



## Research Article

Giorgio Baldinelli\*, Francesco Bianchi, Danilo Costarelli, Francesco D'Alessandro, Flavio Scrucca, Marco Seracini, and Gianluca Vinti

# Innovative techniques for the improvement of industrial noise sources identification by beamforming

<https://doi.org/10.1515/noise-2021-0010>

Received Nov 15, 2020; accepted Feb 18, 2021

**Abstract:** An innovative technique based on beamforming is implemented, at the aim of detecting the distances from the observer and the relative positions among the noise sources themselves in multisource noise scenarios. By means of preliminary activities to assess the optical camera focal length and stereoscopic measurements followed by image processing, the geometric information in the source-microphone direction is retrieved, a parameter generally missed in classic beamforming applications. A corollary of the method consists of the possibility of obtaining also the distance among different noise sources which could be present in a multisource environment. A loss of precision is found when the effect of the high acoustic reflectivity ground interferes with the noise source.

**Keywords:** beamforming, noise sources, industrial plants

## 1 Introduction

Industrial plants play a fundamental role in the economic life of a territory, generating wealth and employment; on the other hand, their environmental impact can be detrimental for the quality of life of the citizenship. Obligations become more stringent when the plants are located in urban or peri-urban areas, close to residential or noise sensitive buildings (schools or hospitals) [1–3]. Limiting to the environment noise, large industrial plants are usually characterized by many sound sources which emissions vary in space and time, having different frequency spectra and characteristics (*e.g.*, presence of impulsive or tonal components). The set of these sources contributes to the global noise level emitted by the plant and received by the surrounding buildings and workplaces, which must comply with the national legal limits. Therefore, the knowledge of the individual sources contribution becomes essential when it is necessary to perform corrective actions on the plant, for instance limiting the emissions of the most disturbing sources or verifying the absence of malfunctioning organs with emissions higher than the standard. Common sound level meters are not suitable for this purpose and other techniques, such as sound intensity, are of limited use for implants of large scale.

In the last couple of decades, however, a new tool has become available for professionals: the acoustic camera, which is essentially a planar array of microphones plus a video camera. Acoustic cameras are becoming cheaper with the years and can host an increasing number of microphones thanks to the development of the technology of low-price MEMS transducers. Using acoustic cameras, graphical maps of sound pressure level or other acoustic descriptors can be obtained, allowing the rough localization of the different sound sources [4–6].

If no geometric information is available on the scenario under investigation, acoustic cameras do not allow to spatially position the noise sources retrieved. Many efforts were spent in this direction, using different techniques [7–9]; the

\***Corresponding Author: Giorgio Baldinelli:** Department of Engineering, University of Perugia, Via Duranti, 67 – 06125 Perugia – Italy; Email: giorgio.baldinelli@unipg.it; Tel.: +39 075 585 3868; Fax: +39 075 585 3697

**Francesco Bianchi:** Department of Physics and Geology, University of Perugia, Via Pascoli – 06123 Perugia – Italy

**Danilo Costarelli, Marco Seracini, Gianluca Vinti:** Department of Mathematics and Computer Science, University of Perugia; Via Vanvitelli, 1 – 06123 Perugia, Italy

**Francesco D'Alessandro:** Italian National Research Council – Institute on Atmospheric Pollution at Italian Ministry of Environment, Land and Sea, Via Cristoforo Colombo, 44 – I, 00147, Rome, Italy

**Flavio Scrucca:** Department of Sustainability – Circular Economy Section, ENEA – Italian National Agency for New Technologies, Energy and Sustainable Economic Development; Lungotevere Thaon di Revel, 76 – 00196, Rome, Italy



method presented in this work is aimed at proposing a simple approach, based on a limited number of measurements and a post-processing analysis to assess the distance among noise sources and their distance from the microphones array.

The spatial placing of the most relevant sound emitters assumes a particular relevance in the phase of noise reduction of industrial sources; an easy way to define their spatial location adds a fundamental information to the standard two-dimensional picture obtained from the beamforming software.

## 2 Beamforming background

Commercially available acoustic cameras use mainly two techniques: acoustic holography and beamforming. Whereas acoustic holography uses sound field data measured over a planar surface close to the source (NAH stands for Near field Acoustic Holography) to estimate a 3D mathematical model of the sound field, beamforming works at medium to long distance from the source and at higher frequencies than acoustic holography [10].

Being beamforming the technique used in this study, some applications and a brief theoretical background are reported in the following.

The acoustic beamforming consists of a spatial filtering technique that permits locating specific sound sources from a variety of other insignificant sources, based on the direction of arrival.

This method uses signal processing algorithms that allows to spatially isolate a sound coming from a precise direction decided by the user, by creating a virtual super-directional microphone that can be oriented starting from a microphone array of known characteristics.

This technique has a wide range of applications [11]: it is the basis of speech recognition; these algorithms are used in audio-video conference rooms where speaker and camera localization is difficult to implement manually. They are used in smartphones and telephones especially when the speaker is used: the caller does not hear his voice back when the voice is transmitted through the loudspeaker. These algorithms also have applications in seismology and they are used to study the wave reflections on the surface of the earth.

Many other fields of interest involve beamforming [12]: it is used as a part of Active Noise Cancellation System in the car where the vehicle takes the voice instructions irrespective of the presence of background noises and music from the infotainment systems by amplifying or gain to

the signal. Beamforming finds also application in the fault detector system where engineers can locate the issues in a component present in an automobile or aircraft without actually opening them completely. It is used to monitor the component and the engine health before getting involved in the overhaul process, thereby reducing a huge amount of maintenance repair and overhaul timings.

This technique makes use of flat and linear arrays placed exclusively in the far-field region and composed of omnidirectional microphones, which do not present attenuation of the signal based on the angle of incidence of the sound.

The beamforming allows obtaining, in short times, an acoustic map of the area under examination, superimposed on the photographic image of the area itself. Thanks to the type of representation offered by this technique, it is therefore possible to see the directions of origin of the noise and then proceed to the identification of the sound sources present in the environment studied quickly and effectively.

Two main beamforming algorithms are used: “delay and sum in the time domain” and “delay and sum in the frequency domain” [12].

### 2.1 Delay and sum beamforming in the time domain

The delay and sum beamforming in the time domain algorithm is the most frequently used and the apparatus that allows its application is described as follows.

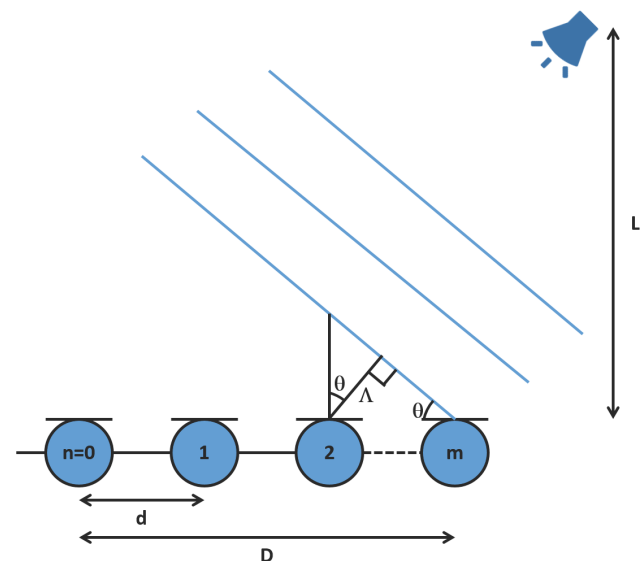


Figure 1: Graphical representation of the linear array and of the source

A linear array of width  $D$  composed of  $m$  omnidirectional microphones arranged at regular intervals  $d$  (where the distance is measured from the central axis of each microphone) is considered (Figure 1).

A block that has the function of temporal delay is connected to each microphone (not specific which type of block as it is sufficient to perform a translation in time of the input signal) and finally the data of all the blocks are connected to a signal adder which outputs the processed signal.

Considering an extended source of sound energy located at a distance  $L$  from the microphone array and placing ourselves in the far field region, we can schematize it as if it were made up of an infinite number of sources placed in an orderly manner over the entire length of the source. Therefore, each microphone will detect the superposition of the signals coming from each angle  $\theta$  to which the points constituting the extended sound source are placed respectively. Initially, we consider only a signal coming from a single fixed angle  $\theta$ , and then implement it for every angle. We assume that the wavefronts seen by the array are flat and that the attenuation of the intensity of the same signal picked up by two different microphones is negligible. The important thing about putting ourselves in far field is that the wavefronts emitted by a single source have the same angle  $\theta$  in each microphone of the array and, therefore, based on the position of the source with respect to the central axis of the array, the same signal will be picked up by each microphone with a time delay dependent on the angle  $\theta$ .

In fact, knowing the position of the source at a certain time  $t$ , the algorithm allows to reconstruct the emitted signal knowing the time delay and performing an appropriate correction in amplitude. The signals of all the  $m$  microphones positioned in the acoustic field perturbed by the source are acquired simultaneously; each signal is then shifted in time to the emission time knowing the path presumably traveled by the sound wave, and finally all  $N$  signals thus obtained are averaged.

The  $\Delta$  delay can be calculated trigonometrically by knowing the wave propagation speed (sound speed  $c$ ) and the distance  $d$  between each microphone using the following formula:

$$\Delta(\theta) = \frac{dsen(\theta)}{c} \Rightarrow \Delta_n(\theta) = \frac{dsen(\theta)}{c}n \quad (1)$$

where  $\Delta_n(\theta)$  is the given signal delay in the  $n$ -th microphone with respect to the signal at microphone  $n = 0$ ; note that the delay is dependent on the angle of incidence of the sound front.

Another variable is introduced:

$$\Lambda(\theta) = dsen(\theta) \Rightarrow \Lambda_n(\theta) = dsen(\theta)n \quad (2)$$

Now, considering a signal coming from a source at the angle  $\theta_0$  and position  $x$

$$F(x, t) = s_{\theta_0} \left( \frac{x}{c} - t \right) \quad (3)$$

the same signal will be picked up by the  $n$ -th microphone like:

$$r_{n,\theta_0}(x, t) = s_{\theta_0} \left( \frac{x + \Lambda_n(\theta_0)}{c} - t \right) \quad (4)$$

The signal coming out from the transducers will travel the line to the delay blocks with a defined  $\Delta_n(\theta)$ , given by the second part of Equation 1. As each signal encounters its own delay block, each of them will be temporally translated so that when the signals exit the blocks they will be temporally aligned at the same time instant:

$$t_{n,\theta_0}(x, t) = s_{\theta_0} \left( \frac{x + \Lambda_n(\theta_0)}{c} - (t + \Delta_n(\theta_0)) \right) \quad (5)$$

Then, each signal will go through the summing block that will join them and, being all temporally aligned, a considerably strengthened signal will be produced:

$$y_{\theta_0}(x, t) = \sum_{n=1}^m s_{\theta_0} \left( \frac{x + \Lambda_n(\theta_0)}{c} - (t + \Delta_n(\theta_0)) \right) \quad (6)$$

Now, focusing the attention to another source at an angle  $\theta_1 \neq \theta_0$ , it will produce on the microphones new signals that will undergo the same process described before. The only difference is that, when they meet the delay blocks, the latter, being set in relation to the angle  $\theta_0$ , will move them in such a way that when they reach the adder they are no longer aligned. Therefore, once the adder combines them together, the generated signal will possess limited intensity:

$$y_{\theta_1}(x, t) = \sum_{n=1}^m s_{\theta_1} \left( \frac{x + \Lambda_n(\theta_1)}{c} - (t + \Delta_n(\theta_0)) \right) \quad (7)$$

Knowing the temporal translation that the delay blocks produce, it is known at what angle  $\theta_0$  the amplified signal arrives. Now, considering an infinity of sources each placed at an angle  $\theta$  between  $\theta_{min}$  and  $\theta_{max}$ , the  $n$ -th microphone directly picks up a signal  $s_n$  which is already the sum of the signals coming from all angles  $\theta$  with  $\theta_{min} < \theta < \theta_{max}$ . In this case the received signal is:

$$r_n(x, t) = s_n \left( \frac{x}{c} - t \right) \quad (8)$$

and the Equation 7 is generalized as follows:

$$y(x, t) = \sum_{n=1}^m s_n \left( \frac{x}{c} - (t + \Delta_n(\theta_0)) \right) \quad (9)$$

Therefore, each microphone of the array picks up the signals coming from every direction but, at the exit, the function  $y(x; t)$  is, in good approximation, only the signal coming from  $\theta_0$ .

## 2.2 Delay and sum beamforming in the frequency domain

One of the improvements that can be implemented to increase the effectiveness of the algorithm is to no longer work in the time domain, moving to the frequency domain, reason for what the method is called frequency domain beamforming [13].

This is done because, in real cases, broadband signals have to be managed and a discrete window that includes a finite bandwidth must be used.

The working principle is almost similar to the one previously seen, but in this case the time delay is applied to the signals in the frequency domain and the signals acquired by the microphones are processed through DFT (Discrete Fourier Transform) and IDFT (Inverse Discrete Fourier Transform) [14].

Considering that DFT works only with discrete signals, each data received by a microphone must be sampled in order to obtain a discrete signal:

$$r_n(x, t) = s_n\left(\frac{x}{c} - t\right) \Rightarrow r_{n,x,t} = s_{n,x,t} \quad (10)$$

The  $m$  signals will pass through  $m$  DFT blocks, and every block will output  $N$  signals with different frequency  $\omega$ , i.e. a frequency spectrum:

$$R_{n,x,\omega} = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} r_{n,x,t} e^{i\frac{\omega}{N}t} \quad (11)$$

Then, every  $\omega$ -th signal coming from the  $n$ -th DFT block will meet a specific block that will add a phase equal to  $\omega\Delta_n(\theta_0)$ ; in fact, in the frequency domain a time delay can be applied by means of a circular translation  $T$ :

$$T_{n,x,\omega} = R_{n,x,\omega} e^{i\omega\Delta_n(\theta_0)} \quad (12)$$

The  $n$  signals with the same frequency  $\omega$  are then summed by the adder in order to obtain:

$$Y_{x,\omega} = \sum_{n=1}^m R_{n,x,\omega} e^{-i\omega\Delta_n(\theta_0)} \quad (13)$$

Finally, the IDFT is applied in order to pass from a discrete to a continuous signal:

$$\begin{aligned} y_{x,t} &= \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} Y_{x,\omega} e^{-i\frac{\omega}{N}t} \\ &= \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \sum_{n=1}^m R_{n,x,\omega} e^{-i\omega\left(\frac{t-\Delta_n(\theta_0)}{N}\right)} \end{aligned} \quad (14)$$

Obtaining the desired signal  $y(x,t)$ :

$$y_{x,t} = y(x, t) \quad (15)$$

## 3 Material and methods

### 3.1 Instruments

The instruments to perform the beamforming measurements is made of a microphone array that allows to retrieve all the parameters to implement the algorithms previously described. The first concept was developed in World War II as a radar antenna, but the father of the microphone array can be considered John Billingsley [14, 15] who proposed an acoustic telescope based on a microphone array to localize in real-time the sound source on full-size jet-engines.

Nowadays, thanks to the technology progress, a more powerful acquisition and computing system allow higher sampling frequencies, longer acquisition times, larger number of microphones and real-time analysis that have strongly improved the instrument performance.

Near-field Acoustic Holography (NAH) and beamforming algorithm are the two main techniques that use the microphone array setup. In principle, the microphone array for each technique should have the specific feature to perform the measurements: NAH can perform source location using the microphone distance (grid) less than half wavelength with good resolution by implementing measures very close to the source; beamforming technique, thanks to the irregular array position, presents good resolution for longer distance to the sound source.

The possibility to match NAH and beamforming in a single array are so called pseudo-random microphone distribution, with a number of microphones ranging from 40 to 60-80-100. The hardware implemented to perform the measurements for this work [16] allows to perform both NAH and beamforming measurements with a pseudo-random array (Figure 2). Its characteristics of are reported in the Table 1:

**Table 1:** Main characteristics of the pseudo-array

Array material	Composite material
<b>Microphones</b>	42
<b>Sample Rate</b>	48 kHz
<b>Mic. Frequency Range</b>	10 Hz to 24 Hz
<b>Operating Range</b>	< 33 dB to 120 dB
<b>Analysis Dynamic Range</b>	Up to 40 dB
<b>Measurement distance</b>	0.2 m to
<b>Resolution</b>	24 bit
<b>Operating temperature</b>	-40°C to +60°C

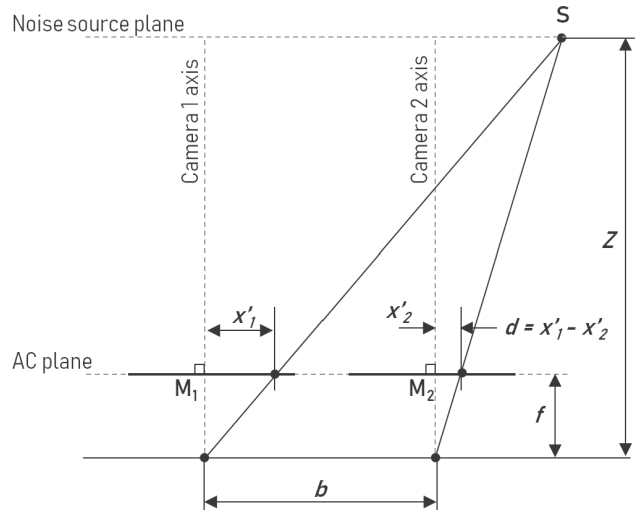


**Figure 2:** Pseudo random microphone array used for the measurements

The new generation of low-cost sensors (MEMS) allow to have a large number of microphones with low power consumption, high SNR=61 dBA, high sensitivity around -26 dBFS and flat frequency response from 60 Hz to 15 kHz. A webcam to retrieve the digital image is placed at the centre of the microphone array.

### 3.2 Distance calculation method

There are many approaches used to calculate the distance of a reference point relative to the position of a camera. These methods can be active and use a signal (such as radio waves, microwaves, infrared, etc.) sent to the object, or they can be passive and only receive information about the target position. Among passive methodologies, the most popular are those relying on stereoscopic measurements based on the use of two cameras displaced one from each other by a known distance. Given the basis of stereo imaging, as an alternative, two or more images taken from a moving camera can also be used to compute distance information [17–19].



**Figure 3:** Stereo imaging basis for the estimation of  $Z$

Therefore, the distance  $Z$  between a noise source and the measurement point where the acoustic camera (AC) is positioned can be determined by performing two different measurements  $M_1$  and  $M_2$  at a distance  $b$  in the plane of the AC (Figure 3). Anyway, a characteristic constructive parameter of the optical camera rigidly connected to the microphones array of the AC has to be known: the focal length  $f$ . It represents the distance between the lens and the image sensor when the subject is in focus, *i.e.* the optical distance from the point where light rays converge to form a sharp image of an object to the digital sensor at the focal plane in the camera. The optical camera included in the used AC has a fixed  $f$ .

Actually, if  $f$  is known, applying basic triangulation, the calculation of  $Z$  can be obtained through the following equation:

$$Z = f \cdot \frac{b}{d} \quad (16)$$

In the equation,  $d$  represents the disparity between the two images expressed in pixels, *i.e.*, at the aim of this work, the distance between the points of maximum sound pressure level where the noise source is localized. The disparity  $d$  can assume different values according to the resolution of the acquired acoustic data and, in particular,  $d$  changes if it is calculated starting from raw or processed data, depending on the typology of the interpolation method used to improve the resolution. Consequently, different estimation of  $Z$  can be found.

The knowledge of  $Z$  also allows to calculate the size of a single pixel, in fact, if the dimension of some reference objects inside the image is known, the size at distance  $Z$ ,  $P_s(Z)$ , is given by the following linear equation:

$$P_s(Z) = P_s(D) \cdot \frac{D}{Z} \quad (17)$$

being  $D$  a depth for which the size of a single pixel  $P_s(D)$  is known. Consequently, knowing  $P_s(Z)$ , simply multiplying it by the number of pixels between two (or more) noise sources, it is possible to obtain an estimation of their real distance in the plane of the acoustic image.

### 3.3 Experimental activities

A first step of the experimental activities was focused on the estimation of the focal length  $f$  of the optical camera included in the AC, since it was missing in the datasheets available and was not possible to retrieve it from a web search. In order to determine  $f$ , on the basis of other characteristic parameters of the camera, the procedure introduced by Zhang [20] was followed. This method requires the camera to detect a planar pattern shown at a certain distance with few different orientations. A specific camera calibrator based on this method is available in MATLAB® [21]. According to the workflow and the instructions given, we used an A3 format checkerboard pattern with one side containing an even number of black and white squares and the other side containing an odd number of them, attached to a planar surface. A number of 15 images at a distance of 1 m – thus ensuring that the checkerboard filled more than the 20% of each image – were captured, placing the checkerboard at an angle less than 45 degrees relative to the camera plane (Figure 4).

The captured images (JPEG format) were added to the calibrator and it was verified that they met the requested specifications, through a specific built-in analysis, before running the calibration procedure. The described methodology, together with the available data, allowed to calculate an  $f$  equal to 2173.6 pixel.

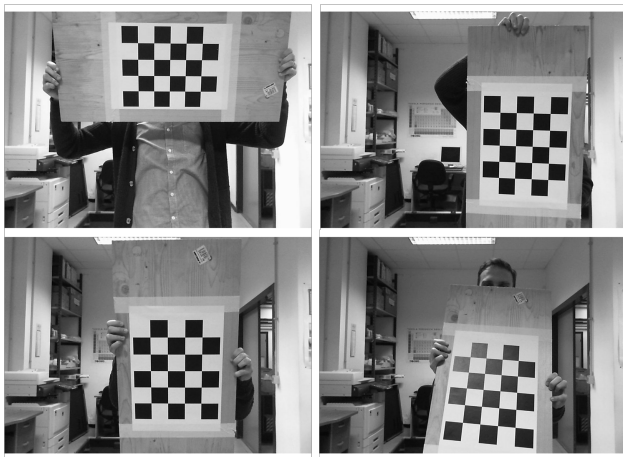


Figure 4: Image acquisition for camera calibration

Once  $f$  was determined, we considered an experimental set up (Figure 7) in an acoustic open field in order to verify the accuracy of the calculation of  $Z$  and the inter-sources distance  $X$  using both raw and processed acoustic data.

The experimental set up included two noise sources ( $S_1$  and  $S_2$ ) consisting of two speakers, respectively positioned on the left and on the right side of the camera optical axis at a distance of 3.00 m, with a resulting total distance between  $S_1$  and  $S_2$  ( $X$ ) equal to 6 m.  $S_1$  was set at an height of 1.50 m, while the height of  $S_2$  was 0.70 m.  $S_1$  and  $S_2$  was tuned to sinusoidal signals of different frequency ( $f_1$  and  $f_2$ ) not integer multiples each other, in order to avoid any summation effects that could result from spurious harmonics superposition. In particular,  $f_1$  was 1,733 Hz and  $f_2$  was 1,000 Hz. The optical camera of the AC was positioned at a height of 1.35 m and the distance between the AC plane and the plane passing for  $S_1$  and  $S_2$  was 12.25 m. The ground of the experimental setup is made of asphalt, with an acoustic reflectivity coefficient around 0.95 at all frequencies [22].

In order to determine  $Z$  according to the method described in the previous section, two measurements ( $M_1$  and  $M_2$ ) were carried out moving the AC on its plane, so that  $b$  was equal to 0.23 m (Figure 5), acquiring both optical and acoustic data.

Acoustic data, expressed as sound pressure levels in dB, were stored in TDMS (Technical Data Management Solution by National Instruments™) files, that were 16 x 21 matrices, and they were also post-processed so to have the

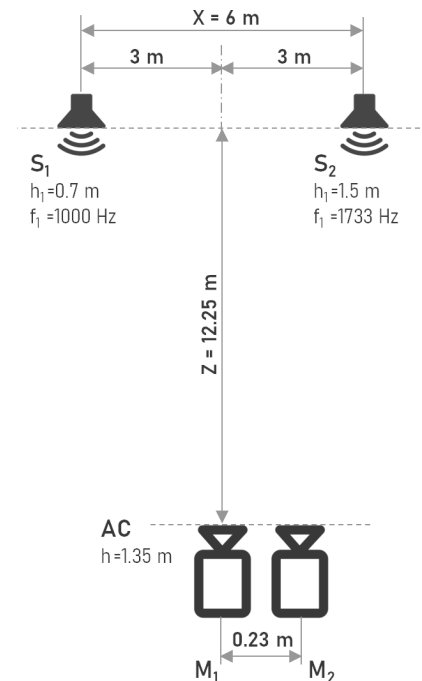
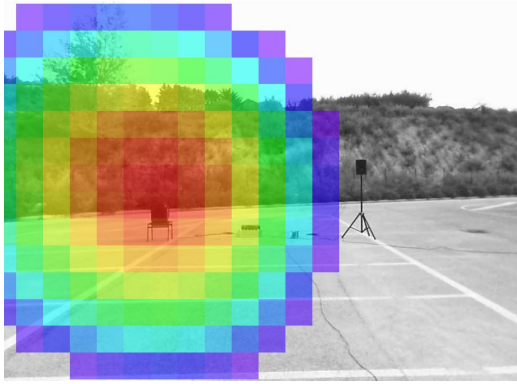
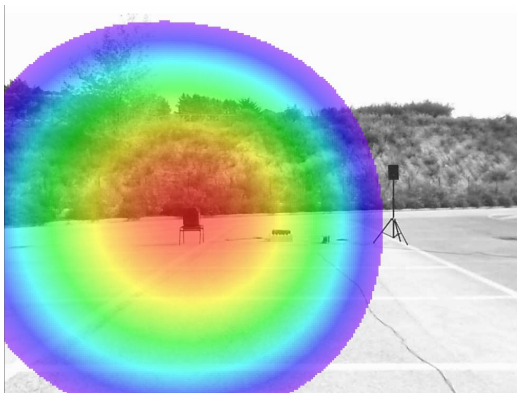


Figure 5: Experimental set up for the measurements



**Figure 6:** Example of an image given by the AC software with raw data



**Figure 7:** Example of an image given by the AC software with interpolated data

data deriving respectively from  $S_1$  and  $S_2$ . Such a construction was made using specific AC software filters and was carried out obtaining other matrices including the spectral components contained in a band of 100 Hz centered to  $f_1$  and  $f_2$ .

There is a correspondence “1 acoustic pixel = 48 x 48 optical pixels” in the raw acoustic data. In fact, as shown in Figure 6, images produced by the AC software without data interpolation are gray scale images of  $1024 \times 780$  pixels, where an “acoustic pixel” corresponds to a square of  $48 \times 48$  optical pixels. Using these raw acoustic data, a first value of  $Z(Z_{raw})$  was calculated.

Subsequently, the resolution of the acoustic data was improved at the aim of assigning a single value of sound pressure to each pixel of the optical acquisition and, therefore, improving the accuracy of  $Z$  estimation. Two standard bivariate interpolation methods were used: the bilinear interpolation and the bicubic interpolation [23, 24], and thus two further values of  $Z$  ( $Z_{bilinear}$  and  $Z_{bicubic}$ ) were calculated. These interpolation methods are also implemented

in the AC post-processing software and through their use it is possible to obtain images like the one of Figure 7.

The three calculated values were then compared to the real known distance between the AC and the sources (12.25 m), in order to evaluate the accuracy of the estimations of the method both using raw and processed acoustic data. Consequently, on the basis of the three  $Z$  values, three values of  $X$  were estimated:  $X_{raw}$ ,  $X_{bilinear}$  and  $X_{bicubic}$ . Obviously, carrying out the calculation,  $S_1$  and  $S_2$  were considered two point-sources with the sound emitting elements located in the center of each speaker.

## 4 Results

The disparity  $d$  is the main output of the experimental measurements and it can be calculated both for sources  $S_1$  and  $S_2$ . In Table 2 are reported the values of disparity, in pixel, for the three levels of image refinement: raw, bilinear and bicubic.

**Table 2:** Disparity calculated for both sources

$d$ (pixel)	$S_1$	$S_2$
Raw	1.00	2.00
Bilinear	47.01	96.00
Bicubic	38.01	82.00

Considering the calculated disparity  $d$ , the above described algorithms to retrieve the distance  $Z$  between the sources plane and the measuring point plane  $M$ , have been applied to the three different matrices obtained from raw data and interpolated data.

The values of the distance  $Z$  have been reported in Table 3. According to the values obtained for  $d$ , even if both the sources are in the same plane parallel to the AC plane, results obtained considering  $S_1$  and  $S_2$  are strongly different. In particular, looking in detail the  $S_2$  results, it clearly emerges that the estimation of  $Z$  is scarcely accurate, as a consequence of the not directivity of  $S_2$ , the high reflectivity coefficient of the asphalt and the height of  $S_2$ , that

**Table 3:** Results obtained with data from  $S_1$

$Z$ (m)	$S_1$	$S_2$
Raw	4.99	2.49
Bilinear	10.63	5.21
Bicubic	13.15	6.10

is a half of the one of  $S_1$ . Given these characteristics, the contribution of reflections occurring on the ground were not negligible and they influenced the post-processing of data.

Considering the  $S_1$  results, instead, it is worth noting that raw data give an estimation of  $Z$  that is not correct, while the bilinear and bicubic interpolation techniques both give a quite accurate estimation. With respect to the known value of the distance  $Z$  between sources and Acoustic Camera (12.25 m), data from  $S_1$  allow to calculate values with percentage difference in the order of about 7-13%. In detail, starting from data interpolated with a bilinear technique, the difference is -13.2% (-1.62 m), while using data interpolated with a bicubic technique the difference is +7.3% (+0.90 m).

Regarding the estimation of the distance  $X$  between the acoustic sources  $S_1$  and  $S_2$ , the analysis of the image allows to obtain the information on the number of pixels between them, which is the necessary data to calculate the Euclidean distance according to equation 17. Using a reference object, in fact, it is possible to define the size of the pixel in the image at a certain distance from AC and consequently to determine the Euclidean distance  $X$ , simply multiplying it by the obtained number of pixels. In Table 4 the number of pixels, the size of each single pixel and the estimation of the Euclidean distance between the sources are reported.

**Table 4:** Results obtained with data from  $S_1$

	$N$ (pixel)	$P_s(Z)$ (mm)	$X$ (mm)
Raw	10.20	496.58	5064.12
Bilinear	489.51	10.56	5169.18
Bicubic	482.61	13.06	6302.82

The better estimation of the distance between the sources is given by the bicubic interpolation technique. The percentage difference calculated with respect to the measured value of 6.05 m (real Euclidean distance): in fact, it is equal to -14.6% for the bilinear interpolation and +4.2% for the bicubic interpolation technique, while the estimation obtained from raw data represents the worst case (-16%).

## 5 Conclusions

The innovative technique proposed gives potentially interesting results on defining both the distance among noise sources and observer and the distances among the noise

sources themselves, in the beamforming methodology for noise sources identification. Results show that the third dimension in the beamforming analysis could be retrieved, with a certain accuracy, moving the microphones array on two positions, while the sound emission remains constant.

The method allows also to obtain the relative real distance among different noise sources arising from a multi-source environment.

A loss of precision is highlighted when the source is positioned close to the ground; a possible reason could be searched in the acoustic reflection of the ground itself, an effect that could be investigated in more details in future analyses. If these effects were confirmed, the data elaboration should take into account of the ground (or eventual vertical obstacles) acoustic properties.

The approach could be improved by means of better mathematic interpolation algorithms, at the aim of increasing the accuracy of the post-processing analysis of data derived from the beamforming technique.

**Funding information:** The authors state no funding involved.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Conflict of interest:** The authors state no conflict of interest.

## References

- [1] Asdrubali F, D'Alessandro F. Innovative Approaches for Noise Management in Smart Cities: a Review. *Curr Pollut Rep.* 2018;4(2):143–53.
- [2] Susanto A, Setyawan DO, Setiabudi F, Savira YM, Listiarini A, Putro EK, et al. GIS-based mapping of noise from mechanized minerals ore processing industry. *Noise Mapp.* 2021;8(1):1–15.
- [3] Aletta F, Guattari C, Evangelisti L, Asdrubali F, Oberman T, Kang J. Exploring the compatibility of “Method A” and “Method B” data collection protocols reported in the ISO/TS 12913-2:2018 for urban soundscape via a soundwalk. *Appl Acoust.* 2019;155:190–203.
- [4] Chiariotti P, Martarelli M, Castellini P. Acoustic beamforming for noise source localization—Reviews methodology and applications. *Mech Syst Signal Process.* 2019;120:422–48.
- [5] Bai MR, Lee J. Industrial noise sources identification by using an acoustic beamforming system. *J Vib Acoust.* 1998;120(2):426–33.
- [6] Alloza P, Vorrhein B. Noise source localization in industrial facilities. *Proc. InterNoise 2019.* Madrid, Spain; 16-19 June 2019.



- [7] Ding H, Lu H, Li C, Jing J, Mei D, Chai G. Localization and identification of three-dimensional sound source with beamforming based acoustic tomography. *Transaction of the ICA 2013*. Montreal, Canada; 2-7 June 2013. <https://doi.org/10.1121/1.4801076>
- [8] Padois T, Gaudreau MA, Marcotte P, Laville F. Identification of noise sources using a time domain beamforming on pneumatic, gas and electric nail guns. *Noise Control Eng J*. 2019;67(1):11–22.
- [9] Leclère Q, Pereira A, Bailly C, Antoni J, Picard C. A unified formalism for acoustic imaging based on microphone array measurements. *Int J Aeroacoust*. 2017;16(4):431–56.
- [10] Christensen J, Hald J, Mørkholt J, Schuhmacher A, et al. A Review of Array Techniques for Noise Source Location. *SAE Technical Paper*. 2003;2003-01-1691.
- [11] Tripathy T. Acoustic Beamforming, Report of the course 85-525 Automotive Applications for Noise, Vibrations and Harshness. University of Windsor; 2007.
- [12] Ferri A. Il metodo del Beam-forming acustico (in Italian). Master thesis. University of Bologna, Italy; 2005.
- [13] Brambilla G, Lo Castro F, Cerniglia A, Verardi P. Visualizzazione 3D del campo acustico mediante un sistema ad array microfonico sferico (in Italian), *Proc Nation Conf Ital Acoust Assoc*. Florence, Italy; 13-15 June 2007.
- [14] Billingsley J, Kinns R. The acoustic telescope. *J Sound Vibrat*. 1976;48(4):485–510.
- [15] Ulf M., History of Acoustic Beamforming. *Berlin Beamforming Conf*. Berlin, Germany; 21- 22 November 2006.
- [16] Amadasi G, Dannat T. Overview of Acoustic Holography and Beamforming techniques for Noise Sources recognition in open/closed environments: choosing the correct theory to apply for sound field reconstruction, *Vibro-Acoustic web site*, <https://www.vibro-acoustic.com>, Retrieved on January 2020.
- [17] Mahammed MA, Melhum AI, Kochery FA. Object distance measurement by stereo vision. *IJSAIT*. 2013;2(2):5–8.
- [18] Patel DK, Bachani PA, Shah NR. Distance measurement system using binocular stereo vision approach. *Int J Eng Res Technol (Ahmedabad)*. 2013;2(12):2461–4.
- [19] Mohamed A, Yang C, Cangelosi A. Stereo vision based object tracking control for a movable robot head. *Proc 4<sup>th</sup> IFAC Conf Intel Contr Automat Sci ICONS 2016*. Reims, France; 1–3 June 2016.
- [20] Zhang ZA. Flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell*. 2014;22(11):1330–4.
- [21] MathWorks® Single Camera Calibrator App. Available at: <https://it.mathworks.com/help/vision/ug/single-camera-calibrator-app.html>
- [22] Cruz RA, Correa CR, Diaz-Ramirez GA. Acoustic properties of concrete modified with an asphalt/styrene butadiene emulsion. *J Phys Conf Ser*. 2019;1247(012036):1–9.
- [23] Keys RG. Cubic Convolution Interpolation for Digital Image Processing. *IEEE Trans Acoust Speech Signal Process*. 1981;29(6):1153–60.
- [24] Lehmann TM, Gönner C, Spitzer K. Survey: interpolation methods in medical image processing. *IEEE Trans Med Imaging*. 1999 Nov;18(11):1049–75.