet, flute, oboe



b) testsignal 2; duet trumpet, trombone

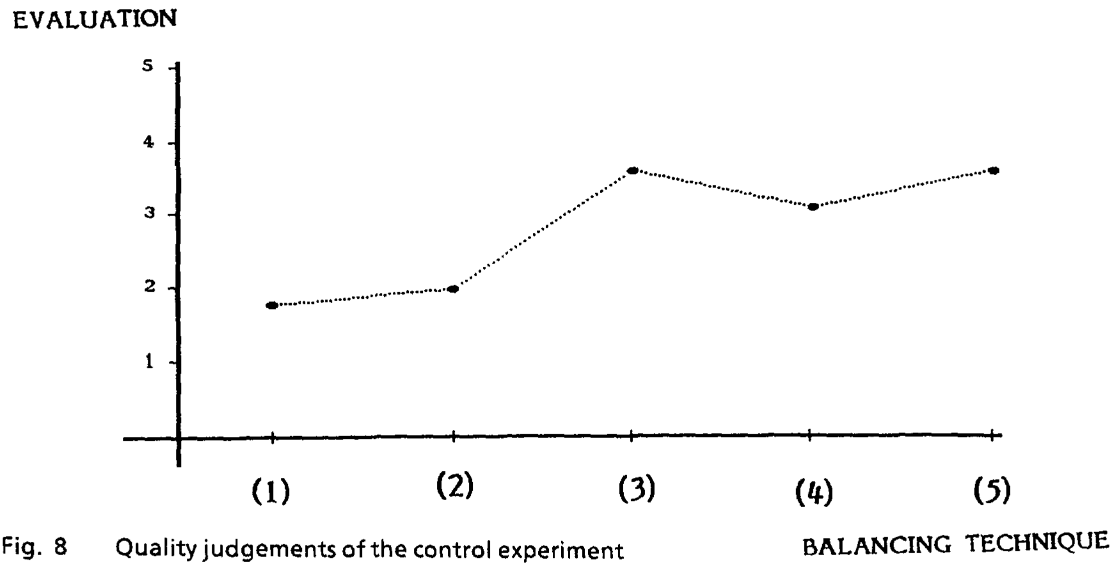
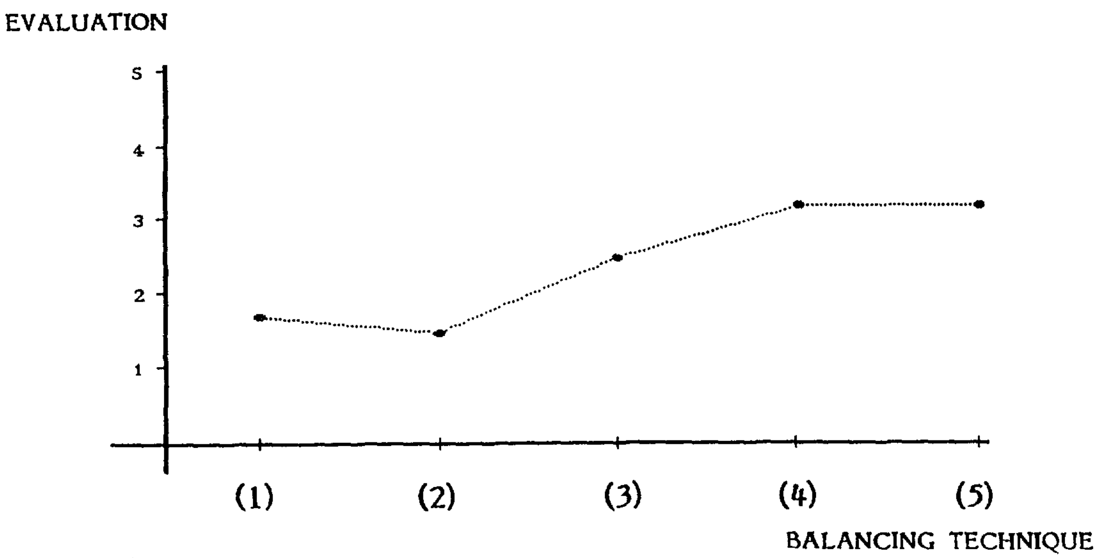


Fig. 8 Quality judgements of the control experiment

a) testsignal 1 (trio)



b) testsignal 2 (duet)

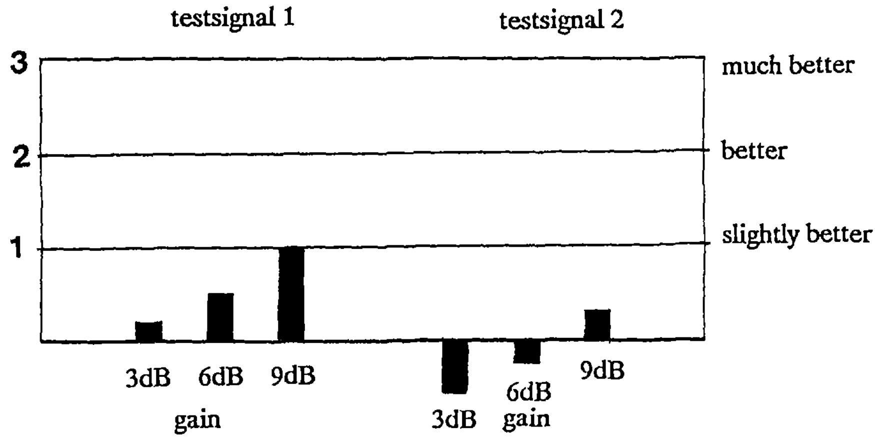
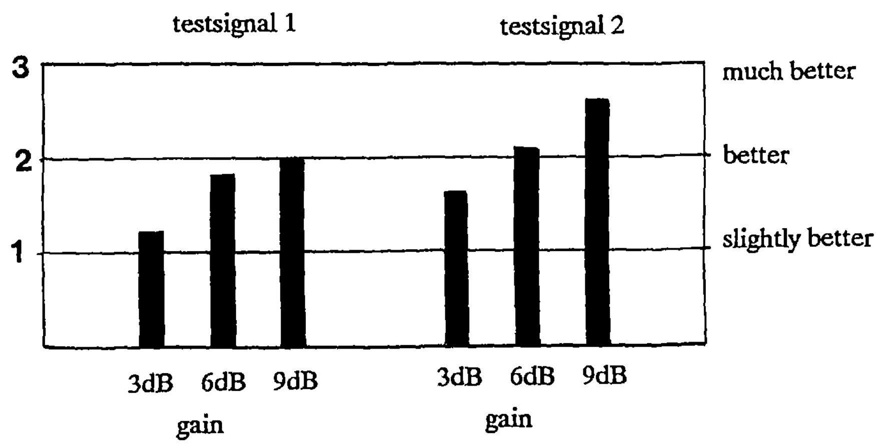


Fig. 9 Results adapted from Fig. 7:

a) quality assessment of delayed intensity balancing (2) in comparison to intensity balancing (1)



b) quality assessment of ORTF-balancing (4) in comparison to intensity balancing (1)