ARCHIVES OF ACOUSTICS
36, 2, 347–363 (2011)
DOI: 10.2478/v10168-011-0026-8

Choosing and Configuring a Stereo Microphone Technique Based on Localisation Curves

Magdalena PLEWA, Piotr KLECZKOWSKI

AGH University of Science and Technology
Department of Mechanics and Vibroacoustics
Al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: magdalena.plewa@gmail.com

(*received February 11, 2011; accepted March 23, 2011*)

Whenever the recording engineer uses stereo microphone techniques, he/she has to consider a recording angle resulting from the positioning of microphones relative to sound sources, besides other acoustic factors. The recording angle, the width of a captured acoustic scene and the properties of a particular microphone technique are closely related. We propose a decision supporting method, based on the mapping of the actual position of a sound source to its position in the reproduced acoustic scene. This research resulted in a set of localisation curves characterising four most popular stereo microphone techniques. The curves were obtained by two methods: calculation, based on appropriate engineering formulae, and experiment consisting in the recording of sources and estimation of the perceived position in listening tests. The analysis of curves brings several conclusions important in the recording practice.

Keywords: microphone techniques, stereo recording, acoustic scene.

1. Introduction

In order to perform a stereo recording, the engineer has a choice between a couple of microphone techniques and some flexibility in choosing the recording position. The latter choice is usually the result of a compromise between a coherent directional image of the sound stage and a required ratio of direct to ambient sound. Some techniques allow to adjust either the distance or the angle between the microphones. While there are many subjective factors influencing the engineer's choices, there are aspects that can be analysed quantitatively but these are hardly ever used in practice. The objective of this work was to verify and extend knowledge in this area and to turn it into a set of useful guidance rules.

When the stereo recording/reproduction process is considered as the mapping of a recorded acoustic scene to a sound image between the loudspeakers it should be analysed as two separate mappings.

The first one transforms physical locations of sound sources into differences of levels and/or time shifts between respective signals in two channels of the microphone system. Thus, the mappings of the localisation of a particular source can be described by two Eqs.:

$$\Delta L = f_L(\alpha, \beta), \tag{1}$$

where ΔL is the difference of levels received by two microphones (often referred to as Interchannel Level Difference – ILD), f_L is a function characteristic for a given microphone system, α is the angle of sound incidence (an acute angle between the axis of the microphone system and the direction of the source, see Fig. 1), β is an angle between the axes of microphones; and:

$$\Delta t = f_t(\alpha, d), \tag{2}$$

where Δt is the difference in arrival time of a sound between two microphones (often referred to as Interchannel Time Difference – ITD), d is the distance between microphones. The actual engineering formulae for (1) and (2) were given by Williams (1984) and Wittek (2002) (see also Sec. 2).

The other pair of mappings transforms ΔL and Δt into subjective perception of a sound image between the loudspeakers. These mappings can only be obtained from perceptual experiments. By convention, the position of a phantom source is measured as the percentage of distance from the middle of the image to either of the speakers, and is referred to as the shift. Thus, a phantom located at each speaker has a shift of 100%. These mappings can be described as:

$$s_L[\%] = \Psi_L(\Delta L) \tag{3}$$

and

$$s_t[\%] = \Psi_t(\Delta t) \tag{4}$$

where s_L and s_t are phantom source shifts resulting from ΔL and Δt respectively, Ψ's are respective perceptual mappings, which are subjective.

Wittek and Thiele (2000) performed their own experiments to evaluate Ψ's and compared them with earlier results obtained by other authors (Wittek, Thiele, 2002). The new results were close to the average of the earlier results and therefore Wittek and Theile accepted them as representative. Both mappings have a similar shape, and they are close to linear, up to the phantom source shift of about 75%. Above that value, a mild nonlinearity of compressive type occurs. In the linear range, the respective coefficients were evaluated as 7.5%/dB and 13%/0.1 ms (Wittek, Thiele, 2002). An important observation was that in the linear range both shifts were additive, i.e.:

$$s[\%] = s_L[\%] + s_t[\%] \tag{5}$$

where s is a total phantom source shift.

Thus, the combination of engineering versions of formulae (1) and/or (2) and experimentally obtained curves Ψ_L and/or Ψ_t allow to calculate the joint mapping

of the angular position of the sound source α into the phantom source shift s for a given microphone system. Such a mapping is referred to as the localisation curve.

The term "recording angle" (RA) refers to the angular sector in front of the microphone system, within which the recorded sources will be spaced between the loudspeakers in a reproduced sound image. For a given RA, the maximum angle of incidence α is equal to RA/2. The limits of the possible choice of RA angle were analysed by WILLIAMS (1987). The RA can be predicted from the localisation curve (WITTEK, THIELE, 2002) and the authors recommend using such an RA which will produce shifts of phantom sources of less than 75%. They argued that the space between -75% and $+75\%$ of shift was the most important imaging area and that localisation curves for different microphone setups did not differ very much within that area.

There is a problem with the calculation of (1) and (2). While the error in engineering calculation of (2) depends mainly on the precision of geometrical data, the f_L function includes directional characteristics of the microphones used, known to be frequency dependent, and therefore engineering calculations of (1) are only approximate. There may also be other sources of errors in these calculations. On the other hand, the estimation of Ψ functions depends on the type of a sound source (BLAUERT, 1997). Therefore, a thorough experimental verification of Eqs. (1)–(4) is an interesting topic for a research. HUGONNET and JOUHANEAU (1987) compared stereo microphone systems by investigating the joint mapping from α to s, using an experimental procedure instead of calculations of (1) and (2). They recorded some test sounds including a violin and short sine waves of several frequencies in a concert hall, and then reproduced the recordings in a stereo system. They measured three values of α differing between sound sources, but the rule for choosing particular values of α was not given. The values of shift were discretised to only three values: 10°, 20° and 30°. They found that the accuracy of angular estimation was higher for coincident microphones.

All experiments estimating Ψ curves known to the authors were performed in the standard stereo setup based on the equilateral triangle. In this work a comprehensive experimental evaluation of localisation curves was performed. Four microphone systems were tested with four sounds – white noise and three instruments, and thus localisation curves for different sound sources were obtained. In order to widen the knowledge on the subject, nonstandard stereo bases with widths of 30° and 90° were also investigated. Both the recordings and reproductions took place in an anechoic chamber. There were two objectives of the experiments: (i) to compare the localisation curves obtained from calculations (1)–(4) with those obtained from strictly controlled and comparable experiments, (ii) to develop a set of guidance rules for choosing and configuring a microphone system so as to obtain a desired recording angle, on the basis of localisation curves obtained from our experiments.

Section 2 presents the calculations of localisation curves, while Sec. 3 presents experimental procedure used to obtain them. In Sec. 4 selected results are presented in the form of plots. Sections 5 and 6 contain conclusions.

2. Calculation of localisation curves

2.1. Microphone systems examined

The following microphone systems were included in the experiments: XY with an angle between microphones β of 90°, giving the theoretical recording angle (RA) of 180° (Wittek, 2002); AB with the distance between the microphones d of 20 cm and estimated RA of 180°; ORTF, with standardized values $\beta = 110°$, $d = 17$ cm and RA estimated at 102°; binaural performed with Neumann KU100 dummy head (Neumann website).

2.2. Calculation of Δt

The layout of microphones and the names of variables used in (6)–(9) are shown in Fig. 1. Interchannel time difference is an effect of different distance from the sound source to microphones. It can be easily calculated from the following relation:

$$\Delta t \text{ [s]} = \Delta t_2 - \Delta t_1 = (R_2 - R_1)/c. \tag{6}$$

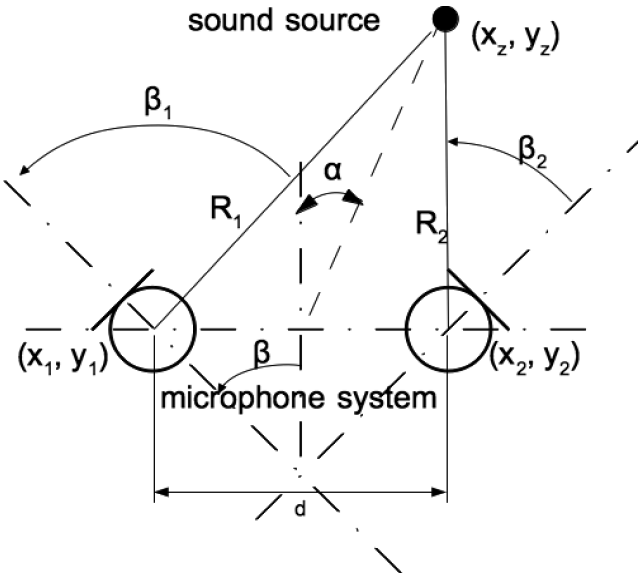

Fig. 1. The general layout of the stereo microphone system and the variables used in formulae.

2.3. Calculation of ΔL

Interchannel level difference is caused by one or both of two factors: different angles of incidence of sound and difference of distance to microphones. Different angles combined with polar sensitivity patterns $r(\beta)$ result in:

$$\Delta L_r = L_{r2} - L_{r1} = 20\log(r_2(\beta_2)/r_1(\beta_1)), \tag{7}$$

where β_1 and β_2 are angles of incidence of sound from the source on individual microphones in a pair, relative to their individual axes, $r_1(\beta_1)$ and $r_2(\beta_2)$ are respective directional sensitivities of two microphones (not included in Fig. 1).

The effect of the difference of distance can be included as:

$$\Delta L_R = L_{R2} - L_{R1} = 20\log(R_2/R_1). \tag{8}$$

Both differences are summed:

$$\Delta L = \Delta L_R + \Delta L_r. \tag{9}$$

Recording angles obtained for AB setups are presented in Table 1. They are close to presented in (WITTEK, 2002).

Table 1. Recording angles obtained for AB setups.

Distance between microphones [cm]	Recording angle
20	>180°
40	100°
70	52°
100	36°
300	12°
500	8°

2.4. Matlab implementation

A Matlab 7 script was developed to perform all calculations according to (6)–(9). The user entered geometrical data about microphones and the sound source and chose the type of directional characteristics. The software generated a matrix consisting of values of r in the α range $\langle 0,\ 90°\rangle$ and the corresponding values of phantom shifts. The selection of results is presented in Figs. 3–14 and 17–18 together with the results of experiments, for comparison.

3. Measurement of localisation curves

The experiments consisted of two parts. In part 1 sounds were recorded from different angles of incidence. In part 2, recorded sounds were reproduced through a stereo system and a group of experts estimated perceived positions of sound sources.

3.1. Recordings

Two pairs of Schoeps CCM4 condenser cardioid microphones (Schoeps website) were used for recording in XY and ORTF systems and a pair of Studio Projects C4 condenser omni-directional microphones (Studio Projects website) was used for the recordings in AB system.

The recordings were performed in an anechoic chamber at the Department of Mechanics and Vibroacoustics of the AGH University of Science and Technology in Cracow. Its critical frequency is 4 Hz.

The sound source was KRK Audio Rokit 5 active monitor (KRK website). The tweeter of the monitor and the microphones were placed at the same horizontal plane. The distance between the tweeter and the center of the microphone system was 1.6 m. The position of the monitor was fixed. The microphones were placed on a rotary table to change the azimuthal position of the sound source relative to microphones. The microphones were connected to Prism Sound MMA 4 XR preamplifier, and then through M-Audio Fast Track Pro interface to a computer with ProTools software. The following sounds were used: the white noise, the snare drum, the trumpet and the acoustic guitar. Sine waves were not used, as preliminary tests indicated that they were difficult to localise for most listeners. The instruments were recorded in a studio by the authors. The four sound samples were arranged in a sequence, where each of them was repeated a number of times. The same sequence of sounds was played and recorded for each of the following directions: 0°, ±5°, ±10°, ±15°, ±30°, ±45°, ±60°, ±75°, ±90°. The choice of angles was based on literature (Blauert, 1997; Moore, 1997; Kleczkowski *et al.*, 2008).

3.2. Estimation of subjective localisation

The organisation of the tests was based on experience gained from the pilot test. They were held in the same anechoic chamber. A pair of Yamaha HS50 active monitors was used for playback purposes. The sounds were played from a computer, placed in a separate control room, through M-Audio Fast Track Pro interface. The listener was seated at the distance of 1.6 m from each of the speakers. Tweeters were positioned at 1.2 m above the steel net floor which was close to the average position of the listeners' ears. The range of possible directions was discretised to 10 values to each side, marked by sheets of paper with numbers from 0 through +10 and −10. The reproduction system arranged for 30° stereo basis is shown in Fig. 2.

Thirty expert listeners participated in the experiment. They were students of Vibroacoustics and Audio Engineering course at the AGH University plus some musicians or audio engineers. All were in the age range from 20 to 30. The test consisted of a training part and a main part. In the training part listeners heard 10 samples of noise randomly distributed in the panorama. In the main part

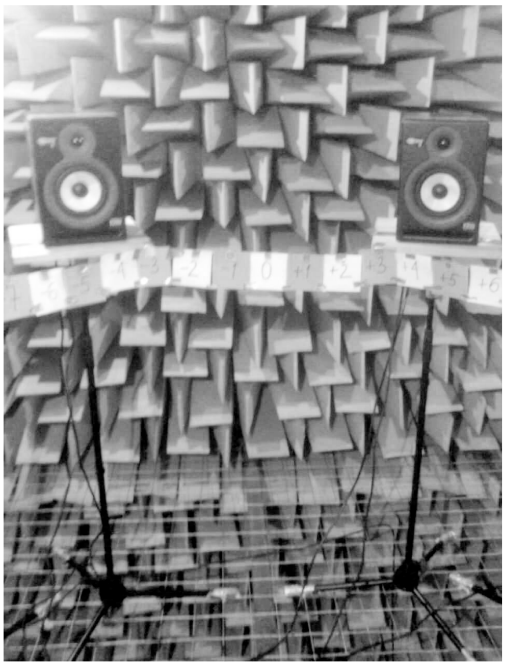

Fig. 2. The reproduction system for the stereo basis of 30°. The sheets of paper marking the directions can be seen.

the subsequent samples were also reproduced randomly from all of the recorded directions, and the number of samples was the same for all directions. The samples were separated by three second pause to give the experts some time to mark the answer in a scoring sheet. Each of the listeners estimated the same sequence of signals.

The main experiment was run in two series, A and B. Each series contained 30 repetitions of each sound. Within a series the number of samples reproduced at negative and positive angles was the same, and that number was doubled for the direction of 0°. Series A contained: AB and ORTF systems at angles in the $\langle 0°, 90° \rangle$ range plus binaural and XY systems recorded in $\langle -90°, 0° \rangle$ range. Series B contained: AB and ORTF systems at angles in the $\langle -90°, 0° \rangle$ range plus binaural and XY systems recorded in $\langle 0°, 90° \rangle$ range. Following the conclusions from the pilot test, each series consisted of two subseries lasting 12 minutes, separated by a 15 minute-break.

4. Results

4.1. Processing of results

The results for negative angles were included in the results for positive angles of the same magnitude, on the basis of symmetry of the hearing system. Thus all

results show the shift of phantom to the right as a function of positive angle of recording. Mean values of different data sets accompanied by respective standard deviations were computed. Mean values can be used as estimations of the subjective shift, while standard deviations are measures of perceptual uncertainty.

Table 2 shows all results, averaged for all subjects and all three stereo bases. It can be noticed, that uncertainty of localisation in AB, binaural and ORTF systems is higher for intermediate values of the input angle, while it is fairly uniform for the XY system. Table 3 presents distances between localisation curves, for the standard 60° stereo basis, as shown in Figs. 4 to 7. The values were obtained by computing mean square distances and then taking square roots. Within each technique, differences between the curves for three stereo bases were also investigated in the same way. The results are shown in Table 4. The distances in Table 4 (1.94% average value) are clearly lower than those in Table 3 (7.25% av. value).

Table 2. Phantom source shift and standard deviation averaged for all stereo bases.

Input angle [°]	Phantom source shift [%]							
	AB		Binaural		ORTF		XY	
	average	st. dev.	average	st. dev.	average	st. dev.	average	st. dev.
0	−3.95	9.81	4.59	10.98	−4.62	11.11	−1.13	11.26
5	11.32	12.36	16.16	11.59	8	11	−1.68	11.47
10	19.21	13.96	29.54	13.4	19.41	13.43	4.44	10.76
15	24.13	16.58	45	14.28	33.35	13.98	8.78	12.11
30	50.64	17.24	79.54	12.65	68.17	14.26	23.44	13.11
45	71.56	16.79	98.33	9.78	89.61	11.53	36.29	12.98
60	84.2	15.22	100.38	8.66	97.78	8.7	51.43	12.51
75	87.63	11.72	101.9	9.07	98.46	7.84	66.67	13.51
90	89.41	11.93	104.21	9.62	99.59	7.07	81.17	11.74

Table 3. Distance between localization curves for different techniques for 60° stereo base. DH stands for "dummy head", i.e. binaural technique.

Techniques	Distance between localization curves [%]
AB/DH	5.82
AB/ORTF	3.73
AB/XY	7.56
DH/ORTF	2.31
DH/XY	13.03
ORTF/XY	11.05

Table 4. Distance between localization curves [%] for different stereo bases.

	AB	Binaural	ORTF	XY
30–60°	1.8	1.4	1.96	1.85
30–90°	2.27	2.09	3.33	2.97
60–90°	0.84	1.09	2.06	1.57

4.2. Selected plots of results

Figures 3 to 14 have their ordinates scaled in % and named "phantom source shift". The scaling could also be in the units of angle, and the ordinate could be called "perceived source position". The first convention was assumed, as it is used in the literature of this subject.

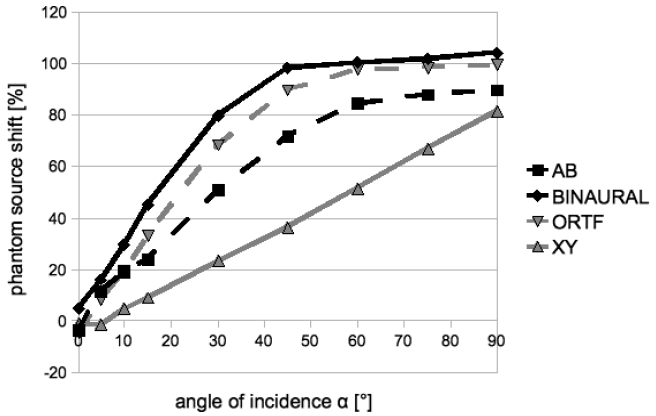

Fig. 3. Phantom source shift versus the angle of sound incidence for different stereo systems averaged for all stereo bases and for all sources.

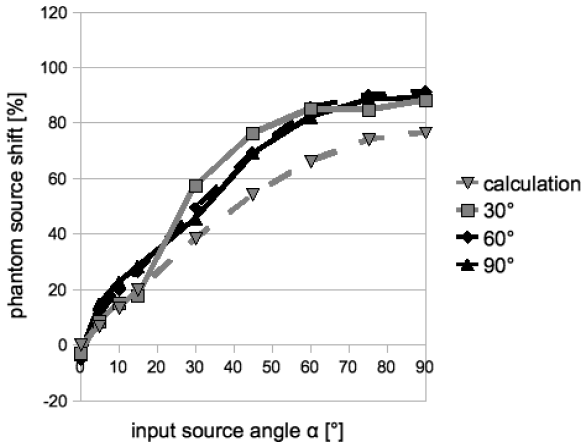

Fig. 4. Localization curves for AB (20 cm) for different bases averaged for all sound sources compared with calculated results.

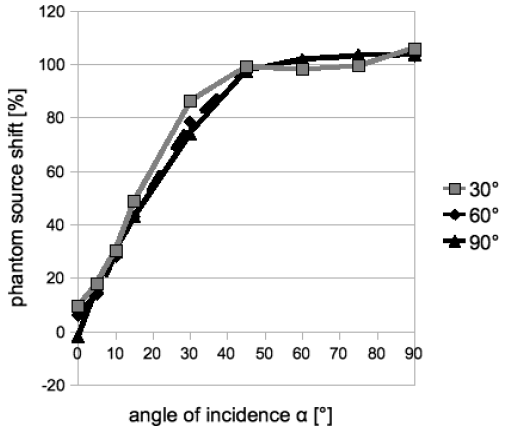

Fig. 5. Localization curves for binaural technique for different bases averaged for all sound sources compared with calculated results.

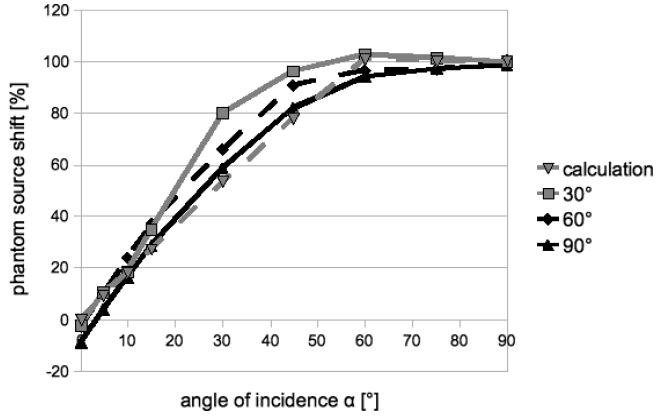

Fig. 6. Localization curves for ORTF for different bases averaged for all sound sources compared with calculated results.

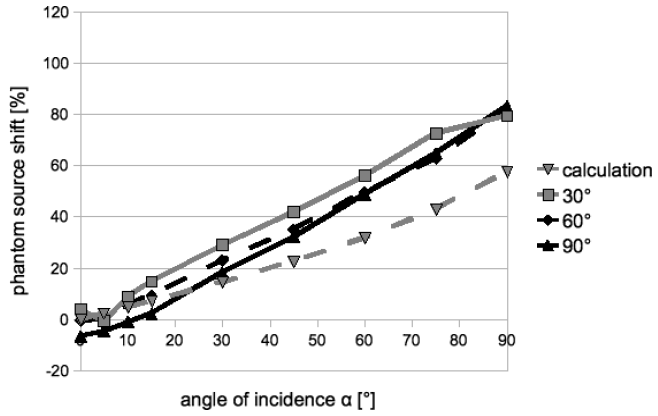

Fig. 7. Localization curves for XY (90°) for different bases averaged for all sound sources compared with calculated results.

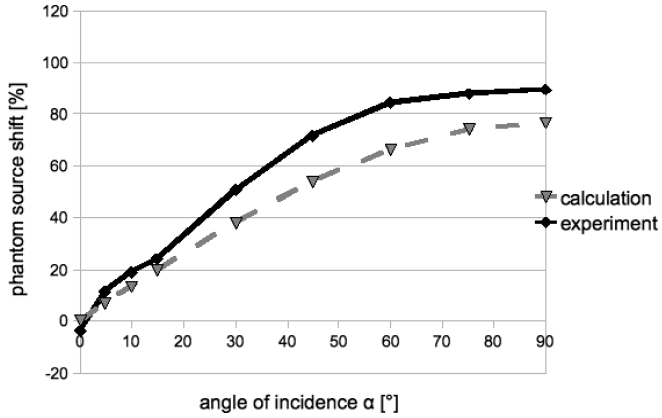

Fig. 8. Comparison between calculated and real localization curves averaged for all sources and bases for AB (20 cm).

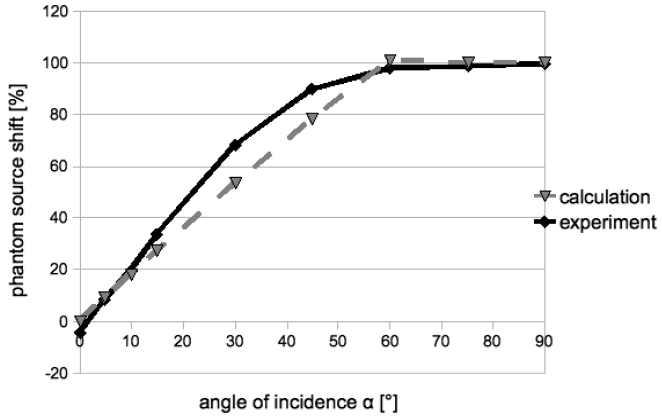

Fig. 9. Comparison between calculated and real localization curves averaged for all sources and bases for ORTF.

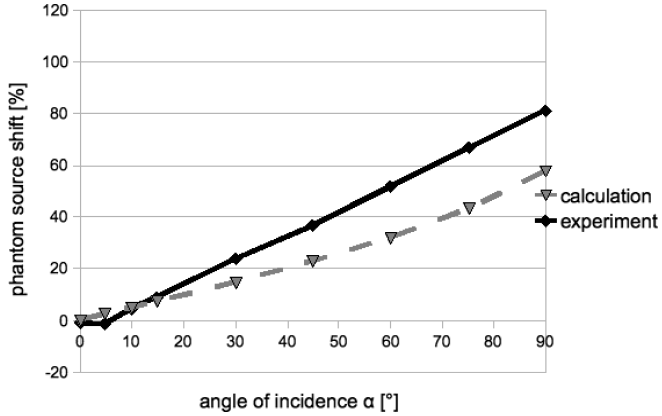

Fig. 10. Comparison between calculated and real localization curves averaged for all sources and bases for XY (90°).

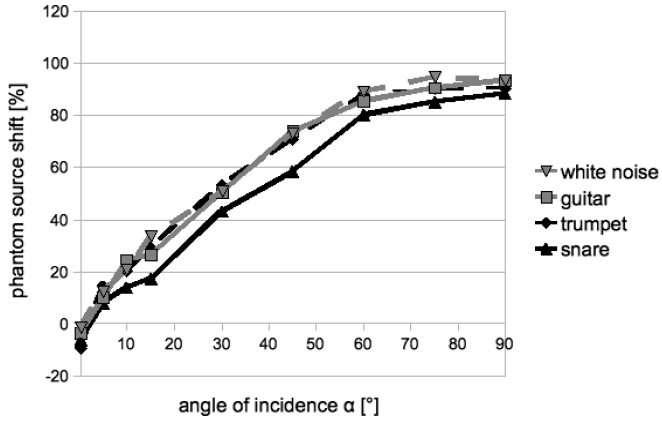

Fig. 11. Localization curves for AB (20 cm) for 60° basis separately for all sound sources.

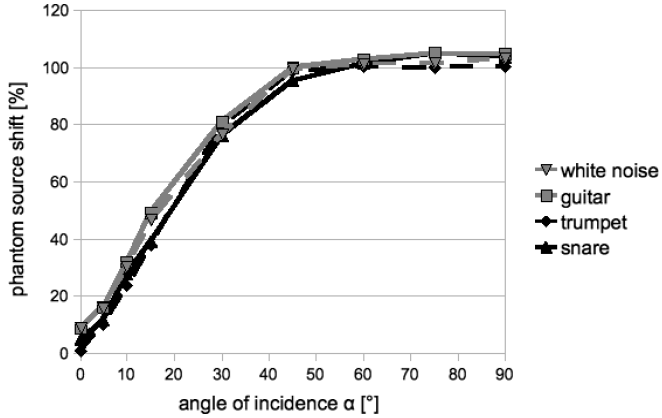

Fig. 12. Localization curves for binaural technique for 60° basis separately for all sound sources.

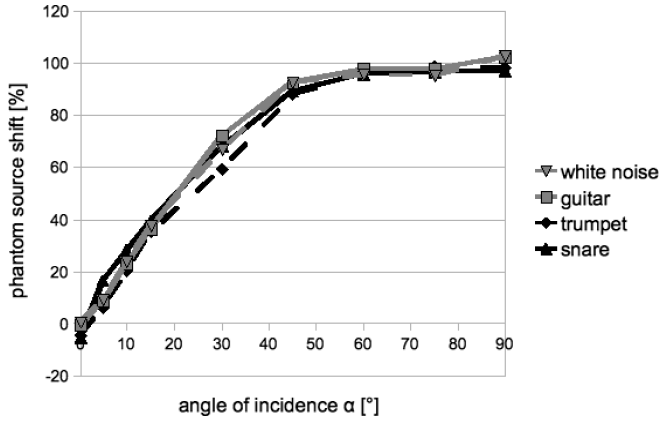

Fig. 13. Localization curves for ORTF for 60° basis separately for all sound sources.

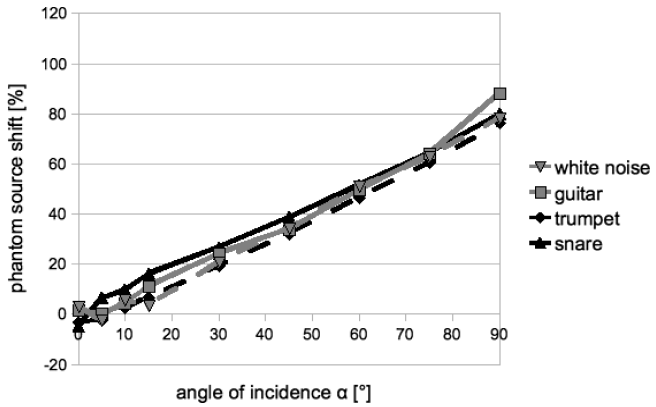

Fig. 14. Localization curves for XY (90°) for 60° basis separately for all sound sources.

Figure 3 shows shifts of the phantom source versus the angle of sound incidence relative to the axis of the microphone system, averaged for all stereo bases and for all sources. In Figs. 4–6 shifts in each stereo basis, averaged for all sources, are presented and compared with shifts obtained from calculations (for 60°). In Fig. 4 shifts in AB for all bases, averaged for all sources, are compared with calculated shifts. Figure 5 presents shifts in the binaural setup for each basis, averaged for all sources. Figure 6 shows shifts in ORTF for each basis, averaged for all sources, and calculated shifts; Fig. 7 presents the same type of results for XY setup. Figures 8–10 show comparison between calculated results and averaged for all sources and bases. The results shown in Figs. 11–14 demonstrate shifts for the standard 60° basis separately for all sound sources, respectively for AB, binaural, ORTF and XY systems.

5. Discussion and conclusions

5.1. General conclusions

The results of calculations and experimental determination of localisation curves bring the following conclusions:

a) localisation curves for particular techniques are different,
b) the results for stereo bases of 30°, 60° and 90° are coherent, as shown in Tables 3 and 4,
c) overall shapes of curves from calculations and obtained experimentally are similar up to some value of the angle of incidence of the sound,
d) results for different sound sources are coherent, no rule of dependence between the type of sound and the shape of localisation curve can be found,
e) recording angles obtained from the experiment are coherent with Williams' results (Williams, 1984, 1987).

An important conclusion is that RA becomes narrower, when the angle between microphones 2β and/or the spacing between them d is increased (Fig. 1). This was shown by our calculations and is presented in Table 1 and Figs. 17 and 18. These calculations were based on the model described by Eqs. (6)–(9). The validity of the model was confirmed by our experimental results presented in Figs. 3–14, therefore they also indirectly prove the above statement. The consequences are outlined in Subsec. 5.2.

The XY system yields the most linear localisation curve in the entire range $\langle 0, 90^\circ \rangle$. This can be attributed to the RA being the widest for this system (see Figs. 3 and 7).

In other techniques the RA was narrower than the range of analysed angles. Wherever the direction of a sound source exceeds RA the source is localised only in one loudspeaker (see Figs. 3–6, 8–9, 11–13). The RA is the narrowest (about $90^\circ = 2 \times 45^\circ$) in the binaural setup (see Fig. 5).

Localisation curves for ORTF and binaural systems are similar, which proves that ORTF is useful as a technique simulating the positions of ears (see Figs. 3 and 6).

Some ripples can be noticed in curves for AB system in the $\langle 0^\circ - 15^\circ \rangle$ range, regardless of the width of the reproduction basis. This partly confirms the findings of Hugonnet and Jouhaneau (1987). AB setup is purely based on ITD. Presumably, small differences in distance between each of the ears and loudspeakers increases uncertainty of localisation (see Fig. 4).

5.2. Application to sound recording practice

The main objective of the proper placement of a microphone system is that the reproduced music scene is spread evenly between the loudspeakers and its ends are placed exactly in the loudspeakers. Some popular notions are shared by engineers, they can also be found in audio literature (Sztekmiler, 2001; Owsinski, 2005). Our results are in contrast with these opinions. Although there is a fundamental difference between these experiences gained in the reverberant field, and ours obtained in free field conditions, but nevertheless our results indicate that popular notions should be treated with caution.

Notion 1: "In coincident and semi-coincident techniques the axis of the microphone system should point at the ends of the music scene." Figure 15 demonstrates this situation, with the assumption of free field conditions. The RA is much wider, so during playback the music scene will be perceived as narrow (Fig. 16). Therefore, when the stereo image evenly distributed between loudspeakers is the objective, we need to adjust the angle and/or distance between microphones or decrease the distance from the music scene so that RA matches our desired area. In the example of Fig. 15 the angle between axes of microphones should be wider than 90°, so that the RA is narrowed to be matched to the width of the music scene. The same mechanism functions with other coincident and semi-coincident

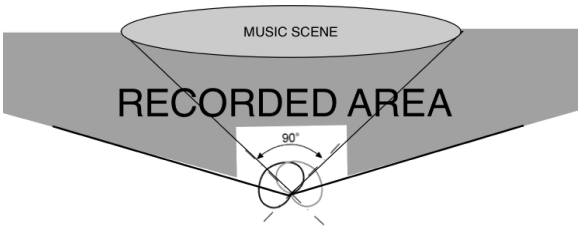

Fig. 15. An XY setup where microphone's axes point at the ends of the music scene.

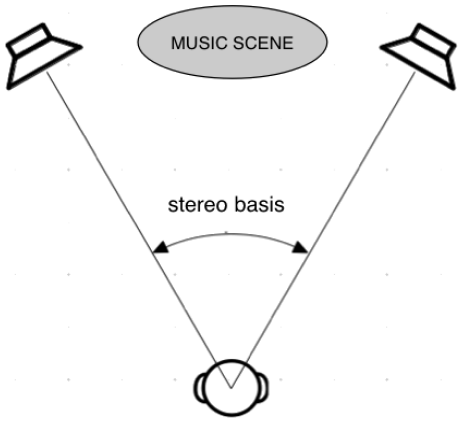

Fig. 16. Reproduction of music scene during playback. Recording angle of the microphone system was too wide so the music scene is perceived as narrow.

techniques (e.g. Blumlein and ORTF). However it should be kept in mind that a sound source placed on the microphones' axis will be perceived as closer than others.

Notion 2: "The distance between microphones in AB system should be about 3 to 5 meters." By referencing Table 1 we can conclude that with such a setup only a narrow area can be reproduced or the stereo image will consist of a narrow central part, then a gap, and many sources located in the very loudspeakers. Should this problem be avoided by increasing the distance from the music scene the microphones would have to be located in the reverberant field. A pair of omnidirectional microphones 3 to 5 meters apart does not work as a stereo microphone technique, where the sense of direction is a primary goal but could be used as far field microphones.

Notion 3: "Localization in spaced techniques is very inconsistent." As we showed in 2, it can be perceived as inconsistent if improperly used.

Notion 4: "The bigger the music scene the wider the angle between microphones' axes and/or the bigger the distance between microphones." The results of our calculations presented in Figs. 17 and 18 demonstrate that in the free field case this is false. The recording of big ensembles usually requires the positioning of a microphone system at an appropriately large distance, but the wider the an-

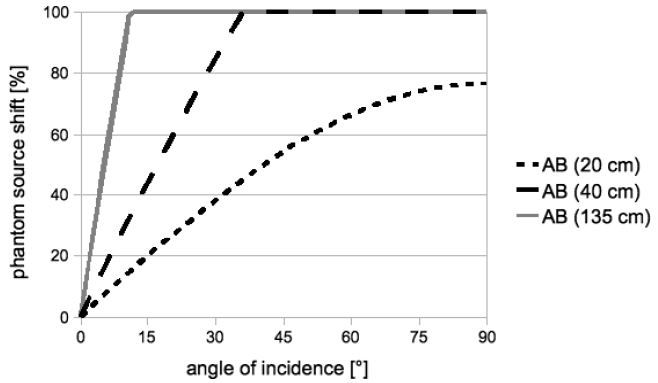

Fig. 17. Localization curves for AB setup calculated for different distances between microphones.

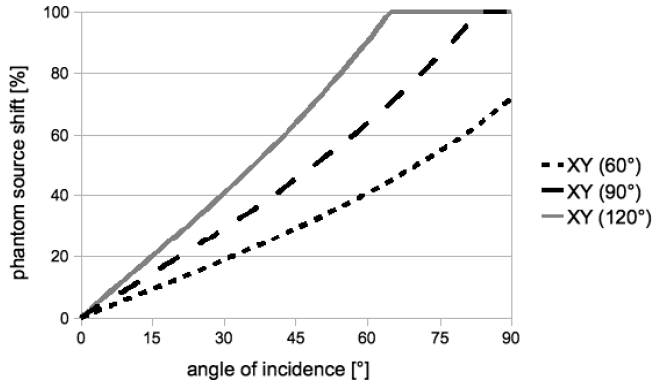

Fig. 18. Localization curves for XY setup calculated for different angles between microphones' axes.

gle between microphones' axes the narrower RA becomes, and so is the musical scene that can be evenly distributed.

Acknowledgments

This research was supported by the AGH University of Science and Technology grant no. 11.11.130.885. The authors would like to thank two anonymous reviewers for their helpful comments.

References

1. Blauert J. (1997), *Spatial hearing: the psychophysics of human sound localization*, MIT Press, Cambridge.

2. Hugonnet C., Jouhaneau J. (1987), *Comparative spatial transfer function of six different stereophonic systems*, AES 82nd Convention, Preprint 2465.

3. Kleczkowski P., Plewa M., Pyda G. (2008), *Localization of a sound source in Double MS recordings*, Archives of Acoustics, **33**, 4 (Supplement), 147–152.
4. KRK – www.krksys.com, access 15.01.2011.
5. Moore B. (2003), *An introduction to the psychology of hearing*, Academic Press, London.
6. Neumann – www.neumann.com, access 15.01.2011.
7. Owsinski B. (2005), *The recording engineer's handbook*, Course Technology PTR, Boston.
8. Schoeps – www.schoeps.de, access 15.01.2011.
9. Studio Projects – www.studioprojectsusa.com, access 15.01.2011.
10. Sztekmiler K. (2008), *The basics of sound reinforcement and audio production* [in Polish: *Podstawy nagłośnienia i realizacji nagrań*], Wydawnictwa Komunikacji i Łączności, Warszawa.
11. Williams M. (1984), *The Stereophonic Zoom: A Practical Approach to Determining the Characteristics of a Spaced Pair of Directional Microphones*, 75nd AES Convention, Preprint 2072.
12. Williams M. (1987) *Unified theory of microphone systems for stereophonic sound recording,* 82nd AES Convention, Preprint 2466.
13. Wittek H., Theile G. (2000), *Investigations into directional imaging using L-C-R stereo microphones*, 21.Tonmeistertagung 2000, Proceedings ISBN 3-598-20362-4, p. 432–454.
14. Wittek H., Theile G. (2002) *The Recording Angle – based on localization curves*, AES 112th Convention Paper, Preprint 5568.
15. Wittek H. (2002) *Image Assistant*, JAVA-Applet and documentation on website: www.hauptmikrofone.de, acces 15.01.2011.