ARCHIVES OF ACOUSTICS
33, 4 (Supplement), 147–152 (2008)

LOCALIZATION OF A SOUND SOURCE IN DOUBLE MS RECORDINGS

Piotr KLECZKOWSKI, Magdalena PLEWA, Grzegorz PYDA

AGH University od Science and Technology
Department of Mechanics and Vibroacoustics
Al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: kleczkow@agh.edu.pl

(*received June 15, 2008; accepted November 3, 2008*)

In recent years new solutions and systems emerge to be used in spatial recording and sound reproduction. The aim of this paper is to investigate and describe sound localization in a horizontal plane in Double MS recordings. Double MS is a spatial microphone technique, which evolved from the intensity MS stereo system. There are many possibilities to decode recorded signals to different audio systems including the Surround 5.1 system. Localization of a sound source in the horizontal plane, which is the most important issue in surround audio systems, is a complicated phenomenon. It is not possible to describe it by using only mathematical models, so we performed a subjective listening test to obtain some quantitative data. The results of the test and the advantages of the MSM technique lead to the conclusion that this microphone system can be used successfully in spatial sound productions.

Keywords: microphone techniques, spatial sound localization, surround audio.

1. Introduction

The quality of different microphone techniques for surround recordings has been investigated by many researchers. A number of papers focusing on the representation of direction by these techniques and the ease of use of Double MS (small number of microphones) inspired us to undertake the presented investigation. The accuracy of sound localization in the horizontal plane is difficult to assess by the measurements of any physical values. Because of that, we conducted subjective listening tests to obtain quantitative results.

1.1. Spatial sound localization

Some specific properties of the auditory system enable the spatial localization of sound. Sound localization in the horizontal plane, is the most important point in this experiment. Effects making possible the localization of the sound source in this plane include interaural time difference, interaural intensity difference, the shape of pinnae, the head movement, and the Haas effect. Their combination results in the mechanism which allows us to localize sounds with different spectra (from simple tones to sounds with very complex spectra) [3, 4].

In surround techniques, a virtual image is formed in the space between loudspeakers and surroundings of the listener. In order to achieve this result, we take advantage of all described earlier properties of human hearing.

A choice of listener's position is of a great importance for correct sound localization. Tens of centimetres of difference in distance from several loudspeakers leads to the intensity distinction and time delays over 1ms cause an impression that the only sound source is the nearest loudspeaker.

1.2. Double MS technique

The Double MS technique evolved from the stereo MS technique [2]. For the first time, it was described in [5]. In this chapter the authors present the idea of Double MS recordings and optimal arrangements of decoding signals to various audio systems.

The Double MS system consists of two MS pairs located in one point but turned in opposite directions (Fig. 1).

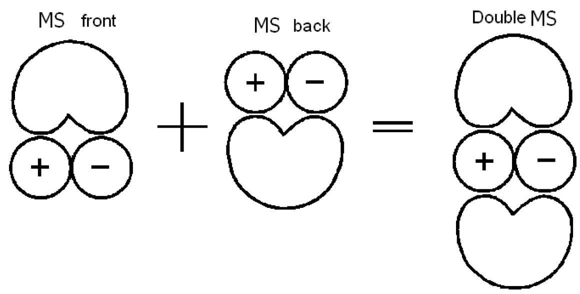

Fig. 1. Graphic presentation of microphones orientation in Double MS system.

Double MS is the intensity technique and because of that the recommended arrangement of microphones is to locate them in one point in space. Three microphones are used in this technique: two cardioids turned in opposite directions (M front and M rear) and one at an angle of 90° with "figure of 8" polar pattern.

To decode signals to the 5.0 system, we have to treat them with Double MS technique like with two MS systems directed forwards and backwards. We decode signals from the front MS to channels L and R, and signals from the rear MS to LS and RS (left and right surround). The central channel is obtained directly from the M_{front} microphone.

In Double MS technique (likewise in MS), an adequate selection of levels of signals from several microphones has the effect on the quality of a recording performed. A crosstalk between channels is the additional factor which influences the quality of recordings. But it is not the crosstalk in classical meaning concerning multi-microphone techniques. The virtual sound source arises from two loudspeakers which emit the coherent sound signal. If the third loudspeaker appears and radiates the same coherent signal (called crosstalk), localization of the virtual source becomes smeared [5].

2. Subjective tests

2.1. Preparation of sound material for tests

The sound material used in tests was recorded in an anechoic room. The system used in the recordings consisted of: a recorder (PC), the microphones – two with cardioid characteristics (Rode NT2) and one figure-8 (sE Electronics Z3300a), a preamplifier (PrismSound MMA-4XR) and an interface (DigiDesign Mbox 2). As the sound source, we used the YAMAHA HS 50 loudspeaker. Before recording, the system was calibrated so that the levels at the output of each microphone were equal for a certain sound level.

All of test signals were recorded in different position of the sound source relative to the microphone system. In each recording, the loudspeaker was located on a circle with the centre in the position of the microphone system. The α angle took a value of $0°$, $\pm 30°$, $\pm 45°$, $\pm 60°$, $\pm 90°$, $\pm 120°$, $\pm 135°$, $\pm 150°$ and $\pm 180°$, respectively.

2.2. The form of tests

The objective of the test was that the listeners determine the direction from which the sound came. They marked their answers on circular diagrams (Fig. 2). The test material consisted of four test signals: the white noise, two synthetic sounds and a gong.

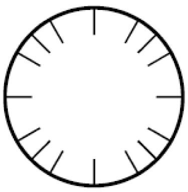

Fig. 2. The diagram used by listeners to mark the direction.

We passed through two series of tests: the first one was applied to signals decoded in a classical way, only by simple crossing to several channels, whereas the second part was composed of recordings decoded in the recommended mode giving the smallest crosstalk between virtual microphones (Table 1) [5].

Table 1. Parameters of recommended decoding of Double MS signals to the 5.0 system [5].

microphone	L	R	C	LS	RS
M	−5.1 dB	−5.1 dB	0.0 dB	−16 dB	−16 dB
S	−2.0 dB	−2.0 dB	$-\infty$	−6.2 dB	−6.2 dB
R	−15.8 dB	−15.8 dB	−11.1 dB	−1.1 dB	−1.1 dB

2.3. The course of tests

The tests took place in a studio of the Academy of Music in Kraków. The audio system was compatible with Surround 5.0. The standard is precisely defined in [1]. The Mackie HR 626 monitors were deployed on the circle with a radius of 3 meters. We did not use an additional device for low frequencies, because the loudspeakers had the sufficient frequency response for the purpose of the tests. 20 listeners, students of the AGH University and the Academy of Music took part in the preliminary test, examining their general ability of directional hearing. The main series of tests was participated by a selected group of 15 listeners. Most of them were audio engineers and academic teachers who have had earlier experience with "surround".

2.4. Test results

The charts presented in Fig. 3 show the dependence of accuracy and the criteria of correctness on the kind of signal. They show the importance of the effect of the type of sound on the effectiveness of the source localization.

a) b)

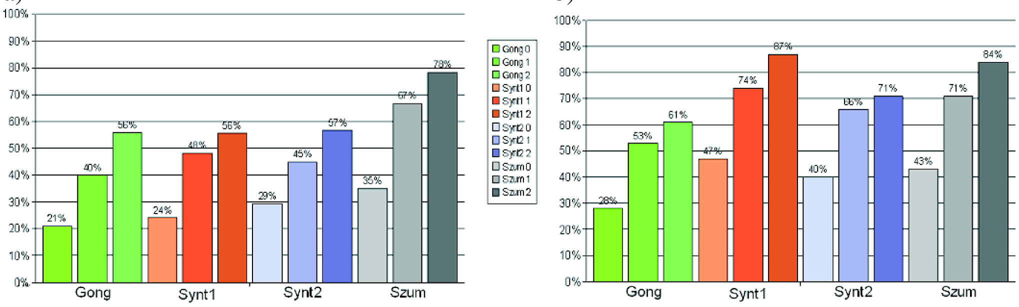

Fig. 3. Results of listening tests. Diagrams present the accuracy of answer versus the type of the signal and the correctness criteria. Shades illustrate precision of the answer (the brightest – answer correct within a discretisation step of 15 degrees, medium – the error of one step allowed, the darkest – the error of two steps allowed: a) the first series of test (classical decoding), b) the second series (recommended decoding).

Table 2 depicts the percentage of correct answers depending on the angle. In the second and forth columns, the results are presented for precise determination of the direction. Values placed in the third and fifth columns are calculated with the assumption that the answers marked on the neighbouring positions on the diagram (Fig. 2) are also accepted.

The errors in localization for different instruments (recommended decoding) have been analysed. The related plot is given in Fig. 4 and the data in Table 3. The "r Pearson" correlation of errors, averaged for each listener, showed good correlation between gong and the other signals. For other pairs of instruments, the correlations obtained had insufficient significance level ($p > 0.05$). The "r Pearson" values for gong with other instruments are shown in Table 4.

Table 2. The angle dependence for different correctness of answer.

angle	classical decoding		recomended decoding	
	answer acuuracy		answer acuuracy	
	exact	± one position	exact	± one position
0°	49%	69%	84%	97%
±30°	45%	59%	32%	71%
±45°	32%	42%	29%	54%
±60°	45%	69%	41%	83%
±90°	29%	37%	31%	69%
±120°	29%	37%	35%	48%
±135°	15%	23%	15%	29%
±150°	31%	37%	20%	38%
±180°	44%	44%	57%	71%

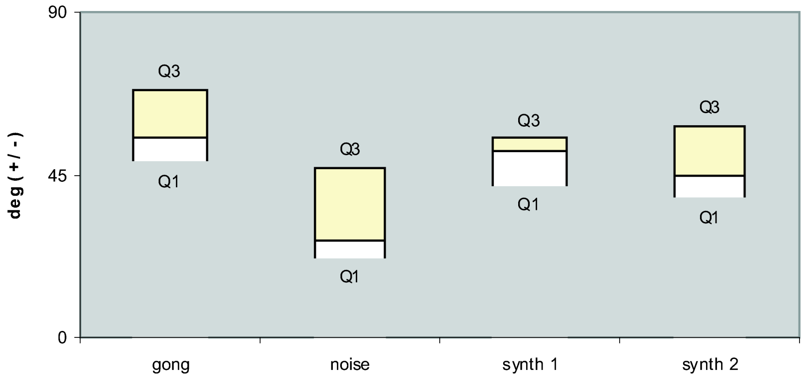

Fig. 4. Box plot of the statistical analysis of errors in the listening test (recommended decoding), referenced to test instruments. Vertical axis is scaled as absolute values of errors (degrees). The minimum error is 0, maximum is 180. The lower edge of each box is equal to the 1st quartile (Q1), its upper edge is equal to the 3rd quartile (Q3). The borderline inside the box marks the median.

Table 3. The numerical data for Fig. 4.

	Q1	Me	Q3
gong	48	55	68
noise	22	27	47
synth 1	42	52	55
synth 2	38	45	58

Table 4. “r Pearson” correlations of errors (averaged for each listener) for gong and other instruments.

r Pearson	gong	significance level
noise	0.503	0.019
synth 1	0.450	0.034
synth 2	0.488	0.023

3. Conclusions

After preliminary analysis of the results, we noticed that listeners tended to localise the sound source at an angle close to, but not exactly matching our assumed grid of angles. We decided to accept the answers missing the correct angle by one and two positions. We observed also an important difference in localization for different sounds. To avoid an influence of the type of sound on the assessment of the system's quality, we analyzed also the results for several kinds of signals.

Though localization of the sound source is used in everyday life, participants of the test that declared sound education on medium and high level gave more correct answers than the others. Classification of the sound coming from the angle of 180° as the 0° direction was the frequent fault.

Although more correct answers were given for sounds coming from the front, it does not testify to better working of the system in these directions. It is caused by the inferior ability of the human hearing to localize sounds from the back.

The response accuracy depended on a test signal. It can be concluded that the best localization is for broadband signals of longer duration. Usage of the recommended decoding method considerably improved the recognition of the direction.

The presented results show that the virtual image produced by the Double MS recording technique and Surround 5.0 playback is realistic enough to be applied in the surround sound recordings. The accuracy of direction mapping shouldn't be measured by an angle, but by a certain space with the centre in the given direction.

Acknowledgment

The authors would like to thank Dr. Mateusz Bień from the Musical Academy in Cracow for making arrangements for the listening tests and for all creative inspirations.

References

[1] AES Technical Council, *Multichannel surround sound systems and operations*, 2001.

[2] Martin G., Mus B., Mus M., *Introduction to Sound Recording*, www.tonmeister.ca, 2004.

[3] Moore B.C.J., *An Introduction to the Psychology of Hearing* [in Polish]: *Wprowadzenie do psychologii słyszenia*, Warszawa–Poznań, 1999.

[4] Rumsey F., *Spatial audio*, Focal Press, Oxford, 2001.

[5] Wittek H., *Double M/S – a Surround recording technique put to test*, www.schoeps.de, 2006.