ACOUSTIC CENTERING OF LITURGICAL MUSIC SINGERS MEASURED BY SURROUNDING SPHERICAL MICROPHONE ARRAYS

Javier García Ruiz[1], Noam R. Shabtai[2], Antonio Pedrero González[3], Michael Vorländer[4]

[1] M.S. student. Universidad Politécnica de Madrid, Madrid, Spain.
garciasruiz@gmail.com
[2] Department of Electrical and Computer Engineering, Ben Gurion University of the Negev, Beer Sheva, Israel.
shabtai.noam@gmail.com
[3] Grupo de Investigación en Acústica Arquitectónica, Universidad Politécnica de Madrid, Madrid, Spain.
antonio.pedrero@upm.es
[4] Institute of Technical Acoustics, RWTH Aachen University, Aachen, Germany.
mvo@akustik.rwth-aachen.de

Abstract

Radiation patterns of acoustic sources are of great importance for a wide range of applications, such as the measurement of the directivity of loudspeakers, sound radiation research of musical instruments and the human voice, auralization, etc. Spherical microphone arrays have recently been used to measure the spatial sound field of musical instruments. These arrays allow for simultaneous measurements of sound pressure at different points on the surface of the sphere and the computation of the spherical harmonics spectrum of the sound source. However, the radiation pattern is affected by the location of the source inside the array. For this reason, source centering algorithms must be applied to reduce the inaccuracies caused by the misalignment of the acoustic center of the source and the geometric center of the array. In this paper, centering techniques were applied to data from different fragments of liturgical songs performed by different singers. This data was obtained from the anechoic recordings of six singers who performed four pieces individually inside a spherical array composed of 31 microphones. The acoustic centering algorithm was applied to this data to obtain the time evolution of the singer's head position with respect to the geometric center of the array. The direction of maximum radiation was then obtained, revealing the orientation of the singer. Finally, a simulation was created to display the evolution of the position of the acoustic center and the orientation of the singer.

Keywords: spherical array, centering algorithm, objective function, aiming algorithm.

PACS no. 43.60.Fg, 43.38.Md

1. Introduction

An understanding of the directivity pattern of an acoustic source is of great importance for a wide range of applications, such as measuring the directivity of loudspeakers, sound radiation research of musical instruments and the human voice for auralization, virtual acoustics, etc. To understand directivity patterns, spherical microphone arrays have recently been used to model the sound field. These arrays allow for the measurements of simultaneous sound pressure at different points on the surface of the sphere and the calculation of the sound source spectrum using spherical harmonics.

However, the radiation pattern that is obtained is affected by the misalignment of the acoustic center of the source and the geometric center of the array. This misalignment may be due to several factors. First, the acoustic center of a sound source does not necessarily coincide with the position of the geometric center of the array, with the particularity that the difference between them is frequency dependent. Moreover, this divergence also occurs when the sound source is not static, for example, when musicians move naturally during their performance. For this reason, an algorithm for acoustic source centering should be applied in order to reduce the inaccuracies caused by momentaneous differences between the geometric center of the array and the acoustic center of the sound source, which would lead to spatial aliasing errors in the radiation pattern.

In this study, acoustic source centering techniques were applied to fragments of liturgical music pieces performed by different singers, recorded using a spherical array of 32 microphones. In the literature, there are several studies on the development and application of acoustic source centering algorithms [1-9]. In this study, two variants of a centering algorithm were used: the first one was developed by Ben Hagai et al. [6], and the second one was proposed by Shabtai and Vorländer [9]. The results shown in this paper were achieved using the algorithm developed by Shabtai and Vorländer, in accordance with their article in which their centering algorithm is compared with the previous one, showing a better convergence for a wide range of frequencies [9].

Shabtai and Vorländer’s [9] algorithm was applied to successive time blocks of the recorded signals, in order to track the position of the acoustic center over time with respect to the geometric center of the array. Furthermore, an algorithm for detecting the direction of maximum sound emission was developed in this study and applied to the same fragments for which the signal centering algorithm was applied.

Finally, an animation was created showing the results of both algorithms simultaneously. Assuming that variations in the position of the acoustic center were mainly caused by head movements of the singer during the performance, and that the direction of maximum sound emission coincides with the pointing direction of the singer's mouth, this animation represents the evolution of the position of the singer's head over time.

2. Anechoic recordings

The data that is analyzed in this paper came from a series of anechoic recordings that were performed in a studio for Pedrero et al. [10]. The recordings are of different liturgical pieces pertaining to the Mozarabic repertoire, performed separately by different singers in an anechoic environment, using a spherical array of microphones as the recording system. The sound signals were recorded simultaneously by an array of 31 microphones, in order to measure the directivity of the singer’s voices in real time. A reference video picturing the indications of the choirmaster was recorded beforehand to achieve proper synchronization between the different singers in the anechoic recordings. The reference video was then played during the recordings using a tablet positioned on a lectern under the light of the singer. Intentionally, singers were not advised with regards to their gestures during their performance. They were only told in which direction they should sing: facing forward and looking toward the positive x-axis. Thus, the functions of directivity obtained are designed to be representative of a normal performance by the musicians, including their natural movements.

2.1. Instrumentation and equipment

The recordings were made in the anechoic chamber of the “Centro de Acústica Aplicada y Evaluación No Destructiva” of the Spanish National Research Council, in Madrid. The interior dimensions of the anechoic chamber are 8 x 5.6 x 8 m (length, width and height).

Recording equipment consists of an array of 32 Sennheiser KE 4-211-2 omnidirectional condenser microphones, distributed uniformly in a spherical shape of 4.2 m in diameter. A structure of light-weight rods fixes the position and the pointing direction of the different microphones.

The singers were placed within the array on a wooden platform of 1 m^2, which is located just above microphone no. 32, so that this microphone is fully shielded and cannot be used for recording. Therefore, the effective number of useful microphones is 31. The height of the platform is placed with the intention of positioning the head of a singer (of average height) at the geometric center of the array.

2.2. Musical material

Eight pieces from the Mozarabic Chant repertoire were recorded by Pedrero et al. [10]. The songs were performed by singers from the *Schola Antiqua* choir, a semi-professional musical group with notable prestige in singing this repertoire. Each of the six singers performed all of the songs individually. A selection of four pieces containing extensive vocal range, a variety of melodic style and diversity in musical form were made for this study.

3. Acoustic source centering

Shabtai and Vorländer [9] developed the acoustic source centering algorithm used in this paper. This algorithm aligns the center of a radiating acoustic source and the physical center of the spherical microphones array. The algorithm was created on the basis that the radiation patterns of an acoustic source can be represented using the pressure values produced by the source on the surface of the sphere that surrounds it. Hence, spherical harmonics are a useful domain for analyzing acoustic radiation patterns.

A pressure function can be represented with its origin shifted from its original position. The problem that arises is that, in such a representation, the function of pressure shifted is generally of infinite order, even though the original centered pressure function was of limited order, and may, therefore, lead to spatial aliasing. For this reason, previous acoustic source centering algorithms used objective functions that minimized the energy of high-order coefficients. The more recent algorithms, however, developed an objective function based on the assumption that when an acoustic source is located at the center of a sphere, the phase shift between all the microphones should be minimal. This objective function was first introduced in the space domain and later in the spherical harmonics domain. Moreover, it was also shown that when this objective function is used in two dimensions, it does not significantly depend on the third coordinate, which indicates better dimensional convexity and more efficient implementation. [9]

3.1. Pre-processing

Data obtained from the recordings had to be processed before applying the acoustic source centering algorithm. Thirty-one WAV files with the signals recorded by each microphone were obtained from the recordings. The first step consisted of applying calibration values provided by the authors of the measures.

Once the data was calibrated, the acoustic source centering algorithm was applied to the signals. It is important to notice that the centering algorithm can only be applied to one specific frequency at a time. Since the input signals contained sung voices, different frequencies appeared at the same time throughout the fragment of interest. Therefore, we decided that the frequency to be analyzed would be the fundamental frequency of the analysis fragment. For this purpose, pitch detection was applied to the fragment of interest.

Pitch detection was applied to the signal collected by microphone no. 30, because this microphone was aligned with the direction towards which the singer directed his voice, resulting in a good signal-noise ratio. The pitch detection function selects a fragment of interest between start and end time values that have been indicated. The selected fragment is divided into time blocks, according to the number of analysis fragments per second indicated. The duration chosen for the time blocks was 0.05 seconds and five time blocks per second were selected. For each one of the time blocks, the fundamental frequency sung by the singer was identified. Fundamental frequencies with anomalous values (compared to the normal recording levels) were not considered frequencies of interest.

3.2. Application centering algorithm

Once the previous steps were applied and data were adjusted, the acoustic source centering algorithm was applied. The radiation pattern is displayed in Figure 1, before and after the centering algorithm was applied, with the results of the objective function displayed in Figure 2, obtained using a series of mathematical operations described in detail in [9].

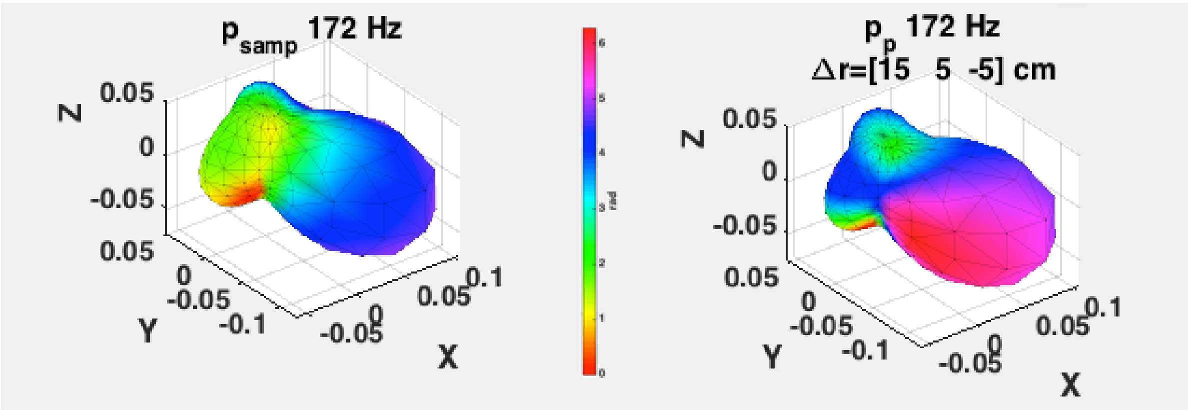


Figure 1. Acoustic radiation pattern of the a) non-centered and b) centered source.

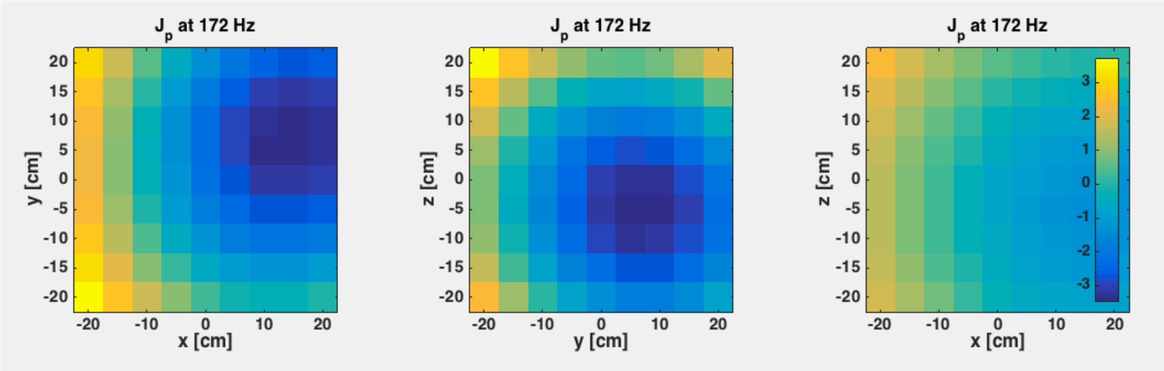


Figure 2. Results of the objective function.

Figures 1 and 2 refer to the time block with the start time 25.8 seconds, where the fundamental frequency detected is 172 Hz. It is sung by singer number four, performing the song called 'Requiem aeternam'.

In Fig. 1.a, the pressure function over the surface of a sphere using balloon plots is shown. The amplitude and the phase are represented there by the distance from the origin and color respectively. In Fig. 2, the results with respect to centering are shown. The three graphs show the results obtained by using the objective function that Shabtai and Vorländer propose in [9]. The graphs represent the results of the objective function in the algorithm in two dimensions where the acoustic center value is shown by planes. The distance between the acoustic center of the source with respect to the geometric center of the array can be estimated from these graphs. On the xy-plane in Fig. 2, the objective function converges in the values identified by dark colors, approximately x=15 cm and y=5 cm. On the yz-plane in Fig. 2, the value of "y" can be confirmed (y=5 cm) and z=-5 cm. Finally, on the xz-plane in Fig. 2, the observed positions of "x" and "z" can be corroborated.

In Fig. 1.b, the radiation pattern of the centered source is shown. Furthermore, the vector that has shifted the acoustic center with respect to the geometric center of the array is presented just above the graph. The value of Δr coincides with the value indicated subsequently in Fig. 2.

In this study, the most relevant data outcome is the value offered by the acoustic center position indicated by Δr. This value is associated with the position of the singer's mouth, and we get the information about the position of the singer's head at each time block analyzed. Analyzing successive time blocks, we create a tracking of head positions throughout the piece performed.

4. Aiming detection

In order to complete the tracking of the singer's head, each of the positions corresponding to the singer's head is associated with a pointing direction for each time block. This means that the position in the directivity diagram where the amplitude is greater must be identified for each time block analyzed. Since maximum position in the directivity pattern is frequency dependent, maximum value of A-weighted broadband sound pressure has been used to identify aiming direction, rather than single-frequency maximum value.

A Matlab function was created to obtain the aiming point. This function performs a Fast Fourier Transform on the 31 microphone signals corresponding to the selected time blocks and calculates the A-weighted values of the results. Subsequently, in order to increase the resolution, an interpolation in the domain of spherical harmonics is performed. Among the interpolated values, the highest is used to identify the aiming direction.

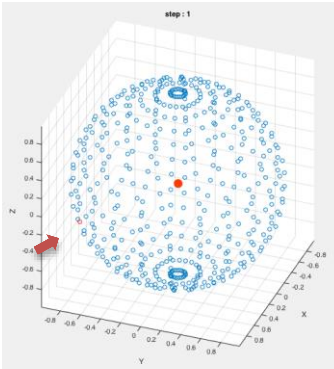


Figure 3. Aiming detection for time block starting at 25.8 seconds.

In Figure 3, the set of points show the sphere that is formed from the interpolated results of the 31 microphones points. Point No. 431 (the red circle) has been identified as the aiming point for this time block analyzed, which starts at 25.8 seconds into the piece ‘Requiem aeternam’, performed by singer number four. The result corresponds to the area in front of the singer, positioned along the negative y-axis and below the singer's head. While this may seem counterintuitive, this position is located where the tablet showing a video of the choirmaster directing was placed.

5. Results

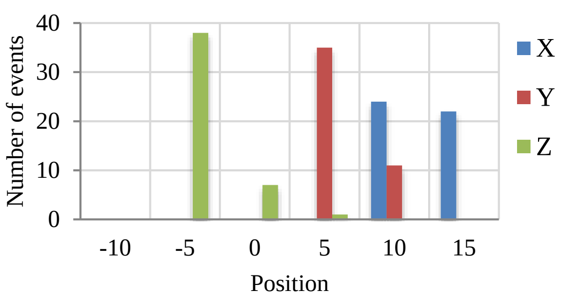
Once the tools were implemented, they have been applied in combination with the recorded data, with the goal of creating a simulation of the singer's head positions and aiming points throughout the fragment of interest. From the different pieces recorded, different fragments were chosen depending on their characteristics.

In Table 1, the results of the first 5 seconds of an analysis made for a fragment of interest of 10 seconds are shown. The song is called ‘Requiem aeternam’ and was performed by singer number 4. This piece was chosen as an example due to its syllabic character.

Table 1. Results for Singer 4 song 6.

ID	Start time of the analysis fragment (s)	Fundamental frequency (Hz)	Acoustic center (cm)			Aiming point	Aiming point coordinates		
			x	y	z		x	y	z
1	24,0	248,1	10	5	-5	432	0,61	-0,71	-0,35
2	24,2	230,6	10	5	-5	432	0,61	-0,71	-0,35
3	24,4	244,4	10	5	-5	431	0,64	-0,74	-0,21
4	24,6	243,5	10	5	-5	431	0,64	-0,74	-0,21
5	24,8	241,6	10	5	-5	431	0,64	-0,74	-0,21
6	25,0	242,6	10	5	-5	431	0,64	-0,74	-0,21
7	25,2	242,6	15	5	-5	431	0,64	-0,74	-0,21
8	25,4	239,8	15	5	-5	431	0,64	-0,74	-0,21
9	25,6	239,8	15	5	-5	431	0,64	-0,74	-0,21
10	25,8	171,8	15	5	-5	431	0,64	-0,74	-0,21
11	26,0	191,1	15	5	-5	431	0,64	-0,74	-0,21
12	26,2	186,5	15	5	-5	431	0,64	-0,74	-0,21
13	26,6	239,8	15	5	0	431	0,64	-0,74	-0,21
14	26,8	245,3	15	5	-5	431	0,64	-0,74	-0,21
15	27,0	267,4	10	5	-5	431	0,64	-0,74	-0,21
16	27,2	255,4	10	5	-5	431	0,64	-0,74	-0,21
17	27,4	246,2	15	5	-5	409	0,41	-0,89	-0,21
18	27,6	239,8	15	5	-5	409	0,41	-0,89	-0,21
19	27,8	209,5	10	5	-5	409	0,41	-0,89	-0,21
20	28,0	214,1	10	5	-5	410	0,39	-0,85	-0,35
21	28,2	252,7	10	5	-5	410	0,39	-0,85	-0,35
22	28,4	256,3	10	5	-5	410	0,39	-0,85	-0,35
23	28,6	255,4	10	5	-5	410	0,39	-0,85	-0,35
24	28,8	242,6	10	5	-5	409	0,41	-0,89	-0,21
25	29,0	191,1	10	10	0	431	0,64	-0,74	-0,21

In Table 1, the fundamental frequencies of each of the time blocks analyzed are shown. For each time block, the acoustic center is indicated by the x, y, and z coordinates in centimeters. These values represent the difference between the acoustic center with respect to the geometric center of the array. The following column indicates the aiming point. The values shown in this column correspond to the points of the sphere surrounding the source which is composed of 484 points. These points were obtained by interpolating the 31 positions that form the spherical array of microphones. The equivalent Cartesian coordinates are indicated for the aiming point.



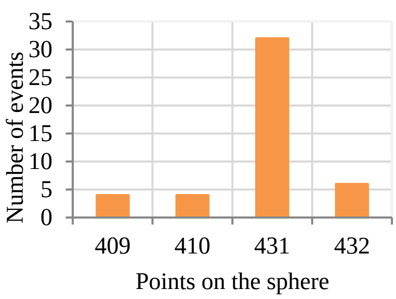


Figure 4. Histograms of positions and aiming points.

In Figure 4, histograms of positions and aiming points of singer number 4 for the 10-seconds fragment analyzed are presented. It can be observed that the singer's head position is located about 5 cm below the center of the array on the z-axis in the largest number of events. The array was placed according to the average height of all six singers. Therefore, in the case of singer number 4, whose height was somewhat lower than the average, the singer’s head position is below the center of the array. Moreover, the head is positioned forward about 10-15 cm on the positive x-axis and 5 cm on the positive y-axis. We attribute this event to the fact that the singers were not given any indication as to their placement or gestures during their performance within the array, but were only instructed to position themselves on a mark on the platform and to sing in a certain direction. It is for this reason that the accuracy of the position in the center of the array in their horizontal plane was exposed to the criteria of the singer and their natural performance movements. It is also significant that a lectern was placed within the anechoic chamber with a device playing a reference video. The video contained the choirmaster’s indications for the singers, in order to synchronize the different performances. This may be evidence of an unconscious instinct by the singers to move themselves closer to the video. This limitation marks the succession of aiming point positions. The set of aiming point positions are divided into four different positions that coincide with the location where the video device is located on the lectern.

Using the results of the application of both tools, a simulation has been performed where the evolution of the singer's head position and the pointing direction are represented. In Figure 5, the specific result for the time block with the start time 25.8 seconds is presented. The singer's head position is represented by a red dot; the pointing direction is the blue line; and, the position where the singer was first instructed to direct his head is marked

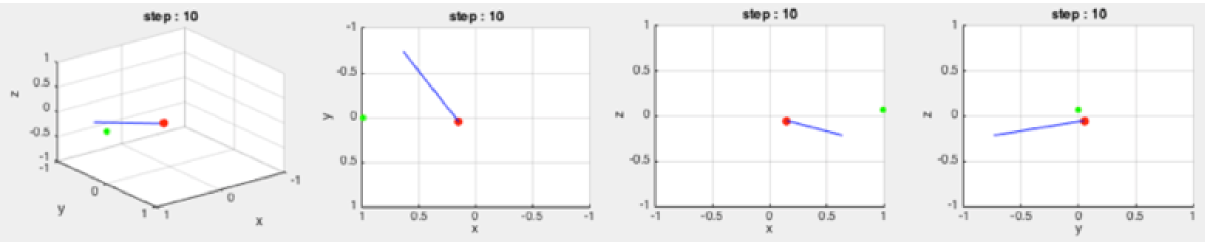


Figure 5. Simulation image for time block starting at 25.8 seconds.

We applied the techniques described above to a total of four different pieces performed individually by six singers. We chose a representative fragment of interest for each of the pieces sung by each singer. The results of the analysis of the 24 fragments are summarized in Figures 6 and 7

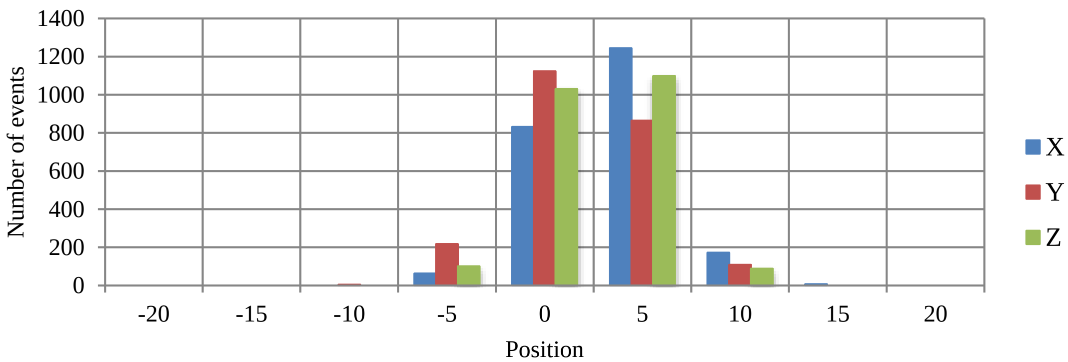
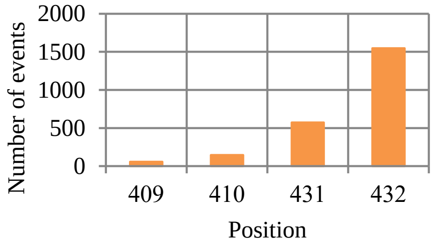


Figure 6. Histogram of the positions obtained for all fragments of interest analyzed.



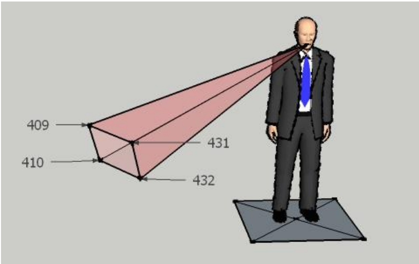


Figure 7. Histogram of aiming points and diagram of points placement for all fragments of interest analyzed.

The set of the singer's most frequently repeated head positions can be observed in Figure 6. Furthermore, the most repeated aiming point seems to correspond, as already mentioned above, to the position of the lectern with the video device.

6. Conclusions

An acoustic source centering algorithm was applied to a number of liturgical music pieces sung by different singers, which had been recorded using a surrounding spherical microphone array. Variations in the position of the acoustic center were detected at the fundamental frequency and were attributed to the natural movements of the singer during performance. The pointing direction was identified on the surface of the sphere surrounding the singer. Finally, the position and orientation of the singer's head was estimated throughout a passage of 10 seconds.

Acknowledgements

The authors wish to thank the members of the Schola Antiqua choir and, especially, its director, Juan Carlos Asensio, for their altruistic collaboration in this project.

References

1. D. Deboy, "Acoustic centering and rotational tracking in surrounding spherical microphone arrays," Master's thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria, 2010.
2. B. Rafaely, "Spatial alignment of acoustic sources based on spherical harmonics radiation analysis," in 2010 4th International Symposium on Communications, Control, and Signal Processing (ISCCSP 2010), Limassol, Cyprus (2010), pp. 313–318.
3. D. Deboy and F. Zotter, "Comparison of acoustic centering maps for radiation capture of musical instruments with spherical microphone arrays," in Proceedings of the 36th German Annual Conference on Acoustics (DAGA 2010), Berlin, Germany (2010), pp. 705–706.
4. D. Deboy and F. Zotter, "Tangential intensity algorithm for acoustic centering," in Proceedings of the 37th German Annual Conference on Acoustics (DAGA 2011), Dusseldorf, Germany (2011), pp. 321–322.

5. M. Pollow, G. K. Behler, and M. Vorländer, “Post-processing and center adjustment of measured directivity data of musical instruments,” in Proceedings Acoustics 2012, Nantes, France (2012), pp. 761–765.
6. I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, “Acoustic centering of sources measured by surrounding spherical microphone arrays,” J. Acoust. Soc. Am. 130, 2003–2015 (2011).
7. N. R. Shabtai, M. Pollow, and M. Vorländer, “Acoustic centering for high-order source directivities,” in Proceedings of the 40th German Annual Conference on Acoustics (DAGA 2014), Oldenburg, Germany (2014), pp. 226–227.
8. N. R. Shabtai and M. Vorländer, “On the convexity of acoustic centering algorithms for the modeling of source directivity patterns,” in 7th Forum Acusticum Krakow, Krakow, Poland (2014).
9. N. R. Shabtai and M. Vorländer, “Acoustic centering of sources with high-order radiation patterns,” J. Acoust. Soc. Am., vol. 137 no. 4, pp. 1947-1961, Apr. 2015.
10. A. Pedrero, M. Pollow, P. Dietrich, G. Behler, M. Vorländer, C. Díaz and A. Díaz-Chyla, “Mozarabic Chant anechoic recordings for auralization purposes,” in Proceedings VIII Congresso Ibero-americano de Acústica, Évora, Portugal (2012).