L8-3

Time/intensity trading stereophony for (HD)TV and audio applications

Ronald M. Aarts

Philips Research Labs. PO Box 80000, 5600JA Eindhoven, The Netherlands

1 Introduction

Increasing effort has been devoted to the improvement of the stereophonic effect eg. [1, 2] and the description of it [3, 4]. Generally it is considered as a serious artefact of the present stereo system that the listener is aware of the stereophonic illusion only in a limited region. This off-centre listening problem becomes even more serious when the distance between the loudspeakers is small compared with the size of the listening area as in multimedia PC monitor and video applications. Hereafter some experiments are described in order to investigate the influence of off-centre listening for normal stereo set-ups and for small base stereo set-ups. Differences between the ear input signals are necessary in order for auditory events to appear laterally, away from the median plane. A frequently used technique for the measurement of the equivalence between interaural time (ITD) and level differences (ILD) are trading experiments, that is asking a subject which time difference is equivalent to a certain level difference. A lot of experiments are reported, however, the major part is performed by using headphones, or if loudspeakers are used by applying an artificial delay between them and leaving the subject in the centre position. It is the aim of the present experiments to apply natural time differences by using a set-up as occurs in practical situations for normal audio listening and for a (HD)TV-sound environment.

2 Experiments

A pair of two-way omnidirectional loudspeakers was used. They were constructed by mounting the drivers coaxially with their axis perpendicular to the horizontal plane. Between the two loudspeakers a third loudspeaker was mounted, hereafter referred to as the centre loudspeaker. The three speakers were mounted at 1.62m, the averaged ear height of the standing subjects. On an arc seven numbered positions for the subject to stand in were marked in such a way that the subject was at 2.19 metres radius from the centre loudspeaker. A signal processor and a small computer were used to control the experiment. The computer drew up a random order of positions for the subject to stand in, in such a way that the experiment was balanced (i.e. the angle φ, between the subject and the axis of the centre loudspeaker had an equal number of occurrences per subject). The difference in volume of the right and left loudspeaker at the start in each run was randomly chosen by the computer. Throughout the experiment the subject kept his head pointing straight forward (i.e. with his shoulders parallel to the wall). The task for the subject was to adjust the level of the right and left loudspeakers until the phantom source was located in the centre position.
A switching facility was included that sent the signal to either the centre loudspeaker or to the left and right loudspeakers, so that at any time during the experiment the subject could switch to the signal coming from the centre loudspeaker as a reference of the centre position (the signals fed to the loudspeakers were, apart from the volume, the same).
Using this box he could control the volume of the left and right loudspeakers, while the total electrical power fed to both loudspeakers was kept constant. The subject could switch, as often as he wanted between the signal coming from the centre loudspeaker only, and back to the left and right loudspeakers, and indicate when he thought the phantom source was at the same position as the centre loudspeaker. The run was then repeated three times at the same listening position, then the experiment leader indicated a new position for the subject and the whole process was repeated until all 7 positions had been taken up.

2.1 Experiment 1

The distance between the loudspeakers was 2.5m. Pink noise was used as stimulus and the level at $\varphi = 0$ at ear height was 78dB. The experiment was carried out in an anechoic chamber, with the three loudspeakers clearly visible. Ten subjects with normal hearing, four female and six male, participated in the experiment. Before the experiments, all subjects' threshold levels were determined by pure tone audiometry. All subjects had a hearing loss of less than 10 dB; the threshold differences between left and right ears of each subject were also less than 10 dB.

2.1.1 Results

The level differences between the right (R) and left (L) loudspeakers resulting from this experiment are plotted in Fig. 1 (X-marks). The vertical bars indicate the variance over all the 10 subjects with the four replications per subject.
In looking at the data two features appear. Firstly, the variance per angle is rather large. This is caused by the large variability among the subjects. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The R^2 value was 99%, indicating that almost all of the variance can be explained by the factors angles and subjects and their interaction each appearing as highly significant ($p < 10^{-10}$). The variance of each individual subject was remarkably low. They can reproduce the task very well, but do not agree with each other.
The second feature is the deviation of 2dB at the centre position, as can been seen in Fig. 1. One would expect this to be 0dB. At first a psychological reason was expected, because the experiment was performed in the corner of the anechoic chamber with the corner being at the left hand of the subject. However, in the third experiment the set-up was made symmetrical while the difference remained. This level difference is reported also in [5], however they considered it to be an experimental artefact. The grand mean of all the 280 observations appears to be very close to 0 dB, as one might aspect for equal deviations to right and left from the median plane, indicating that there is a good symmetry between the left and right plane, except for the centre position itself. All subjects reported a broadening of the phantom source for increasing angles from the centre position. They noticed also that at the outmost locations, the space between the loudspeakers was filled by the image.

2.2 Experiment 2

The same set-up and conditions as in the previous experiment were used. The only difference was the stimulus. A female voice was used as the stimulus, the duration of the excerpt was about 20 sec. and was then repeated over and over. The results are depicted in Fig. 1 (O-marks). The Figure shows that the results are very similar to those obtained from the pink noise stimulus.

2.3 Experiment 3

The aim of this experiment was to study the time intensity trading at a very close listening distance and a relatively small loudspeaker base as would occur in Multi Media PC and TV-applications. The conditions were adapted to the monitor application. A 16:9 aspect ratio monitor with a 36" screen was used, without displaying a picture. This was done to make the localization more precise than with a picture, because it is known that when the seen and heard locations of an object are made discrepant the sound is perceived near or at the seen location. The centre loudspeaker was stuck in the middle of the screen and the left and right loudspeaker were mounted 1m apart. The experiment was performed in a studio using the same female voice as in experiment 2. The subjects were seated in a chair at five different positions, on a line parallel to the screen at a distance of 1m. Each position was repeated four times. Nine subjects out of the set of the previous experiments performed the test. The level differences between the right and left loudspeakers were the outcome of the experiment and are plotted in Fig. 2. Again a large variance can been observed and the bias at the centre position of almost 2dB appears. The data were analyzed using ANOVA, and gave similar results as in the previous experiments.

3 Discussion

The data of the three experiments can be converted to exclude the intensity difference caused by different pathlengths from left and right-loudspeaker to the head, r_{-1} and r_1 respectively. This is done by adding the term $20\log(r_{-1}/r_1)$ to the R-L values. According to the time-intensity trading concept, the thus resulting level differences are caused by time differences only. They are plotted in Fig. 3, where $\Delta T = (r_1 - r_{-1})/c$. For comparison some curves found in literature are depicted in Fig. 3. It is known that many factors affect the trading relation. The ILD and the ITD interact in a complex fashion, some of these features of this interaction may be predicted on physiological grounds. The nature of the stimulus employed and the level at which it is presented are probably the most influential. However, there are also important psychological factors. One of these can be conflicting intersensory information. When spatial information in two modalities is available, the information in one of the modalities is partly or wholly discarded in favour of that in the other. This can be important for sound for television, but can also have influence on the trading experiments; at large angles the subjects can see only one loudspeaker when they are facing forward. Possibly that the experimental set-up can effect the results even when the ILD's and ITD's are the same as in another experiment. This was the motivation to use a third loudspeaker in the centre, so it could serve both as an acoustical and as a visual reference.

4 Conclusions

The results for pink noise and the female voice as stimuli are very similar. Time intensity trading beyond about 1ms is not possible, the phantom source then broadens with increasing time difference. The results exhibit a large variability between the subjects, however a low variance within the subjects. The trading ratios for (HD)TV applications at a close viewing distance and a small loudspeaker base, are different than for a wider stereo base as for normal audio applications.

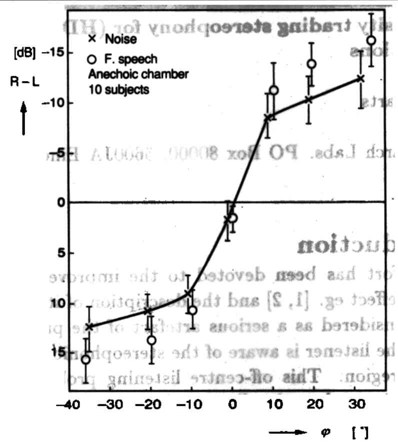

Fig. 1 Required level difference between right and left loudspeakers, as function of the subjects' locations, to perceive the phontom image in the centre.

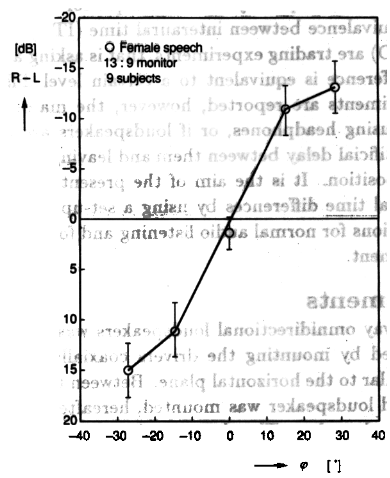

Fig.2 Required level difference between right and left loudspeakers, as function of the subjects' locations, to preceive the phantom image in the centre of a TV monitor.

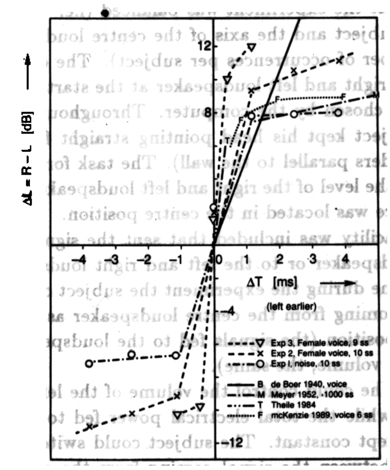

Fig. 3 Time intensity trading results of the experiments in comparison with other experiments from the literature.

References

[1] B.B. Bauer. Broadening the area of stereophonic perception. *J. Audio Eng. Soc.*, 8(2):91–94, April 1960.

[2] P. Goossens. Position independent stereo for HDTV by time/intesity trading. Technical report, Philips Research Labs, 1990. Nat.Lab. Report 6431.

[3] K. de Boer. *Stereofonische geluidsweergave (Stereophonic sound reproduction).* PhD thesis, Technical University Delft, December 1940.

[4] J. Blauert. *Spatial hearing: The Psychophysics of Human Sound Localizaton.* The MIT Press, 1983.

[5] A.R. McKenzie and I.H. Flindell. Time/intensity trading in stereophony. In *Proceedings I.O.A.*, pages 117–133. Institute of Acoustics (UK), 1989. Volume 11, part 7.