A Human Head Shaped Array of Microphones and Cameras for Automotive Applications

Daniel Pinardi
Dept. of Engineering and Architecture
University of Parma
Parma, Italy
daniel.pinardi@unipr.it

2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA) | 978-1-6654-0998-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/I3DA48870.2021.9610879

***Abstract* - Nowadays, a growing interest in the recording and reproduction of spatial audio was observed. However, despite many microphone arrays were developed in the last years, there are still few solutions for Noise, Vibration and Harshness (NVH) applications at low frequency. In this paper, a new array of microphones and cameras is presented, for recording both acoustic and visual spatial information. It can be used for the spatial analysis and visualization of the sound field and to perform recordings that can be rendered in a virtual reality (VR) environment. The system was optimized for the low frequency range, as most of the available solutions have proved unsatisfactory for frequencies below 400 Hz. Moreover, the system is cost-effective if compared to other existing products designed for similar applications. The spatial performance of the array is evaluated in comparison with the current state of the art systems. Finally, a field application is presented. The new head shaped microphone array demonstrated its effectiveness for evaluating the performance of an automotive Active Noise Control (ANC) system.**

Keywords—Active Noise Control, automotive, beamforming, microphone array, Noise Vibration and Harshness, spatial audio

I. Introduction

The use of microphone arrays for spatial audio recordings dates to the seventies, when the Soundfield™ Microphone [1] was proposed, relying on the Ambisonics theory [2]: a method for representing the complete acoustical spatial information in one point of the space by means of a Spherical Harmonics (SH) [3], [4] expansion of the signals recorded by the capsules of the array. Microphone arrays equipped with four capsules, such as the Soundfield™, can encode the First Order Ambisonics (FOA), which exhibits a reduced spatial resolution. To overcome this limit, the first order approach was later extended to High Order Ambisonics (HOA), which requires more microphones and hence more channels.

Since 1978, many four channels microphone arrays came to market and several higher order compact spherical microphone arrays were built too, with a number of capsules ranging between 8 and 252 [5], [6], [7], [8], [9], [10]. Recently, larger arrays specially designed for VR and NVH applications came to the market as well, equipped with a high number of channels and integrating a video recording system: VisiSonics with 64 capsules and five lenses (Fig. 1, left), Sonicam with 64 Micro Electro-mechanical System (MEMS) microphones and nine lenses (Fig. 1, center) and Bruel&Kjaer with 36 or 50 capsules and 12 lenses (Fig. 1, right). Nevertheless, all of them still presents some limitations: Bruel&Kjaer arrays have poor video resolution and employ low sensitivity capsules; the Sonicam array can record only first order Ambisonics together with the panoramic video; the VisiSonics array has a video resolution limited to 3k and it is the most expensive solution available on the market.

Fig. 1. Three microphone arrays: VisiSonics (left), Sonicam (center) and Bruel&Kjaer (right)

In this paper, a new microphone array is presented to overcome the mentioned problems. The proposed solution features 32 capsules, allowing to encode HOA up to fourth order. The size of the array allows a good spatial resolution at low frequency, making it suitable for NVH applications. It comprises a circular array of eight cameras, which can record high resolution panoramic video. Such spatial audio-video recordings can be reproduced in a VR environment employing Head Mounted Displays (HMDs) with head tracking and binaural decoding [11], [12]. The array has been calibrated, so that true values of sound pressure levels (SPL) can be calculated. Finally, the developed system is cost-effective, thanks to the usage of 3D-printing technology.

The paper is organized as follow. Section II provides a technical description of the design. In section III the beamforming algorithm employed in this work is explained. Section IV describes the procedure developed for calibrating the microphone array. Section V describes the numerical simulation of the array, section VI the evaluation of the performance and finally section VII presents a field application of the new system.

II. Design of the Head-Shaped Array

A. Panoramic Video Recording System

Compact panoramic video recorders are available as consumer products: they are made of two hemispherical lenses, provide up to 8k resolution and can be employed next or above a microphone array. However, this solution entails the introduction of an offset between acoustic and optical centers, resulting in a mismatch between the sound field and the visual background.

Therefore, a distributed array of cameras has been integrated to ensure a perfect coincidence of the acoustic and optical centers. It consists in eight cameras having cubic shape, arranged along a circle (Fig. 2). The cameras are positioned as close as possible each other to increase the superimposition of the Field-of-View (FOV), which results in a better stitching of the recorded video. This led to a radius of the camera array equal to $r = 93\,mm$. Each camera has a

FOV of 72° and a resolution of 1440x1920 when mounted rotated by 90° to extend the vertical coverage. Thus, the resolution of the stitched video is 7200x1920, or 13.8 Mpixel, which is satisfactory for VR reproduction. Current VR systems are based on HMDs with a resolution of 1440x1600 per eye and a FOV of 110°. Therefore, the minimum video resolution required to take advantage of all the quality of the HMDs is 4700x2600 (12.2 Mpixel), which is, for instance, more than twice the capability of the VisiSonics (4.7 Mpixel).

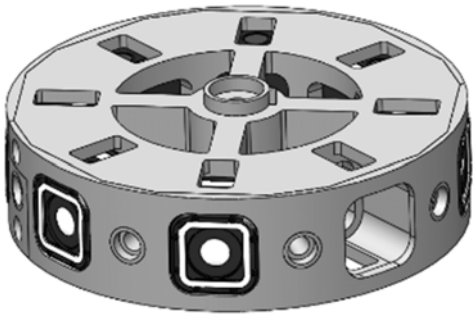

Fig. 2. 3D CAD model of the cylindrical array of eight cameras.

However, two drawbacks must be accepted. Firstly, a distributed array of cameras entails manual stitching, which is time-consuming and requires additional software. In addition, the shape of the microphone array cannot be spherical, due to the presence of the cylindrical housing for the video recording system. Therefore, it is not possible to characterize the array relying on a theoretical solution.

B. Microphone Array Design

It is well known that the distance between the capsules determines the theoretical frequency limits (i.e., f_{min} and f_{max}) of beamforming [13]. Capsules are too close at low frequency and too far at high frequency:

$$f_{min} = \frac{1}{4} \cdot \frac{c}{d_{max}} \quad (1)$$

$$f_{max} = \frac{1}{4} \cdot \frac{c}{d_{min}} \quad (2)$$

where c is the speed of sound ($343\ m/s$ in air at $T = 20°C$); d_{max} and d_{min} are the maximum and minimum distance between capsules, respectively. Relations (1) and (2) make it clear that the larger the array, the better its capability to work at low frequencies. In the proposed system, capsules are arranged so that $d_{max} = 222.2\ mm$, therefore we get from (1) a theoretical $f_{min} \cong 400\ Hz$. The minimum spacing between capsules is $d_{min} = 34\ mm$, therefore we get from (2) a theoretical $f_{max} \cong 2.5\ kHz$. Such frequency range of optimal usability is satisfactory for NVH applications.

The choice of having a cylindrical array of cameras resulted in a non-spherical design. It was opted to close the array with two hemispherical caps having the same radius of the cylindrical part. Given these dimensions, the microphone array is comparable to a human head, so hereafter we will refer to it with the name "Head-Shaped Array" (HSA). The array has a diameter of $186\ mm$ and a total height of $235\ mm$. The system was realized with 3D printing technology employing ULTEM material. In the bottom, it is provided with a thread so that it can be mounted on a standard microphone stand (Fig. 3, left), or, using a kind of "neck", on a standard dummy torso (Fig. 3, right), a very common solution adopted for experimental measurements inside vehicles.

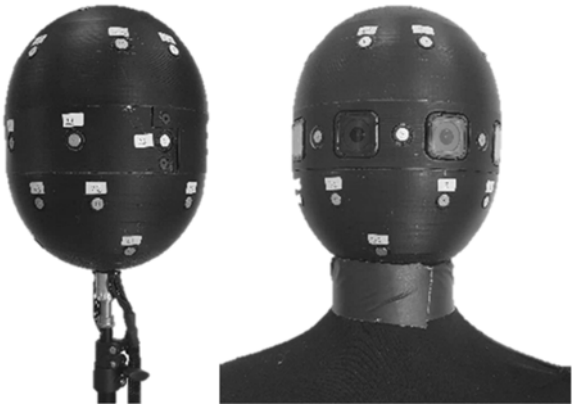

Fig. 3. Head-Shaped Array, mounted on a microphone stand (left) or in a dummy torso (right).

The capsules are electret omnidirectional microphones, "PRIMO – EM172", having the frequency response of Fig. 4 and the following characteristics: $SNR = 80\ dB$, $Self-noise = 14\ dB(A)$, $SPL_{max} = 119\ dB$.

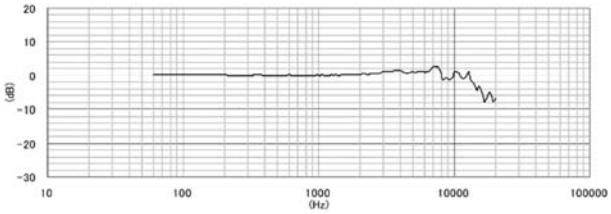

Fig. 4. Primo EM172 capsules frequency response

The number of capsules is $M = 32$, arranged along five rings equally spaced at five different elevations with 4, 8, 8, 8, and 4 microphones each one, from bottom to top respectively. Along the azimuthal plane, all capsules are equally spaced by an angle of $11.25°$ to maximize the horizontal resolution. In the central cylinder, which is the camera housing, one capsule is placed between each pair of cameras. Since the shape is not spherical, the distribution of the capsule over the surface is not uniform. However, an increased density of microphones on the azimuthal plane enhanced the spatial resolution in this region of the space, similarly to the human auditory system.

III. Beamforming Theory

The capsules of a microphone array detect the sound pressure and such set of signals is usually known as "A-format". Then, the operation of beamforming converts these signals in a complex combination of sound pressure and particle velocity. Depending on the target of the beamforming, different spatial audio formats are obtained, typically Ambisonics (also known as "B-format") or Spatial PCM Sampling (SPS, also known as "P-format") [14]. The conversion can be approached in two different ways. When a linear processing is used, a matrix of filters synthesizes the virtual microphones from the capsule signals. In case of parametric processing, the raw capsule signals are employed for performing a parametric spatial analysis of the sound scene, extracting the single source signals and their locations, and then using the theoretical SH formulas for recreating the B-format signals [15], [16], [17], [18], [19].

In this paper, a linear processing was employed, leaving parametric solutions to future developments. Beamforming is obtained with a matrix of Finite Impulse Response (FIR) filters, computed by inverting the array response with a regularized Kirkeby inversion [20]:

$$H_{m,v,k} = \left[C^*_{m,d,k} \cdot C_{m,d,k} + \beta_k \cdot I\right]^{-1} \cdot \left[C^*_{m,d,k} \cdot A_{d,v} \cdot e^{-j\pi k}\right] \quad (3)$$

where $m = [1, \dots, M]$ are the capsules; $v = [1, \dots, V]$ are the virtual microphones; k is the frequency index; $d = [1, \dots, D]$ are the Directions of Arrival (DoA) of the sound waves; the matrix C is the complex response of each capsule m for each direction d; the matrix A defines the frequency independent amplitude of the target directivity patterns; $e^{-j\pi k}$ introduces a latency that ensures filters causality; $\cdot$ is the dot product; I is the identity matrix; $[\,]^*$ denotes the conjugate transpose; $[\,]^{-1}$ denotes the pseudo-inverse; $\beta_{[k]}$ is a frequency-dependent regularization parameter [21], which apply a frequency constraint to the inversion, thus preserving the Signal-to-Noise Ratio (SNR) in the frequency ranges where beamforming is physically limited by the distance between the capsules.

The regularization parameter $\beta_{[k]}$ was defined as follow: $\beta_i = 0.01$ for the in-band range, $\beta_o = 1$ for the out-band range, $\Delta f = 0.3$ octaves for the transition bandwidth, $f_{low} = 20\ Hz$ for the lower transition frequency, and $f_{high} = 6\ kHz$ for the higher transition frequency.

In this work, SPS format is employed. It consists in a set of unidirectional virtual microphones pointing outwards in all directions. To sample uniformly the sound field, it was used a set of $V = 32$ virtual microphones having fourth order cardioid directivity. Such directivity is preferable with respect to the hypercardioid, since the latter causes the presence of a rear lobe, while the cardioid is not affected by any side or rear lobes (Fig. 5). The target directivity as a function of the angle ϑ between the aiming direction of the virtual microphone and the DoA of the sound wave, is given by:

$$A(\vartheta) = [0.5 + 0.5 \cdot \cos(\vartheta)]^4 \quad (4)$$

For optimizing the uniform coverage of the space, spherical design [22] can be used to define the directions of the virtual microphones. In this case, a spherical design of order *T=7* was used, since this choice maximizes the order *T* with the given number of virtual microphones, that is *V=32*.

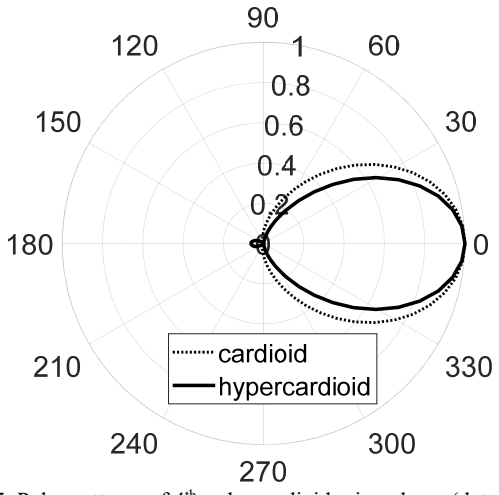

Fig. 5. Polar patterns of 4th order cardioid microphone (dotted line) and 4th order hypercardioid microphone (solid line).

IV. Array Calibration

The calibration of the microphone is important when noise analysis is performed. Although it is always considered for measurement microphones, which are supplied with a factory calibration, it is ignored for the microphone arrays. For this reason, a procedure was developed to calibrate the HSA. An almost flat point source ($\pm 1.5\ dB$ in the range $38\ Hz - 20\ kHz$), a Genelec Studio Monitor type 8351A, was place in an anechoic room. The equivalent level $L_{eq\,(SLM)}$ was measured with a calibrated Sound Level Meter (SLM) at 1 m distance while playing a pink noise filtered in the 1 kHz octave band. Then, the same signal was recorded with the microphone array in place of the SLM.

For the SPS format case, the array recording is converted into SPS by convolving with the matrix H_{SPS} calculated with (3). Then, a calibration factor cf_{SPS} [dB] is calculated, imposing that the energetic sum of all the SPS signals must be equal to the calibrated SPL measured by the SLM:

$$L_{eq\,(SLM)} = cf_{SPS} + \sum_{v=1}^{V} L_{eq,v} \quad (5)$$

where $L_{eq,v}$ is the equivalent level of each virtual microphone of the SPS format.

For the Ambisonics case, the array recording of the pink noise was encoded into SH, by convolving with the matrix H_{SH} calculated with (3). Then, another calibration factor cf_{SH} [dB] is calculated, imposing that the level of the omnidirectional channel W (SH number 0) must be equal to the calibrated SPL measured by the SLM:

$$L_{eq\,(SLM)} = cf_{SH} + L_{eq,SH=0} \quad (6)$$

where $L_{eq,SH=0}$ is the equivalent level of the SH number 0, following the standard AmbiX [23].

In both cases, the calibration can be applied by adding the calibration factors cf_{SPS} or cf_{SH} to the recordings performed with the microphone array and encoded with the filtering matrices H_{SPS} or H_{SH}. The fundamental aspect is that the array is calibrated together with the beamforming matrix. Therefore, the calibration is valid for the specific encoding matrices H_{SPS} or H_{SH}. If they are changed, the calibration must be recalculated.

Thanks to this method, it was possible to obtain calibrated color maps of the spatial distribution of SPL at the microphone array position (see Section VII), which are coherent in level with the Ambisonics signals employed for the head tracked playback. This feature is ignored also in most advanced solutions for sound field visualization [24], [25].

V. Numerical Simulation of the Array

The response of the array, namely the matrix C in (3), provides the complex information of sound pressure at the capsules for many DoA of the sound waves. There are three possible approaches for deriving the C matrix: theoretical solution of the waves diffraction equations [26], [27], [28], experimental measurement in anechoic room [29], [30] and numerical simulation, which was employed in this work.

First, the numerical method was verified by comparing it with the theoretical solution for the case of plane waves diffracted by a rigid sphere. The theoretical solution was found in [31]. The numerical simulation was performed in COMSOL Multiphysics with Finite Elements Method (FEM) and employing a Plane Wave Radiation (PWR) condition. The same geometry was introduced in both cases, a sphere of

84 mm diameter and 32 capsules. The simulation was calculated for the front direction, in the frequency range $10\,Hz - 3.5\,kHz$ with a resolution of $10\,Hz$. The SPL spectra are compared in Fig. 6, for the capsules number 1, 5, 13, 17, 21, 24, 29 and 32. One can note that the two solutions are almost identical. This result ensures that the numerical method is reliable for calculating the spatial array response, i.e., the *C* matrix, of a microphone array.

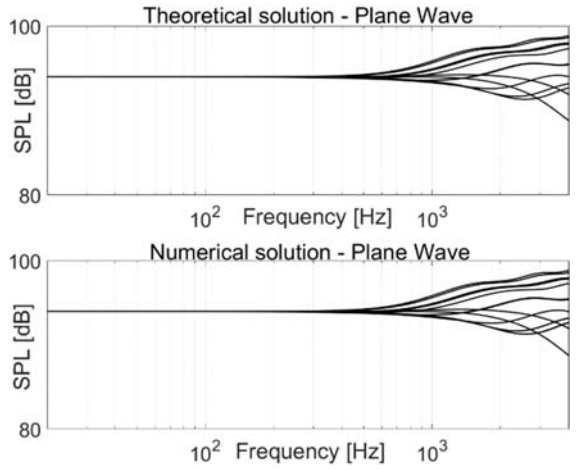

Fig. 6. SPL spectra compared up to 3.5 kHz between theoretical and numerical solutions.

Then, a numerical simulation of the HSA was performed in the frequency range $10\,Hz - 6\,kHz$, with a resolution of $10\,Hz$. In this case, the calculation was repeated for many DoA of the plane wave. The geometry of the DoA employed for characterizing the array corresponds to a spherical design of order $T = 21$, which has $D = 240$ directions [32]. Since the solution is performed in frequency domain, the *C* matrix required to solve (3) is directly obtained.

VI. Performance Evaluation

The performance of the HSA was compared against two existing solutions. One is the Eigenmike32™ (EM), a spherical array of 84 mm diameter with 32 capsules, considered the reference equipment for spatial audio recording in the last decade. The second one is the Bruel&Kjaer array, a spherical array of 195 mm diameter with 36 capsules and 12 cameras (B&K36 in the following), which is mainly designed for NVH applications, as the video resolution is too low for VR reproduction.

For evaluating the spatial performance of each system, the SPS format and a metric defined by the author was used [33]. The metric uses two parameters: Directivity Factor Q and Half Power Beam Width BW. Directivity Factor Q is defined as:

$$Q_{[k]} = \frac{I_{\max\,[k]}}{I_{0\,[k]}} \tag{7}$$

where I_{max} is the magnitude of the sound intensity vector in the direction of the maximum and I_0 is the average of the magnitude of sound intensity over the sphere. The BW parameter is defined as:

$$BW_{[k]} = 2 \cdot \angle(\vec{S}_{\max\,[k]} - \vec{S}_{max-3\,dB,[k]}) \tag{8}$$

thus, the BW is twice the angle between the direction of maximum sensitivity $\vec{S}_{max}$ of the virtual microphone and the direction having sensitivity $3dB$ below the maximum, $\vec{S}_{max-3\,dB}$. For a super-directive unidirectional virtual microphone, it is desired to have a high Q and a low BW.

For calculating the effective directivity A' obtained for each virtual microphone *v* in each direction *d*, the matrices *C* and *H* are convolved, thus multiplied in time domain:

$$A'_{[d,v,k]} = C_{[d,m,k]} \cdot H_{[v,m,k]} \tag{9}$$

The ideal result would be $A' = A$. For the two spherical arrays, the EM and the B&K36, the theoretical solution [31] was used to obtain the matrix *C* and thus to derive the beamforming matrices *H* for the SPS format. In these cases, the frequency range $20\,Hz - 20\,kHz$ was covered. The two parameters, Q and BW, were averaged in 1/3 octave bands and over the *V* virtual microphones. The results are shown in Fig. 7, superimposed for the three arrays.

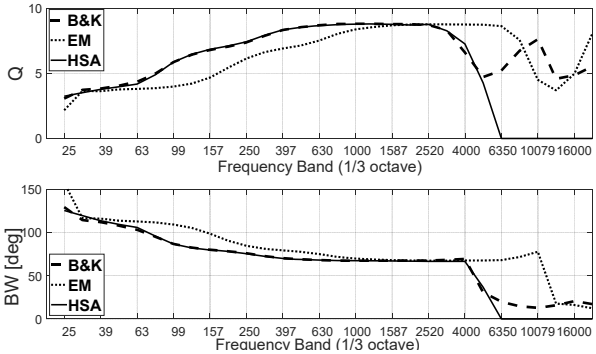

Fig. 7. Directivity Factor Q and Beam Width BW for the new HSA, the Eigenmike32™ and the Bruel&Kjaer with 36 channels.

Being the smallest, the EM shows a lower value of Q and a larger value of BW in the range $50\,Hz - 1.5\,kHz$. On the opposite, beamforming is much better than the other two systems above 2.5 kHz. Regarding the B&K36 and the HSA, they perform almost identically in the frequency range $20\,Hz - 4\,kHz$. For the HSA, the directivity Q starts to decrease above 2.5 kHz, as in (2). At low frequency, it is possible to see that Q decreases and BW increases below $400\,Hz$, as in (1). It is not possible to have beamforming above 4 kHz, therefore having calculated the numerical simulation up to 6 kHz did not cause any loss of information.

VII. Field Applications

The described system was compared with the EM in two field tests: a laboratory recording with an artificial noise source and the performance evaluation of an active noise cancelling (ANC) system installed in a car.

Regarding the first experiment, a pink noise signal of 30 s was played through a Genelec™ Studio Monitor 8351A and recorded by placing the microphone arrays in front of the loudspeaker at 2 meters distance, one at a time. Then, color maps at various octave bands were computed with a graphical interpolation of the SPL values calculated for each of the 32 beams of the SPS format. At first, the octave band centered at 2 kHz was considered (Fig. 8). In this frequency range, the arrays perform equally, according to Fig. 7. Then, the octave band centered at $500\,Hz$ was considered (Fig. 9). In this frequency range, the HSA performs better, and such result matches with Fig. 7 too. Finally, the octave band centered at $31.5\,Hz$ was studied (Fig. 10). One can note that the visualization of the sound field provided by the HSA is coherent with the source, even if the analyzed frequency

range, $22\ Hz - 44\ Hz$, is outside the optimal range of the system. Instead, the result provided by the EM is distorted and misaligned with respect to the loudspeaker, showing that beamforming is not working at all in such a low frequency range.

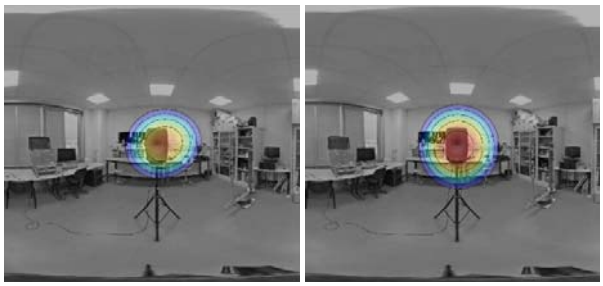

Fig. 8. SPL color maps in the octave band centered at $2\ kHz$, Eigenmike32™ (left) and HSA (right).

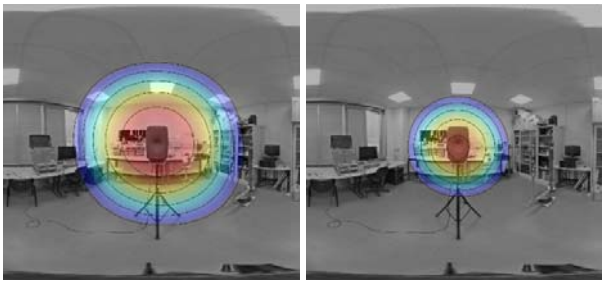

Fig. 9. SPL color maps in the octave band centered at $500\ Hz$, Eigenmike32™ (left) and HSA (right).

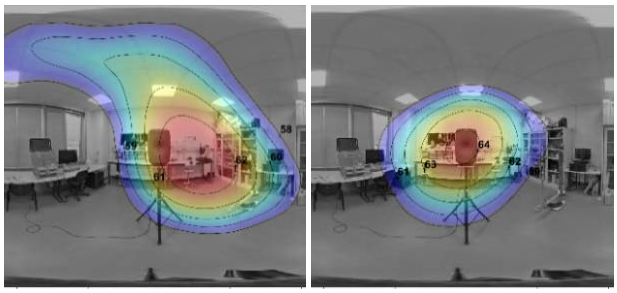

Fig. 10. SPL color maps in the octave band centered at $31.5\ Hz$, Eigenmike32™ (left) and HSA (right).

The second test consists in the performance evaluation of an automotive ANC system. For this application, solutions capable to work at very low frequencies are extremely desirable, as these systems are usually active in the range $50\ Hz - 500\ Hz$ [34]. The usage of analysis techniques based on the spatial information of the sound field can improve these systems, e.g., identifying the noise source positions and thus suggesting the optimal placement of the loudspeakers or providing advanced diagnostic tool to assess their effectiveness [35].

A C-segment car equipped with a Road Noise Cancellation (RNC) system, developed by ASK Industries S.p.A. [34], was measured with both arrays, mounted one at a time on a mannequin torso positioned on the passenger seat. Test recordings of $50\ s$ were taken on a straight rough road at $55\ km/h$ with the RNC system initially switched off (RNC-off) and then switched on (RNC-on). SPL color maps filtered in the range $70\ Hz - 330\ Hz$ were calculated in the two working conditions, with the previously described method. Results are shown in Fig. 11 and Fig. 12 for the EM, in Fig. 13 and Fig. 14 for the HSA.

Fig. 11. SPL color map $70\ Hz - 330\ Hz$, Eigenmike32™, RNC-off.

Fig. 12. SPL color map $70\ Hz - 330\ Hz$, Eigenmike32™, RNC-on.

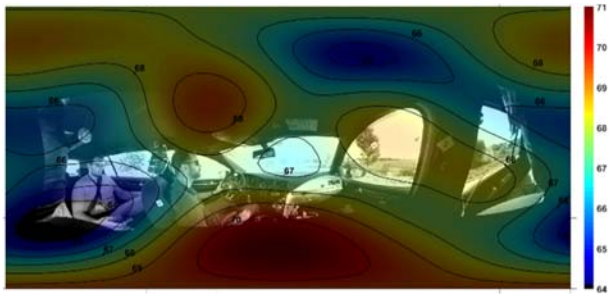

Fig. 13. SPL color map $70\ Hz - 330\ Hz$, HSA, RNC-off.

Fig. 14. SPL color map $70\ Hz - 330\ Hz$, HSA, RNC-on.

With the system switched off (RNC-off) and employing the HSA, roof and windows reflections can be seen (Fig. 13). In the same condition, those details are not visible with the EM (Fig. 11): a large spot appears in the middle, misleading the sound propagation analysis. When the system is switched on (RNC-on) the HSA shows more precisely the cancellation effectiveness: it is possible to identify the position of the largest residual energy contribution (Fig. 14). Also in this case, the EM is not accurate (Fig. 12): energy is spread all over the whole map, which does not provide useful information.

VIII. Conclusions

A non-spherical array of microphones and cameras has been presented. It proved to be effective for analyzing the sound field at low frequencies, making it suitable for Noise,

Vibration and Harshness applications. At the same time, most of the shortcomings identified in the existing products have been solved: high resolution video recording system, sound pressure level calibration, and affordable cost.

The so-called Head Shaped Array works in the frequency range $20\ Hz - 4\ kHz$ but provides optimal beamforming in the range $400\ Hz - 2.5\ kHz$, as expected from the theoretical calculations carried out in the design phase. This performance is comparable to the Bruel&Kjaer array of 36 capsules and 12 lenses, a product similar in the design and application, but with lower video resolution and higher cost.

The new system was used in two field tests characterized by low frequency noise: a laboratory experiment with a loudspeaker and a road noise active cancellation system installed inside a car. In both cases, color maps of calibrated sound pressure levels were calculated by employing the Sound PCM Sampling spatial audio format. The visualization of the sound field obtained with the Head Shaped Array is accurate, proving that the system can be successfully employed for NVH analysis.

ACKNOWLEDGMENT

The author wants to express his grateful to ASK Industries SpA and ARGO Tractors for financial and technical support.

REFERENCES

[1] M. A. Gerzon, «The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound,» *50th AES Convention,* 1975.

[2] M. A. Gerzon, "Ambisonics. Part two: Studio techniques," *Studio Sound,* vol. 17, no. 10, p. 60, 1975.

[3] N. M. Ferres, "An Elementary Treatise on Spherical Harmonics and Subjects Connected with them," *London: Macmillan and Co.,* 1877.

[4] A. Farina, "Explicit formulas for High Order Ambisonics," August 2017. [Online]. Available: http://www.angelofarina.it/Aurora/HOA_explicit_formulas.htm.

[5] R. Gonzalez, J. Pearce and T. Lokki, "Modular Design for Spherical Microphone Arrays," *Proc. of the AES International Conference,* vol. 2018, pp. 140 - 146, August 2018.

[6] S. Moreau, J. Daniel and S. Bertet, "3D sound field recording with higher order ambisonics-objective measurements and validation of spherical microphone," *120th AES Convention,* 2006.

[7] A. Farina, A. Capra, L. Conti, P. Martignon and F. Fazi, "Measuring spatial impulse responses in concert halls and opera houses employing a spherical microphone array," *19th International Congress On Acoustics,* September 2007.

[8] N. Peters and A. W. Schmeder, "Beamforming using a spherical microphone array based on legacy microphone characteristics," *Proc. of the International Conference on Spatial Audio (ICSA),* 2011.

[9] S. Sakamoto, S. Hongo, T. Okamoto, Y. Iwaya and Y. Suzuki, "Sound-space recording and binaural presentation system based on a 252-channel microphone array," *Acoustical Science and technology,* vol. 36, no. 6, pp. 516 - 526, 2015.

[10] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. on IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* vol. 2, pp. 1781 - 1784, 2002.

[11] A. Farina, D. Pinardi, M. Binelli, M. Ebri and L. Ebri, "Virtual reality for subjective assessment of sound quality in cars," *144th AES Convention,* 2018.

[12] M. Binelli, D. Pinardi, T. Nili and A. Farina, "Individualized HRTF for playing VR videos with Ambisonics spatial audio on HMDs," *AES International Conference on Audio for Virtual and Augmented Reality,* 2018.

[13] B. Rafely, Fundamentals of Spherical Array Processing, Springer-Verlag Berlin Heidelberg, 2015.

[14] A. Farina, M. Binelli, A. Capra, E. Armelloni, S. Campanini and A. Amendola, "Recording, Simulation and Reproduction of Spatial Soundfields by Spatial PCM Sampling (SPS)," *International Seminar on Virtual Acoustics,* 2011.

[15] S. Delikaris-Manias and V. Pulkki, "Parametric Spatial Filter Utilizing Dual Beamformer and SNR-Based Smoothing," in *AES 55th International Conference: Spatial Audio*, 2014.

[16] V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," in *Proc. 28th AES International Conference*, Pitea, 2006.

[17] V. Pulkki, "Spatial Sound Reproduction with Directinal Audio Coding," *Journal of the AES,* vol. 55, no. 6, pp. 503-516, 2007.

[18] S. Berge and N. Barrett, "High Angular Resolution Planewave Expansion (HARPEX)," in *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, 2010.

[19] V. Pulkki, A. Politis, M. V. Laitinen, J. Vilkamo and J. Ahonen, "First-order directional audio coding (DirAC)," in *Parametric Time-Frequency Domain Spatial Audio*, 2017, pp. 89-138.

[20] O. Kirkeby, F. Orduna, P. A. Nelson and H. Hamada, "Inverse filtering in sound reproduction," *Measurement and Control,* vol. 26, no. 9, pp. 261 - 266, November 1993.

[21] H. Tokuno, O. Kirkeby, P. A. Nelson and H. Hamada, "Inverse filter of sound reproduction systems using regularization," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences,* Vols. E80-A, no. 5, pp. 809 - 820, 1997.

[22] R. H. Hardin and N. J. A. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions," *Discrete and Computational Geometry,* vol. 15, pp. 429-441, 1996.

[23] C. Nachbar, F. Zotter and E. Deleflie, "Ambix - A suggested Ambisonics format," *Ambisonics Simposium,* 2011.

[24] L. McCormack, S. Delikaris-Manias, A. Politis, D. Pavlidi, A. Farina, D. Pinardi and V. Pulkki, "Applications of spatially localized active-intensity

vectors for sound-field visualization," *Journal of the Audio Engineering Society,* vol. 67, no. 11, pp. 840-854, 2019.

[25] L. McCormack, «Spatial Audio Real-time Applications (SPARTA),» Aalto University, [Online]. Available: http://research.spa.aalto.fi/projects/sparta_vsts/.

[26] L. McCormack, S. Delikaris-Manias, A. Farina, D. Pinardi and V. Pulkki, "Real-time conversion of sensor array signals into spherical harmonic signals with applications to spatially localised sub-band sound-field analysis," *144th AES Convention,* 2018.

[27] Institute of Electronic Music and Acoustics, "IEM Plug-in Suite," [Online]. Available: https://plugins.iem.at/download/.

[28] J. W. S. Rayleigh, The Theory of Sound - Volume II, New York: Dover Publications, 1896.

[29] A. Farina, A. Capra, L. Chiesi and L. Scopece, "A Spherical Microphone Array for Synthesizing Virtual Directive Microphones in Live Broadcasting and in Post Production," *40th International Conference: Spatial Audio: Sense the Sound of Space,* 2010.

[30] A. Farina, S. Campanini, L. Chiesi, A. Amendola and L. Ebri, "Spatial Sound Recording With Dense Microphone Arrays," *55th AES Conference,* August 2014.

[31] A. Politis, "Acoustical Spherical Array Processing Library," Department of Signal Processing and Acoustics, Aalto University, Finland, 2016. [Online]. Available: http://research.spa.aalto.fi/projects/spharrayproc-lib/spharrayproc.html#59.

[32] D. Pinardi, "Spherical t-Designs for Characterizing the Spatial Response of Microphone Arrays," in *I3DA - International Conference on Immersive and 3D Audio*, Bologna, 2021.

[33] D. Pinardi e A. Farina, «Metrics for Evaluating the Spatial Accuracy of Microphone Arrays,» in *International Conference on Immersive and 3D Audio*, Bologna, 2021.

[34] A. Costalunga, C. Tripodi, L. Ebri, M. Vizzaccaro, L. Cattani, E. Ugolotti and T. Nili, "Experimental results on active road noise cancellation in car interior," *144th AES Convention,* May 2018.

[35] D. Pinardi, L. Ebri, C. Belicchi, A. Farina e M. Binelli, «Direction Specific Analysis of Psychoacoustics Parameters inside Car Cockpit: A Novel Tool for NVH and Sound Quality,» in *SAE Technical Papers*, Graz, 2020.

D. Pinardi received the M.S. (Cum Laude) degree in mechanical engineering from University of Parma, Italy, in July 2016, with a thesis on loudspeaker modelling. In March 2020, he got the Ph.D. degree in industrial engineering with a thesis on the design of microphone, hydrophone and camera arrays for spatial audio recording.

He is a research assistant of Prof. Angelo Farina at the Dept. of Engineering and Architecture from 2016, mainly specialized in spatial audio. Topics of his interest are design of sensor arrays, acoustic simulations and 3D auralization, applied to automotive field and underwater acoustics.