The accuracy and precision of localisation of instruments and vocals found in a typical pop or rock band in an Ambisonics-based virtual environment

Marko Horvat, Frane Anzulovic, Dominik Kisic, Kristian Jambrosic

THE ACCURACY AND PRECISION OF LOCALISATION OF INSTRUMENTS AND VOCALS FOUND IN A TYPICAL POP OR ROCK BAND IN AN AMBISONICS-BASED VIRTUAL ENVIRONMENT

Marko Horvat Frane Anzulović Dominik Kisić Kristian Jambrošić
University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
marko.horvat@fer.hr

ABSTRACT

This paper aims to explore the degree of accuracy and precision of the localization of instruments and vocals that are typically found in contemporary pop or rock bands, when the mix of a recording of such a band is to be made for an Ambisonics-based virtual environment, rather than for the usual two-channel stereo reproduction. For this purpose, a dry recording of a single song was acquired, containing the tracks with drums, bass guitar, lead guitar, keyboards and vocal. A short excerpt of the recording and, consequently, of the corresponding tracks was cut out from the original recording. A virtual environment was created based on Ambisonics encoding. Within this environment, seven positions in the front horizontal half-plane were determined for the placement of the source. Each of the shortened tracks was used as a source signal on all seven positions and the encoding was made in first-, second- and third-order Ambisonics. A series of listening tests was made, in which the subjects were asked to localize the source with its content reproduced to them through a loudspeaker system capable of reproducing Ambisonics signal up to the third order. The results of the tests were analysed in the context of how accurately and precisely the listeners could localize the sources, with regard to source position, the content of the source, and the order of the encoded signal.

1. INTRODUCTION

Ambisonics is a form of spatial audio (re)production proposed by Gerzon [1,2]. What makes if different from other systems designed for this purpose is that the audio information is processed in two stages; i.e. the encoding stage and the decoding stage. In the encoding stage, a sound environment can be recorded or synthesized through a set of functions known as spherical harmonics. The highest order of spherical harmonics used in the encoding stage determines the final encoding order of the Ambisonics-encoded material. As a rule, higher order encoding should improve localization accuracy and precision. On the other hand, lower order encoding introduces diffusivity into the overall sound image, and the localization accuracy and precision deteriorates, causing a spatial blur [3]. In the decoding stage, the audio data can be converted and adapted for practically any kind of reproduction system with a reasonably regular configuration, ranging from mono and stereo systems to 5.1, 7.1 or similar setups, all the way to full-3D reproduction. Decoding for binaural listening is available as well. Therefore, Ambisonics offers unprecedented flexibility, compared to all other audio (re)production systems. The highest order of the Ambisonics audio a sound reproduction system will be able to reproduce directly depends on the number of available loudspeakers. With N being the desired Ambisonics order, the minimum number of loudspeakers for 2D reproduction is $2N+1$, and for 3D reproduction this number should be $(N+1)^2$.

To date, many different subjective assessments regarding the performance of Ambisonics systems have been made. In [4], source perception in higher-order Ambisonics systems is investigated in light of localization accuracy, perceived source width and the overall realistic impression. The stimuli includes speech, pure tones and a musical environment synthesized with multiple instruments. The results speak in favour of third-order systems, regarding their quality, size and localization. In [5], localization accuracy was investigated from the viewpoint of different microphones capable of recording first- and higher-order Ambisonics audio. The results show that the localization accuracy depend both on the Ambisonics order and on the direction of arrival of a sound. In [6], localization has been investigated using Ambisonics decoders of different orders and with different methods of spatial smoothing. The listening test was made both in and out of the sweet spot of the reproduction system. The results show an improvement of localization as the order increases, and better localization in the sweet spot. The method of spatial smoothing also has an influence on localization, with the in-phase decoder as the worst option.

This paper presents the investigation of localization accuracy and precision from the possible viewpoint of a composer or a mixing engineer, whose task is to place instruments and vocals into an Ambisonics-encoded auditory scene.

In light of this, contemporary music has been chosen as the stimuli for the experiment, rather than hypothetic signals such as pure tones, pulses or broadband noise. Namely, the samples of instruments and vocals typically found in a pop or rock band were used. The purpose of the research is to examine whether the position of real musical sources in a spatial audio mix is perceived correctly, i.e. at the positions they were placed to by the mixing engineer.

The independent variables in the experiment are the type of source, the position of the source, and the Ambisonics order of the encoding used to place the source to the required position. The dependent variable is the perceived azimuthal position of the source, as reported by the listeners.

2. THE LISTENING TEST

2.1 Setup

The listening experiment was conducted in the Auralization laboratory of the Faculty of electrical engineering and computing in Zagreb. The laboratory hosts a spatial audio reproduction system based on Ambisonics coding and decoding scheme.

The listening room, in which the loudspeaker system is mounted, is separated from the control room, where the rest of the hardware is located. A two-way audio connection has been established between the two rooms, consisting of a loudspeaker and a microphone in the listening room (used by the listener), and a microphone and a pair of headphones in the control room (used by the operator). This setup completely prevents the formation of acoustic feedback. At this time, no video communication is installed.

The system consists of 16 loudspeakers in a 4-8-4 configuration. The horizontal plane is covered with eight evenly spaced loudspeakers (45° apart, starting at 0° as the frontal direction). The extension to three dimensions is achieved by placing four loudspeakers above the horizontal plane, namely at the elevation of 40° and the azimuths of 45°, 135°, -45° and -135°. The lower half-space is covered by another four loudspeakers at the elevation of -40°, at azimuths stated above. All loudspeakers are positioned on a sphere with the radius of 1.65 meters. The described configuration is capable of reproducing third order Ambisonics audio in the horizontal plane, and second order Ambisonics audio in three dimensions.

Due to inherent limitations of the spatial audio reproduction system, the "sweet spot" of the system is small, allowing only one listener at a time to take the test. Therefore, only one fully rotatable barstool is permanently installed in the room. The barstool was chosen instead of a regular chair for its height, so that the listener's head would ultimately be located in the horizontal plane.

The software part of the system consists of REAPER DAW software [7] with IEM Plug-in Suite [8], an open source package dedicated to Ambisonics processing.

The setup of the experiment is shown in Figure 1. Only the horizontal plane is shown, because the virtual sources were placed in the horizontal plane only. The green circles represent the positions of loudspeakers in the horizontal plane. The blue circles show the positions (azimuths) of virtual sources, namely, 90° to the left, then 60° and 30° to the front left, then 0° straight in front of the listener, then -30° and -60° to the front right, and, finally, -90° to the right of the listener. As shown in the figure, the defined setup yields three virtual sources that coincide with actual loudspeakers (90°, 0°, and -90°), and four that do not (60°, 30°, -30°, and -60°).

To facilitate the assessment of the perceived azimuthal direction of the virtual source, a paper tape (in red) was stretched along the circumference of the horizontal circle that contains the loudspeakers. To avoid a skewed distribution of the reported azimuths for the two outermost virtual sources, the tape covered the extended azimuthal range from 120° to -120°. The markings on the tape were made in 1° resolution, and were made to resemble a measuring tape tool; i.e. the 10°-marks were made across the full width of the tape, the 5°-marks were made across the 2/3rd of the width, and the rest were made across the 1/3rd of the width of the tape. To avoid confusing the listeners, the values indicated on the tape ranged from 0° at azimuth 120° to 240° at azimuth -120°, thus increasing from left to right and avoiding negative values altogether. The values reported by the listeners were transformed into the standard coordinate system in the data processing stage.

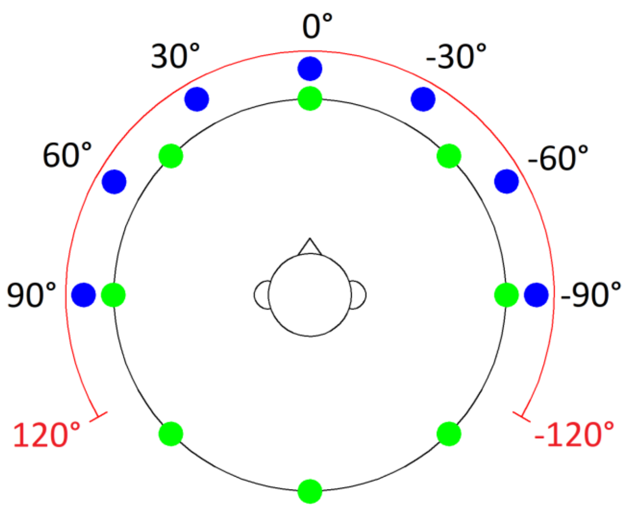

Figure 1. The positions of the loudspeakers in the horizontal plane of the reproduction system (green), the positions of the virtual sources (blue); the range of azimuths used by the listeners (red)

2.2 Stimuli

An excerpt from a raw recording of a rock band was found in a freely available online database [9]. Originally, it consisted of 12 tracks. Seven tracks covered the drums, and the rest were the bass guitar, two electric guitars, the piano, and the vocal. The instruments and the vocal were recorded one by one, so their recordings were quite clean, with no spill coming from other instruments. To yield the final stimuli that was used in the listening experiment, the drums tracks were mixed together to the final mono downmix, and all the other tracks were taken “as is”. Only one electric guitar was selected. In the end, five different mono source samples were obtained: the bass guitar, the drums, the electric guitar, the piano, and the vocal. All samples were cut to 10 seconds, with a short fade-in and fade-out.

To form a test sequence for the listening test, each of the five source samples was encoded as the first, second or third order Ambisonics signal at one of seven defined positions in the horizontal plane. Therefore, the test sequence consisted of 105 encoded samples. Five additional training samples were added to the beginning of the test sequence. The test sequence was formed so that each 10-second sample was followed by 10 seconds of silence, thus yielding the total length of the test sequence of 37 minutes. To avoid systemic errors, the samples were randomly ordered within the test sequence, and three different sequences were made.

2.3 Listeners

A total of 20 listeners took part in the listening test, of which 5 were women, and 15 were men. All listeners have a background in audio, sound, and music. The overall average age of the whole group is 32 years. However, it should be mentioned that 11 listeners were students in the age range from 20 to 23 years old, while the rest of the group consisted of faculty employees, most of which are in their late 30s and 40s.

Initially, all listeners reported to have normal hearing. To test their clams, the responses of each listener were tested by assessing the histogram of all the reported values. The values were first transformed to yield the deviation of the reported azimuth from the real azimuth of a virtual source. For a person with normal hearing, the expectation was to obtain a symmetric distribution around zero, as shown in Figure 2.

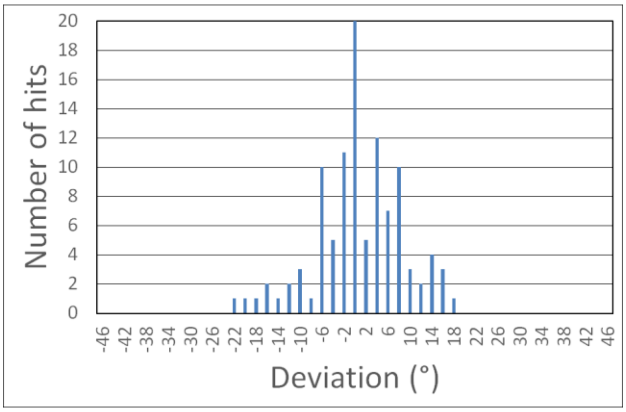

Figure 2. The distribution of the deviation of the reported azimuth from the real azimuth of the virtual source for a listener with normal hearing

However, during this analysis, the histogram was obtained for one of the listeners as shown in Figure 3, suggesting that the perception of direction for this listener is consistently shifted to the right by an average of 12°. The listener ultimately reported to have tinnitus in their left ear, and their test results were discarded.

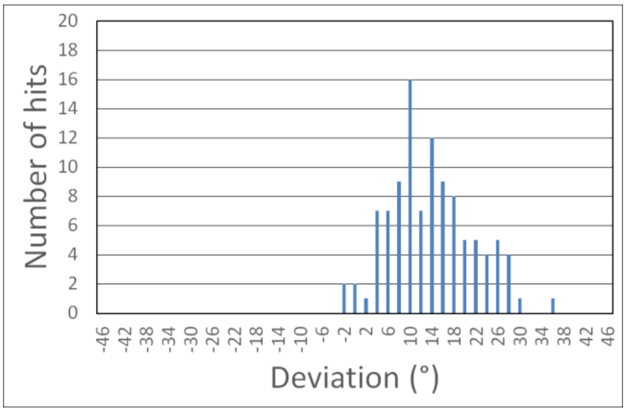

Figure 3. The distribution of the deviation of the reported azimuth from the real azimuth of the virtual source for a listener with impaired hearing

2.4 Procedure

The listeners were tested individually. At the beginning of the test, the listeners were informed on the nature of the experiment. It was asked of them to try to locate the presented sound sources as accurately as possible, using the tape measure with marked angles.

They were advised to close their eyes while performing the task and to point their hand in the desired direction. This way, all the distractions in the listening room were avoided. They were allowed to rotate on the barstool. Once they were sure that the chosen direction is the right one, they could open their eyes and simply read the angle value on the tape in the direction in which their hand was pointing. To allow the listeners to concentrate solely on their task, they were not required to record the values by themselves in any way, e.g. by writing them down, thus leaving their hands free. Instead, they were asked to use the audio communication system by uttering the values during the 10-second period of silence between two sound samples. The operator in the control room constantly monitored the open audio link to the listening room, and repeated the values back to the listener. If the read-back value was the equal to the one originally uttered by the listener, the operator wrote it down into the test form. If not, the listener warned and corrected the operator, and the correct value was written down.

Before engaging in the test, five training samples were presented to the listeners, thus familiarizing them with the test procedure. Simultaneously, the listeners were acquainted with all five types of instruments/vocals that would be repeatedly used in the rest of the test.

The entire test was performed with no breaks. Afterwards, the listeners gave the feedback to the operator, claiming without exception that the test was not physically or mentally demanding.

3. RESULTS AND DISCUSSION

As 20 listeners evaluated 105 different sound samples (5 instruments x 7 source positions x 3 Ambisonics encoding orders), a total of 2100 observations was collected.

The raw observations were azimuthal angles in the interval [0°, 240°], increasing clockwise. To match the standard coordinate system (x-axis points to the front, y-axis points to the left), the data was transformed to match the original azimuth range of [-120°, 120°], that increases in the counterclockwise direction.

The data was further transformed so that is reflects the deviation of the perceived azimuth from the assigned azimuth of a source. A negative value of the deviation implies that the listener perceives the sound source to the left of its "true" position, whereas a positive value of the deviation implies that the perceived position of a source is to the right of its assigned position.

The defined dependent variable that represents the described deviation was used to test both the accuracy and the precision of the localization of the sound sources, as displayed by the listeners. The mean value was taken as the measure of accuracy, and the variance as the measure of spread was taken as the descriptor of precision. Ideally, if all the listeners were able to locate the sound source perfectly, the mean value of the deviation would be zero, and the variance would be equal to zero as well.

3.1 Normality of the data

In the initial stage, the data was checked for normality by means of the Lilliefors normality test. Not only was each of the 105 test cases checked, but also the data grouped by instrument (5 groups), azimuth (7 groups), and order (3 groups). The normality tests revealed that the observations are normally distributed in about 65% of the tested cases and groups. The ones that did not pass the test were inspected visually by means of histograms, and it was determined that there are no severe violations of the normal distribution.

3.2 Accuracy of localization

In the second stage, the localization accuracy was tested by means of a one-sampled t-test. First the two-tailed test was made, with the null hypothesis was that the sample mean of the deviation for each of the 105 cases was not significantly different from zero, against the alternative hypothesis that it is. The results of this test reveal that the null hypothesis can be rejected at the 0.05 significance level in 47 out of 105 cases. Next, both lower tail and upper tail one-tailed tests were made as well. The null hypothesis was the same as before. The alternative hypothesis for the lower tail test was that the sample mean is lower than zero, indicating a significant perceived shift of the source to the left of its true position. Similarly, the alternative hypothesis for the upper tail test was that the sample mean is larger than zero, indicating a significant perceived shift of the source to the right of its true position. To match the two-tailed test, these tests were made at the 0.025 significance level. The results show a statistically significant shift to the left in 14 cases, and to the right in 33 cases. The count of significant deviations from zero, made according to instrument, is the following: 8 cases for bass and vocal, 10 cases for drums and piano, and 11 cases for the guitar. The same count made taking into account the direction of arrival reveals the following: one case for 90°, 11 cases for 60°, 15 cases for 30°, 9 cases for 0° degrees, 2 cases for -30°, 6 cases for -60°, and 3 cases for -90°. The count made according to the encoding order is the following: 12 cases for first order, 17 cases for second order, and 18 cases for third order.

The results indicate that the accuracy of localization does not depend much on the type of sound, i.e. the instrument/vocal. On the other hand, the direction of arrival of sound, as well as the encoding order seem to have an influence on the localization accuracy.

The surprising finding is that the first order Ambisonics encoding is apparently the most accurate one, and that the source placed right in front of the listener, i.e. at 0°, will significantly deviate from that direction, as perceived by the listener. Additionally, all cases with sources placed at 30° show a significant deviation to the right from that direction, while most of the cases with sources at 60° indicate a significant deviation to the left.

To explain these findings, the accuracy indicated by the means has to be observed in light of precision indicated by the standard deviations. Specifically, if the spread of data represented by standard deviation is large, even a large deviation from the true direction might turn out not to be statistically significant. On the other hand, a small spread of data indicated by a small standard deviation will make even a rather small deviation from the true direction statistically significant.

To accompany the results of the t-test, the sample means for all 105 cases are shown in Table 1, and the corresponding standard deviations are shown in Table 2.

The data confirms that the frontal direction generally has the smallest spread of observations, yielding a number of cases with small mean deviation declared as statistically significant. The comparison of standard deviations by encoding order reveals that the first order is indeed the least precise, having the largest spread of observations. As stated above, this can lead to a misleading conclusion that it is the most accurate.

Given the fact that the sources placed at 30° subjectively significantly deviate from that direction in all cases leads to the conclusion that there is a possibility that some kind of systemic error has occurred. Detailed examinations of the listening setup have found nothing that supports this conclusion.

		Instrument				
Order	Direction	Bass	Drums	Guitar	Piano	Vocal
1	-90	-0.8	-3.0	**-6.4**	**-5.7**	-5.0
1	-60	1.1	-2.1	-3.5	-0.2	-2.9
1	-30	0.7	2.5	1.5	-0.9	3.1
1	0	1.9	**4.3**	**4.9**	3.1	2.6
1	30	**8.0**	**8.1**	**11.0**	**10.9**	**9.1**
1	60	-0.3	0.6	**7.0**	**6.6**	-0.4
1	90	-0.6	**-5.3**	5.8	3.8	0.0
2	-90	**-3.0**	-2.1	-1.1	-3.9	-2.6
2	-60	2.2	**3.3**	3.8	**2.7**	3.2
2	-30	-0.9	-0.3	2.7	-0.2	**8.1**
2	0	**4.3**	2.8	**4.5**	**4.9**	**2.7**
2	30	**7.4**	**8.4**	**11.7**	**8.0**	**9.4**
2	60	**-15.1**	**-9.6**	**-5.4**	-3.2	**-7.5**
2	90	-1.9	-0.8	-0.9	2.9	-1.1
3	-90	0.1	0.6	-2.2	-1.7	-1.9
3	-60	3.0	**7.9**	**5.3**	**5.2**	**8.5**
3	-30	**-2.6**	0.1	1.9	0.6	-1.2
3	0	2.7	**3.5**	**4.6**	**3.1**	-0.4
3	30	**10.0**	**9.7**	**10.4**	**7.6**	**7.0**
3	60	**-7.4**	**-12.7**	**-7.7**	**-6.4**	**-11.2**
3	90	-1.4	-2.1	-0.4	0.8	-1.5

Table 1. The means of the deviation from true direction of arrival; values in bold indicate statistically significant cases.

3.3 Precision of localization

In the third stage of statistical analysis, attention was turned to evaluating the precision of localization. To achieve this, the hypothesis of equal variances was tested by means of the Levene's test for equality of variances. This test is reasonably tolerant to deviations from normality and to the existence of outliers. The null hypothesis is that the variances of all tested groups are equal. In this particular case, the null hypothesis states that the independent variable (instrument, direction, or encoding order) has no influence on the precision of localization indicated by the variance of the data. The alternative hypothesis is that the variances are not equal, i.e. that at least one group is significantly different in this respect from the other ones.

The Levene's test was performed on the data grouped by instrument, direction, and encoding order. The results of this test are shown graphically in Figures 4, 5 and 6.

		Instrument				
Order	Direction	Bass	Drums	Guitar	Piano	Vocal
1	-90	10.6	11.9	**12.1**	**10.3**	11.8
1	-60	8.7	11.3	11.2	12.8	13.6
1	-30	8.2	6.7	7.4	9.3	9.2
1	0	9.3	**8.9**	**7.1**	10.1	6.6
1	30	**10.8**	**12.5**	**12.7**	**12.8**	**14.1**
1	60	9.7	9.5	**10.3**	**10.8**	13.3
1	90	13.5	**6.1**	13.5	15.2	10.8
2	-90	**6.3**	5.1	5.6	9.4	6.5
2	-60	8.6	**6.5**	10.5	**5.6**	10.6
2	-30	7.1	6.6	5.9	6.2	**12.1**
2	0	**5.4**	6.3	**6.6**	**6.4**	**4.8**
2	30	**10.7**	**9.4**	**8.2**	**9.8**	**9.5**
2	60	**12.2**	**9.3**	**9.0**	10.0	**10.0**
2	90	5.3	5.1	5.5	6.8	7.9
3	-90	3.4	5.8	5.4	3.9	5.2
3	-60	8.4	**5.9**	**8.1**	**7.2**	**11.8**
3	-30	**5.4**	5.1	6.4	5.6	6.2
3	0	8.8	**4.0**	**5.3**	**3.9**	7.5
3	30	**8.5**	**9.2**	**6.9**	**10.2**	**7.8**
3	60	**7.5**	**7.0**	**10.9**	**7.8**	**8.7**
3	90	3.6	5.0	5.9	5.9	3.8

Table 2. The standard deviations of the deviation from the true direction of arrival; values in bold indicate statistically significant cases to accompany the means in Table 1.

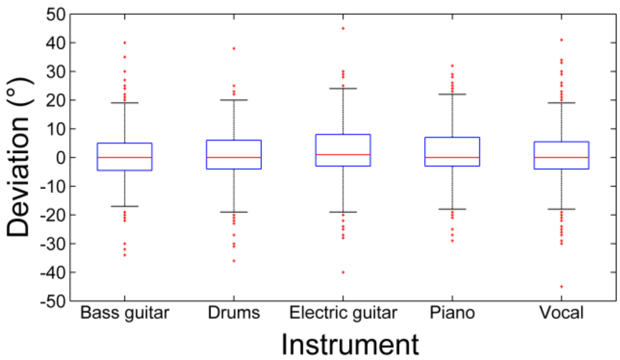

Figure 4. The boxplot of the data grouped by instrument to complement the corresponding Levene's test

The Levene's test on the data grouped by instrument did not prove to be statistically significant at the 0.05 level, indicating that the null hypothesis that the variances of these five groups are equal cannot be rejected. Therefore, the conclusion of the test is that the type of intrument does not have an influence on the precision of its localization in the given setup. The boxplot shown in Figure 4 supports this conclusion.

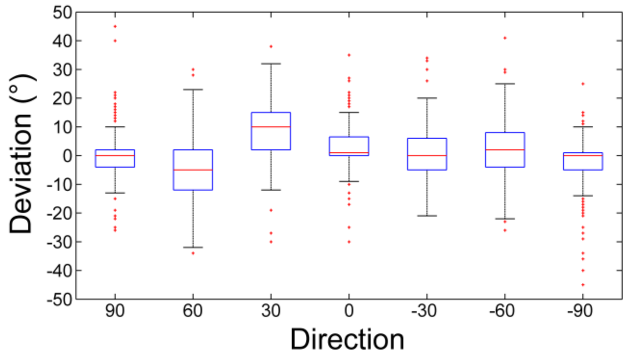

Figure 5. The boxplot of the data grouped by direction to complement the corresponding Levene's test

The Levene's test on the data grouped by direction was statistically significant at the 0.05 level, indicating that the assigned direction has a strong influence on the precision of localization. The boxplot shown in Figure 5 reveals that all three assigned directions that coincide with the directions of actual loudspeakers (90°, 0° and -90°) have a much smaller spread of observations than the remaining four directions (60°, 30°, -30° and -60°). This finding confirms the theory that lies behind Ambisonics systems and their predicted behaviour.

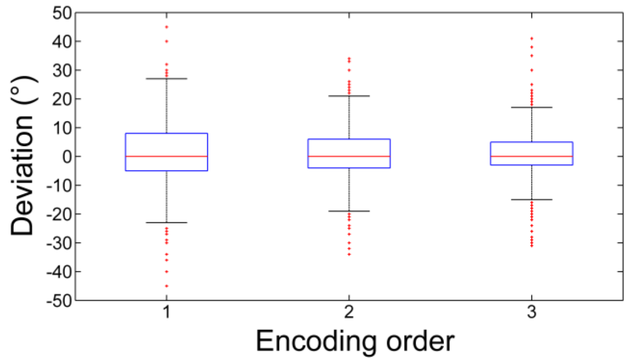

Figure 6. The boxplot of the data grouped by encoding order to complement the corresponding Levene's test

The Levene's test on the data grouped by encoding order was statistically significant at the 0.05 level, indicating that the encoding oder has an influence on the precision of localization. The boxplot shown in Figure 6 the spread of observations is smaller for a higher encoding order, indicating that the precision of localization indeed increases with increasing encoding order. This finding is also in agreement with the behaviour of Ambisonics-based systems.

3.4 Limitations

The limitations of the study are mainly connected to the data collection method and to the limitations imposed by the laboratory setup. The listeners were asked to close their eyes and to open them again only after they had chosen the perceived direction of the source. Nevertheless, they often reported that them being able to see the loudspeakers was a bit distracting. Therefore, to improve similar estimations in the future, a cylinder of suffcicient size, made of acoustically transparent fabric, shall be placed around the listener position. In this manner, a visual barrier will be created that will prevent the listeners to see any part of the sound reproduction system. The data collection method that utilizes the paper tape with indicated angles can only be used in the horizontal plane, with no possibility of extending it to all three dimensions by assessing the elevation as well. Therefore, a more adequate method needs to be devised before any localization experiments are to be made in full 3D space. Finally, many listeners reported the sound coming from above the horizontal plane, even though the position of the sources was set to horizontal plane only. As it was beyond the scope of this study to study this effect, and the laboratory setup was not equipped for assessing the perceived elevation of the source, the listeners were advised to report only the perceived azimuth.

4. CONCLUSIONS

The research presented in this paper was focused on examining whether or not accuracy and precision of source localization in an Ambisonics-based mix is affected by the type of the source, its assigned direction and the Ambisonics order used in the encoding stage. The hypothetical mix is made of sounds of instruments and vocals typically found in a rock band, and the sources are placed in the front half of the horizontal plane.

The results reveal that the bass guitar and the vocal are the sounds that are the easiest to locate accurately, followed by the drums and the piano, while the electric guitar is the one most difficult to place to its intended position.

The analysis of accuracy dependent of the true position of the source yielded a surprising result: the sounds coming from the loudspeakers directly to the left and to the right of the listener were the one that were located the most accurately. In general, the sounds coming from the front right quadrant were located accurately, while the sounds coming from the left quadrant showed considerable deviation of their perceived direction from the assigned one. Moreover, the sounds coming from the direction 30° had the tendency of deviating to the right, i.e. towards the frontal direction of 0°, whereas the sound coming from the direction 60° had the tendency of deviating to the left, i.e. towards the left side direction of 90°.

The accuracy dependent on the encoding order yields another surprising finding that the first order seems to yield the most accurate localization.

The precision of localization does not yield a statistically significant difference between the instruments involved in the experiment.

The influence of the assigned direction of the source on the precision of localization is clearly visible. The source positions that coincide with actual loudspeakers were located with significantly better precision than those placed in directions that do not coincide with any loudspeakers.

The encoding order has a significant influence on the precision of localization, in the sense that the precision improves with increasing encoding order.

Future research will strive towards extending the possible directions of arrival to all three dimensions. For such an experiment, a reliable way of assessing both the azimuth and the elevation needs to be devised.

5. ACKNOWLEDGMENTS

The authors would like to thank the creators of the IEM Plugin Suite as an invaluable resource and tool, and to express their gratitude to Mike Senior of Cambridge Music Technology for creating the vast database of raw audio material and making it freely available.

6. REFERENCES

[1] M. A. Gerzon: "Periphony: With-Height Sound Reproduction", *Journal of the Audio Engineering Society*, Vol. 21, No. 1, pp. 2–10, 1973.

[2] M. A. Gerzon: "Practical Periphony: The Reproduction of Full-Sphere Sound", in Audio Engineering Society Convention 65, Audio Engineering Society, 1980.

[3] T. Carpentier: "Ambisonics Spatial Blur", in Audio Engineering Society Convention 142, Audio Engineering Society, 2017.

[4] A. J. Horsburgh, R. E. Davis, M. Moffat, and D. Fraser Clark: "Subjective Assessments of Higher Order Ambisonic Sound Systems in Varying Acoustical Conditions", in Audio Engineering Society Convention 133, Audio Engineering Society, 2012.

[5] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel: "Investigation on Localisation Accuracy for First and Higher Order Ambisonics Reproduced Sound Sources", *Acta Acustica United with Acustica*, Vol. 99, No. 4, pp. 642-657, 2013.

[6] M. Frank, F. Zotter, and A. Sontacchi: "Localization Experiments Using Different 2D Ambisonics Decoders", 25th Tonmeistertagung – VDT International Convention, 2008.

[7] REAPER software, https://www.reaper.fm/

[8] IEM Plugin Suite, https://plugins.iem.at/

[9] Cambridge Music Technology - The 'Mixing Secrets' Free Multitrack Download Library, http://www.cambridge-mt.com/ms/mtk/